From Asymptotic to Finite-Size Security in Decoy-State Quantum Key Distribution

by

Lars Kamin

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Physics (Quantum Information)

Waterloo, Ontario, Canada, 2026

© Lars Kamin 2026

Examining Committee Membership

The following served on the Examining Committee for this thesis. The decision of the Examining Committee is by majority vote.

External Examiner:	Frédéric Dupuis Associate Professor, Department of Computer Science, University of Montreal
Supervisor(s):	Norbert Lütkenhaus Professor, Institute for Quantum Computing and Department of Physics & Astronomy, University of Waterloo
Internal Member:	Kevin Resch Professor, Institute for Quantum Computing and Department of Physics & Astronomy, University of Waterloo
Internal-External Member:	Graeme Smith Associate Professor, Institute for Quantum Computing and Department of Applied Mathematics, University of Waterloo
Other Member(s):	Thomas Jennewein Professor, Department of Physics, Simon Fraser University and Institute for Quantum Computing and Department of Physics & Astronomy, University of Waterloo

Author's Declaration

This thesis consists of material all of which I authored or co-authored: see Statement of Contributions included in the thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Statement of Contributions

This thesis is based on the following articles of which I am the first or co-first author.

Contributions from Ref. [1]

I was the principal researcher and author of the manuscript. All data analysis, programming and figure generation were performed by me. Norbert Lütkenhaus provided conceptual guidance, continuous mentorship and critical feedback on multiple drafts of the manuscript.

Contributions from Ref. [2]

I contributed equally with Amir Arqand. My contributions were in the numerical algorithm, the extension to decoy-state protocols, especially the new tight decoy-state methods and the comparisons with other proof techniques. My co-authors mainly derived the theoretical foundations.

Contributions from Ref. [3]

I was the principal researcher and author of the manuscript, except Sections VIII – IX. Those were done by my co-author, Devashish Tupkary. All data analysis, programming and figure generation were performed by me. Norbert Lütkenhaus provided conceptual guidance and critical feedback on multiple drafts of the manuscript.

Contributions from Ref. [4]

I was the principal researcher and author of the manuscript. I was the principal investigator and lead author, responsible for the theoretical formulation, numerical implementation, data analysis, figure preparation, and manuscript writing. John Burniston implemented parts of the Frank–Wolfe algorithm, derived the required gradient expressions, and developed the initial perturbation methods later extended by me. Ernest Y.-Z. Tan provided conceptual guidance and critical feedback on multiple drafts of the manuscript, and developed the methods in appendices A and E of the manuscript.

Abstract

Quantum Key Distribution (QKD) promises information-theoretic security, yet bridging the gap between theoretical proofs and practical implementations, specifically those operating with finite resources and imperfect devices against general coherent attacks, remains a critical challenge. This thesis develops a spectrum of efficient security proof techniques within the composable security framework, calculating key rates for both fixed- and variable-length protocols while accounting for realistic imperfections.

We begin by addressing detection setups through an extension of a squashing map, the flag-state squasher, used for reducing the infinite-dimensional Hilbert spaces of optical elements to finite dimensions. This extension accommodates arbitrary passive linear optical setups while allowing for the inclusion of detection inefficiencies and dark counts in the security analysis.

Subsequently, we advance the analysis of decoy-state protocols and introduce two major improvements. First, we reformulate the decoy-state analysis to recover no-decoy key rates, tightening the optimization. Second, we derive a unified framework that performs the key rate optimization and decoy analysis in a single step. This enables the bounding of the relevant entropies with arbitrary precision in the finite-size regime and successfully recovers the Devetak-Winter formula in the asymptotic limit.

Furthermore, we improve the security analysis for generic QKD protocols against independent and identically distributed (IID) collective attacks. Our refined analysis yields finite-size corrections proportional to detected rather than transmitted signals and, by developing sharper concentration inequalities, achieves significantly improved finite-size scaling.

Finally, leveraging the marginal constrained entropy accumulation theorem (MEAT), we establish a flexible numerical Rényi security framework against coherent attacks for both fixed- and variable-length protocols. This approach consistently outperforms existing reference proof techniques, including those based on entropic uncertainty relations, providing significantly higher key rates for both qubit and practically relevant decoy-state protocols. Moreover, we present finite-size key rates for generic QKD protocols accounting for realistic intensity and phase imperfections. Overall, this thesis provides the necessary theoretical framework to bridge the gap between idealized models and experimental reality, offering a scalable path toward secure quantum communication under realistic conditions, as demonstrated by the application of these techniques in experimental collaborations.

Acknowledgements

First of all, I would like to thank my supervisor, Professor Norbert Lütkenhaus, for the opportunity to carry out my doctoral research at the Institute for Quantum Computing at the University of Waterloo and for his continuous support throughout my PhD. I also appreciate the stimulating research opportunities and resources he has cultivated within the OQCT group, which has been a great home for my academic development.

I would like to express my gratitude to Prof. Ernest Y.-Z. Tan. His scientific guidance and deep knowledge were indispensable to my studies. I am deeply thankful for his time and for the many insightful discussions that shaped the direction of my research.

Moreover, my thanks go to the members of the OQCT research group. I especially thank Amir Arqand, John Burniston, Florian Kanitschar, Aodhan Corrigan, Zhiyao Wang, Jerome Wiesemann, and Shlok Nahar for the inspiring discussions and supportive atmosphere we shared. I also acknowledge my co-authors: Prof. Norbert Lütkenhaus, Prof. Ernest Y.-Z. Tan, Amir Arqand, John Burniston, Ian George, and Devashish Tupkary.

Additionally, I would like to thank Aodhan Corrigan, Zhiyao Wang, Jerome Wiesemann, and Amir Arqand for their feedback on early drafts of this thesis.

I would also like to thank my advisory committee and examiners: Prof. Michele Mosca, Prof. Thomas Jennewein, Prof. Kevin Resch, Prof. Graeme Smith, Prof. Frédéric Dupuis, and Prof. Norbert Lütkenhaus.

Finally, I would like to thank my parents, my friends, and especially my partner, Jacqueline Powichrowski, for their love, encouragement, and support.

Table of Contents

Examining Committee Membership ii

Author's Declaration iii

Statement of Contributions iv

Abstract v

Acknowledgements vi

List of Figures xiii

List of Tables xxi

1 Introduction 1

1.1 Outline of the Thesis 4

2 Background 7

2.1 Technical Definitions and Notation 7

2.2 Distance Measures and Norms 13

2.3 Entropic Quantities 14

2.3.1 Min- and Smooth Min-Entropy 15

2.3.2 Rényi Entropies 17

2.4 Concepts from Convex Optimization 18

3 Quantum Key Distribution Background **21**
3.1 Security Definitions 21
3.2 Protocol Description 24
3.2.1 Decoy-State Protocols 28
3.3 Source Replacement Scheme 29
3.4 Shield Systems 30
3.5 Leftover Hashing Lemmas 31

4 Squashing Maps **36**
4.1 Preliminaries 36
4.2 Qubit Squasher 37
4.3 Flag-State Squasher 38

5 Flag-State Squasher for Passive Linear Optical Detection Setups **41**
5.1 General Structure of Subspace Estimation 41
5.2 Defining a Suitable POVM Element for Passive Linear Optical Setups 43
5.2.1 Motivating Example: Perfect Passive 4-State Analyzer 47
5.3 Subspace Estimation for Arbitrary Passive Linear Optical Setups . . . 50
5.3.1 Including Detection Losses 54
5.3.2 Including Dark Counts 64
5.4 Discussion of Previous Results 65
5.5 Detector Characterization for Flag-State Squasher 66
5.6 Conclusion 68

6 Asymptotic Key Rates with Decoy States **69**
6.1 Standard Decoy Methods 70
6.2 Improved Decoy Methods 73
6.3 Extension of Asymptotic Key Rate with Differing Intensities 75

6.4 Key Rate Optimization Problem including Flag-state Squashing 78
6.5 Resulting Subspace Bounds for Decoy-State Methods 78
6.6 Example: Biased Passive WCP 6-State 79
6.6.1 Flag-State Squasher 80
6.6.2 Decoy Analysis 82
6.6.3 Final Key Rate Optimization 82
6.6.4 Results with Equal Intensities 83
6.6.5 Results with Differing Intensities 84
6.7 Conclusion 88

7 Finite-Size Security for Prepare-and-Measure Protocols under IID Attacks 89

7.1 Improved Prepare-and-Measure Finite-size 89
7.1.1 Acceptance Test and Resulting Entropy Bounds 90
7.1.2 Security Proof against IID Attacks for Generic P&M Protocols 98
7.2 Decoy-State Methods in Finite-Size Regime 102
7.2.1 Source Replacement Scheme 102
7.2.2 Asymptotic Scenario for Decoy-State Methods 104
7.2.3 Finite-Size Scenario for Decoy-State Methods 104
7.3 Optimizing Security Parameters 111
7.4 Optimizing Expected Key Rates for Fixed Length Protocols 113
7.5 Examples for Collective Attacks 115
7.5.1 Qubit BB84 Protocol with Loss 116
7.5.2 Decoy BB84 Protocol 118
7.6 Variational Bounds for Small Security Parameters 119
7.7 Conclusion 122

8 Entropy Accumulation Theorem-based QKD Proofs **124**
8.1 Background on EAT-based Proofs 124
8.2 Formalizing QKD protocols in MEAT 126
8.2.1 Source Replacement Scheme 126
8.2.2 Announcements and Conditioning on Test and Generation rounds 128
8.2.3 Defining the QKD channel 129
8.2.4 Fixed-Length Security 132
8.3 Variable-Length Protocols in MEAT 135
8.3.1 Properties of Variable-Length Key Rates 139
8.3.2 Optimality of the Tradeoff Function 140
8.4 Prepare-and-Measure Protocols in GEAT 141

9 Security of Prepare-and-Measure Protocols via GEAT **147**
9.1 Qubit BB84 with Loss 148
9.2 Decoy BB84 Protocol with Improved Analysis 150
9.3 Conclusion 152

10 Rényi Security Framework **154**
10.1 Calculating Versions of $h_\alpha^{\uparrow,\mathrm{QKD}}$ 154
10.2 Computing the Normalization Constant κ 162
10.2.1 Reformulating the Optimization Problem for κ 166
10.3 Squashing Maps 168
10.4 Source Maps 172
10.5 Qubit BB84 and Comparisons with other Proof Techniques 175
10.6 Extension to Block-Diagonal States 179
10.6.1 Bounding the Entropy 181
10.6.2 Bounding the Constraints 182
10.6.3 Reformulated Optimization Problems 184

10.7 Decoy-state Protocols . . . 185
10.7.1 Simplifications of the Optimization Problems . . . 186
10.7.2 Tightness of Block-Diagonal and Decoy Methods . . . 190
10.7.3 Example: Active Decoy BB84 . . . 191
10.8 Device Imperfections and Implications . . . 198
10.9 Intensity Imperfections . . . 200
10.9.1 Example: Biased Passive BB84 with Intensity Imperfections . . . 205
10.10 Phase Imperfections . . . 208
10.10.1 Definition of Block-Tagged States . . . 210
10.10.2 Approximate Diagonalization with Cutoff . . . 212
10.10.3 Key Rate Construction . . . 215
10.10.4 Example: 4-6 Protocol with Phase Imperfections . . . 218
10.11 Conclusion . . . 221

11 Conclusion and Outlook **223**

References **226**

A Additional Proofs for Rényi Framework **236**
A.1 Convexity Proofs . . . 236
A.1.1 Convexity of Quantities related to Petz Conditional Rényi entropy 237
A.1.2 Perturbed Objective Function . . . 238
A.1.3 Joint Convexity of h_α^{QKD} . . . 241
A.1.4 Joint Convexity of Normalization Constants $\kappa^{\downarrow}$ and $\kappa^{\supset}$ . . . 245
A.2 Proofs for Phase Imperfections . . . 248

B Code Availability **257**

C Kraus Operators for Various Example Protocols **258**

C.1 Kraus Operators for Example Protocols in Asymptotic Regime 258

C.1.1 Kraus Operators BB84 Protocol 258

C.1.2 POVM Elements and Kraus Operators Six-State Protocol . . . 259

C.2 Kraus Operators for Example Protocols in IID Regime 261

C.2.1 Qubit BB84 261

C.2.2 Decoy BB84 263

C.3 Kraus Operators and States for Example Protocols in Rényi Security Framework 264

C.3.1 Conditional POVM Construction 264

C.3.2 Qubit BB84 268

C.3.3 Active Decoy BB84 269

C.3.4 Passive Decoy BB84 269

C.3.5 Passive Decoy 4-6 Protocol 272

List of Figures

4.1 Schematic representation of the qubit squasher from Ref. [62, 63]. (Figure adapted from [63]) 37

5.1 A detection setup using only passive linear optics, including *losses*, represented as a circuit implementing a linear mode transformation U. For a typical 4-state analyzer used in optical BB84 protocols, $N_{\text{in}} = 2$ represents the incoming signal modes, while $N_{\text{out}} = 8$ accounts for the four detectors and their corresponding loss modes. . . 43

5.2 A perfect detection setup using only passive linear optics represented as a circuit implementing a linear mode transformation U. (Figure 5 from Ref. [1]) 45

5.3 Schematic representation of the decomposition of Lemma 5.3.2 of the unitary transformation U of a detection setup into two consecutive ones, W_{tot} and $(T_d \oplus T_l)$. The unitary W_{tot} does not mix the signal modes in system Sig with the modes in system Env_1 which thus always remain in the vacuum state. This is indicated by lines continuing through W_{tot}. 55

6.1 Example of a perfect passive detection setup for a six-state analyser for polarization encoding. We label beamsplitters with blue lines, polarizing beamsplitters with additional blue boxes, polarization rotators with green boxes and detectors with yellow paraboloids. The first two beamsplitters divert the incoming light such that with probabilities $p_{\mathrm{HV}}, p_{\mathrm{AD}}$ and p_{RL} the polarization is measured in the horizontal/vertical, anti-diagonal/diagonal and right-/left circular basis after appropriate rotation by the polarization rotators. Finally, the polarizing beamsplitters split the light with equal probability in each polarization basis, such that e.g. horizontal polarized light is measured in the b_1 mode. The submatrix G includes all elements in the red dashed box to a semi-unitary. While this example illustrates an ideal setup, imperfections such as biased splitting ratios or misalignment can be naturally captured within this framework simply by adjusting the corresponding matrix elements of the semi-unitary G. (Figure 6 from Ref. [1]) 80

6.2 Comparison of secret key rates of the passive unbiased six-state protocol over the distance between numerical results for three intensities using the flag-state squasher with $N_B = 1$ presented in this work (orange diamonds) and results from [62, 63] (green crosses). (adapted Figure 7 from Ref. [1]) 84

6.3 Comparison of secret key rates of the passive biased ($p_z = 0.8$, $p_x = p_y = 0.1$) six-state protocol over the distance between our numerical results for one decoy intensity (orange stars) with numerical results from Ref. [71] using two decoy intensities (green circles). (adapted Figure 8 from Ref. [1]) 85

6.4 Comparison of secret key rates of the 6-state protocol without decoy intensities over distance between our numerical results with $N_B = 1$ (orange circles), $N_B = 8$ (purple circles), numerical results from [23] (yellow triangles) and analytical results from [10] (blue line), demonstrating that our numerical methods successfully recover the known analytical scaling. (adapted Figure 9 from Ref. [1]) 86

6.5 Comparison of secret key rates of the 6-state protocol with one decoy intensity over the distance between the rates using a bit bias (red diamonds) and no bit bias (green crosses). (Figure 10 from Ref. [1]) 87

6.6 Comparison of secret key rates *per single photon* of the 6-state protocol with one decoy intensity over the distance between the rates using a bit bias (red diamonds) and no bit bias (green crosses). (Figure 11 from Ref. [1]) 88

7.1 One dimensional representation of the construction of the feasible set from the expected frequency $\bar{F}$ with equal upper and lower ν. (Figure 1 from Ref. [3]) 92

7.2 Plot of upper, lower and independent bounds on the expected key rate of the *fixed-length* qubit BB84 protocol for *IID attacks* with a channel loss of 10 dB loss and a total number of signals of $n = 10^9$ over the tolerated fluctuation parameter with $t_k = t_{\text{sift}} \equiv t$. (Figure 2 from Ref. [3]) 115

7.3 Secret key rate for *IID attacks* of the *fixed-length* qubit BB84 protocol against channel loss for varying number of total signals sent $n = 10^4, \ldots, 10^9$. (Figure 3 from Ref. [3]) 117

7.4 Comparison of the *fixed-length* qubit BB84 secret key rates for *IID attacks* between 1-norm constraints of [43] and our entrywise constraints of Theorem 7.1.1. The key rate is plotted against channel loss with a varying number of total signals sent n. (Figure 4 from Ref. [3]) 118

7.5 Secret key rate for *IID attacks* of the *fixed-length* decoy-state BB84 protocol using two decoy intensities against channel loss with a varying number of total signals sent $n = 10^6, \ldots, 10^{11}$. (Figure 5 from Ref. [3]) 120

7.6 Optimal signal intensities for *IID attacks* of the *fixed-length* decoy-state BB84 protocol using two decoy intensities against channel loss with a varying number of total signals sent $n = 10^6, \ldots, 10^{11}$. . . . 121

8.1 Representation of the tensor product structure of the MEAT theorem [41, Corollary 4.1]. (Figure 1 from Ref. [4]) 127

8.2 Circuit representation of the concatenated isometries from Definition 8.2.1 defining the QKD channel $\mathcal{M}_{\mathrm{QKD}}$. The channel maps the bipartite input state ρ_{AB} to the final classical registers I, $\widehat{C}$, and S by applying the isometries $V_{\mathtt{test}}$, V_{meas}, $V_{\mathcal{G}}$, V_{ϕ}, and $V_{\mathcal{Z}}$ and tracing out the complementary systems $(ABXY\tilde{T}FS_Q)$. The isometry $V_{\mathtt{test}}$ applies the projectors $\sqrt{\Pi^x}$ and creates the value x in register F. Note that the isometries $V_{\mathrm{meas}} \otimes |t\rangle\langle t|_F$ and $V_{\mathcal{G}} \otimes |g\rangle\langle g|_F$ are depicted sequentially for clarity only. They act as a single block but commute due to their mutually orthogonal support on the flag register F. 130

9.1 Key rate comparison for the qubit BB84 protocol with loss, between the GEAT analysis (solid lines), and the IID setting (dashed lines). In all cases, the signals are depolarized with $p_{\mathrm{hon}}^{\mathrm{depol}} = 0.01$, the security parameter is $\varepsilon_{\mathrm{sec}} = 10^{-8}$, and for simplicity we show only the results for unique acceptance. The testing probability γ and the Rényi parameter α are optimized for the EAT analysis, and the former was also optimized in the IID analysis. (Figure 1 from Ref. [2]) 149

9.2 Key rate comparison for the qubit BB84 protocol with loss for a realistic acceptance set, between the GEAT analysis (solid lines), and the EUR-based analysis in [24] incorporating improvements from [83] (dashed lines). In all cases, the signals are depolarized with $p_{\mathrm{hon}}^{\mathrm{depol}} = 0.01$ and the security and completeness parameters are $\varepsilon_{\mathrm{sec}} = 10^{-8}$ and $\varepsilon^{\mathrm{com}} = 10^{-3}$, respectively. The testing probability γ and the Rényi parameter α are again optimized for the EAT analysis, and the former was also optimized in the EUR analysis. The dashed lines based on the EUR results from Ref. [24] show the expected key rates. (Bottom of Figure 2 from Ref. [2]) 150

9.3 Key rate comparison for the decoy-state protocol, between the GEAT analysis in this work (solid lines), and the IID setting (dashed lines). In all cases, fully phase-randomized weak coherent pulses with intensities $\{\mu_{\text{sig}} = 0.9, \mu_2 = 2 \times 10^{-2}, \mu_3 = 10^{-3}\}$ were used. Signals are misaligned with $\theta_{\text{hon}}^{\text{misalign}} = \text{asin}(0.1)$, the photon number cutoff is $N_{\text{ph}} = 10$, the security parameter is $\varepsilon_{\text{sec}} = 10^{-8}$, and for simplicity we show only the results for unique acceptance. The testing probability γ and the Rényi parameter α are optimized for the EAT analysis, and the former was also optimized for the IID analysis. (Figure 3 from Ref. [2]) 152

9.4 Key rate comparison for the decoy-state protocol for a realistic acceptance set, between the GEAT analysis (solid lines), and the EUR-based analysis in [24] incorporating improvements from [83] (dashed lines). In all cases, two decoy signals are used with intensities $\mu_2 = 2 \times 10^{-2}$ and $\mu_3 = 10^{-3}$, while the actual signal has the intensity of $\mu_{\text{sig}} = 0.9$, and the signals are chosen uniformly at random in each round. Signals are misaligned with $\theta_{\text{hon}}^{\text{misalign}} = \text{asin}(0.1)$, the photon number cutoff is $N_{\text{ph}} = 10$, the security parameter is chosen to be $\varepsilon_{\text{sec}} = 10^{-8}$, and the completeness parameter for the realistic tolerance intervals is $\varepsilon^{\text{com}} = 10^{-3}$. The testing probability γ and the Rényi parameter α are optimized for each data point. We find that for this specific protocol, our approach appears to perform better at low loss values, while the EUR-based approach appears to perform better at high loss values. The dashed lines based on the EUR results from Ref. [24] show the expected key rates. (Bottom of Figure 4 from Ref. [2]) . . . 153

10.1 Schematic representation of the Rényi entropy hierarchy. An arrow indicates that the source entropy upper-bounds the target. The question mark signifies an indeterminate ordering. 157

10.2 Results for variable-length secret key rates of the qubit BB84 protocol plotted against the channel loss in dB and a constant depolarization with $p_{\text{hon}}^{\text{depol}} = 0.03$ with a varying number of total signals sent $n = 10^5, \ldots, 10^{10}$. The security parameters, and error correction efficiency were chosen as $\varepsilon_{\text{PA}} = \varepsilon_{\text{EV}} = \frac{1}{2}10^{-80}$ and $f_{\text{EC}} = 1.1$. As a comparison, we showcase results using the EUR approach of [24] incorporating the improvements from [83]. For both approaches, we optimize the probability γ of using the X-basis in each round, and for our work we additionally optimize the Rényi parameter α. (Figure 2 from Ref. [4]) 177

10.3 Comparison of the variable-length secret key rates for the qubit BB84 protocol. We compare our approach against the EUR [24] and postselection technique [27], plotted against channel loss in dB for a constant depolarization of $p_{\text{hon}}^{\text{depol}} = 0.03$. The lines with circles indicate the key rates for $n = 10^{10}$ signals sent, while the lines with triangles show the minimum total signals sent required to achieve a non-zero key rate. As before, the security parameters and error correction efficiency were chosen as $\varepsilon_{\text{PA}} = \varepsilon_{\text{EV}} = \frac{1}{2}10^{-80}$ and $f_{\text{EC}} = 1.1$. For all approaches, we optimize the testing probability γ of using the X-basis in each round; for our work, we additionally optimize the Rényi parameter α. 178

10.4 Results for variable-length secret key rates of the active decoy BB84 protocol with *two* decoy intensities plotted against the channel loss in dB with a constant misalignment angle of $\theta_{\text{hon}}^{\text{misalign}} = 0.03$ for a varying number of total signals sent $n = 10^5, \ldots, 10^{11}$. The security parameters, and error correction efficiency were chosen as $\varepsilon_{\text{PA}} = \varepsilon_{\text{EV}} = \frac{1}{2}10^{-80}$ and $f_{\text{EC}} = 1.1$. As a comparison, we showcase results using the EUR approach of [24] incorporating the improvements from [83]. For both approaches, we optimize the probability γ of using the X-basis in each round and the signal intensity μ_s. Additionally, for our work we optimize the Rényi parameter α. (Figure 3 from Ref. [4]) 194

10.5 Optimal signal intensities for of the variable-length active decoy BB84 protocol with *two* decoy intensities plotted against the channel loss in dB with a constant misalignment angle of $\theta_{\text{hon}}^{\text{misalign}} = 0.03$ for a varying number of total signals sent $n = 10^5, \ldots, 10^{11}$. The security parameters and error correction efficiency were chosen as $\varepsilon_{\text{PA}} = \varepsilon_{\text{EV}} = \frac{1}{2}10^{-80}$ and $f_{\text{EC}} = 1.1$. 195

10.6 Results for variable-length secret key rates of the active decoy BB84 protocol with *one* decoy intensity plotted against the channel loss in dB with a constant misalignment angle of $\theta_{\text{hon}}^{\text{misalign}} = 0.03$ for a varying number of total signals sent $n = 10^5, \ldots, 10^{11}$. The security parameters, and error correction efficiency were chosen as $\varepsilon_{\text{PA}} = \varepsilon_{\text{EV}} = \frac{1}{2}10^{-80}$ and $f_{\text{EC}} = 1.1$. As a comparison, we showcase results using the EUR approach of Refs. [23, 25]. For both approaches, we optimize probability γ of using the X-basis in each round and the signal intensity μ_s. Additionally, for our work we optimize the Rényi parameter α. (Figure 4 from Ref. [4]) 196

10.7 Comparison of the variable-length secret key rates for the decoy-state BB84 protocol with *two* decoy intensities. We compare our approach against the EUR [24] and postselection techniques [3], plotted against channel loss in dB for a constant misalignment angle of $\theta_{\text{hon}}^{\text{misalign}} = 0.03$. Lines with circles indicate the key rates for $n = 10^{11}$ signals sent, while lines with triangles show the minimum total signals required to achieve a non-zero key rate. As before, the security parameters and error correction efficiency are set to $\varepsilon_{\text{PA}} = \varepsilon_{\text{EV}} = \frac{1}{2}10^{-80}$ and $f_{\text{EC}} = 1.1$. For all approaches, we optimize the testing probability γ (using the X-basis) and the signal intensity μ_s; for our work, we additionally optimize the Rényi parameter α. . . 197

10.8 Results for variable-length secret key rates of the passive decoy BB84 protocol with *two* decoy intensities and intensity imperfection plotted against the channel loss in dB with a varying number of total signals sent $n = 10^6, 10^8, 10^{10}$. The security parameters, and error correction efficiency were chosen as $\varepsilon_{\text{PA}} = \varepsilon_{\text{EV}} = \frac{1}{2}10^{-80}$ and $f_{\text{EC}} = 1.1$. (Figure 5 from Ref. [4]) 207

10.9 Results for variable-length secret key rates of the passive decoy BB84 protocol with *one* decoy intensity and intensity imperfection plotted against the channel loss in dB with a varying number of total signals sent $n = 10^6, 10^8, 10^{10}$. The security parameters, and error correction efficiency were chosen as $\varepsilon_{\text{PA}} = \varepsilon_{\text{EV}} = \frac{1}{2}10^{-80}$ and $f_{\text{EC}} = 1.1$. (Figure 6 from Ref. [4]) 209

10.10 Results for variable-length secret key rates of the passive decoy-state 4-6 protocol with *one* decoy intensity and phase imperfections plotted against the channel loss in dB with a varying number of total signals sent $n = 10^6, 10^8, 10^{10}$ and different quality factors $q = 1$ and $q = 0.99$. The security parameters, and error correction efficiency were chosen as $\varepsilon_{\text{PA}} = \varepsilon_{\text{EV}} = \frac{1}{2}10^{-80}$ and $f_{\text{EC}} = 1.1$. As a comparison, we showcase results from [3], which applies the postselection technique [27], for $n = 10^{10}, 10^{12}, 10^{14}$ marked with (PS). For our work we optimize the Rényi parameter α. (Figure 7 from Ref. [4]) 220

List of Tables

2.1 List of notations. 8

Chapter 1

Introduction

A quantum key distribution (QKD) protocol aims to establish an information-theoretically secure key between two parties, typically Alice and Bob, in the presence of an adversarial eavesdropper, Eve, by exploiting the fundamental laws of quantum mechanics. The protocol operates with Alice and Bob connected via an insecure quantum channel and an authenticated classical channel. Eve is permitted to interfere with the quantum signals in any way allowed by quantum mechanics, and, crucially, the adversary is not restricted by any computational assumptions, unlike in classical cryptography. However, Eve cannot alter or impersonate messages exchanged through the authenticated classical channel.

QKD is special because its security is based on fundamental laws of nature. The no-cloning theorem [5] states that one cannot perfectly copy unknown non-orthogonal quantum states, thereby preventing an adversary from making copies of the transmitted quantum signals for later access or decoding. This theorem captures the essence of QKD's security claim. Furthermore, QKD is considered forward secure in the sense that a key generated during a QKD protocol remains as secure as it was when established, as an adversary can only successfully attack the key during its creation phase.

The goal of a QKD security proof is to demonstrate that the final shared key is secure regardless of Eve's attack strategy. It is essential that the security analysis accounts for *all* possible attacks Eve could perform in the channel. This is challenging because it must address the most general forms of attacks, often referred to as coherent attacks. Some earlier QKD works focused only on specific, less general classes of attacks, such as those where Eve attacks the transmitted states in an independent and identically distributed (IID) manner across the rounds, known as IID collective attacks. Such restricted attacks generally do not capture the full scope of Eve's capabilities. For safe deployment of QKD, security must

be proven against the entire class of coherent attacks, rather than being restricted to IID collective attacks. Moreover, any physical implementation of QKD involves only a finite number of rounds, requiring that the security proof account for finite-size effects compared to the asymptotic limit of infinitely many signals.

The first QKD protocol, now known as the BB84 protocol, was developed in 1984 by Bennett and Brassard [6]. This was followed a few years later by the E91 protocol, presented by Ekert in 1991 [7]. Both relied on perfect qubit sources, which are still challenging to implement practically. It is far more common to use weak coherent pulses (WCP). However, WCP sources inevitably emit multi-photon signals, making them susceptible to the photon-number-splitting (PNS) attack [8, 9]. In a PNS attack, the adversary splits off one photon from every multi-photon signal, measures it, and forwards the remaining part. Consequently, Eve gains full information about any multi-photon signal, necessitating a significant reduction in the final secure key length to maintain security [10].

To counter the PNS attack, decoy-state protocols were subsequently developed [11–14]. These protocols still use fully phase-randomized WCP sources but transmit signals using multiple decoy intensities. By collecting additional data from signals transmitted with different intensities, decoy-state protocols aim to estimate Eve's influence on the single-photon signals, as these are inherently secure even under a PNS attack. Decoy-state protocols remain the most practically relevant QKD protocols today.

To prove the security of QKD protocols, various methods have been developed, including:

- Phase error correction-based proofs [15–19].
- Entropic uncertainty relations (EURs) [20] and their applications to QKD [21–25].
- The postselection technique [26, 27].
- The entropy accumulation theorem (EAT) and subsequent variants [28–31] with applications to QKD [2, 32–34].

Although achieved through very different means, all these security proofs ultimately establish a similar bound on the *achievable* secure key length:

$$l \geq n\underline{h} - \text{corrections}, \tag{1.0.1}$$

where l is the secure key length, n is the number of signals sent, $\underline{h}$ is a lower bound on some entropy of the state generated by the QKD protocol, and the corrections informally represent the combined finite-size penalties and error correction costs.

A key tradeoff exists between the finite-size key rates and the flexibility of the proof techniques. For instance, while the phase error correction and EUR techniques generally offer good finite-size performance, they often require the protocol analysis to be reducible to a scenario close to the original qubit BB84 protocol [35], thereby limiting their applicability. Conversely, the postselection technique is highly flexible, applying to a wide range of protocols, but typically exhibits poor finite-size performance. Versions of EAT developed in [28–31] offer better finite-size key rates than the postselection technique [2, 32] while retaining some flexibility, but the rates are often still worse than those from phase error correction or EURs. Furthermore, these EAT versions typically impose technical restrictions on the protocol, such as repetition rate restrictions.

The connection between secrecy (the output being close in trace distance to a perfectly correlated uniform shared key) and entropy is often established by the leftover hashing lemma (LHL) [36]. The LHL guarantees a small trace distance if the chosen key length l is smaller than the smooth min-entropy of the global state before the privacy amplification step. Traditionally, the LHL for smooth min-entropies [36] is used, requiring lower bounds on the n-round smooth min-entropy of the protocol's output state, as seen in works like [22, 36, 37]. Recently, an alternative LHL based on Rényi entropies was presented in Ref. [38].

Despite these theoretical advancements, bridging the gap to practical implementations remains a critical challenge. Current QKD setups operate with finite resources and imperfect devices. Furthermore, questions originating from recent experimental collaborations, such as the Reference Frame Independent Quantum Key Distribution (ReFQ) project, which is part of the Quantum EncrYption and Science Satellite (QEYSSat) mission [39], and the Drone- and Vehicle-Based QKD project [40], have highlighted the critical need to model a combination of realistic hardware imperfections simultaneously, rather than in isolation. Most notably, both projects required the correct modelling of their detection setups and intensity fluctuations.

In this thesis, we address these challenges by developing a robust numerical framework that fills critical gaps in the security analysis of practical QKD protocols. While we provide foundational refinements to the analysis of IID collective attacks, our primary achievement is establishing a highly efficient security proof against coherent attacks. To directly answer the demands of experimental implementations, we first address realistic detection setups by extending the flag-state squasher. This extension rigorously incorporates detection inefficiencies and dark counts, yielding a generic method applicable to arbitrary linear optical detection setups, including commonly used examples of passive 4-state analyzers. Furthermore, we develop tight reformulations of decoy-state protocols that can significantly reduce experimental complexity and recover several limiting cases that were previously unattainable.

Leveraging the marginal constrained entropy accumulation theorem (MEAT) [41], we unify these tools to establish a generic numerical framework capable of evaluating finite-size secure key rates for arbitrary QKD protocols. Crucially, this framework allows us to simultaneously account for a combination of realistic hardware imperfections, with a specific focus on the phase and intensity fluctuations observed in practical sources. For phase imperfections, this constitutes the first finite-size security proof to date; for intensity fluctuations, we demonstrate that our approach can tolerate substantial margins of error that current experiments can already comfortably achieve. Remarkably, even while capable of accommodating these imperfections, our approach yields the highest finite-size key rates to date, significantly outperforming existing reference techniques.

1.1 Outline of the Thesis

Chapters 2 to 4: Background, Quantum Key Distribution Background and Squashing Maps
In these chapters, we lay out the required concepts and notations of the thesis. Furthermore, we introduce QKD formally and describe squashing methods.

Chapter 5: Flag-State Squasher for Passive Linear Optical Detection Setups
In this chapter, we build on a general squashing method, the so-called flag-state squasher from Ref. [42] which is often used to reduce optical detection setups described in infinite dimensions to finite dimensions. This squashing method, however, requires a bound on the weight inside some untouched subspace. This subspace typically corresponds to the $\leq n$-photon subspace and in this case the weight corresponds to the probability of less than or equal to n photons arriving in Bob's detector.

The critical part is finding this weight $p(\leq n)$ purely based on Bob's observations, m_{obs}, and properties of his detection setup:

$$p(\leq n) \geq 1 - \frac{m_{\text{obs}}}{c_n}, \tag{1.1.1}$$

where c_n is a constant depending only on Bob's detection setup. We derive such a bound for arbitrary passive linear optical detection setups, providing a rigorous path to incorporate detection inefficiencies and dark counts into security proofs. This generalization was particularly motivated by the requirements of current experimental collaborations, including the Drone- and Vehicle-Based QKD project [40] and the ReFQ project.

Chapter 6: Asymptotic Key Rates with Decoy States
Previously, an unresolved problem of the decoy-state analysis was that it could not recover the results from Ref. [10] and therefore, was not tight in both the resulting key rates and the number of intensities used. Moreover, even the tighter decoy-state methods presented in [23], which use additional information on the vacuum error rate, do not recover the key rate of a single weak coherent pulse completely, although they are already an important step towards it.

To resolve this gap, in this chapter we will present a refined decoy-state analysis which fully recovers the no-decoy WCP rate of Ref. [10].

Chapter 7: Finite-Size Security for Prepare-and-Measure Protocols under IID Attacks
We derive a refined security analysis for IID collective attacks. With previous results such as [43], the IID key length would scale as

$$l \leq nh_{\mathrm{vN}} - \sqrt{n} \cdot \text{const.} - \text{corrections}, \tag{1.1.2}$$

where h_{vN} is a lower bound on the von-Neumann entropy and corrections informally represent the remaining finite-size penalties and error correction costs. Most importantly, the correction is proportional to $\sqrt{n}$ which scales with the total number of signals sent, whereas the von-Neumann entropy scales with the total number of detected signals. We remove this discrepancy and prove a key length expression that scales with the detected signals in all terms.

Additionally, we present a refined statistical analysis for the so-called feasible set required to find the lower bound h_{vN} on the von-Neumann entropy.

Chapter 8: Entropy Accumulation Theorem-based QKD Proofs
In this chapter, we introduce the marginal constrained entropy accumulation theorem (MEAT) and show how a QKD protocol can be formalized in this framework. The MEAT relies on Rényi entropies, a generalization of the von-Neumann entropy, and proves a lower bound on the entropy of the form

$$\widetilde{H}_{\alpha}^{\uparrow}(S_1^n|EI_1^n)_{\rho|\mathrm{accept}} \geq h_{\alpha}^{\uparrow,\mathrm{QKD}} - \text{corrections}. \tag{1.1.3}$$

Effectively, $h_{\alpha}^{\uparrow,\mathrm{QKD}}$ corresponds to the final key length and the MEAT has been shown to provide very tight key rates, [41]. Applying the MEAT, we prove finite-size security against coherent attacks for fixed and variable-length protocols. Additionally, we draw a

connection to previous results with the generalized entropy accumulation theorem (GEAT).

Chapter 9: Security of Prepare-and-Measure Protocols via GEAT
In this chapter, we simply present the results from Ref. [2] based on the GEAT as they will achieve strictly lower key rates than those based on the MEAT.

Chapter 10: Rényi Security Framework
In this chapter, we derive a framework that allows for an efficient computation of Rényi entropies, especially in the context of the MEAT, which is applicable to both fixed- and variable-length protocols. Furthermore, we extend this to decoy-state protocols, proving a tight reformulation which achieves the Devetak-Winter formula in the asymptotic limit. Moreover, we extend our finding to two examples of source imperfections of specific interest. We show that one can still achieve high finite-size key rates in the presence of a fluctuating laser source for decoy-state protocols and present the first finite-size security proof against coherent attacks for decoy-state protocols with WCP sources that are not fully phase-randomized.

These results directly address the specific imperfections observed in experiments like the ReFQ project, demonstrating our framework's capacity to provide rigorous security proofs for practical implementations. Additionally, our methods greatly simplify the calculation of variable-length key rates by completing the most demanding computational steps before the start of the protocol.

Chapter 11: Conclusion and Outlook
Finally, we draw a conclusion by presenting open questions and present possible directions for future extensions of the frameworks introduced in this thesis.

Chapter 2

Background

2.1 Technical Definitions and Notation

In this section, we review several key concepts required for security proofs in quantum key distribution. We start with registers and operator spaces and finish this section with several properties of relevant entropic quantities.

Informally, many tasks in quantum information concern the information content stored in physical objects. These physical objects store the information using a set of symbols, often called an alphabet. Registers are the mathematical abstraction of such devices. In the following definition we make this abstraction precise.

Definition 2.1.1 (Register and Alphabet [44, Definition 2.1])**.** An *alphabet* Σ is a finite non-empty set. A *register* X is either one of the following two objects:

1. An alphabet Σ,
2. An n-tuple $X = (Y_1, ..., Y_n)$, where n is a positive integer and $Y_1, \ldots, Y_n$ are registers.

To any register X with an associated alphabet Σ, we associate a complex Hilbert space $\mathbb{C}^{\Sigma}$ of dimension $|\Sigma|$.

Under this definition, one can view registers in terms of alphabets, or conversely, view alphabets as the state space of a register. For our purposes, it is more natural to define registers primarily and refer to their associated alphabets. We typically denote the alphabet

Symbol	*Definition*
log	Base-2 logarithm
H	Base-2 von Neumann entropy
$\overline{H}_\alpha^\uparrow, \overline{H}_\alpha^\downarrow, \widetilde{H}_\alpha^\uparrow, \widetilde{H}_\alpha^\downarrow$	Base-2 Rényi entropies
$\lfloor\cdot\rfloor$ (resp. $\lceil\cdot\rceil$)	Floor (resp. ceiling) function
$\|\cdot\|_p$	Schatten p-norm
supp	Support of an operator or a probability distribution
$\mathcal{L}(A)$	Set of linear operators on register A
$\mathrm{Herm}(A)$	Set of Hermitian operators on register A
$\mathrm{Pos}(A)$	Set of positive semidefinite operators on register A
$\mathrm{S}_=(A)$ (resp. $\mathrm{S}_\leq(A)$)	Set of normalized (resp. subnormalized) states on register A
A_1^n	Abbreviated notation for registers $A_1 \dots A_n$
$\vec{A}$	Abbreviated notation for registers $A_1 \dots A_n$ generated from an independent and identically distributed process

Table 2.1: List of notations.

of a register with the same letter in a calligraphic font; for example, a register Z possesses the associated alphabet $\mathcal{Z}$.

Using this notation, on registers $Z_1, \dots, Z_n$ (in short Z_1^n), we can define frequency distributions in the following sense.

Definition 2.1.2 (Frequency distributions)**.** Let $Z_1, \dots Z_n$ be registers with a common alphabet $\mathcal{Z}$. Then, for a string $z_1^n \in \mathcal{Z}^n$, $\mathrm{freq}_{z_1^n}$ denotes the following frequency distribution on $\mathcal{Z}$:

$$\mathrm{freq}_{z_1^n}(z) \coloneqq \frac{\text{number of occurrences of } z \text{ in } z_1^n}{n}. \tag{2.1.1}$$

Next, we define the operator norm and bounded operators.

Definition 2.1.3 (Linear Operators, Operator Norm and Bounded Operators)**.** Let $\mathcal{H}$ be a Hilbert space and $A : \mathcal{H} \to \mathcal{H}$ an operator on $\mathcal{H}$. With $\mathcal{L}(\mathcal{H})$, we denote the space of operators which are linear. Furthermore, for $A \in \mathcal{L}(\mathcal{H})$, we define the *operator norm* as

$$\|A\|_\infty \coloneqq \sup_{\substack{|\psi\rangle \in \mathcal{H}, \\ \langle\psi|\psi\rangle = 1}} \|Ax\|. \tag{2.1.2}$$

The linear operator A is called *bounded* if $\|A\|_\infty < \infty$ and the set of bounded operators on $\mathcal{H}$ is given by

$$\mathcal{B}(\mathcal{H}) \coloneqq \{A \in \mathcal{L}(\mathcal{H}) \mid \|A\|_\infty < \infty\}. \tag{2.1.3}$$

For physical devices, Hermitian and positive operators are most important as they possess real or positive eigenvalues,respectively, and thus, correspond to observables.

Definition 2.1.4 (Hermitian and Positive Operators)**.** Let $\mathcal{H}$ be a Hilbert space. An operator $H \in \mathcal{B}(\mathcal{H})$ is *Hermitian* or self-adjoint if $H = H^\dagger$, where $H^\dagger$ is the adjoint operator of H. We denote the set of Hermitian operators by

$$\mathrm{Herm}(\mathcal{H}) \coloneqq \{H \in \mathcal{B}(\mathcal{H}) \mid H = H^\dagger\}. \tag{2.1.4}$$

An operator $P \in \mathrm{Herm}(\mathcal{H})$ is *positive semidefinite* (or *positive definite*) if $\langle\psi| P |\psi\rangle \geq 0$ (or $\langle\psi| P |\psi\rangle > 0$) for all $|\psi\rangle \in \mathcal{H}$. Equivalently, we will write $P \geq 0$ or $P > 0$, respectively. We denote the set of positive semidefinite operators by

$$\mathrm{Pos}(\mathcal{H}) \coloneqq \{P \in \mathrm{Herm}(\mathcal{H}) \mid P \geq 0\}. \tag{2.1.5}$$

For a Hilbert space associated with a register Q we will denote these sets with the shorthand notation $\mathcal{B}(Q)$, $\mathrm{Herm}(Q)$ and $\mathrm{Pos}(Q)$.

Definition 2.1.5 (Trace)**.** Let $\mathcal{H}$ be a Hilbert space and $A \in \mathcal{B}(\mathcal{H})$. For any orthonormal basis of $\mathcal{H}$ $\{|u_i\rangle\}_{i \in I}$, where I is an arbitrary index set, we define the trace as

$$\mathrm{Tr}[A] \coloneqq \sum_{i \in I} \langle u_i| A |u_i\rangle . \tag{2.1.6}$$

Definition 2.1.6 (Trace Norm and Trace-Class Operators [45, Definition 4.32])**.** Let $\mathcal{H}$ be a Hilbert space and $A \in \mathcal{B}(\mathcal{H})$ a bounded operator on $\mathcal{H}$. The operator A is called *trace-class* if there exits a orthonormal basis $\{|u_n\rangle\}_{n \in I}$, where I is an arbitrary index set, such that

$$\mathrm{Tr}[A] = \sum_{n \in I} \langle u_n| |A| |u_n\rangle < \infty, \tag{2.1.7}$$

where $|A| \coloneqq \sqrt{A^\dagger A}$. The *trace norm* of A is defined as

$$\|A\|_1 \coloneqq \mathrm{Tr}[|A|]. \tag{2.1.8}$$

Finally, the set of trace-class operators on $\mathcal{H}$ is defined by

$$\mathcal{B}_1(\mathcal{H}) \coloneqq \{A \in \mathcal{B}(\mathcal{H}) \mid \|A\|_1 < \infty\}. \tag{2.1.9}$$

Remark 2.1.7. For $A \in \mathcal{B}_1(\mathcal{H})$, the value of $\mathrm{Tr}[A]$ in Definition 2.1.5 does not depend on the choice of the orthonormal basis [45, Proposition 4.36].

Next, we focus on defining the state space and density operators.

Definition 2.1.8 (State space, state vector and pure states)**.** Let $\mathcal{H}$ be a complex n-dimensional Hilbert space. If a physical system Q can be described by $\mathcal{H}$, $\mathcal{H}$ is called the *state space* associated with system Q. A unit vector $|\psi\rangle$ in $\mathcal{H}$ is called a *state vector* and the evolution of $|\psi\rangle$ completely describes the system, when $|\psi\rangle$ represents the initial conditions for the system. These states are also called *pure states*, as the state is known precisely.

In contrast to pure states are mixed states, where the actual state is unknown and can only decompose as a statistical mixture. For example, if we are only interested in the evolution of a subsystem, the state of the subsystem might not remain in a pure state. Therefore, pure states are not the most general class of states describing a physical system and we will define mixed states in terms of density operators below.

Definition 2.1.9 (Ensemble of pure states)**.** Let $\{|\psi_i\rangle\}_{i\in I}$ be a set of pure states, where I is an arbitrary countable index set. Furthermore, let $\{p_i\}_{i\in I}$ be a set of probabilities, which sum up to unity, $\sum_{i\in I} p_i = 1$. The set $\{p_i, |\psi_i\rangle\}_{i\in I}$, is called an *ensemble of pure states*.

Remark 2.1.10. An ensemble of pure states can always be extended such that the set $\{|\psi_i\rangle\}$ forms a basis by extending the set to a basis with states $|\psi_j\rangle$ and setting respective coefficients to 0.

Such an ensemble of pure states will still not be the most general way of representing the state of a physical system, as the example of a subsystem suggested above. Therefore, we define mixed states as a mixture of positive semidefinite operators with unit trace and the sets of normalized and subnormalized density operators.

Definition 2.1.11 (Mixed States and Density operators)**.** We define the sets of *normalized* and *sub-normalized* density operators on a Hilbert space $\mathcal{H}$ by

$$\mathrm{S}_{\leq}(\mathcal{H}) \coloneqq \{\rho \in \mathcal{B}_1(\mathcal{H}) \mid \rho \geq 0, \mathrm{Tr}[\rho] \leq 1\}, \tag{2.1.10}$$

$$\mathrm{S}_{=}(\mathcal{H}) \coloneqq \{\rho \in \mathcal{B}_1(\mathcal{H}) \mid \rho \geq 0, \mathrm{Tr}[\rho] = 1\}. \tag{2.1.11}$$

Furthermore, we define *mixed states* in the following sense. Let $\{p_i\}_{i\in I}$ be a set of probabilities, which sum up to unity, $\sum_{i\in I} p_i = 1$ and let $\{\rho_i\}$ be a set of normalized density operators, i.e. $\rho_i \in \mathrm{S}_{=}\,\mathcal{H}$. The state ρ given by

$$\rho = \sum_{i\in I} p_i \rho_i \tag{2.1.12}$$

is called a mixed state.

As before we use the shorthand notations of $\mathrm{S}_{=}(Q)$ and $\mathrm{S}_{\leq}(Q)$ for the sets of normalized and subnormalized density operators on the Hilbert space associated with a register Q.

Finally, the spectral theorem, see e.g. [46, Theorem B.136], allows us to write any density operator as a mixture of pure states.

Theorem 2.1.12 (Spectral Decomposition for density operators)**.** *Let $\rho \in \mathrm{S}_{\leq}(\mathcal{H})$. Then, ρ can be written as*

$$\rho = \sum_{i\in I} p_i \left|\psi_i\right\rangle\!\left\langle\psi_i\right|, \tag{2.1.13}$$

where $\{|\psi_i\rangle\}$ forms a basis of $\mathcal{H}$ and $\{p_i, |\psi_i\rangle\}$ is an ensemble of pure states.

Proof. The statement follows immediately from the spectral theorem, see e.g. [46, Theorem B.136], by noting that density operators are trace-class and thus compact. □

Definition 2.1.13 (Classical-Quantum State)**.** A state $\rho \in \mathrm{S}_{\leq}(CQ)$ is said to be a *classical-quantum state* (with respect to a specified basis on C), in short cq-state, if it is in the form

$$\rho_{CQ} = \sum_{c} \lambda_c \left|c\right\rangle\!\left\langle c\right| \otimes \sigma_c, \tag{2.1.14}$$

for some normalized states $\sigma_c \in \mathrm{S}_{=}(Q)$ and weights $1 \geq \lambda_c \geq 0$, with $|c\rangle$ a basis on C. In most circumstances, we will not explicitly specify this "classical basis" of C, leaving it to be implicitly defined by context. It may be convenient to absorb the weights λ_c into the states σ_c, writing them as subnormalized states (defined below) $\omega_c = \lambda_c\sigma_c \in \mathrm{S}_{\leq}(Q)$ instead.

Definition 2.1.14 (Normalized and sub-normalized conditional states)**.** For a cq-state $\rho \in \mathrm{S}_{\leq}(CQ)$, written in the form $\rho_{CQ} = \sum_c |c\rangle\langle c| \otimes \omega_c$ for some $\omega_c \in \mathrm{S}_{\leq}(Q)$, and an event Ω defined on the register C, we will define the corresponding *sub-normalized conditional state* as

$$\rho_{CQ\wedge\Omega} := \sum_{c\in\Omega} \lambda_c \, |c\rangle\langle c| \otimes \omega_c. \tag{2.1.15}$$

Additionally, we define the *normalized conditional state* $\rho_{CQ|\Omega}$ as

$$\rho_{CQ|\Omega} := \frac{\mathrm{Tr}[\rho_{CQ}]}{\mathrm{Tr}[\rho_{CQ\wedge\Omega}]}\rho_{CQ\wedge\Omega}. \tag{2.1.16}$$

Given the above definitions, we may sometimes write ρ_{CQ} with the slightly more suggestive notation

$$\rho_{CQ} = \sum_c \Pr[c] \, |c\rangle\langle c| \otimes \rho_{Q|c}, \tag{2.1.17}$$

where $\Pr[c]$ describes the probability distribution on C, and $\rho_{Q|c}$ denotes the conditional state on Q corresponding to the register C taking value c.

Remark 2.1.15. We may sometimes write the cq-state ρ_{CQ} equivalently as

$$\rho_{CQ} = \sum_c \rho(c) \, |c\rangle\langle c| \otimes \rho_{Q|c}, \tag{2.1.18}$$

where $\rho(c)$ describes the probability of $C = c$ according to ρ. Furthermore, we may indicate the probability distribution on C induced by ρ as the tuple

$$\boldsymbol{\rho}_C := (\rho(1), \rho(2), \dots). \tag{2.1.19}$$

Finally, we define positive operator-valued measures, as they will allows us to characterize measurements of quantum systems.

Definition 2.1.16 (POVM)**.** Let $\mathcal{H}$ be a Hilbert space and Σ an alphabet. A set of positive operators $\{M_k \in \mathrm{Pos}(\mathcal{H}) \mid k \in \Sigma\}$ is called a *positive operator-valued measure* (POVM) if

$$\sum_{k\in\Sigma} M_k = \mathbb{I}_{\mathcal{H}}, \tag{2.1.20}$$

and each element M_k is called a POVM element.

Remark 2.1.17. All POVM elements satisfy $0 \leq M_k \leq \mathbb{I}_{\mathcal{H}}$.

The POVM element M_k is associated with outcome k such that the probability of observing this outcome is

$$\Pr(k) = \mathrm{Tr}[M_k \rho], \tag{2.1.21}$$

which is simply Born's rule.

2.2 Distance Measures and Norms

In this section we define different distance measures following [47, Chapter 3]. Before we start with the definition of the actual distance measures, we require the definition of functions of positive semidefinite operators.

Definition 2.2.1 (Functions of Operators [47, Section 2.5])**.** Let $f : \mathbb{R}_{>0} \to \mathbb{R}$ be continuous and $M \in \mathrm{Pos}(\mathcal{H})$ with spectral decomposition

$$M = \sum_i \lambda_i \left|u_i\right\rangle\!\left\langle u_i\right| . \tag{2.2.1}$$

We define $f(M)$ as

$$f(M) \coloneqq \sum_{i:\lambda_i \neq 0} f(\lambda_i) \left|u_i\right\rangle\!\left\langle u_i\right| , \tag{2.2.2}$$

i.e. we define $f(M)$ on the support of M.

Definition 2.2.2 (Schatten Norms [47, Definition 3.2])**.** For any $A \in \mathcal{B}(\mathcal{H})$ and $p \geq 1$, we define the *Schatten p-norm* as

$$\left\|A\right\|_p \coloneqq \left(\mathrm{Tr}[|A|^p]\right)^{\frac{1}{p}} , \tag{2.2.3}$$

where, as before, $|A| \coloneqq \sqrt{A^\dagger A}$.

Remark 2.2.3. For $p = 1$ we recover the trace norm and for $p = \infty$ we recover the operator norm.

Definition 2.2.4 (Fidelity [47, Definition 3.6])**.** For $\rho, \sigma \in \mathrm{S}_=(\mathcal{H})$, we define the *fidelity* between ρ and σ as

$$F(\rho, \sigma) \coloneqq \left(\mathrm{Tr}\left[\left|\sqrt{\rho}\sqrt{\sigma}\right|\right]\right)^2 . \tag{2.2.4}$$

By Uhlmann's theorem, see e.g. [48], the fidelity of normalized states ρ and σ in system A can be equivalently viewed as the maximum inner product between two purifications, i.e.

$$F(\rho,\sigma) = \max_{\substack{|\phi\rangle_{AR},\, \mathrm{Tr}_R[|\phi\rangle\langle\phi|_{AR}]=\rho \\ |\psi\rangle_{AR},\, \mathrm{Tr}_R[|\psi\rangle\langle\psi|_{AR}]=\sigma}} |\langle\phi|\psi\rangle|^2 \tag{2.2.5}$$

In order to make use of fidelity as a distance measure for subnormalized states, we define the generalized fidelity. It extends the notion of fidelity to subnormalized states.

Definition 2.2.5 (Generalized Fidelity [47, Definition 3.7])**.** For $\rho,\sigma \in \mathrm{S}_{\leq}(\mathcal{H})$, we define the *generalized fidelity* between ρ and σ as

$$F_*(\rho,\sigma) \coloneqq \left(\mathrm{Tr}\left[\left|\sqrt{\rho}\sqrt{\sigma}\right|\right] + \sqrt{(1-\mathrm{Tr}[\rho])(1-\mathrm{Tr}[\sigma])}\right)^2. \tag{2.2.6}$$

Definition 2.2.6 (Purified Distance [47, Definition3.2])**.** For $\rho,\sigma \in \mathrm{S}_{\leq}(\mathcal{H})$, we define the *purified distance* between ρ and σ as

$$P(\rho,\sigma) \coloneqq \sqrt{1-F_*(\rho,\sigma)}. \tag{2.2.7}$$

2.3 Entropic Quantities

In any of the QKD security proofs we will see in the following chapters, proving security always reduces to finding bounds on some type of entropy. In this section, we give the respective definitions of each entropy and present some of their relevant properties.

The simplest example is the conditional von Neumann entropy. In line with the definitions of the Rényi entropies we will define the conditional von Neumann entropy in terms of a divergence. Specifically, we define the von Neumann entropy in terms of the relative entropy, first studied by Umegaki [49].

Definition 2.3.1 (Relative Entropy)**.** Let $\rho,\sigma \in \mathrm{S}_{\leq}(A)$ and $\rho \neq 0$. We define the *relative entropy* as

$$D(\rho\|\sigma) \coloneqq \frac{\mathrm{Tr}[\rho\,(\log(\rho)-\log(\sigma))]}{\mathrm{Tr}[\rho]}, \tag{2.3.1}$$

whenever $\mathrm{supp}(\rho) \subseteq \mathrm{supp}\,\sigma$ and ∞ otherwise.

This relative entropy is one example of a divergence, and later in Definition 2.3.9, we will see more examples. Since it is typically referred to as the relative entropy we will adopt this notion. With this definition of the relative entropy at hand, we can define the von Neumann entropy.

Definition 2.3.2 (Conditional von Neumann Entropy)**.** Let $\rho \in \mathrm{S}_{=}(AB)$. We define the *conditional von Neumann entropy* as

$$H(A|B) \coloneqq \max_{\sigma_B \in \mathrm{S}_{=}(B)} -D(\rho_{AB} \| \mathbb{I}_A \otimes \sigma_B)\,. \tag{2.3.2}$$

whenever $\mathrm{supp}(\rho) \subseteq \mathrm{supp}\,\sigma$ and ∞ otherwise.

Remark 2.3.3. The optimization in Eq. (2.3.2) is achieved for $\sigma_B = \rho_B$, see e.g. [50]. This lets us write the conditional von Neumann entropy in the following equivalent forms

$$\begin{aligned} H(A|B)_\rho &= -D(\rho_{AB} \| \mathbb{I}_A \otimes \rho_B) \\ &= H(AB)_\rho - H(B)_\rho\,. \end{aligned} \tag{2.3.3}$$

2.3.1 Min- and Smooth Min-Entropy

Next, we focus on defining the smooth min-entropy which requires us to first define the min-entropy. Intuitively, for classical X and Y, the min-entropy $H_{\min}(X|Y)$ can be understood as the logarithm of the probability of guessing X with access to Y. Alternatively, it can be viewed as the minimum amount of entropy contained in X after an adversary had access to Y.

Definition 2.3.4 (Min-Entropy as in [47, Definition 6.2])**.** Let $\rho_{AB} \in \mathrm{S}_{\leq}(AB)$. The min-entropy of A conditioned on B of the state ρ_{AB} is

$$H_{\min}(A|B)_\rho := \sup_{\sigma_B \in \mathrm{S}_{\leq}(B)} \sup\{\lambda \in \mathbb{R} : \rho_{AB} \leq 2^{-\lambda} \mathbb{I}_A \otimes \sigma_B\} \tag{2.3.4}$$

Note in [47] $e^{-\lambda}$ was used to indicate $2^{-\lambda}$.

However, for QKD purposes this worst-case estimate described above is often too pessimistic and so in Ref. [51], a smoothed version was introduced. Therefore, we first define the so-called smoothing ball which specifies the region the smoothing is performed over. We note here that the original version in Ref. [51] was defined in terms of trace distance, but here we choose the more modern approach of purified distance.

Definition 2.3.5 (Smoothing Ball as in [47, Definition 6.4])**.** Let $\rho \in \mathrm{S}_{\leq}(A)$ and let $0 \leq \varepsilon < \sqrt{\mathrm{Tr}[\rho]}$. We define the ε-smoothing ball of states in $\mathrm{S}_{\leq}(A)$ around ρ as

$$B_\varepsilon(\rho) := \{\tau \in \mathrm{S}_{\leq}(A) : P(\tau, \rho) \leq \varepsilon\}, \tag{2.3.5}$$

where $P(\cdot, \cdot)$ is the purified distance of Definition 2.2.6.

With these tools, we can define the smooth min-entropy.

Definition 2.3.6 (Smooth Min-Entropy as in [47, Definition 6.5])**.** Let $\rho \in \mathrm{S}_{\leq}(A)$ and $0 \leq \varepsilon < \sqrt{\mathrm{Tr}[\rho]}$. We define the ε-smooth min-entropy of A conditioned on $\mathrm{S}_{\leq}(A)$ around ρ as

$$H_{\min}^{\varepsilon}(A|B)_{\rho} := \sup_{\tilde{\rho} \in B_{\varepsilon}(\rho_{AB})} H_{\min}(A|B)_{\tilde{\rho}}, \tag{2.3.6}$$

These smoothed entropies have the following useful properties with respect to subnormalized conditional states.

Lemma 2.3.7 (Lemma 10 as in [37])**.** *Let $\rho_{ABXY} \in \mathrm{S}_{\leq}(ABXY)$ be classical on X and Y, let $\Omega : \mathcal{X} \times \mathcal{Y} \to \{0,1\}$ be an event with $\Pr[\Omega]_{\rho} > 0$. Then for $\varepsilon \in \left[0, \sqrt{\Pr[\Omega]_{\rho}}\right)$, we have*

$$H_{\min}^{\varepsilon}(AX|BY)_{\rho \wedge \Omega} \geq H_{\min}^{\varepsilon}(AX|BY)_{\rho}, \tag{2.3.7}$$

$$H_{\min}^{\varepsilon}(AX|B)_{\rho \wedge \Omega} \geq H_{\min}^{\varepsilon}(AX|B)_{\rho}. \tag{2.3.8}$$

Proof. See the proof in [37, Lemma 10]. □

Additionally, when conditioning on a classical register Z one can bound the smooth min-entropy $H_{\min}^{\varepsilon}(A|BZ)$ with the entropy of the worst case conditional state $H_{\min}^{\varepsilon}(A|B)_{\rho^{(z)}}$. This property will be very useful for one of the security proofs we present later in Chapter 7, but it fits at this stage. We originally proved this in [3, Lemma 18].

Lemma 2.3.8 (Conditioning on Classical Register)**.** *Let $\rho_{ABZ} \in \mathrm{S}_{=}(ABZ)$ be classical on Z. Then,*

$$H_{\min}^{\varepsilon}(A|BZ)_{\rho} \geq \min_{z \ s.t. \ \Pr(z) \neq 0} H_{\min}^{\varepsilon}(A|B)_{\rho^{(z)}}. \tag{2.3.9}$$

Proof. First, we will construct an explicit state in the ε-ball of ρ_{ABZ} and then show that it yields the claimed lower bound on the smooth min-entropy.

By proposition 6.1 in Ref. [47], for each $z \in \mathcal{Z}$, there exists an optimizer $\sigma^{(z)} \in B_{\varepsilon}(\rho^{(z)})$ such that $H_{\min}^{\varepsilon}(A|B)_{\rho^{(z)}} = H_{\min}(A|B)_{\sigma^{(z)}}$. Next, let us define the state

$$\sigma := \sum_{z \in \mathcal{Z}} \Pr(z) \, |z\rangle\langle z| \otimes \sigma^{(z)} \in \mathrm{S}_{\leq}(ABZ), \tag{2.3.10}$$

which we will show to be contained in $B_\varepsilon(\rho)$. For this purpose, consider the purified distance between ρ and σ

$$P(\rho,\sigma) = P\left(\sum_{z\in\mathcal{Z}} \Pr(z)\,|z\rangle\langle z|\otimes\rho^{(z)}, \sum_{z\in\mathcal{Z}} \Pr(z)\,|z\rangle\langle z|\otimes\sigma^{(z)}\right) \tag{2.3.11}$$

$$\leq \sup_{z\in\mathcal{Z}} P\left(|z\rangle\langle z|\otimes\rho^{(z)}, |z\rangle\langle z|\otimes\sigma^{(z)}\right) \tag{2.3.12}$$

$$= \sup_{z\in\mathcal{Z}} P\left(\rho^{(z)},\sigma^{(z)}\right) \leq \varepsilon, \tag{2.3.13}$$

where for the first inequality [47, eq. (3.59)] was used and for the third line we used the multiplicativity of the fidelity, which still applies to the generalized fidelity. The last inequality results from $\sigma^{(z)} \in B_\varepsilon(\rho^{(z)})$. Hence, we found a state $\sigma \in B_\varepsilon(\rho)$.

Next, we will use $\sigma \in B_\varepsilon(\rho)$ to bound the smooth min-entropy of ρ by

$$H_{\min}^\varepsilon(A|BZ)_\rho = \sup_{\tilde\rho\in B_\varepsilon(\rho_{ABZ})} H_{\min}(A|BZ)_{\tilde\rho} \geq H_{\min}(A|BZ)_\sigma \tag{2.3.14}$$

$$\geq \min_{z \text{ s.t. } \Pr(z)\neq 0} H_{\min}(A|B)_{\sigma^{(z)}} \tag{2.3.15}$$

$$= \min_{z \text{ s.t. } \Pr(z)\neq 0} H_{\min}^\varepsilon(A|B)_{\rho^{(z)}}, \tag{2.3.16}$$

where for the third inequality the property of the min-entropy [47, eq. (6.25)] was applied and the last equality follows from each $\sigma^{(z)}$ being the optimizer such that $H_{\min}^\varepsilon(A|B)_{\rho^{(z)}} = H_{\min}(A|B)_{\sigma^{(z)}}$. □

2.3.2 Rényi Entropies

The smooth entropy in the previous section was used in the very first finite-size security proofs. However, very recently a different kind of entropies, known as Rényi entropies, were shown to be very useful in QKD applications. We will also make use of them in later sections, thus we introduce them here for reference. We again follow Ref. [47].

Definition 2.3.9 (Rényi Divergences [47, Def. 4.3 and Def. 4.5])**.** Let $\alpha \in (0,1)\cup(1,\infty)$, and $\rho,\sigma \in \mathrm{S}_{\leq}(A)$ with $\mathrm{Tr}[\rho] \neq 0$, the *sandwiched quantum Rényi divergence* (or minimal quantum Rényi divergence) is defined as

$$\widetilde{D}_\alpha(\rho\|\sigma) := \frac{1}{\alpha-1}\log\frac{\mathrm{Tr}\left[\left(\sigma^{\frac{1-\alpha}{2\alpha}}\rho\sigma^{\frac{1-\alpha}{2\alpha}}\right)^\alpha\right]}{\mathrm{Tr}[\rho]}, \tag{2.3.17}$$

for $(\alpha < 1 \wedge \rho \not\perp \sigma) \vee \operatorname{supp}(\rho) \subseteq \operatorname{supp}\sigma$, and ∞ otherwise.

The *Petz quantum Rényi divergence* is defined as

$$\overline{D}_\alpha(\rho\|\sigma) \coloneqq \frac{1}{\alpha-1}\log\frac{\operatorname{Tr}[\rho^\alpha\sigma^{1-\alpha}]}{\operatorname{Tr}[\rho]}, \tag{2.3.18}$$

whenever $\alpha < 1$ and ρ, σ are not orthogonal or $\vee\operatorname{supp}(\rho) \subseteq \operatorname{supp}\sigma$, and ∞ otherwise. In any statement that applies to both divergences, we will write $\mathbb{D}_\alpha$ to denote either divergence.

Remark 2.3.10. Both of these Rényi divergences converge back to the relative entropy, Definition 2.3.1, in the limit of $\alpha \to 1$. However, for all $\alpha \neq 1$ it holds that

$$\overline{D}_\alpha(\rho\|\sigma) \geq \widetilde{D}_\alpha(\rho\|\sigma) \tag{2.3.19}$$

for all $\rho, \sigma \in \mathrm{S}_{\leq}(A)$.

Similar to Definition 2.3.2, we define the conditional Rényi entropies in terms of the divergences. However, in the case of Rényi entropies the optimized and unoptimized versions do not agree.

Definition 2.3.11 (Conditional Rényi Entropies [47, Def. 5.2])**.** For $\alpha \geq 0$ and $\rho_{AB} \in \mathrm{S}_=(AB)$ the *quantum conditional Rényi entropies* are defined as:

$$\overline{H}^{\downarrow}_\alpha(A|B)_\rho \coloneqq -\overline{D}_\alpha(\rho_{AB}\|I_A \otimes \rho_B) \tag{2.3.20}$$

$$\overline{H}^{\uparrow}_\alpha(A|B)_\rho \coloneqq \sup_{\sigma_B \in \mathrm{S}_=(B)} -\overline{D}_\alpha(\rho_{AB}\|I_A \otimes \sigma_B) \tag{2.3.21}$$

$$\widetilde{H}^{\downarrow}_\alpha(A|B)_\rho \coloneqq -\widetilde{D}_\alpha(\rho_{AB}\|I_A \otimes \rho_B) \tag{2.3.22}$$

$$\widetilde{H}^{\uparrow}_\alpha(A|B)_\rho \coloneqq \sup_{\sigma_B \in \mathrm{S}_=(B)} -\widetilde{D}_\alpha(\rho_{AB}\|I_A \otimes \sigma_B) \tag{2.3.23}$$

We refer to these conditional entropies as *Petz* (first two) and *sandwiched* (last two) conditional Rényi entropies. Again, in any statement that applies to all conditional Rényi entropies, we use $\mathbb{H}_\alpha$.

2.4 Concepts from Convex Optimization

Finally, we require a few definitions from convex optimization theory. Here we roughly follow Ref. [52] and refer the reader to it for further details.

Definition 2.4.1 (Convex Set)**.** Let V be a vector space over $\mathbb{R}$ and $C \subseteq V$ a subset. The set C is called *convex* if for all $x, y \in C$ and $\lambda \in [0,1]$ it holds that

$$\lambda x + (1-\lambda)\, y \in C. \tag{2.4.1}$$

Similarly, we define convex and concave functions.

Definition 2.4.2 (Convex and Concave Functions)**.** Let V_1 and V_2 be vector spaces over $\mathbb{R}$ and $f : V_1 \to V_2$ a function. We call the function f *convex* if for all $x, y \in C$ and $\lambda \in [0,1]$ it holds that

$$f(\lambda x + (1-\lambda)\, y) \leq \lambda f(x) + (1-\lambda)\, f(y), \tag{2.4.2}$$

and *concave* if

$$f(\lambda x + (1-\lambda)\, y) \geq \lambda f(x) + (1-\lambda)\, f(y). \tag{2.4.3}$$

Remark 2.4.3. Equivalently, one can define a concave function by requiring that $-f$ is convex.

Certain transformations preserve the convexity of a function. One example are affine transformations.

Lemma 2.4.4 (Composition with affine functions preserves convexity; Section 3.2.2 in [52])**.** *Let V_1 and V_2 be vector spaces over $\mathbb{R}$, $C_1 \subseteq V_1$, $C_2 \subseteq V_2$ be convex sets, $f : C_1 \to \mathbb{R}$ a convex function on C_1, and $g : C_2 \to C_1$ an affine function on C_2, then the function $h = f \circ g : C_2 \to \mathbb{R}$ is convex on C_2.*

Additionally,the composition of convex and concave functions can return a convex function. The following conditions state certain operations which guarantee this.

Lemma 2.4.5 (Scalar Composition Rules)**.** *Let V_1, V_2 and V_3 be vector spaces over $\mathbb{R}$ and $g : V_1 \to V_2$, $h : V_2 \to V_3$ be two functions. Define $f : V_1 \to V_3$ as*

$$f \coloneqq h \circ g. \tag{2.4.4}$$

Then,

1. *f is convex if h is convex and non-decreasing, and g is convex,*
2. *f is convex if h is convex and non-increasing, and g is concave,*
3. *f is concave if h is concave and non-decreasing, and g is concave,*

4. f is concave if h is concave and non-increasing, and g is convex.

For a convex function f one can define the optimization problem

$$\min_{x \in C} f(x) \tag{2.4.5}$$

for some convex set C. This optimization problem can then be solved efficiently with convex optimization solvers. One very simple example of an algorithm to solve this optimization problem for differentiable functions is the Frank-Wolfe algorithm [53].

Definition 2.4.6 (Frank-Wolfe Algorithm)**.** Let V be a vector space over the field $\mathbb{R}$, and $C \subseteq V$ a convex subset. Let $f : C \to \mathbb{R}$ be a convex and differentiable function at all points in S. The convex optimization problem,

$$\min_{\mathbf{x} \in S} f(\mathbf{x}), \tag{2.4.6}$$

can be numerically solved by the following algorithm.

Algorithm 1 Frank-Wolfe optimization [53]

1: Let $\epsilon > 0$, $\mathbf{x}_0 \in S$ and set $i = 0$.
2: Compute step direction $\Delta\mathbf{x} \coloneqq \arg\min_{\Delta\mathbf{x}} \langle \boldsymbol{\nabla} f(\mathbf{x}_i), \Delta\mathbf{x} \rangle$ subject to $\Delta\mathbf{x} + \mathbf{x}_i \in S$.
3: If $|\langle \boldsymbol{\nabla} f(\mathbf{x}_i), \Delta\mathbf{x} \rangle| < \epsilon |f(\mathbf{x}_i)|$ then STOP.
4: Compute step size $\lambda \coloneqq \arg\min_{\lambda \in [0,1]} f(\mathbf{x}_i + \lambda \Delta\mathbf{x})$.
5: Set $\mathbf{x}_{i+1} = f(\mathbf{x}_i + \lambda \Delta\mathbf{x})$ and $i \leftarrow i + 1$. Go to step 2.

Chapter 3

Quantum Key Distribution Background

In this chapter, we give all required definitions for quantum key distribution (QKD) protocols and security proofs. We start by giving the formal definition of security. Then, we will discuss two different types of QKD protocols, entanglement-based (EB) and prepare-and-measure (P&M) QKD protocols. Finally, we conclude with different types of leftover hashing lemmas that will be a central part for any security proof in this thesis.

3.1 Security Definitions

In a QKD protocol, two trusted parties, typically called Alice and Bob, aim to generate a shared secret key that is unknown or decoupled from any adversary, typically Eve, who is allowed to perform any attack permitted by quantum mechanics. Thus, the final goal of a QKD protocol is to generate two strings K_A and K_B,corresponding to Alice's and Bob's keys, respectively, that are identical and decoupled from Eve's knowledge stored in her register E. In this section, we formalize this notion of security in the composable security framework of Ref. [54, Sec. III.B]. Here composable refers to the property that a key produced by such a protocol can safely be used in a following protocol.

The steps of a QKD protocol consist of possibly several quantum phases in which Alice and Bob send or measure quantum signals, classical communication over an authenticated channel, testing steps, error correction to ensure that Alice and Bob reach the same key and a step called privacy amplification. We will discuss the exact steps of a QKD protocol later in Protocol 1 and Protocol 2.

Besides the distinction between P&M and EB protocols, one needs to distinguish between *fixed-* and *variable*-length protocols. Fixed-length protocols always output final keys of fixed length and perform a testing step during the protocol called *acceptance test* in which Alice and Bob decide if they continue or abort the protocol based on a subset of their observations, called test data.

On the other hand, variable-length protocols output keys of varying length, depending on the quality of the observations contained in the same test data. Variable-length protocols decide the final key length during their testing stage.

In preparation for the following security definitions let $\sigma_{K_A K_B E}$ denote the output state after n rounds of a QKD protocol, where K_A, K_B again denote Alice and Bob's final key registers, and E denotes Eve's side information at the end of the QKD protocol, and includes all classical announcements. For fixed-length protocols the acceptance event Ω_{acc} includes passing the acceptance decision during the acceptance test and any testing stage, such as error verification. Again, we will discuss the details later.

Intuitively, the security definition states that the output state $\sigma_{K_A K_B E}$ of the QKD protocol is ε-close to perfect uniformly random and correlated keys of length l

$$\mathbb{K}^l_{K_A K_B} = \sum_{k \in \{0,1\}^l} \frac{|kk\rangle\langle kk|_{K_A K_B}}{2^l}, \tag{3.1.1}$$

while Eve decouples completely.

Definition 3.1.1 (ε-Security of Fixed-Length Protocols)**.** Let Ω_{acc} be the event of a QKD protocol not aborting. Then, a fixed-length QKD protocol is ε_{sec}-*secure* if for all input states $\sigma_{A^n B^n}$, the resulting output state $\sigma_{K_A K_B E}$ satisfies

$$\frac{1}{2}\Pr[\Omega_{\text{acc}}] \left\| \sigma_{K_A K_B E|\Omega_{\text{acc}}} - \mathbb{K}^l_{K_A K_B} \otimes \sigma_{E|\Omega_{\text{acc}}} \right\|_1 \leq \varepsilon_{\text{sec}}, \tag{3.1.2}$$

where $\mathbb{K}^l_{K_A K_B}$ denotes the state of a perfect uniform shared key of length l

$$\mathbb{K}^l_{K_A K_B} = \sum_{k \in \{0,1\}^l} \frac{|kk\rangle\langle kk|_{K_A K_B}}{2^l}. \tag{3.1.3}$$

Furthermore, a fixed-length QKD protocol is ε_{sct}-*secret* if

$$\frac{1}{2}\Pr[\Omega_{\text{acc}}] \left\| \sigma_{K_A E|\Omega_{\text{acc}}} - \mathbb{U}^l_{K_A} \otimes \sigma_{E|\Omega_{\text{acc}}} \right\|_1 \leq \varepsilon_{\text{sct}}, \tag{3.1.4}$$

where $\mathbb{U}_{K_A}^l$ is a fully mixed state of dimension 2^l and ε_{cor}-*correct* if

$$\Pr[K_A \neq K_B \wedge \Omega_{\text{acc}}] \leq \varepsilon_{\text{cor}}. \tag{3.1.5}$$

Moreover, a QKD protocol that is ε_{cor}-correct and ε_{sct}-secret, is $\varepsilon_{\text{sec}} = \varepsilon_{\text{sct}} + \varepsilon_{\text{cor}}$ secure [54].

The distinction between security, secrecy and corrections is very convenient as it allows one to focus on secrecy and correctness separately.

Next, for variable-length protocols the security definition is slightly different. Here we have to consider the average over any possible key length.

Definition 3.1.2 (ε-Security of Variable-Length Protocols)**.** Let $\Omega_{\text{len}=m}$ be the event of a QKD protocol generating a final key of length $m \in \mathbb{N}_0$. A variable-length QKD protocol is ε_{sec}-*secure* if for any input state $\sigma_{A^nB^n}$ the resulting output state $\sigma_{K_AK_BE}$ satisfies

$$\frac{1}{2}\sum_{m=0}^{\infty}\Pr[\Omega_{\text{len}=m}]\left\|\sigma_{K_AK_BE|\Omega_{\text{len}=m}} - \mathbb{K}_{K_AK_B}^m \otimes \sigma_{E|\Omega_{\text{len}=m}}\right\|_1 \leq \varepsilon_{\text{sec}}, \tag{3.1.6}$$

where $\mathbb{K}_{K_AK_B}^m$ denotes the state of a perfect uniform shared key of length m:

$$\mathbb{K}_{K_AK_B}^m = \sum_{k\in\{0,1\}^m} \frac{|kk\rangle\langle kk|_{K_AK_B}}{2^m}. \tag{3.1.7}$$

Furthermore, a variable-length QKD protocol is ε_{sct}-*secret* if

$$\frac{1}{2}\sum_{m=0}^{\infty}\Pr[\Omega_{\text{len}=m}]\left\|\sigma_{K_AE|\Omega_{\text{len}=m}} - \mathbb{U}_{K_A}^m \otimes \sigma_{E|\Omega_{\text{len}=m}}\right\|_1 \leq \varepsilon_{\text{sct}}, \tag{3.1.8}$$

where $\mathbb{U}_{K_A}^m$ is a fully mixed state of dimension 2^m, and ε_{cor}-*correct* if

$$\Pr[K_A \neq K_B \wedge \Omega_{\text{acc}}] \leq \varepsilon_{\text{cor}}. \tag{3.1.9}$$

A variable-length QKD protocol that is ε_{cor}-correct and ε_{sct}-secret, is $\varepsilon_{\text{sec}} = \varepsilon_{\text{sct}} + \varepsilon_{\text{cor}}$ secure [55].

Remark 3.1.3. For P&M protocols one can restrict the input states to those which satisfy the marginal constraint $\text{Tr}_{B^n}[\sigma_{A^nB^n}] = \tau_A^{\otimes n}$ for some state τ_A defined by Alice's choices. In Section 3.3 we will explain why this restriction is valid for P&M protocols.

Remark 3.1.4. Fixed-length security follows from variable-length security, by selecting only one length in the sum in Eq. (3.1.6) and identifying this with the fixed key length l of Protocol 1; all other events are combined to one single abort event. Furthermore, secrecy and correctness imply security in the same way.

3.2 Protocol Description

Now, we focus on the specific steps of a QKD protocol we consider in this work. We start by stating a description of a generic prepare-and-measure protocol and then formulate the protocol description for an entanglement-based protocol.

In a prepare-and-measure (P&M) protocol Alice prepares certain quantum states $\{\sigma_i\}_{i=1\ldots d_A}$ and sends each of them to Bob with some probability. Bob then measures his incoming quantum states, thus the name prepare-and-measure protocol.

Protocol 1. Prepare-and-Measure Protocol.
Parameters:

$n \in \mathbb{N}_0$	:	Total number of rounds
$l \in \mathbb{N}_0$	:	Length of final key
$\{\sigma_i\}_{i=1\ldots d_A}$	:	States sent by Alice
d_A	:	Number of different signal states
$\mathcal{S}$	:	alphabet of raw-key registers, usually $\{0, 1, \perp\}$
$\{M_i^B\}_{i=1\ldots d_B}$	:	POVM elements acting on Bob's system describing his measurements outcomes
d_B	:	Number of Bob's POVM elements
γ	:	Probability that Alice chooses a round to be a test round
S_{acc}	:	Convex acceptance set of accepted frequencies
λ_{EC}	:	Number of bits exchanged during error correction step
ε_{PA}	:	Security parameter contribution from privacy amplification
ε_{EV}	:	Security parameter contribution from error verification
Ω_{AT}	:	Event of passing the acceptance test
Ω_{EV}	:	Event of passing the error verification
$\Omega_{\text{acc}} = \Omega_{\text{AT}} \wedge \Omega_{\text{EV}}$	:	Event of the protocol not aborting

Protocol steps:

1. For each round $i \in \{1, 2, \ldots, n\}$, Alice and Bob perform the following steps:[1]

[1]Note that in contrast to the analysis in e.g. [33], here we do *not* limit the repetition rate of the protocol such that Bob performs the i^{th} measurement before Alice sends the $(i+1)^{\text{th}}$ state — in our protocol, the states can be sent and measured in any time-ordering with respect to each other, apart from the trivial physical constraint that Bob can only perform the i^{th} measurement after Alice sends the i^{th} state. In fact, Alice can even send out all her signal states before Bob performs any measurements.

(a) **State preparation and transmission:** In each round, with probability γ Alice independently chooses it to be a *test round* or *generation round*. Next, Alice prepares one of d_A possible signal states $\{\sigma_k\}_{k=1\ldots d_A}$, according to some distribution (which could depend on the choice of a test or generation round). She stores the label for her choice of the signal state in a classical register X_i with alphabet $\mathcal{X}$, and computes a classical register C_i^A with alphabet $\mathcal{C}^A$ for public announcement (including for instance the test/generation decision). Finally, Alice sends the signal state to Bob via a quantum channel.

(b) **Measurements:** Bob measures his received states described by a POVM with POVM elements $\{M_k^B\}_{k=1\ldots d_B}$, and stores his results in a classical register Y_i with alphabet $\mathcal{Y}$. Furthermore, he computes a classical register C_i^B with alphabet $\mathcal{C}^B$ for public announcement.

2. **Public announcement:** Alice and Bob announce their values $C_1^A \ldots C_n^A, C_1^B \ldots C_n^B$. For each i, Alice and Bob can perform some further public classical communication (which can be two-way) based only on the values $X_i Y_i C_i^A C_i^B$ and local randomness; let I_i be a register denoting all such public communication with alphabet $\mathcal{I}$. Then Alice and Bob both compute a value $\widehat{C}_i$ that is found by applying some deterministic function ϕ on I_i [2]. These values $\widehat{C}_1^n$ will later be used in the acceptance test (to decide whether to abort) or the variable length decision (to decide the final key length). We require that it is set to a fixed symbol $\widehat{C}_i = \texttt{gen}$ whenever Alice chose the round to be a generation round (informally, this corresponds to the fact that the acceptance test or variable length decision later will only depend on data from the test rounds). We also suppose that all the classical registers $\widehat{C}_i$ are isomorphic to each other, with a common alphabet $\widehat{\mathcal{C}}$.

   For notational convenience in our later analysis, let T_1^n denote registers that are set to $T_i = \texttt{test}$ in test rounds and otherwise set to $T_i = I_i$. Let $\tilde{T}_1^n$ denote a copy of T_1^n that Eve holds (which she can indeed compute since she has access to the public communication I_i).

3. **Sifting and key map:** For each i, Alice applies a sifting step[3] and/or key map procedure based on her raw data X_i and the public announcements I_i (and local randomness, if needed), to produce a classical register S_i.

[2] The map ϕ maps all announcements in a generation round to a special generation flag $|\texttt{gen}\rangle_{\widehat{\mathcal{C}}}$.

[3] For the purposes of this work, when a round is sifted out, we take this to mean Alice sets S_i to a fixed value in the alphabet of S_i (say, 0). However, it should be possible to modify this to a version where such rounds are actually discarded (i.e. not included in the privacy amplification step) by using the techniques in [55].

4. **Acceptance test (parameter estimation) or variable length decision:** In a fixed length protocol, Alice and Bob compute the frequency distribution $\mathbf{F}^{\text{obs}}$ (see Definition 2.1.2) of the observed string of values $\hat{c}_1^n$ on the registers $\widehat{C}_1^n$. Then they accept if $\mathbf{F}^{\text{obs}} \in S_{\text{acc}}$ where S_{acc} is the predetermined acceptance set, and abort the protocol otherwise. We call the event of passing this stage Ω_{AT}.

   Alternatively, in a variable-length protocol, they compute the variable key length l_{var} by computing a particular function of $\hat{c}_1^n$ (see Theorem 8.3.3 for further details).

5. **Error correction and verification:** In a fixed-length protocol, Alice and Bob publicly communicate λ_{EC} bits for error correction, whereas in a variable-length protocol they communicate $\lambda_{\text{EC}}(\hat{c}_1^n)$ bits. Next, Bob uses those bits together with his data $Y_1^n I_1^n$ to produce a guess $\overline{S}_1^n$ for Alice's string S_1^n. This is followed by an error-verification step, where Alice sends a 2-universal hash of S_1^n with length $\lceil \log(1/\varepsilon_{\text{EV}}) \rceil$ to Bob, who compares it to the hash of his guess and accepts if the hashes match (and otherwise aborts). Let register L contain the full data communicated in this step and we call the event of passing the error correction procedure Ω_{EV}.

6. **Privacy amplification:** In a fixed length protocol, Alice randomly chooses a 2-universal hash function from some family (with fixed output length l) announces it and applies it to her string S_1^n producing her final key K_A of *fixed* length l. Bob then also applies the hash function to his guess $\overline{S}_1^n$, producing his final key K_B of *fixed* length l.

   Alternatively, in a variable-length protocol, Alice and Bob still randomly choose a 2-universal hash function from some family. However, this hash function now maps n bits to l_{var} bits. Then, they apply the hash function to their strings S_1^n and $\overline{S}_1^n$, producing the final keys K_A and K_B of length l_{var}.

In contrast, in entanglement-based (EB) protocols Alice and Bob receive either half of an entangled quantum state and measure their respective half. One can convert between both pictures and in the next Section 3.3, we will relate the above P&M protocol to an equivalent EB protocol. However, it is still instructive to discuss the steps of a generic EB protocol here as well.

Protocol 2. Entanglement-based Protocol.
Parameters:

$n \in \mathbb{N}_0$	:	Total number of rounds
$l \in \mathbb{N}_0$	:	Length of final key
$\|\Psi\rangle_{A_1^n B_1^n \tilde{E}}$	:	Initial joint state of Alice, Bob and Eve
$\{M_i^A\}_{i=1\ldots d_A}$	:	POVM elements acting on Alice's system describing her measurements outcomes
d_A	:	Number of different signal states
$\mathcal{S}$	:	alphabet of raw-key registers, usually $\{0, 1, \perp\}$
$\{M_i^B\}_{i=1\ldots d_B}$	:	POVM elements acting on Bob's system describing his measurements outcomes
d_B	:	Number of Bob's POVM elements
γ	:	Probability that Alice chooses a round to be a test round
S_{acc}	:	Convex acceptance set of accepted frequencies
λ_{EC}	:	Number of bits exchanged during error correction step
ε_{PA}	:	Security parameter contribution from privacy amplification
ε_{EV}	:	Security parameter contribution from error verification
Ω_{AT}	:	Event of passing the acceptance test
Ω_{EV}	:	Event of passing the error verification
$\Omega_{\text{acc}} = \Omega_{\text{AT}} \wedge \Omega_{\text{EV}}$	:	Event of the protocol not aborting

Protocol steps:

1. **State preparation and transmission:** Eve prepares the pure state $|\psi\rangle_{A_1^n B_1^n \tilde{E}}$. The registers A_1^n, B_1^n and $\tilde{E}$ are given to Alice, Bob and Eve, respectively.

2. For each round $i \in \{1, 2, \ldots, n\}$, Alice and Bob perform the following steps:

   (a) **Measurements:** Alice and Bob measures their received states described by the POVMs with POVM elements $\{M_k^A\}_{k=1\ldots d_A}$ and $\{M_k^B\}_{k=1\ldots d_B}$, respectively. They store their results in classical registers X_i and Y_i with alphabets $\mathcal{X}$ and $\mathcal{Y}$. Furthermore, they compute a classical registers C_i^A and C_i^B with alphabets $\mathcal{C}^A$ and $\mathcal{C}^B$, respectively, for public announcement.

The rest of the protocol is identical to Protocol 1.

3.2.1 Decoy-State Protocols

In practical implementations of QKD protocols, weak coherent pulse (WCP) sources are often used due to their common availability. These sources emit fully phase-randomized weak coherent pulses, which are coherent mixture of n-photon states. However, WCP sources come with a caveat as they are susceptible to the photon-number-splitting (PNS) attack [9]. In such an attack, an adversary could gain full information about any multi-photon state ($n > 1$) without revealing themselves. Several techniques have been developed to mitigate this issue, one of which is the use of decoy-state protocols, first introduced in Refs. [11, 13, 14].

In decoy-state protocols, the goal is to bound Eve's influence on the single-photon signals, as only these are secure. To achieve this, Alice sends each signal with multiple different intensities. The intensity used for key generation is typically called the signal intensity, while all other intensities are termed decoy intensities. Some versions of decoy-state protocols require only a subset of all signals to be sent with the decoy intensities, as will be discussed below. Due to the structure of the states emitted by fully phase-randomized WCP sources, Eve cannot distinguish the signal intensity chosen for a particular signal by merely observing the photon number of that signal. Thus, Eve's optimal attack strategy is to base her attack purely on the photon number. The additional observations from the decoy intensities then allow one to bound Eve's influence on the single-photon signals.

Below, we present a decoy-state version of the P&M protocol defined in the previous section. Again, Alice uses a source emitting fully phase-randomized weak coherent pulses and sends a subset of her signals with different intensities. Signals from WCP sources are mixed states, which is why we allowed mixed states to be sent out in the P&M protocol above. The state preparation and announcements still fit the general framework laid out in Protocol 1, but here we provide slightly more details. In terms of the protocol steps, we replace step 1. (a) of Protocol 1 with (a') stated below.

Protocol 3 (Decoy-State Prepare-and-Measure Protocol)**.** All parameters are chosen as in Protocol 1. The intensity μ_s stands for the signal intensity and is used primarily for key generation and $\{\mu_2, \mu_3, \dots\}$ are the decoy intensities. For a decoy-state version of Protocol 1, only step (1.) is replaced with step (1′.) from below. Furthermore, Alice additionally announces the particular intensity choice used for each signal sent, which is still contained in register C_A.

1′. **State preparation and transmission:** In each round Alice decides with probability $1 - \gamma$ whether she sends a test or generation round. In the case of a generation round

Alice prepares one of d_A states $\{|\sigma_k\rangle\}_{k=1\ldots d_A}$ with intensity μ_s. For test rounds Alice selects the intensity μ_i with probability $p(\mu_i|\text{test})$. She stores a label for her signal state in register X_i. Depending on her choices Alice computes her announcement and stores it in register C_A. Finally, Alice sends the signal state to Bob via a quantum channel.

Remark 3.2.1. Results from Refs. [22–24] require all intensities to be sent in key generation rounds, potentially resulting in suboptimal rates. In contrast, all of our proof techniques presented in this thesis allow us to consider protocols in which *only* the signal intensity is used in key generation rounds.

3.3 Source Replacement Scheme

The source replacement scheme [56, 57] is very common technique to map a P&M protocol to an equivalent entanglement based protocol. We will now describe the construction in more detail. Assume Alice prepares the states

$$\{|\phi_x\rangle_{A'}\}_{x=1\ldots d_A}, \tag{3.3.1}$$

each with probability p_j and sends them through a quantum channel $\mathcal{E} \in \text{CPTP}(A', B)$. If Alice stores here signal choice in a classical register X the state after the channel would have the form

$$\nu_{XB} = \sum_x |x\rangle\langle x|_X \otimes \mathcal{E}\left(|\phi_x\rangle\langle\phi_x|_{A'}\right). \tag{3.3.2}$$

Alternatively, Alice could prepare the pure state

$$|\psi\rangle_{AA'} = \sum_x \sqrt{p_x}\,|x\rangle_A\,|\phi_x\rangle_{A'}\,, \tag{3.3.3}$$

keep her system A and send the system A' through the channel $\mathcal{E}$ to Bob. This results in the state

$$\rho_{AB} = \left(\mathcal{I} \otimes \mathcal{E}\right)\left(|\psi\rangle\langle\psi|_{AA'}\right), \tag{3.3.4}$$

which we will call the pre-measurement state, as will become obvious soon.

Now, if Alice measures her system A in the computational basis, $\{|x\rangle_A\}$, the post measurement state posses the correct probability distribution and encodes the correct state in Bob's system. To see this assume Alice measures $|y\rangle\langle y|$ which yields

$$\text{Tr}_A[|y\rangle\langle y|_A\,\rho_{AB}] = p(y)\mathcal{E}\left(|\phi_y\rangle\langle\phi_y|_{A'}\right), \quad \text{and} \quad \text{Tr}[|y\rangle\langle y|_A\,\rho_{AB}] = p(y). \tag{3.3.5}$$

Thus, in this equivalent picture Alice still sends the state $|\phi_y\rangle$ with probability $p(y)$ to Bob and one could represent a P&M protocol as an entanglement based protocol through this construction.

However, it is important to note that Alice's system is required to remain untouched as it will never pass through the channel $\mathcal{E}$. Hence, if one aims to minimize the key rate over the pre-measurement state ρ_{AB} one needs to maintain this requirement which is typically done by demanding that

$$\mathrm{Tr}_B[\rho_{AB}] = \mathrm{Tr}_{A'}[|\psi\rangle\langle\psi|_{AA'}]. \tag{3.3.6}$$

Therefore, due to the source replacement scheme for P&M protocols, in the security Definitions 3.1.1 and 3.1.2 we can restrict the input states $\sigma_{A^nB^n}$ such that

$$\mathrm{Tr}_B^n\left[\sigma_{A^nB^n}\right] = \bigotimes_{i=1}^{n} \tau_A^{(i)}, \tag{3.3.7}$$

for some states $\tau_A^{(i)}$.

3.4 Shield Systems

A concept related to the source replacement scheme is that of shield systems [58], often used when Alice sends mixed states. One notable example are decoy state protocols, where Alice's source emits weak coherent pulses, which are mixed states.

If Alice's source emits mixed states, one can purify those states using a shield system. For instance, consider the scenario where Alice sends the states σ_x for $x = 1, \ldots, d_A$. One can define a special purifying system A_S with dimension $\geq \dim(A')$ such that each state σ_x can be purified. This allows us to write

$$|s_x\rangle_{A_SA'}, \tag{3.4.1}$$

with $\mathrm{Tr}_{A_S}[|s_i\rangle\langle s_i|] = \sigma_x$. Then, applying the source replacement scheme, one can view Alice's state preparation as her preparing the state

$$|\psi\rangle_{AA_SA'} = \sum_{x=1}^{d_A} \sqrt{p_x}\,|x\rangle_A\,|s_x\rangle_{A_SA'}, \tag{3.4.2}$$

and measuring her system A with the POVM elements $|x\rangle\langle x|$ as described above. The shield system is unique because it remains inaccessible to both Alice and Eve after the state preparation.

3.5 Leftover Hashing Lemmas

Leftover hashing lemmas (LHLs) such as [51, Corollary 5.6.1] or [38, Theorem 9] are essential parts for many QKD security proofs. They connect the extractable key in the secrecy definition Eq. (3.1.4) with an appropriate conditional entropy. Thus, intuitively they give us a measure of how long the final secret key can be.

This final secret key will be the output of a hash function chosen at random. Typically one uses a two-universal hash function. Thus, we start by defining them below.

Definition 3.5.1 (Two-Universal Hashing)**.** A family $\mathcal{F}_{\mathcal{A}\to\mathcal{Z}}$ of functions from a set $\mathcal{A}$ to $\mathcal{Z}$ with distribution $P_{\mathcal{F}}$ is called 2-universal if for all $x, x' \in \mathcal{A}$ with $x \neq x'$ it holds

$$\Pr_{\mathcal{F}_{\mathcal{A}\to\mathcal{Z}}}[f(x) = f(x')] \leq \frac{1}{|\mathcal{Z}|}. \tag{3.5.1}$$

In preparation for the leftover hashing lemmas, let us define the channel $\mathcal{R} : AS^H \to AS^H Z$ for any state $\rho_{AS^H} \in \mathrm{S}_{\leq}(AS^H)$ as

$$\mathcal{R}[\rho] \coloneqq \sum_{\substack{a\in\mathcal{A},\\ h\in\mathcal{F}}} |h(a)\rangle\langle h(a)|_Z \otimes |a,h\rangle\langle a,h|_{AS^H}\, \rho\, |a,h\rangle\langle a,h|_{AS^H}\,, \tag{3.5.2}$$

and the channel $\mathcal{R}_h : A \to Z$ for any ρ_A and $h \in \mathcal{F}$ as

$$\mathcal{R}_h[\rho] \coloneqq \sum_{a\in\mathcal{A}} |h(a)\rangle\langle h(a)|_Z \langle a|\, \rho_A\, |a\rangle\,. \tag{3.5.3}$$

Depending on the choice of hash function stored in S^H, the channel $\mathcal{R}$ applies this hash function to every element in A and stores the output in a separate register Z. Equivalently, the channel $\mathcal{R}_h$ applies the particular hash function h to all elements in A and creates the new register Z. Finally, let us also define the channel $\mathcal{U}$, creating a perfect uniform key

$$\mathcal{U}[\rho] = \frac{\mathbb{I}_Z}{|\mathcal{Z}|}\,\mathrm{Tr}[\rho]. \tag{3.5.4}$$

With these definitions at hand we can state the following LHL based on the smooth-min entropy. This lemma was first proven in [51, Corollary 5.6.1] and was later generalized and streamlined in Ref. [59, Theorem 6] by using the modern definition of purified distance for smoothing, which we also employ in this thesis. However, Ref. [59] uses a slightly different distance measure from a uniform distribution than the one required for the QKD security statements in Definition 3.1.1 and Definition 3.1.2. Thus, we present the version from Ref. [37], which uses both the purified distance for smoothing and the appropriate distance measure from a uniform distribution.

Theorem 3.5.2 (Smooth-min Leftover Hashing Lemma (Adapted from [37, Proposition 9])). *Let $\sigma_{AE} \in \mathrm{S}_{\leq}(AE)$ be a classical-quantum state, and $\varepsilon \in \left[0, \sqrt{\mathrm{Tr}[\sigma_{AE}]}\right)$. Let $\mathcal{F}_{\mathcal{A}\to\mathcal{Z}}$ be a set of two-universal hash functions with $\mathcal{Z} = \{0,1\}^l$, $l \in \mathbb{N}_0$. Moreover, let $\rho_{S^H} = \frac{1}{|\mathcal{F}|}\sum_{h\in\mathcal{F}_{\mathcal{A}\to\mathcal{Z}}} |h\rangle\langle h|_{S^H}$, which corresponds to choosing the hash function h uniformly at random. Then, we have*

$$\begin{aligned}\frac{1}{2}\,\mathbb{E}_h \left\|\left(\mathcal{R}_h - \mathcal{U}\right)(\sigma_{AE})\right\|_1 &= \frac{1}{2}\left\|\omega_{ZS^HE} - \frac{1}{|\mathcal{Z}|}\mathbb{I}_Z \otimes \omega_{S^HE}\right\| \\ &\leq \frac{1}{2}2^{-\frac{1}{2}\left(H_{\min}^{\varepsilon}(A|E)_\sigma - l\right)} + 2\varepsilon,\end{aligned} \tag{3.5.5}$$

where $\omega_{ZS^HE} = \mathrm{Tr}_A\left[\mathcal{R}\left(\sigma_{AE}\otimes\rho_{S^H}\right)\right]$.

Proof. See [37, Appendix B]. □

As we will see later, this LHL enables one to prove security for QKD protocols by finding a lower bound on the smooth min-entropy. In contrast to this original version of the LHL, the Rényi version states similar bounds in terms of a Rényi entropy and thus, enables one to prove security of QKD protocols fully based on Rényi entropies.

The leftover hashing lemma in Ref. [38] was written for families of λ-randomizing channels $\mathcal{R}_h$ defined in terms of hash functions h. Therefore, let us formally define those λ-randomizing channels.

Definition 3.5.3 (Randomizing family of hash functions (adapted from [38, Definition 4])). A family of channels $\{\mathcal{R}_h : h \in \mathcal{F}_{\mathcal{A}\to\mathcal{Z}}\}$ with $\mathcal{F}_{\mathcal{A}\to\mathcal{Z}}$ a family of hash functions with probability distribution p, is called λ-randomizing if for all classical-quantum states $\rho_{AE} \in L(AE)$, it holds

$$\mathbb{E}_h \left\|\left(\mathcal{R}_h - \mathcal{U}\right)(\rho_{AE})\right\|_1 \leq \lambda \|\rho_{AE}\|_1. \tag{3.5.6}$$

In Ref. [38, Theorem 9], the following leftover hashing lemma was shown.

Theorem 3.5.4 (Rényi Leftover Hashing Lemma (Adapted from [38, Theorem 9])). *Let $\{\mathcal{R}_h : h \in \mathcal{F}_{\mathcal{A}\to\mathcal{Z}}\}$ be a family of λ-randomizing hash functions and let $\sigma_{AE} \in \mathrm{S}_{=}(AE)$ be a classical-quantum state. Then, for $\alpha \in (1,2)$ we have*

$$\frac{1}{2}\,\mathbb{E}_h \left\|\left(\mathcal{R}_h - \mathcal{U}\right)(\sigma_{AE})\right\|_1 \leq 2^{\frac{2(1-\alpha)}{\alpha}} 2^{\frac{1-\alpha}{\alpha}\left(\widetilde{H}_\alpha^{\uparrow}(A|E)_\sigma - 2\log(\lambda) - \log(|\mathcal{Z}|)\right)}. \tag{3.5.7}$$

Proof. See [38, Theorem 9]. □

Additionally, Ref. [38] proves that two-universal hash functions as 1-randomizing, but the definition for two-universal hash functions was chosen with an equality instead of an inequality. Additionally, we note that the proof of [38, Theorem 9] only requires the λ-randomizing property for normal operators. Therefore, we show that two-universal hash functions defined as in Definition 3.5.1 are λ-randomizing on the set of normal operators.

Lemma 3.5.5 (Two-universal hashing is $\sqrt{2}$-randomizing)**.** *Any family of two-universal hash functions is $\sqrt{2}$-randomizing on the set of normal operators.*

Proof. For brevity, let us define the hermitian preserving map $\mathcal{N}_h$ as $\mathcal{N}_h \coloneqq \mathcal{R}_h - \mathcal{U}$. Then, Lemma 5 of [38] showed that if

$$\mathbb{E}_h \|\mathcal{N}_h(\rho_A)\|_2^2 \leq \lambda^2 \|\rho_A\|_2^2 \tag{3.5.8}$$

for any $\rho_A \in L(A)$, then the family of hash functions is λ-randomizing. The same holds if we restrict our input set to normal operators and, thus we use this to prove our case.

Next, we note that the proof of [38, Lemma 7] is still valid for positive semidefinite operators when we use Definition 3.5.1 for two-universal hashing. The only step that needs to be altered is the step bounding the collision probability $\Pr[h(a) = h(a')]$ for $a \neq a'$. For $\rho_A = \sum_a p(a) |a\rangle\langle a| \in \mathrm{Pos}(A)$ the exact step reads

$$\sum_{a,a' \in \mathcal{A}} \Pr[h(a) = h(a')] p(a) p(a') \leq \sum_a p(a)^2 + \frac{1}{|\mathcal{Z}|} \sum_{a \neq a'} p(a) p(a'), \tag{3.5.9}$$

which still holds since all $p(a)$ are non-negative. Thus, we have the following intermediate result

$$\mathbb{E}_h \|\mathcal{N}_h(\rho_A)\|_2^2 \leq \|\rho_A\|_2^2, \tag{3.5.10}$$

for all $\rho_A \in \mathrm{Pos}(A)$.

The rest of the proof progresses with the following steps. First, we extend the result to hermitian operators and then we extend the claim further to normal operators.

Thus, let $M \in \mathrm{Herm}(A)$ be arbitrary. By the Jordan-decomposition [60, Proposition 4.1.13.] one can write M as $M = M_+ - M_-$, where both $M_\pm$ are positive semidefinite. Thus, we find

$$\|\mathcal{N}_h(M)\|_2^2 = \|\mathcal{N}_h(M_+) - \mathcal{N}_h(M_-)\|_2^2 \leq 2\|\mathcal{N}_h(M_+)\|_2^2 + 2\|\mathcal{N}_h(M_-)\|_2^2, \tag{3.5.11}$$

where we used the triangle inequality together with $2ab \leq a^2 + b^2$ for any real a, b. Hence, since the expectation value is linear, we find

$$\mathbb{E}_h \|\mathcal{N}_h(M)\|_2^2 \leq 2 \left(\mathbb{E}_h \|\mathcal{N}_h(M_+)\|_2^2 + \mathbb{E}_h \|\mathcal{N}_h(M_-)\|_2^2 \right). \tag{3.5.12}$$

Now, using Eq. (3.5.10) we can bound each of the terms to find

$$2\left(\mathbb{E}_h \|\mathcal{N}_h(M_+)\|_2^2 + \mathbb{E}_h \|\mathcal{N}_h(M_-)\|_2^2\right) \leq 2\left(\|M_+\|_2^2 + \|M_-\|_2^2\right) = 2\|M\|_2^2, \tag{3.5.13}$$

where in the last step we used $\|M_+\|_2^2 + \|M_-\|_2^2 = \|M\|_2^2$ by properties of the 2-norm. Thus, we reach the next intermediate result

$$\mathbb{E}_h \|\mathcal{N}_h(M_A)\|_2^2 \leq 2\|M_A\|_2^2, \tag{3.5.14}$$

for all $M_A \in \mathrm{Herm}(A)$.

Finally, consider $T \in L(A)$ a normal operator. Due to the Toeplitz decomposition [60, Theorem 4.1.2], we can decompose T as

$$T = A + iB, \tag{3.5.15}$$

where both A and B are hermitian operators and commute. Additionally, we note that $[\mathcal{N}_h(A), \mathcal{N}_h(B)] = [\mathcal{R}_h(A), \mathcal{R}_h(B)] = 0$, since $\mathcal{R}_h$ maps any operator to a diagonal operator and thus they commute.

Therefore, a simple calculation shows

$$\mathcal{N}_h(T)^\dagger \mathcal{N}_h(T) = \mathcal{N}_h(A)^\dagger \mathcal{N}_h(A) + \mathcal{N}_h(B)^\dagger \mathcal{N}_h(B), \tag{3.5.16}$$

since all other terms cancel and hence

$$\|\mathcal{N}_h(T)\|_2^2 = \|\mathcal{N}_h(A)\|_2^2 + \|\mathcal{N}_h(B)\|_2^2. \tag{3.5.17}$$

Now, we can apply our second intermediate result for hermitian operators Eq. (3.5.14) to find

$$\mathbb{E}_h \|\mathcal{N}_h(T)\|_2^2 = \mathbb{E}_h \|\mathcal{N}_h(A)\|_2^2 + \mathbb{E}_h \|\mathcal{N}_h(B)\|_2^2 \tag{3.5.18}$$

$$\leq 2\left(\|A\|_2^2 + \|B\|_2^2\right) = 2\|T\|_2^2, \tag{3.5.19}$$

where for the final equality we used [61, Corollary 1]. Therefore, in summary, we showed

$$\mathbb{E}_h \|(\mathcal{R}_h - \mathcal{U})(\rho_A)\|_2^2 \leq 2\|\rho_A\|_2^2 \tag{3.5.20}$$

for all normal operators ρ_A in $L(A)$ and thus by Lemma 5 of [38] two-universal hash functions are $\sqrt{2}$-randomizing on the set of normal operators. □

Now, that we established that even with the weaker Definition 3.5.1 of two-universal hashing we maintain the λ-randomizing property, although with a different factor, we can state the following corollary.

Corollary 3.5.6 (Rényi Leftover Hashing Lemma for Two-Universal Hashing)**.** *Let σ_{AE} be a classical-quantum state, and $\mathcal{F}_{\mathcal{A}\to\mathcal{Z}}$ be a set of two-universal hash functions with $\mathcal{Z} = \{0,1\}^l$, $l \in \mathbb{N}_0$. Moreover, let $\rho_{S^H} = \frac{1}{|\mathcal{F}|}\sum_{h\in\mathcal{F}_{\mathcal{A}\to\mathcal{Z}}} |h\rangle\langle h|_{S^H}$, which corresponds to choosing the hash function h uniformly at random. Then, for $\alpha \in (1,2)$ we have*

$$\begin{aligned}\frac{1}{2}\,\mathbb{E}_h \left\|\left(\mathcal{R}_h - \mathcal{U}\right)(\sigma_{AE})\right\|_1 &= \frac{1}{2}\left\|\omega_{ZES^H} - \frac{1}{|\mathcal{Z}|}\mathbb{I}_Z \otimes \omega_{ES^H}\right\|_1 \\ &\leq 2^{\frac{2(1-\alpha)}{\alpha}} 2^{\frac{1-\alpha}{\alpha}\left(\widetilde{H}^{\uparrow}_{\alpha}(A|E)_\sigma - 1 - l\right)} = 2^{\frac{1-\alpha}{\alpha}\left(\widetilde{H}^{\uparrow}_{\alpha}(A|E)_\sigma + 1 - l\right)},\end{aligned} \tag{3.5.21}$$

where again $\omega_{ZS^HE} = \mathrm{Tr}_A\left[\mathcal{R}\left(\sigma_{AE}\otimes\rho_{S^H}\right)\right]$.

Proof. This follows immediately from [38] inserting

$$2\log(\lambda) = 2\log\left(\sqrt{2}\right) = 1. \tag{3.5.22}$$

□

Remark 3.5.7. As shown in Corollary 3.5.6, two-universal hash functions satisfying only the inequality in Definition 3.5.1 remain applicable for privacy amplification, albeit with an additional one-bit loss in the resulting secret key length.

Chapter 4

Squashing Maps

In this chapter we highlight the relevant background for squashing methods in preparation for the following chapter in which we derive a version of the flag-state squasher for detection setups comprised passive linear optics.

4.1 Preliminaries

For a detection setup used in an optical implementation of QKD, one can represent the measurements by a POVM on an infinite-dimensional Hilbert space. Here we will identify this infinite-dimensional with $\mathcal{H}_{B_M}$. A squashing map acts on $\mathcal{H}_{B_M}$ and maps it to a finite-dimensional Hilbert space $\mathcal{H}_B$, which can be a qubit space, but in general will not. Mathematically, a squashing map Λ acts as

$$\Lambda : \mathrm{S}_{=}\left(\mathcal{H}_{B_M}\right) \to \mathrm{S}_{=}\left(\mathcal{H}_B\right), \tag{4.1.1}$$

where now $\mathcal{H}_B$ is *finite-dimensional*. The motivation behind such a mapping could be manifold. Most notably for numerical security proofs, a squashing map is required in order to treat the infinite-dimensional full space $\mathcal{H}_{B_M}$ appropriately. For analytical security proofs a squashing map could be used to reduce the problem to previously shown security of a qubit protocol.

Now, let us be more precise in what we require from a squashing map. Hence, for the discussion of this section, let us define the POVM elements $F^{(k)}_{B_M}$ and $F^{(k)}_{B}$ indexed by the outcomes k acting on $\mathcal{H}_{B_M}$ and $\mathcal{H}_B$, respectively. We require that any potential squashing

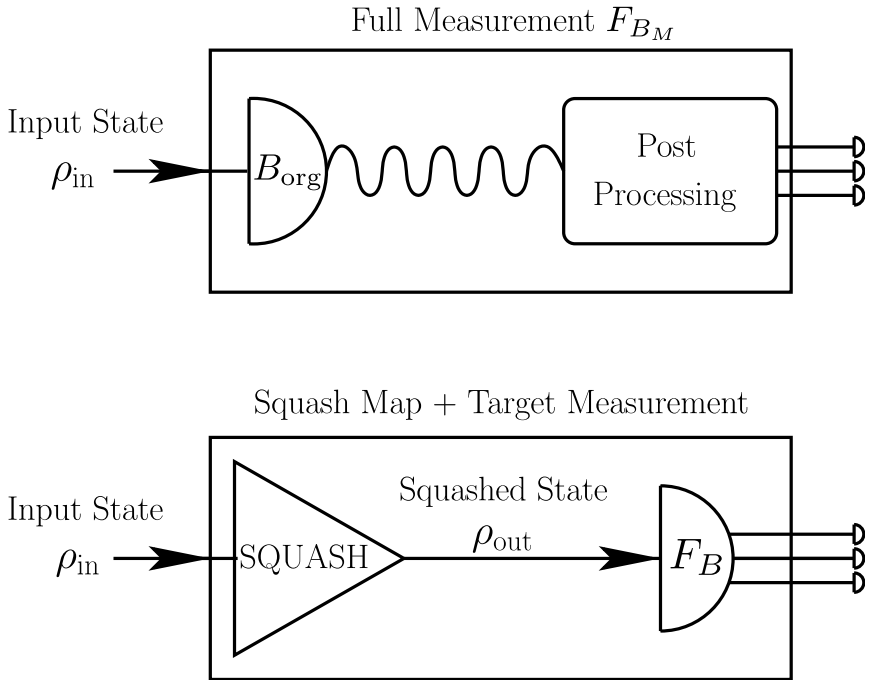


Figure 4.1: Schematic representation of the qubit squasher from Ref. [62, 63]. (Figure adapted from [63])

map Λ keeps the observations invariant under its action, i.e.

$$\text{Tr}\left[\rho_{\text{in}} F_{B_M}^{(k)}\right] = \text{Tr}\left[\Lambda(\rho_{\text{in}}) F_B^{(k)}\right], \tag{4.1.2}$$

for all input states ρ_{in} and indices k. One can equivalently phrase this as an requirement on the dual map $\Lambda^\dagger$, which CP and unital,

$$F_{B_M}^{(k)} = \Lambda^\dagger\left(F_B^{(k)}\right), \tag{4.1.3}$$

for all k. In Refs. [62, 63] a method for verifying the existence of a squashing map was derived.

4.2 Qubit Squasher

In Refs.[62, 63] a qubit-based squashing model was derived and the main idea is represented in Fig. 4.1. As stated before, an optical detection setup is represented by infinite dimensional POVM elements. In line with the previous discussion, let us denote the original or

basic POVM corresponding to all possible outcomes as $F_{B_{\text{org}}}$ and it is represented by $F_{B_{\text{org}}}$ in Fig. 4.1. These basic outcomes often undergo a classical post-processing $\mathcal{P}$ and for the existence of this squashing map are even required to be post-processed. Let us denote the effective *full* measurement after the post-processing as F_{B_M},

$$F_{B_M}^{(k)} = \sum_l \mathcal{P}_{kl} F_{B_{\text{org}}}^{(l)}. \tag{4.2.1}$$

Next, as stated above, we want to squash this full measurement F_{B_M} such that we can describe the statistics with an equivalent target POVM F_B on a finite-dimensional subspace such that Eqs. (4.1.2) and (4.1.3) are satisfied. This immediately tells us that the post-processing needs to result in a full POVM F_{B_M} that has equal elements as the target POVM F_B. The target POVM F_B is often chosen such that it acts on a Hilbert space comprised of a qubit and a vacuum space, hence the name qubit squasher.

For optical implementations of QKD protocols using threshold detectors, the squashing maps presented in Refs. [62, 63] can be applied under certain conditions. For example, they exist for a biased and unbiased active 4-state analyzer, typically used in a BB84 protocol. However, for a six-state analyzer the squashing map does not exist. One remedy is possible, one can add 16.67% of random noise, which essentially makes this construction unusable in a QKD protocol.

For passive detection setups, the situation is slightly different. Here, the qubit squasher exists for both a 4-state and 6-state analyzer, but only for unbiased setups. This limitation is resolved by the flag-state squasher introduced in the following section. Moreover, for a perfectly unbiased configuration, it is again necessary to add noise to the 6-state setup [63, Section VII].

4.3 Flag-State Squasher

As mentioned above, the qubit squasher does not exist for every detection setup. Hence, one often uses the alternative of the flag-state squasher, as presented in [42]. This squashing map always exists, is much more flexible and allows the treatment of imperfections in the detection setup unlike the squashing map from [62, 63], which in general only exists in idealized scenarios. For example, in [42] the flag-state squasher was applied to a detection efficiency mismatch between detectors.

The flag-state squasher still applies the mapping of Eq. (4.1.1) from the full space $\mathcal{H}_{B_M}$ to some *finite-dimensional* target space $\mathcal{H}_B$, which is however not the qubit space anymore.

Similarly, it is still required that Eqs. (4.1.2) and (4.1.3) are satisfied for all input states ρ_{in} and POVM elements k.

This squashing map critically depends on Bob's POVM elements possessing a block-diagonal structure, i.e. all of Bob's POVM elements can be written as

$$F_{B_M}^{(k)} = F_{B_M,\leq N_B}^{(k)} \oplus F_{B_M,>N_B}^{(k)}, \tag{4.3.1}$$

and act upon a Hilbert space $\mathcal{H}_{B_M} = \mathcal{H}_{\leq N_B} \oplus \mathcal{H}_{>N_B}$, where N_B is a free parameter. Typically N_B represents a photon number cutoff for Bob's detection setup, which will be the case in this thesis. In general the block-diagonal structure could be in any other degree of freedom as well.

Exploiting the block-diagonal structure of Bob's full POVM $F_{B_M}^{(k)}$, the flag-state squasher leaves the subspace with less than or equal to N_B photons unchanged and replaces all contributions with higher photon numbers than N_B with a flag. Thus, an arbitrary input state

$$\rho = \begin{pmatrix} \rho_{ss} & \rho_{sl} \\ \rho_{ls} & \rho_{ll} \end{pmatrix} \tag{4.3.2}$$

gets transformed by Λ as

$$\Lambda(\rho) = \begin{pmatrix} \rho_{ss} & 0 \\ 0 & \sum_{k=1}^{J} \operatorname{Tr}\left[\rho_{ll} F_{B_M,>N_B}^{(k)}\right] |k\rangle\langle k| \end{pmatrix} \tag{4.3.3}$$

where, the flags $|k\rangle\langle k|$ for each observation k are orthogonal. Hence, all events corresponding to a photon number above N_B are known to Eve. Equivalently, the full POVM F_{B_M} is transformed as

$$F_{B_M}^{(k)} = F_{B_M,\leq N_B}^{(k)} \oplus F_{B_M,>N_B}^{(k)} \mapsto F_B^{(k)} = F_{B_M,\leq N_B}^{(k)} \oplus |k\rangle\langle k|, \tag{4.3.4}$$

for each k.

The existence of such a squashing map is formalized in the following theorem adapted from [42, Theorem 1].

Theorem 4.3.1 (Flag-State Squasher adapted from [42, Theorem 1])**.** *Suppose that we have a POVM with elements $F_{B_M}^{(k)}$, where $k \in \{1, \ldots, J\}$, such that each element can be written in a block-diagonal form $F_{B_M,\leq N_B}^{(k)} \oplus F_{B_M,>N_B}^{(k)}$, with an associated Hilbert space structure given by $\mathcal{H}_{\leq N_B} \oplus \mathcal{H}_{>N_B}$. Then there exists a completely positive trace preserving (CPTP) map Λ (referred to as a squashing map) from $\mathcal{H}_{\leq N_B} \oplus \mathcal{H}_{>N_B}$ to $\mathcal{H}_{\leq N_B} \oplus \mathcal{H}_J$, where*

the dimension $dim(\mathcal{H}_J) = J$*, such that* $\mathrm{Tr}\left(\rho F_{B_M}^{(k)}\right) = \mathrm{Tr}\left(\Lambda(\rho) F_B^{(k)}\right) \ \ \forall \rho \in \mathcal{H}_{\leq N_B} \oplus \mathcal{H}_{>N_B}$ *with*

$$F_B^{(k)} = F_{B_M, \leq N_B}^{(k)} \oplus |k\rangle\langle k| \,, \tag{4.3.5}$$

where the states $|k\rangle$ *form an orthonormal basis of* $\mathcal{H}_J$*.*

Proof. See the proof of Theorem 1 in Ref. [42]. □

Thus, the flag-state squasher always exists if Bob's POVM is block-diagonal. However, this construction is only useful for QKD if we have a non-trivial bound on the weight inside the $\leq N_B$ subspace, since Eve could always steer all photons into the $> N_B$-photon subspace and Bob would still observe the same statistics. In this case everything about Bob's observations would be known to Eve and one would achieve zero key rate. Therefore, a critical piece of the flag-state squasher is finding a tight lower bound on the weight inside the $\leq N_B$-photon subspace of Bob's system.

Chapter 5

Flag-State Squasher for Passive Linear Optical Detection Setups

This chapter is based on Section VII and Appendices A and B of our work in Ref. [1]. We derive a subspace bound for the flag-state squasher introduced in Section 4.3, applicable to arbitrary passive linear optical detection setups, including detector losses and dark counts. As previously discussed, the squashing methods of Refs. [62, 63] are limited to perfect setups. While the flag-state squasher presented in [42] can accommodate arbitrary detection setups, it requires sophisticated numerical optimizations to derive subspace bounds. Therefore, the motivation for deriving such a subspace bound lies in the necessity for squashing methods that are adaptable to general detection setups. Most importantly, our bound derived in this chapter can be obtained analytically or through a single singular value decomposition.

5.1 General Structure of Subspace Estimation

Later in this chapter, we use Bob's multi-click POVM element to bound the weight inside the subspace. Here, we describe how, given a suitable POVM element M, one can derive an analytical lower bound on the weight inside the $\leq N_B$-photon subspace of Bob's system, providing an analytical counterpart to the numerical results presented in Ref. [42]. As in Section 4.3, we assume that Bob's POVM elements are block-diagonal with respect to the Fock basis (i.e., one block per individual photon number). We denote by $M_{\tilde{n}}$ the $\tilde{n}$-photon block of M in terms of the photon number *arriving* at Bob's setup.

Theorem 5.1.1 (Subspace Bound for Generic Block-Diagonal POVM Element (adapted from [64, App. A])). *Let $M \in \mathrm{Pos}(\mathcal{H}_B)$ be a POVM element that is block-diagonal in the Fock basis, that is,*

$$M = \bigoplus_{\tilde{n}=0}^{\infty} M_{\tilde{n}}, \tag{5.1.1}$$

with the associated expectation value $m_{\mathrm{obs}|x}$ conditioned on the state ρ_x. Then, the probability $p(\tilde{n} \leq N_B|x)$ for this state ρ_x is lower-bounded by

$$p(\tilde{n} \leq N_B|x) \geq 1 - \frac{m_{\mathrm{obs}|x} - c_{\leq N_B}}{c_{\geq N_B+1} - c_{\leq N_B}} =: p_L(\tilde{n} \leq N_B|x), \tag{5.1.2}$$

where the constants $c_{\leq N_B}$ and $c_{\geq N_B+1}$ are defined as

$$c_{\leq N_B} := \min_{\tilde{n} \leq N_B} c_{\tilde{n}}, \qquad c_{\geq N_B+1} := \min_{\tilde{n} \geq N_B+1} c_{\tilde{n}}. \tag{5.1.3}$$

Proof. The expectation value $m_{\mathrm{obs}|x}$, conditioned on the state ρ_x, can be expressed in terms of the contributions of each photon number $\tilde{n}$ arriving at Bob's setup as follows:

$$m_{\mathrm{obs}|x} = \sum_{\tilde{n}=0}^{\infty} p(\tilde{n}|x) \Pr\left[m_{\mathrm{obs}}|\tilde{n}, x\right] = \sum_{\tilde{n}=0}^{\infty} p(\tilde{n}|x) \operatorname{Tr}\left[M_{\tilde{n}} \rho_x^{(\tilde{n})}\right], \tag{5.1.4}$$

where $\rho_x^{(\tilde{n})}$ is the normalized conditional state in the $\tilde{n}$-photon subspace. For each $\tilde{n} \in \mathbb{N}_0$, we define $c_{\tilde{n}}$ as a lower bound on $\operatorname{Tr}\left[M_{\tilde{n}} \rho^{(\tilde{n})}\right]$ for all possible states $\rho^{(\tilde{n})}$ in this subspace, such that $0 \leq c_{\tilde{n}} \leq \min_{\rho^{(\tilde{n})}} \operatorname{Tr}\left[M_{\tilde{n}} \rho^{(\tilde{n})}\right]$. Thus, the expectation value $m_{\mathrm{obs}|x}$ satisfies

$$m_{\mathrm{obs}|x} \geq \sum_{\tilde{n}=0}^{\infty} p(\tilde{n}|x) c_{\tilde{n}}. \tag{5.1.5}$$

Splitting this sum into contributions where $\tilde{n} \leq N_B$ and $\tilde{n} > N_B$, and subsequently rearranging for $p(\tilde{n} \leq N_B|x)$, yields the desired lower bound on the weight inside the subspace. □

Remark 5.1.2. We again highlight that we use $\tilde{n}$ to denote the photon number *arriving* at Bob's setup, which, due to Eve's interaction, may differ from Alice's photon number n.

Therefore, by applying the theorem above to the observation $m_{\mathrm{obs}|x}$ and the lower bounds $c_{\tilde{n}}$ corresponding to a given POVM element M, we can derive a lower bound on the weight within the $\leq N_B$-photon subspace. The following section details the remaining tasks: identifying a suitable POVM element M and determining the constants $c_{\tilde{n}}$.

5.2 Defining a Suitable POVM Element for Passive Linear Optical Setups

Now, we aim to choose and define a suitable detectable event and associated POVM element M for the subspace estimation, relying exclusively on passive linear optics in the detection setup. We assume threshold detectors here, but the analysis can still be used for number resolving detectors, although some changes are required.

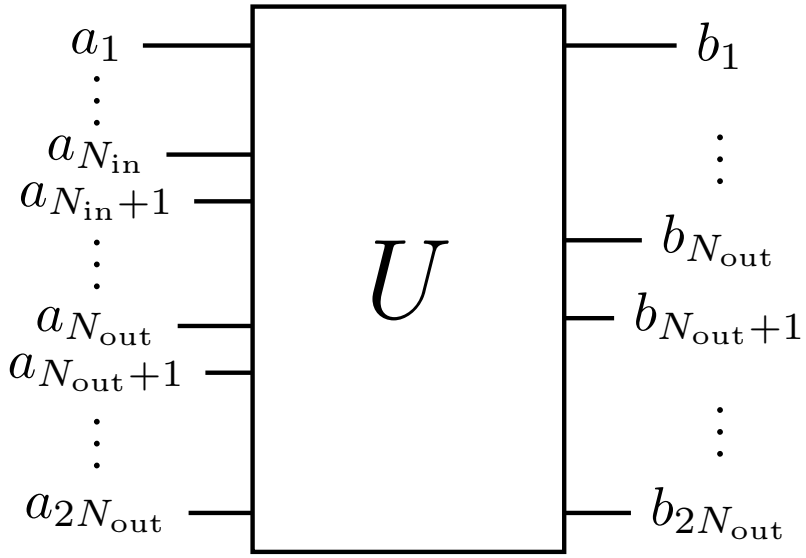


Figure 5.1: A detection setup using only passive linear optics, including *losses*, represented as a circuit implementing a linear mode transformation U. For a typical 4-state analyzer used in optical BB84 protocols, $N_{\text{in}} = 2$ represents the incoming signal modes, while $N_{\text{out}} = 8$ accounts for the four detectors and their corresponding loss modes.

We consider a generic detection setup comprised solely of passive linear optical elements. Therefore, let us assume we have N_{in} incoming (non-vacuum) signal modes and N_{out} modes corresponding to the modes measured in the detectors. Furthermore, we model any losses by additional beamsplitters splitting each outgoing mode into two, thus leading to another set of N_{out} modes corresponding to the losses of the respective detector (which are traced out and not measured). Hence, our model includes N_{in} incoming signal and $2N_{\text{out}}$ outgoing modes.

Since a generic passive linear optical setup is inherently described by a unitary transformation [65], the difference between the N_{in} signal inputs and the $2N_{\text{out}}$ total outputs

must be accounted for by physical vacuum modes. By including these $2N_{\text{out}} - N_{\text{in}}$ incoming *vacuum* modes, which enter through the unused ports of the setup, we can represent the complete detection setup by the linear mode transformation

$$b_i = \sum_{j=1}^{2N_{\text{out}}} U_{ij} a_j, \tag{5.2.1}$$

where $\{a_j\}$ are the incoming and $\{b_i\}$ are the outgoing modes, respectively, and U is a unitary matrix. A schematic representation of such a mode transformation can be seen in Fig. 5.1. For example, if polarization modes span the input space, then a_H and a_V, the annihilation operators of the horizontal and vertical polarization, would be the only incoming (non-vacuum) mode operators.

Hereafter, we treat the outgoing modes b_i as functions of the incoming modes a_j, determined by the linear mode transformation given in Eq. (5.2.1).

For simplicity, we first consider ideal detectors with no losses. This allows us to disregard both the additional outgoing modes associated with losses and the incoming vacuum modes $N_{\text{out}} + 1, \ldots, 2N_{\text{out}}$ also associated with detector losses. We label the detected outgoing modes as b_1 to $b_{N_{\text{out}}}$, and the corresponding incoming modes as a_1 to $a_{N_{\text{out}}}$, where N_{in} of these are non-vacuum. This scenario is illustrated in Fig. 5.2.

Then, assuming threshold detectors, a single click in only detector i, expressed in terms of the outgoing modes, is given by the POVM element

$$F_{\text{sc}}^i = \sum_{n=1}^{\infty} \frac{1}{n!} \left(b_i^\dagger\right)^n |0\rangle\langle 0| \left(b_i\right)^n, \tag{5.2.2}$$

and the POVM element of a no-click event in any detector is simply

$$F_{\text{no}} = |0\rangle\langle 0|. \tag{5.2.3}$$

We observe that these POVM elements implicitly depend on the unitary matrix U, since the outgoing modes b_i are determined by the transformation of the incoming modes via U. Additionally, we note that only the components of U acting on the incoming signal modes are physically relevant to our analysis, which we make rigorous in the following lemma. For this purpose, let us partition the unitary matrix U of a passive linear optical detection setup as

$$U = \left(\begin{array}{c|c} G & U' \end{array} \right). \tag{5.2.4}$$

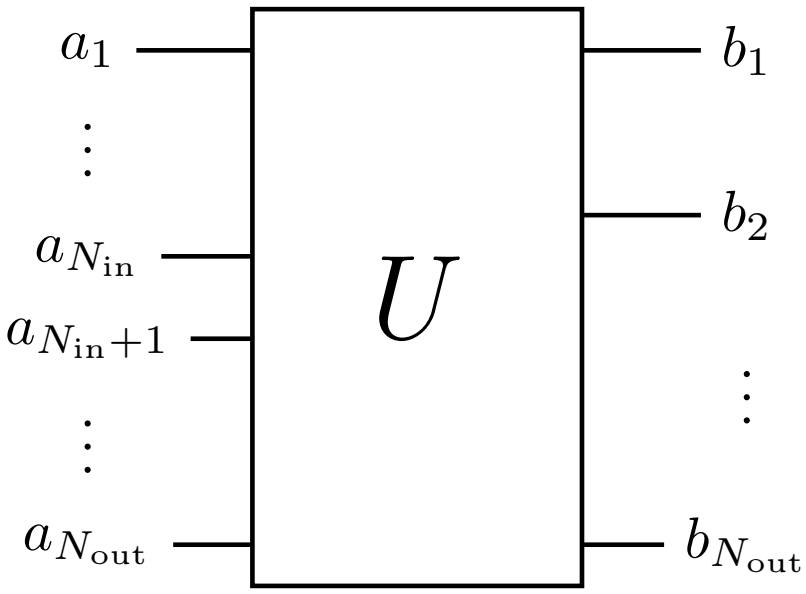


Figure 5.2: A perfect detection setup using only passive linear optics represented as a circuit implementing a linear mode transformation U. (Figure 5 from Ref. [1])

Lemma 5.2.1. *Assume the unitary matrix U of a passive linear optical detection setup is partitioned as*

$$U = \left(\begin{array}{c|c} G & U' \end{array} \right). \tag{5.2.5}$$

Then, for any state ρ defined exclusively on the incoming signal modes N_{in} (with vacuum in all other modes), the expectation value of the sum single clicks with $\tilde{n}$ photons depends only on the $N_{\text{out}} \times N_{\text{in}}$ submatrix G, that is,

$$\begin{aligned} &\operatorname{Tr}\left[\sum_{i=1}^{N_{\text{out}}} \frac{1}{\tilde{n}!}\left(b_i^\dagger\right)^{\tilde{n}} |0\rangle\langle 0| \left(b_i\right)^{\tilde{n}} \left(\rho \otimes |0\rangle\langle 0|\right)\right] \\ &\quad = \sum_{i=1}^{N_{\text{out}}} \operatorname{Tr}\left[\frac{1}{\tilde{n}!}\left(\sum_{j=1}^{N_{\text{in}}} G_{ij}^* a_j^\dagger\right)^{\tilde{n}} |0\rangle\langle 0| \left(\sum_{k=1}^{N_{\text{in}}} G_{ik} a_k\right)^{\tilde{n}} \rho\right]. \end{aligned} \tag{5.2.6}$$

Proof. Consider the expectation value of a single term for an arbitrary state ρ defined only on the incoming signal modes,

$$\operatorname{Tr}\left[\frac{1}{\tilde{n}!}\left(b_i^\dagger\right)^{\tilde{n}} |0\rangle\langle 0| \left(b_i\right)^{\tilde{n}} \left(\rho \otimes |0\rangle\langle 0|\right)\right]. \tag{5.2.7}$$

Expanding this expression using Eq. (5.2.1) and the multinomial theorem yields

$$
\begin{aligned}
&\mathrm{Tr}\left[\frac{1}{\tilde{n}!}\left(\sum_{j=1}^{N_{\mathrm{out}}} U_{ij}^{*} a_j^{\dagger}\right)^{\tilde{n}} |0\rangle\langle 0| \left(\sum_{k=1}^{N_{\mathrm{out}}} U_{ik} a_k\right)^{\tilde{n}} (\rho \otimes |0\rangle\langle 0|)\right] \\
&= \frac{1}{\tilde{n}!} \sum_{\substack{k_1+\cdots+k_{N_{\mathrm{out}}}=\tilde{n} \\ l_1+\cdots+l_{N_{\mathrm{out}}}=\tilde{n}}} \binom{\tilde{n}}{k_1,\ldots,k_{N_{\mathrm{out}}}} \binom{\tilde{n}}{l_1,\ldots,l_{N_{\mathrm{out}}}} \\
&\quad \cdot \mathrm{Tr}\left[\prod_{t,s=1}^{N_{\mathrm{out}}} \left(U_{it}^{*} a_t^{\dagger}\right)^{k_t} |0\rangle\langle 0| \left(U_{is} a_s\right)^{l_s} (\rho \otimes |0\rangle\langle 0|)\right] \qquad (5.2.8) \\
&= \sum_{\substack{k_1+\cdots+k_{N_{\mathrm{out}}}=\tilde{n} \\ l_1+\cdots+l_{N_{\mathrm{out}}}=\tilde{n}}} \sqrt{\binom{\tilde{n}}{k_1,\ldots,k_{N_{\mathrm{out}}}} \binom{\tilde{n}}{l_1,\ldots,l_{N_{\mathrm{out}}}}} \\
&\quad \cdot \mathrm{Tr}\left[\prod_{t,s=1}^{N_{\mathrm{out}}} \left(U_{it}^{*}\right)^{k_t} U_{is}^{l_s} |k_1,\ldots,k_{N_{\mathrm{out}}}\rangle\langle l_1,\ldots,l_{N_{\mathrm{out}}}| (\rho \otimes |0\rangle\langle 0|)\right] \qquad (5.2.9) \\
&= \mathrm{Tr}\left[\frac{1}{\tilde{n}!}\left(\sum_{j=1}^{N_{\mathrm{in}}} U_{ij}^{*} a_j^{\dagger}\right)^{\tilde{n}} |0\rangle\langle 0| \left(\sum_{k=1}^{N_{\mathrm{in}}} U_{ik} a_k\right)^{\tilde{n}} \rho\right]. \qquad (5.2.10)
\end{aligned}
$$

Consequently, only the first N_{in} columns of U, i.e. G, are relevant, and for any ρ it holds that

$$
\begin{aligned}
&\mathrm{Tr}\left[\sum_{i=1}^{N_{\mathrm{out}}} \frac{1}{\tilde{n}!}\left(b_i^{\dagger}\right)^{\tilde{n}} |0\rangle\langle 0| \left(b_i\right)^{\tilde{n}} (\rho \otimes |0\rangle\langle 0|)\right] \\
&\quad = \sum_{i=1}^{N_{\mathrm{out}}} \mathrm{Tr}\left[\frac{1}{\tilde{n}!}\left(\sum_{j=1}^{N_{\mathrm{in}}} G_{ij}^{*} a_j^{\dagger}\right)^{\tilde{n}} |0\rangle\langle 0| \left(\sum_{k=1}^{N_{\mathrm{in}}} G_{ik} a_k\right)^{\tilde{n}} \rho\right].
\end{aligned} \qquad (5.2.11)
$$

□

This lemma is essential for the subspace estimation, as demonstrated below. In the ideal, lossless case, we use the multi-click POVM element M^{mult} for the subspace estimation, which we define as

$$
M^{\mathrm{mult}}(G) \coloneqq \mathbb{I} - F_{\mathrm{no}} - \sum_{i=1}^{N_{\mathrm{out}}} F_{\mathrm{sc}}^{i}. \qquad (5.2.12)
$$

Assuming a photon-number cutoff of $N_B = 1$ on Bob's side (i.e., targeting single-photon signals), we choose M^{mult} because it is only triggered by multi-photon events. Most importantly, by Lemma 5.2.1, this POVM element depends exclusively on G.

Next, following Theorem 5.1.1, the subspace estimation additionally requires the $\tilde{n}$-photon blocks of M^{mult}, which for all $\tilde{n} \geq 1$ are given by

$$\begin{aligned} M^{\text{mult}}_{\tilde{n}}(G) &= \Pi_{\tilde{n}} M^{\text{mult}}(G) \Pi_{\tilde{n}} \\ &= \Pi_{\tilde{n}} - \sum_{i=1}^{N_{\text{out}}} F^{i}_{\text{sc},\tilde{n}} \\ &= \Pi_{\tilde{n}} - \sum_{i=1}^{N_{\text{out}}} \frac{1}{\tilde{n}!} \left(b_i^\dagger\right)^{\tilde{n}} |0\rangle\langle 0| \left(b_i\right)^{\tilde{n}} . \end{aligned} \tag{5.2.13}$$

Finally, to apply Theorem 5.1.1, our objective is to establish a lower bound $c_{\tilde{n}} \leq \text{Tr}\left[M^{\text{mult}}_{\tilde{n}} \rho^{(\tilde{n})}\right]$ for all states $\rho^{(\tilde{n})}$. To achieve this, we derive an upper bound on the expectation value of the sum of single-click POVM elements in the preceding equation. Here, we once again utilize the input-output mode relations specified in Eq. (5.2.1) alongside Lemma 5.2.1. Let us denote this upper bound as $u_{\tilde{n}}$, which is defined as

$$u_{\tilde{n}} = \max_{\rho^{(\tilde{n})}} \text{Tr}\left[\sum_{i=1}^{N_{\text{out}}} F^{i}_{\text{sc},\tilde{n}} \rho^{(\tilde{n})}\right], \tag{5.2.14}$$

and relates to the lower bound $c_{\tilde{n}}$ by $c_{\tilde{n}} = 1 - u_{\tilde{n}}$.

The derivation of our final bound proceeds as follows. First, we note that each summand in the maximization of $u_{\tilde{n}}$ corresponds to a single-click event. Then, we group Bob's single-click POVM elements into blocks containing the same number of elements as the incoming modes. For each block separately, one can find a simple upper bound via a singular value decomposition that depends only on submatrices of G. To motivate this approach, we discuss the example of a perfect passive 4-state analyzer in the following section.

5.2.1 Motivating Example: Perfect Passive 4-State Analyzer

As an initial example, and to develop intuition about our objectives, let us consider a perfect passive 4-state receiver. In this case, the submatrix G is given by

$$G = \begin{pmatrix} \sqrt{p_z} & 0 \\ 0 & \sqrt{p_z} \\ \sqrt{\frac{p_x}{2}} & \sqrt{\frac{p_x}{2}} \\ \sqrt{\frac{p_x}{2}} & -\sqrt{\frac{p_x}{2}} \end{pmatrix}, \tag{5.2.15}$$

where the input modes correspond to H and V and the detectors 1 to 4 at the output are ordered to correspond to the polarizations H, V, D, and A. Additionally, note that due to Eq. (5.2.6), each inner product

$$\mathrm{Tr}\left[F^i_{\mathrm{sc},n}\left(\rho \otimes |0\rangle\langle 0|\right)\right], \tag{5.2.16}$$

depends solely on the i-th row of G. Consequently, knowing G allows us to bound all single-click probabilities in terms of its elements. For example, consider detector 1 (corresponding to H):

$$\begin{aligned} &\mathrm{Tr}\left[F^1_{\mathrm{sc},n}\left(\rho \otimes |0\rangle\langle 0|\right)\right] \\ &= \mathrm{Tr}\left[\frac{1}{n!}\left(\sum_{j=1}^{2} G^*_{1j} a_j^\dagger\right)^n |0\rangle\langle 0| \left(\sum_{k=1}^{2} G_{1k} a_k\right)^n \rho\right] \\ &= \mathrm{Tr}\left[\frac{p_z^n}{n!}\left(a_H^\dagger\right)^n |0\rangle\langle 0| \left(a_H\right)^n \rho\right] \leq p_z^n, \end{aligned} \tag{5.2.17}$$

where, in the last step, we maximize the expression by choosing $\rho = |n,0\rangle\langle n,0|_{HV}$, i.e., all n photons in the H mode.

As stated above to minimize $\mathrm{Tr}\left[M^{\mathrm{mult}}_{\tilde{n}} \rho^{(\tilde{n})}\right]$, we aim to maximize

$$\sum_{i=1}^{4} \mathrm{Tr}\left[F^i_{\mathrm{sc},\tilde{n}} \rho^{(\tilde{n})}\right] \tag{5.2.18}$$

over all possible input states $\rho^{(\tilde{n})}$. However, finding the state that maximizes this sum is challenging. As an initial step, let us maximize each term separately. Using Eq. (5.2.17), we obtain

$$\max_{\rho^{(\tilde{n})}} \sum_{i=1}^{4} \mathrm{Tr}\left[F^i_{\mathrm{sc},\tilde{n}} \rho^{(\tilde{n})}\right] \leq \sum_{i=1}^{4} \max_{\rho^{(\tilde{n})}} \mathrm{Tr}\left[F^i_{\mathrm{sc},\tilde{n}} \rho^{(\tilde{n})}\right] = 2p_z^{\tilde{n}} + 2p_x^{\tilde{n}}. \tag{5.2.19}$$

This follows because each $F^i_{\mathrm{sc},n}$ corresponds to a rank-one projector. Consequently, a lower bound $c_{\tilde{n}} \leq \min_{\rho^{(\tilde{n})}} \mathrm{Tr}\left[M^{\mathrm{mult}}_{\tilde{n}} \rho^{(\tilde{n})}\right]$ is given by

$$c_{\tilde{n}} = 1 - 2p_z^{\tilde{n}} - 2p_x^{\tilde{n}}. \tag{5.2.20}$$

For $\tilde{n} = 1$, this yields $c_{\tilde{n}} = -1$, which is unsuitable for subspace estimation. Furthermore, with $p_z = p_x = \frac{1}{2}$ and $\tilde{n} = 2$, we find $c_{\tilde{n}} = 0$, providing only trivial bounds in subspace population estimation. Thus, as expected, maximizing each term individually does not yield a useful bound on $c_{\tilde{n}}$ for small photon numbers.

However, if we partition the sum such that we combine both single-click POVMs in the Z-basis (1 & 2) and both single-click POVM elements in the X-basis (3 & 4), we have rank two operators and find

$$\max_{\rho^{(\tilde{n})}} \mathrm{Tr}\left[\left(F^1_{\mathrm{sc},\tilde{n}} + F^2_{\mathrm{sc},\tilde{n}}\right)\rho^{(\tilde{n})}\right] + \max_{\rho^{(\tilde{n})}} \mathrm{Tr}\left[\left(F^3_{\mathrm{sc},\tilde{n}} + F^4_{\mathrm{sc},\tilde{n}}\right)\rho^{(\tilde{n})}\right] = p_z^{\tilde{n}} + p_x^{\tilde{n}}. \tag{5.2.21}$$

Note that for $p_z = p_x = \frac{1}{2}$ this amounts to $c_{\tilde{n}} = 1 - \frac{1}{2^{\tilde{n}-1}} \geq 0$ for all $\tilde{n} \geq 1$. Inserting this result into Eq. (5.1.2) and noticing that $\min_{\tilde{n}\leq N_B} c_{\tilde{n}} = c_0 = 0$ yields

$$p_L(\tilde{n} \leq N_B|x) = 1 - \frac{m_{\mathrm{mult}|x}}{1 - p_z^{N_B+1} - p_x^{N_B+1}}, \tag{5.2.22}$$

which is a useful lower bound on the $\leq N_B$-photon subspace for any $N_B \geq 1$.

Alternatively, one could also partition the sum such that we combine POVM elements from different basis, i.e. (1 & 3) and (2 & 4). In this case for simplicity consider $p_z = p_x = \frac{1}{2}$ only and we find

$$\begin{aligned} &\max_{\rho^{(n)}} \mathrm{Tr}\left[\left(F^1_{\mathrm{sc},\tilde{n}} + F^3_{\mathrm{sc},\tilde{n}}\right)\rho^{(\tilde{n})}\right] + \max_{\rho^{(n)}} \mathrm{Tr}\left[\left(F^2_{\mathrm{sc},\tilde{n}} + F^4_{\mathrm{sc},\tilde{n}}\right)\rho^{(\tilde{n})}\right] \\ &= \frac{\sqrt{2+\sqrt{2}}^{\tilde{n}}}{2^{\tilde{n}-1}} > \frac{1}{2^{\tilde{n}-1}}. \end{aligned} \tag{5.2.23}$$

This shows that the previous partition of the sum was indeed better, although chosen completely arbitrarily. One fact which distinguished the previous one from this partition mixing the bases was that now the POVM elements in each partition were not orthogonal anymore. Having only orthogonal POVM elements in each sum allows us to detect any multi-photon state having weight in both H and V (or equivalently D and A) at the same time. However, having for example the combination $F^H_{\mathrm{sc},\tilde{n}} + F^D_{\mathrm{sc},\tilde{n}}$ allows for states like $|1,1\rangle_{HV}$ to still result in a single click with non-vanishing probability and remain unrecognized as a multi-photon state.

For any partitioning of the sum including a higher number of single-click POVM elements than the number of incoming modes (here $N_{\mathrm{in}} = 2$), it will in general not be obvious how to estimate a bound analytically. Therefore, we will also restrict the general theorem of Theorem 5.3.1 to this choice of blocking.

The simple example presented above also shows the dependence of the resulting bound $c_{\tilde{n}}$ on the partitioning of the sum, if the full sum cannot be maximized in a single step. However, as we have also seen, a partitioning by blocks of the number of incoming modes can already suffice to achieve a useful lower bound for the subspace estimation.

Now, to generalize the idea of this example, we will apply a partitioning with respect to the number of incoming signal modes N_{in}, such that each partition of the sum contains at most N_{in} modes. There will be $k_{\max} = \lfloor \frac{N_{\text{out}}}{N_{\text{in}}} \rfloor$ possible blocks involving N_{in} incoming and outgoing modes. If N_{out} is not a multiple of N_{in}, then there is also one additional block involving the remaining $N_{\text{out}} - k_{\max} N_{\text{in}}$ outgoing modes. Let us formalize this as follows, let $\mathcal{I}$ label the index set $1, \ldots, N_{\text{out}}$ and let each $\mathcal{I}_k$, $k = 1, \ldots, k_{\max}$, be a subset with $|\mathcal{I}_k| = N_{\text{in}}$. Finally, let $\mathcal{I}_{k_{\max}+1}$ contain $N_{\text{out}} - k_{\max} N_{\text{in}}$ indices. Then,

$$\mathcal{I} = \bigcup_{k=1}^{k_{\max}+1} \mathcal{I}_k. \tag{5.2.24}$$

Connecting this to our previous example of a 4-state receiver, the considered partitionings were

$$\mathcal{I} = \{1,2\} \cup \{3,4\} \quad \text{and} \quad \mathcal{I}' = \{1,3\} \cup \{2,4\}, \tag{5.2.25}$$

and it is immediately obvious that there is one more partitioning one could consider.

Additionally, each set $\mathcal{I}_k$ directly corresponds to the rows of G which are involved in the mode transformation due to Eq. (5.2.6) for that part of the sum.

This fact allows us to view the results we have shown for the 4-state receiver above from a different perspective as well. As we will see in the proof of Theorem 5.3.1, each summand in Eq. (5.2.21) was in fact the maximum singular value of the corresponding block of G. The same holds true if we chose to partition according to $\mathcal{I}'$, mixing the bases.

5.3 Subspace Estimation for Arbitrary Passive Linear Optical Setups

This intuition can be generalized to an arbitrary G and the resulting lower bound on the subspace is summarized in the next Theorem 5.3.1.

Theorem 5.3.1 (Subspace estimation for perfect passive linear optical detection setup)**.** *Let $U \in \mathbb{C}^{N_{\text{out}} \times N_{\text{out}}}$ be a unitary matrix describing the mode transformation of a perfect (no losses) passive linear optical setup as in Eq. (5.2.1), and label the first N_{in} columns of U as $G \in \mathbb{C}^{N_{\text{out}} \times N_{\text{in}}}$. By using the POVM element*

$$M^{\text{mult}}(G) := \mathbb{I} - |0\rangle\langle 0| - \sum_{i=1}^{N_{\text{out}}} F_{sc}^{i}. \tag{5.3.1}$$

of Eq. (5.2.12) and assuming threshold detectors, the $\tilde{n} \leq N_B$-subspace can be bounded by

$$p_L(\tilde{n} \leq N_B|x) = 1 - \frac{m_{\text{mult}|x}}{c_{\geq N_B+1}}, \tag{5.3.2}$$

where $c_{\geq N_B+1} := \min_{\tilde{n}\geq N_B+1} c_{\tilde{n}}$. The constants $c_{\tilde{n}}$ depend on the partitions $\mathcal{I} = \bigcup_{k=1}^{k_{max}+1} \mathcal{I}_k$, where $k_{max} = \lfloor \frac{N_{\text{out}}}{N_{\text{in}}} \rfloor$. These partitions are of size $|\mathcal{I}_k| = N_{\text{in}}$ for $k = 1, \ldots, k_{max}$ and $|\mathcal{I}_{k_{max}+1}| = N_{\text{out}} - k_{max}N_{\text{in}}$. The constants $c_{\tilde{n}}$ are defined as

$$\begin{aligned} c_{\tilde{n}} &= \max_{\mathcal{I}=\cup\mathcal{I}_k} \left(1 - \sum_{k=1}^{k_{max}+1} \mu_{\max}\left(\mathcal{I}_k\right)^{2\tilde{n}}\right), \quad \tilde{n} > 0, \\ c_0 &= 0, \end{aligned} \tag{5.3.3}$$

where $\mu_{\max}(\mathcal{I}_k)$ is the maximum singular value of the $N_{\text{in}} \times N_{\text{in}}$ block of G containing all rows labelled by $\mathcal{I}_k$.

Proof. For simplicity, consider the trivial partition

$$\begin{aligned} \mathcal{I} = &\{1, \ldots, N_{\text{in}}\} \cup \cdots \cup \{(k_{\max} - 1) N_{\text{in}} + 1, \ldots, k_{\max}N_{\text{in}}\} \\ &\cup \{k_{\max}N_{\text{in}} + 1, \ldots, N_{\text{out}}\} \end{aligned} \tag{5.3.4}$$

$$=: \mathcal{I}_1 \cup \cdots \cup \mathcal{I}_{k_{\max}+1}, \tag{5.3.5}$$

since none of the following steps will make any use of the particular partition $\mathcal{I}$. This choice is only for convenience.

Furthermore, define the operators P_k, corresponding to the columns of G contained in $\mathcal{I}_k$ (i.e. the k-th block), as

$$P_k := \sum_{i=(k-1)N_{\text{in}}+1}^{kN_{\text{in}}} \frac{1}{\tilde{n}!} \left(\sum_{j=1}^{N_{\text{in}}} G_{ij}^* a_j^\dagger\right)^{\tilde{n}} |0\rangle\langle 0| \left(\sum_{k=1}^{N_{\text{in}}} G_{ik} a_k\right)^{\tilde{n}} \tag{5.3.6}$$

for $k = 1, \ldots, k_{\max}$, and

$$P_{k_{\max}+1} := \sum_{i=k_{\max}N_{\text{in}}+1}^{N_{\text{out}}} \frac{1}{\tilde{n}!} \left(\sum_{j=1}^{N_{\text{in}}} G_{ij}^* a_j^\dagger\right)^{\tilde{n}} |0\rangle\langle 0| \left(\sum_{k=1}^{N_{\text{in}}} G_{ik} a_k\right)^{\tilde{n}} \tag{5.3.7}$$

for the remaining terms. If the output modes are multiples of the input modes, then $P_{k_{\max}+1}$ will be empty.

Then, for any $\rho^{(\tilde{n})}$

$$\mathrm{Tr}\left[\sum_{i=1}^{N_{\mathrm{out}}} \frac{\left(b_i^\dagger\right)^{\tilde{n}}}{\tilde{n}!} |0\rangle\langle 0| \left(b_i\right)^{\tilde{n}} \left(\rho^{(\tilde{n})} \otimes |0\rangle\langle 0|\right)\right] = \sum_{k=1}^{k_{\max}+1} \mathrm{Tr}\left[P_k \rho^{(\tilde{n})}\right], \tag{5.3.8}$$

and hence,

$$\max_{\rho^{(n)}} \mathrm{Tr}\left[\sum_{i=1}^{N_{\mathrm{out}}} \frac{\left(b_i^\dagger\right)^{\tilde{n}}}{\tilde{n}!} |0\rangle\langle 0| \left(b_i\right)^{\tilde{n}} \left(\rho^{(\tilde{n})} \otimes |0\rangle\langle 0|\right)\right] \leq \sum_{k=1}^{k_{\max}+1} \max_{\rho^{(n)}} \mathrm{Tr}\left[P_k \rho^{(\tilde{n})}\right], \tag{5.3.9}$$

Now, we will make use of our partitioning in G as well. Again, note that due to Lemma 5.2.1 and the particular partitioning $\mathcal{I}$ each P_k only depends on the k-th $N_{\mathrm{in}} \times N_{\mathrm{in}}$ block Λ_k of G, i.e.

$$G = \left(\begin{array}{c} \Lambda_1 \\ \hline \vdots \\ \hline \\ \Lambda_{k_{\max}} \\ \hline \\ \Lambda_{k_{\max}+1} \end{array}\right), \tag{5.3.10}$$

where the last block $\Lambda_{k_{\max}}$ has size $(N_{\mathrm{out}} - k_{\max} N_{\mathrm{in}}) \times N_{\mathrm{in}}$.

One can view each Λ_k as a separate mode transformation for each block. However, each block Λ_k is not a unitary transformation, still, this construction let us perform a SVD on each Λ_k separately. Therefore, for each $k = 1, \ldots, k_{\max}$ it holds $\Lambda_k = U_k \Sigma_k V_k^\dagger$, where both U_k and V_k are now unitary matrices of dimension $N_{\mathrm{in}} \times N_{\mathrm{in}}$ and Σ_k contains the singular values on the diagonal, i.e. $\Sigma_k = \mathrm{diag}(\sigma_1, \ldots \sigma_{N_{\mathrm{in}}})$. We will consider $\Lambda_{k_{\max}+1}$ separately afterwards.

We will proceed with an arbitrary $k \in \{1, \ldots k_{\max}\}$, and note that one can absorb $V_k^\dagger$ into the definition of the incoming signal modes via $\vec{c} = V_k^\dagger \vec{a}$.

Hence, while stating the dependence on the modes explicitly in square brackets, it holds

$$\mathrm{Tr}\left[P_k[\Lambda_k \vec{a}] \rho^{(\tilde{n})}[\vec{a}]\right] = \mathrm{Tr}\left[P_k[U_k \Sigma_k \vec{c}] \rho^{(\tilde{n})}[\vec{c}]\right]. \tag{5.3.11}$$

The linear mode transformation U_k induces a unitary transformation Ω_U on the Hilbert

space. Therefore, one can equivalently write

$$\begin{aligned}\mathrm{Tr}\left[P_k[U_k\Sigma_k\vec{c}]\rho^{(\tilde{n})}[\vec{c}]\right] &= \mathrm{Tr}\left[\Omega_U P_k[\Sigma_k\vec{c}]\ \Omega_U^\dagger\rho^{(\tilde{n})}[\vec{c}]\right] \\ &= \mathrm{Tr}\left[P_k[\Sigma_k\vec{c}]\ \Omega_U^\dagger\rho^{(\tilde{n})}[\vec{c}]\Omega_U\right].\end{aligned} \tag{5.3.12}$$

Since this can equivalently be viewed as a unitary acting on ρ, the maximization can be recast as

$$\max_{\rho^{(n)}}\mathrm{Tr}\left[P_k\rho^{(\tilde{n})}\right] = \max_{\tau^{(n)}}\mathrm{Tr}\left[\bar{P}_k\tau^{(\tilde{n})}\right], \tag{5.3.13}$$

where for simplicity $\bar{P}_k := P_k[\Sigma_k\vec{c}]$ is the POVM element before the mode transformation by U_k and $\tau^{(\tilde{n})} = \Omega_U^\dagger\rho^{(\tilde{n})}[\vec{c}]\Omega_U$.

Now, for $k = k_{\max}+1$ (or equivalently $P_{k_{\max}+1}$) the same procedure can be applied after we extend both the unitary $U_{k_{\max}+1}$ and $\Sigma_{k_{\max}+1}$ appropriately. If we extend $U_{k_{\max}+1}$ with a block of an identity of dimension $l = N_{\mathrm{out}} - k_{\max}N_{\mathrm{in}}$ and call the resulting matrix $\bar{U}_{k_{\max}+1}$ it has the form

$$\bar{U}_{k_{\max}+1} = U_{k_{\max}+1}\oplus\mathbb{I}_{l\times l}. \tag{5.3.14}$$

We need to extend $\Sigma_{k_{\max}+1}$ with zeros to reach

$$\bar{\Sigma}_{k_{\max}+1} = \begin{pmatrix}\Sigma_{k_{\max}+1}\\ \hline 0_{l\times N_{\mathrm{in}}}\end{pmatrix} \tag{5.3.15}$$

Then, it holds

$$\bar{U}_{k_{\max}+1}\bar{\Sigma}_{k_{\max}+1}V_{k_{\max}+1}^\dagger = \begin{pmatrix}\Lambda_{k_{\max}+1}\\ \hline 0_{l\times N_{\mathrm{in}}}\end{pmatrix}, \tag{5.3.16}$$

and the resulting mode transformation is still equivalent. We can now view $\bar{\Sigma}_{k_{\max}+1}$ as a square matrix containing singular values on the diagonal out of which some are equal to zero. Let us note that the unitary acting on the Hilbert space, $\Omega_{\bar{U}}$, generated by $\bar{U}_{k_{\max}+1}$ acts as $\Omega_{\bar{U}} = \Omega_U\otimes\mathbb{I}$. Therefore, we can still apply the same procedure as for any other $k = 1\ldots k_{\max}$.

In summary, we can write each $\bar{P}_k$ as

$$\bar{P}_k = \sum_{i=(k-1)N_{\mathrm{in}}+1}^{kN_{\mathrm{in}}}\frac{\sigma_i^{2\tilde{n}}}{\tilde{n}!}\left(c_i^\dagger\right)^{\tilde{n}}|0\rangle\langle 0|\left(c_i\right)^{\tilde{n}}, \tag{5.3.17}$$

where again the c_i are the incoming modes including the unitary $V_k^\dagger$ and the σ_i are the singular values of the block Λ_k of G (including zeros for $k_{\max}+1$).

Finally, the trace $\mathrm{Tr}\left[P_k \rho^{(\tilde{n})}\right]$ is maximized by

$$\max_{\rho^{(n)}} \mathrm{Tr}\left[P_k \rho^{(\tilde{n})}\right] = \max_{\tau^{(n)}} \mathrm{Tr}\left[\bar{P}_k \tau^{(\tilde{n})}\right] = \max_{1 \leq i \leq N_{\mathrm{in}}} \sigma_i^{2n} =: \mu_k^{2n}, \tag{5.3.18}$$

where for convenience we defined the singular value achieving the maximum as μ_k. Thus, in total we find

$$\min_{\rho^{(n)}} \mathrm{Tr}\left[M_{\tilde{n}} \rho^{(\tilde{n})}\right] \geq 1 - \sum_{k=1}^{k_{\max}+1} \mu_k^{2n}. \tag{5.3.19}$$

As stated at the beginning of the proof, all of the steps from above are still valid for any arbitrary partition $\mathcal{I}$. Hence, we can also maximize over all possible partitions, which results in the theorem statement. □

We will present an example using the flag-state squasher together with the weight estimation from above in Section 6.6.4. Next, while the above analysis assumes ideal detectors, we now extend it to include detector losses and dark counts, which are critical for practical implementations.

5.3.1 Including Detection Losses

After covering the perfect detection setup, we consider the situation with losses in the setup. We will reduce this situation to the lossless case by finding an equivalent transformation which first introduces the losses and then acts with an equivalent lossless detection setup on the input states.

As discussed earlier, we model a lossy setup with a unitary U having N_{in} incoming signal modes and additional $2N_{\mathrm{out}} - N_{\mathrm{in}}$ incoming vacuum modes. The output of the detection setup is then described by N_{out} detection and N_{out} loss modes, as shown in Fig. 5.1. For brevity, let us also label the system of the incoming N_{in} signal modes by Sig, the system of vacuum modes $N_{\mathrm{in}} + 1, \ldots N_{\mathrm{out}}$ by Env_1 and the remaining incoming modes by Env_2. As a first step of our subspace estimation, we will recast this unitary U with an equivalent unitary, $\tilde{U}$, that acts equally on the incoming signal modes. The special property of $\tilde{U}$ will be that it decomposes into two consecutive unitaries W_{tot} and $(T_d \oplus T_l)$. A schematic representation of this decomposition is shown in Fig. 5.3.

The first unitary, W_{tot} corresponds to the losses and most notably does not mix the incoming signal modes highlighted by Sig in Fig. 5.3 with the modes in Env_1. The remaining unitaries T_d and T_l act independently on each block of modes, performing the mode transformation on the detected and loss modes. Since this construction does not mix any part

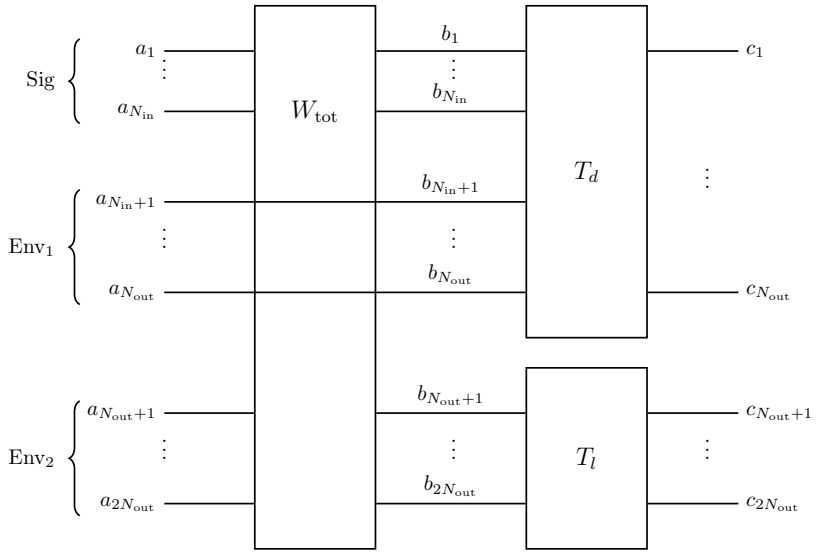


Figure 5.3: Schematic representation of the decomposition of Lemma 5.3.2 of the unitary transformation U of a detection setup into two consecutive ones, W_{tot} and $(T_d \oplus T_l)$. The unitary W_{tot} does not mix the signal modes in system Sig with the modes in system Env_1 which thus always remain in the vacuum state. This is indicated by lines continuing through W_{tot}.

of the signal modes with the vacuum modes in Env_1, such the modes in Env_1 will remain in the vacuum state, the input modes to the unitary T_d recover the same situation as in the lossless case. This decomposition is rigorously formalized in the following Lemma 5.3.2.

Lemma 5.3.2 (Detector decomposition)**.** *Let $U \in \mathbb{C}^{2N_{\text{out}} \times 2N_{\text{out}}}$ be a unitary matrix. Furthermore, label the first N_{in} columns of U as G. Then, there exist $V_d, V_l \in \mathbb{C}^{N_{\text{out}} \times N_{\text{in}}}$ and $W \in \mathbb{C}^{2N_{\text{in}} \times N_{\text{in}}}$ semi-unitaries ($W^\dagger W = \mathbb{I}$), such that G can be written as*

$$G = (V_d \oplus V_l)\, W. \tag{5.3.20}$$

Additionally, there exists a unitary matrix $\tilde{U}$ with equal first N_{in} columns G, which can be written as

$$\tilde{U} = \begin{pmatrix} T_d & 0 \\ 0 & T_l \end{pmatrix} \cdot \left(\begin{array}{c|c} W_d \oplus \frac{\mathbb{I}}{\sqrt{2}} & \\ W_l \oplus \frac{\mathbb{I}}{\sqrt{2}} & W' \end{array} \right) := V_{tot} \cdot W_{tot}, \tag{5.3.21}$$

where W' is any matrix completing W_{tot} to a unitary, and both $V_{tot} \in \mathbb{C}^{2N_{\text{out}} \times 2N_{\text{out}}}$ and $W_{tot} \in \mathbb{C}^{2N_{\text{out}} \times 2N_{\text{out}}}$ are unitary matrices.

Proof. The proof will proceed in two stages. First, we will construct the decomposition for G through a polar decomposition. Then we extend G to the full unitary matrix $\tilde{U}$ by applying QR decompositions.

Therefore, let us start by rewriting G as

$$G = \left(\begin{array}{c|c|c} \vec{b}_1 & \dots & \vec{b}_{N_{\text{in}}} \\ \hline \vec{c}_1 & \dots & \vec{c}_{N_{\text{in}}} \end{array}\right), \tag{5.3.22}$$

where $\vec{b}_i, \vec{c}_i \in \mathbb{C}^{N_{\text{out}}}$. Define the vectors $\vec{v}_i$ and $\vec{w}_i$ for $1 \dots N_{\text{in}}$ by

$$\vec{v}_i := \frac{\vec{b}_i}{\beta_i}, \quad \vec{w}_i := \frac{\vec{c}_i}{\gamma_i}, \tag{5.3.23}$$

$$\text{where } \beta_i := \left\|\vec{b}_i\right\|, \quad \gamma_i := \|\vec{c}_i\|, \tag{5.3.24}$$

and using those vectors, define the matrices

$$V_d' := \left(\begin{array}{c|c|c} | & & | \\ \vec{v}_1 & \dots & \vec{v}_{N_{\text{in}}} \\ | & & | \end{array}\right) \text{ and } V_l' := \left(\begin{array}{c|c|c} | & & | \\ \vec{w}_1 & \dots & \vec{w}_{N_{\text{in}}} \\ | & & | \end{array}\right). \tag{5.3.25}$$

Then, we can rewrite G as

$$G = \begin{pmatrix} V_d' & 0 \\ 0 & V_l' \end{pmatrix} \begin{pmatrix} \operatorname{diag}(\vec{\beta}) \\ \operatorname{diag}(\vec{\gamma}) \end{pmatrix}, \tag{5.3.26}$$

where $\vec{\beta}$ and $\vec{\gamma}$ are vectors having β_i and γ_i as their components. These matrices are not semi-unitaries yet, so we take the polar decomposition, see [60, Theorem 7.3.1], of V_d' and V_l' yielding

$$V_d' = V_d P_d \quad \text{and} \quad V_l' = V_l P_l, \tag{5.3.27}$$

where V_d, V_l are semi-unitaries and P_d, P_l are positive semidefinite. To simplify the notation, let us define $W_d := P_d \operatorname{diag}(\vec{\beta})$ and $W_l := P_l \operatorname{diag}(\vec{\gamma})$. This allows us to write G as

$$G = \begin{pmatrix} V_d & 0 \\ 0 & V_l \end{pmatrix} \begin{pmatrix} W_d \\ W_l \end{pmatrix} =: \begin{pmatrix} V_d & 0 \\ 0 & V_l \end{pmatrix} \cdot W \tag{5.3.28}$$

By definition, again see [60, Theorem 7.3.1], V_d and V_l are semi-unitaries. Also, W is semi-unitary since

$$\begin{aligned} W^\dagger W &= \operatorname{diag}(\vec{\beta}) P_d^\dagger P_d \operatorname{diag}(\vec{\beta}) + \operatorname{diag}(\vec{\gamma}) P_l^\dagger P_l \operatorname{diag}(\vec{\gamma}) \\ &= \operatorname{diag}(\vec{\beta}) (V_d')^\dagger V_d' \operatorname{diag}(\vec{\beta}) + \operatorname{diag}(\vec{\gamma}) (V_l')^\dagger V_l' \operatorname{diag}(\vec{\gamma}) \\ &= G^\dagger G \\ &= \mathbb{I}_{N_{\text{in}}}, \end{aligned} \tag{5.3.29}$$

where we used the fact $P_{d/l}^\dagger P_{d/l} = P_{d/l}^2 = (V'_{d/l})^\dagger V'_{d/l}$, by the polar decomposition. Hence, we found the desired decomposition of G, which finishes the first part of the proof.

Next, we proceed with the second part, extending G to the claimed unitary $\tilde{U}$ with the same first N_{in} columns (or equivalently G). We will use the QR-decomposition, see [60, Theorem 2.1.14], to show that any of the claimed extensions of the semi-unitaries to unitaries indeed exist. Therefore, applying a QR-decomposition we can write

$$V_d = Q_d R \quad \text{and} \quad V_l = Q_l R, \tag{5.3.30}$$

where both Q_d and Q_l are unitary $N_{\text{out}} \times N_{\text{out}}$ matrices and the matrix R is given such that it is composed of an identity and a block of zeros O of size $(N_{\text{out}} - N_{\text{in}}) \times N_{\text{in}}$, i.e.

$$R = \begin{pmatrix} \mathbb{I}_{N_{\text{in}}} \\ \hline O \end{pmatrix}. \tag{5.3.31}$$

With that define the matrix V_{tot} as

$$V_{\text{tot}} = \begin{pmatrix} Q_d & 0 \\ 0 & Q_l \end{pmatrix}, \tag{5.3.32}$$

which is a unitary matrix because both Q_d and Q_l are unitary. Now, we first extend W to

$$W_{\text{ext}} := \begin{pmatrix} W_d \oplus \frac{\mathbb{I}_{N_{\text{out}}-N_{\text{in}}}}{\sqrt{2}} \\ W_l \oplus \frac{\mathbb{I}_{N_{\text{out}}-N_{\text{in}}}}{\sqrt{2}} \end{pmatrix} \tag{5.3.33}$$

and then take another QR decomposition of W_{ext} such that

$$W_{\text{ext}} = W_{\text{tot}} \cdot \begin{pmatrix} \mathbb{I}_{N_{\text{out}}} \\ \hline O_{N_{\text{out}}} \end{pmatrix}, \tag{5.3.34}$$

where $O_{N_{\text{out}}}$ is a matrix containing only zeros of size $N_{\text{out}} \times N_{\text{out}}$. This shows that there indeed exists a matrix W' such that

$$W_{\text{tot}} = \left(\begin{array}{c|c} W_d \oplus \frac{\mathbb{I}}{\sqrt{2}} & \\ W_l \oplus \frac{\mathbb{I}}{\sqrt{2}} & W' \end{array} \right), \tag{5.3.35}$$

is a unitary matrix. Finally, the unitary $\tilde{U} := V_{\text{tot}} \cdot W_{\text{tot}}$ results in the same G, since

$$\begin{aligned} V_{\text{tot}} W_{\text{tot}} \cdot \begin{pmatrix} \mathbb{I}_{N_{\text{in}}} \\ 0 \end{pmatrix} &= \begin{pmatrix} Q_d & 0 \\ 0 & Q_l \end{pmatrix} \cdot \left(\begin{array}{c} W_d \\ 0 \\ \hline W_l \\ 0 \end{array}\right) \\ &= \begin{pmatrix} Q_d \begin{pmatrix} W_d \\ 0 \end{pmatrix} \\ Q_l \begin{pmatrix} W_l \\ 0 \end{pmatrix} \end{pmatrix} = \begin{pmatrix} V_d \cdot W_d \\ V_l \cdot W_l \end{pmatrix} = G, \end{aligned} \tag{5.3.36}$$

which completes the proof. □

Next, we will use this lemma and especially the property that W_{tot} does not mix the modes in the Sig system with the vacuum modes in the system Env_1 to recover the lossless case. Hence, the goal is to argue that the input to the unitary T_d is exactly of the form previously considered in Theorem 5.3.1. For this purpose, let us define the mode operators in the following way, let $\vec{a}$ be the incoming modes (including the N_{in} signal modes), $\vec{b}$ the transformed modes after the loss, i.e. after applying W_{tot}, and $\vec{c}$ be the modes after also V_{tot} has been applied, i.e.

$$\vec{c} = V_{\text{tot}} \vec{b} = V_{\text{tot}} W_{\text{tot}} \vec{a}. \tag{5.3.37}$$

Due to the decomposition of Lemma 5.3.2 one can recast the mode transformation by U as an equivalent transformation due to $\tilde{U}$ acting equally on the incoming modes in system Sig. This allows for a decomposition into two steps. Let $\Omega_{W_{\text{tot}}}$ be the unitary transformation on the Hilbert space implementing the linear mode transformation W_{tot}. Then, the state in the systems Sig, Env_1 after the first mode transformation W_{tot} is

$$\begin{aligned} \rho_{\text{Sig,Env}_1} &= \text{Tr}_{\text{Env}_2} \left[\Omega_{W_{\text{tot}}} \left(\rho_{\text{Sig}} \otimes |0\rangle\langle 0|_{\text{Env}_1 \text{Env}_2} \right) \Omega_{W_{\text{tot}}}^{\dagger} \right] \\ &= \sigma_{\text{Sig}} \otimes |0\rangle\langle 0|_{\text{Env}_1}, \end{aligned} \tag{5.3.38}$$

where $\sigma_{\text{Sig}} = \Psi(\rho_{\text{Sig}})$ for a CPTP map Ψ defined by this mode transformation. Additionally, we used the fact that W_{tot} does not transform any modes of system Sig into Env_1 or vice versa. With the definitions from above the state σ_{Sig} is defined in terms of the mode operators $\vec{b}$. As one part to recover the lossless situation, we will give the adversary control over this channel Ψ, which will be formalized in Theorem 5.3.3.

By construction, any non-vacuum state in system Sig after the mode transformation W_{tot} will lead to a detection. Therefore, define new equivalent POVM elements on systems

Sig, Env_1 in terms of the modes transformed due to V_d, resulting from Lemma 5.3.2. Then, the POVM element M^{mult} used for the subspace estimation will be written as before, but now depends on V_d instead of G, i.e.

$$M^{\text{mult}}(V_d) := \mathbb{I} - |0\rangle\langle 0| - \sum_{i=1}^{N_{\text{out}}} F_{\text{sc}}^i, \tag{5.3.39}$$

where each POVM element F_{sc}^i is defined as

$$F_{\text{sc}}^i = \sum_{\tilde{n}=1}^{\infty} \frac{\left(c_i^\dagger\right)^{\tilde{n}}}{\tilde{n}!} |0\rangle\langle 0| \left(c_i\right)^{\tilde{n}}. \tag{5.3.40}$$

It is important to note that now each of these mode operators c_i is given by

$$c_i = \sum_{j=1}^{N_{\text{out}}} (T_d)_{ij} b_j, \tag{5.3.41}$$

where T_d is the expanded unitary version of V_d. Again, note that only the first N_{in} columns of T_d matter due to Eq. (5.2.6), which is V_d, hence the definition of M^{mult} in terms of V_d only.

Combining both observations, we have reduced the loss case to the perfect detector case and can apply the previous bound on V_d, which is summarized in the next Theorem 5.3.3.

Theorem 5.3.3 (Subspace estimation for linear optical detection setup with loss)**.** *Let $U \in \mathbb{C}^{2N_{\text{out}} \times 2N_{\text{out}}}$ be a unitary matrix describing the mode transformation of a passive linear optical setup with losses as in Eq. (5.2.1). Let $V_d \in \mathbb{C}^{N_{\text{out}} \times N_{\text{in}}}$ be the semi-unitary resulting from applying Lemma 5.3.2 to U. By using the POVM element*

$$M^{\text{mult}}(V_d) := \mathbb{I} - |0\rangle\langle 0| - \sum_{i=1}^{N_{\text{out}}} F_{sc}^i. \tag{5.3.42}$$

of Eq. (5.3.39), defined in terms of V_d, and assuming threshold detectors, the $\tilde{n} \leq N_B$-subspace can be bounded by

$$p_L(\tilde{n} \leq N_B|x) = 1 - \frac{m_{\text{mult}|x}}{c_{\geq N_B+1}}, \tag{5.3.43}$$

where $c_{\geq N_B+1} := \min_{\tilde{n}\geq N_B+1} c_{\tilde{n}}$. *The constants* $c_{\tilde{n}}$ *depend on the partitions* $\mathcal{I} = \bigcup_{k=1}^{k_{max}+1} \mathcal{I}_k$, *where* $k_{max} = \lfloor \frac{N_{\text{out}}}{N_{\text{in}}} \rfloor$. *These partitions are of size* $|\mathcal{I}_k| = N_{\text{in}}$ *for* $k = 1, \ldots, k_{max}$ *and* $|\mathcal{I}_{k_{max}+1}| = N_{\text{out}} - k_{max} N_{\text{in}}$. *The constants* $c_{\tilde{n}}$ *are defined as*

$$\begin{aligned} c_{\tilde{n}} &= \max_{\mathcal{I}=\cup\mathcal{I}_k} \left(1 - \sum_{k=1}^{k_{max}+1} \mu_{\max}^{2\tilde{n}} (\mathcal{I}_k) \right), \quad \tilde{n} > 0, \\ c_0 &= 0, \end{aligned} \tag{5.3.44}$$

where $\mu_{\max}(\mathcal{I}_k)$ *is the maximum singular value of the* $N_{\text{in}} \times N_{\text{in}}$ *block of* V_d, *containing all rows labelled by* $\mathcal{I}_k$.

Proof. At first, let us represent U in its equivalent form $\tilde{U}$ using Lemma 5.3.2 as

$$\tilde{U} = \begin{pmatrix} T_d & 0 \\ 0 & T_l \end{pmatrix} \cdot \left(\begin{array}{c|c} W_d \oplus \frac{\mathbb{I}}{\sqrt{2}} & W' \\ W_l \oplus \frac{\mathbb{I}}{\sqrt{2}} & \end{array} \right) =: V_{\text{tot}} \cdot W_{\text{tot}}. \tag{5.3.45}$$

As described above, let us also label the mode operators in the following way, let $\vec{a}$ be the incoming modes (including the N_{in} signal modes), $\vec{b}$ the transformed modes after the loss, i.e. after applying W_{tot}, and $\vec{c}$ be the modes after also V_{tot} has been applied, i.e.

$$\vec{c} = V_{\text{tot}} \vec{b} = V_{\text{tot}} W_{\text{tot}} \vec{a}. \tag{5.3.46}$$

Thus, since $\tilde{U}$ and U agree on the first N_{in} columns, it holds $\tilde{U}\vec{a}|_{N_{\text{in}}} = U\vec{a}|_{N_{\text{in}}}$ for the incoming signal modes. As mentioned before, for brevity, let us label the system of the incoming N_{in} signal modes by Sig, the system of vacuum modes $N_{\text{in}}+1, \ldots, N_{\text{out}}$ by Env_1 and the remaining incoming modes by Env_2.

Next, consider the full multi-click POVM element M^{full} acting on the full system of output modes $\vec{c}$, which has the form

$$M^{\text{full}} = M^{\text{mult}} \otimes \mathbb{I}_{\text{Env}_2}, \tag{5.3.47}$$

where M^{mult} acts on the N_{out} detected modes and with the identity on the N_{out} loss modes in system Env_2. Here we suppressed the dependence of M^{mult} on V_d for brevity. The full multi-click POVM element has this form because a multi-click event does not depend on the state in the loss modes of system Env_2.

Finally, for any unitary transformation V acting on the modes, let Ω_V be the corresponding transformation on the Hilbert space. Then, exploiting the structure of the mode

transformation $\tilde{U}$, for any state ρ_{sig} defined on the N_{in} signal input modes in the Sig system, we find

$$\begin{aligned}
&\operatorname{Tr}\left[M^{\text{full}}\Omega_U\left(\rho_{\text{sig}}\otimes|0\rangle\langle 0|_{\text{Env}_1\text{Env}_2}\right)\Omega_U^\dagger\right] && (5.3.48)\\
&=\operatorname{Tr}\left[M^{\text{full}}\Omega_{\tilde{U}}\left(\rho_{\text{sig}}\otimes|0\rangle\langle 0|_{\text{Env}_1\text{Env}_2}\right)\Omega_{\tilde{U}}^\dagger\right] && (5.3.49)\\
&=\operatorname{Tr}\left[M^{\text{full}}\Omega_{V_{\text{tot}}}\Omega_{W_{\text{tot}}}\left(\rho_{\text{sig}}\otimes|0\rangle\langle 0|_{\text{Env}_1\text{Env}_2}\right)\Omega_{W_{\text{tot}}}^\dagger\Omega_{V_{\text{tot}}}^\dagger\right] && (5.3.50)\\
&=\operatorname{Tr}\left[\left(\Omega_{T_d}^\dagger\otimes\Omega_{T_l}^\dagger\right)\left(M^{\text{mult}}\otimes\mathbb{I}_{\text{Env}_2}\right)\left(\Omega_{T_d}\otimes\Omega_{T_l}\right)\Omega_{W_{\text{tot}}}\left(\rho_{\text{sig}}\otimes|0\rangle\langle 0|_{\text{Env}_1\text{Env}_2}\right)\Omega_{W_{\text{tot}}}^\dagger\right] && (5.3.51)\\
&=\operatorname{Tr}\left[\left(\Omega_{T_d}^\dagger M^{\text{mult}}\Omega_{T_d}\otimes\mathbb{I}_{\text{Env}_2}\right)\Omega_{W_{\text{tot}}}\left(\rho_{\text{sig}}\otimes|0\rangle\langle 0|_{\text{Env}_1\text{Env}_2}\right)\Omega_{W_{\text{tot}}}^\dagger\right] && (5.3.52)\\
&=\operatorname{Tr}\left[\Omega_{T_d}^\dagger M^{\text{mult}}\Omega_{T_d}\operatorname{Tr}_{\text{Env}_2}\left[\Omega_{W_{\text{tot}}}\left(\rho_{\text{sig}}\otimes|0\rangle\langle 0|_{\text{Env}_1\text{Env}_2}\right)\Omega_{W_{\text{tot}}}^\dagger\right]\right]. && (5.3.53)
\end{aligned}$$

Next, let us define the CPTP-map Ψ as

$$\Psi(\rho_{\text{sig}}):=\operatorname{Tr}_{\text{Env}_1\text{Env}_2}\left[\Omega_{W_{\text{tot}}}\left(\rho_{\text{sig}}\otimes|0\rangle\langle 0|_{\text{Env}_1\text{Env}_2}\right)\Omega_{W_{\text{tot}}}^\dagger\right], \tag{5.3.54}$$

which satisfies

$$\operatorname{Tr}_{\text{Env}_2}\left[\Omega_{W_{\text{tot}}}\left(\rho_{\text{sig}}\otimes|0\rangle\langle 0|_{\text{Env}_1\text{Env}_2}\right)\Omega_{W_{\text{tot}}}^\dagger\right]=\Psi(\rho_{\text{sig}})\otimes|0\rangle\langle 0|_{\text{Env}_1}, \tag{5.3.55}$$

because the vacuum state remains invariant under the mode transformation W_{tot} and additionally, the mode transformation does not mix the modes in the systems Sig and Env_1. Inserting the definition of Ψ, yields in summary

$$\begin{aligned}
&\operatorname{Tr}\left[M^{\text{full}}\Omega_{\tilde{U}}\left(\rho_{\text{sig}}\otimes|0\rangle\langle 0|_{\text{Env}_1\text{Env}_2}\right)\Omega_{\tilde{U}}^\dagger\right] && (5.3.56)\\
&=\operatorname{Tr}\left[\Omega_{T_d}^\dagger M^{\text{mult}}\Omega_{T_d}\left(\Psi\left(\rho_{\text{in}}\right)\otimes|0\rangle\langle 0|_{\text{Env}_1}\right)\right], && (5.3.57)
\end{aligned}$$

which is an expression very similar in form to the lossless case apart from the additional CPTP map Ψ.

Next, we restrict our attention to the action of M^{full} on the detected modes, denoted by M^{mult}. By making the dependence on V_d explicit, the POVM element after the mode transformation Ω_{T_d} is given by

$$\Omega_{T_d}^\dagger M^{\text{mult}}\Omega_{T_d}=M^{\text{mult}}(V_d)=\mathbb{I}-|0\rangle\langle 0|-\sum_{i=1}^{N_{\text{out}}}F_{\text{sc}}^i, \tag{5.3.58}$$

where, consistent with our previous labelling convention, each F_{sc}^{i} is defined as

$$F_{\text{sc}}^{i} = \sum_{\tilde{n}=1}^{\infty} \frac{\left(c_i^{\dagger}\right)^{\tilde{n}}}{\tilde{n}!} |0\rangle\langle 0| \left(c_i\right)^{\tilde{n}}. \tag{5.3.59}$$

Here, the operators c_i (representing the detected modes after the transformation T_d) are related to the initial operators b_j by

$$c_i = \sum_{j=1}^{N_{\text{out}}} (T_d)_{ij} b_j. \tag{5.3.60}$$

Thus, analogous to the lossless scenario, c_i is a function of the modes b_j, but is now connected via the submatrix V_d of T_d rather than G.

We require the $\tilde{n}$-photon component of $M^{\text{mult}}(V_d)$, which is given by

$$M_{\tilde{n}}^{\text{mult}}(V_d) = \Pi_{\tilde{n}} - \sum_{i=1}^{N_{\text{out}}} \frac{1}{\tilde{n}!} \left(c_i^{\dagger}\right)^{\tilde{n}} |0\rangle\langle 0| \left(c_i\right)^{\tilde{n}}. \tag{5.3.61}$$

Combining the previous steps, we can solve for the constants $c_{\tilde{n}}$. In this context, $\tilde{n}$ represents the photon number after the application of W_{tot}. The minimization is given by

$$\min_{\substack{\rho_{\text{sig}} \\ \text{s.t. } \Psi(\rho_{\text{in}})=\tau^{(\tilde{n})}}} \text{Tr}\left[M_{\tilde{n}}^{\text{mult}}(V_d) \left(\tau^{(\tilde{n})} \otimes |0\rangle\langle 0|_{\text{Env}_1}\right)\right] \tag{5.3.62}$$

$$\geq \min_{\tau^{(\tilde{n})}} \text{Tr}\left[M_{\tilde{n}}^{\text{mult}}(V_d) \left(\tau^{(\tilde{n})} \otimes |0\rangle\langle 0|_{\text{Env}_1}\right)\right] = c_{\tilde{n}}. \tag{5.3.63}$$

We can now apply Theorem 5.3.1 to $M_{\tilde{n}}^{\text{mult}}(V_d)$ to perform the optimization over $\tau^{(\tilde{n})}$ and reach the theorem statement. □

A brief remark regarding the application of this theorem to QKD security proofs is necessary.

Remark 5.3.4. In a security proof, the quantum channel Ψ, defined by

$$\Psi(\rho_{\text{sig}}) := \text{Tr}_{\text{Env}_1\text{Env}_2}\left[\Omega_{W_{\text{tot}}} \left(\rho_{\text{sig}} \otimes |0\rangle\langle 0|_{\text{Env}_1\text{Env}_2}\right) \Omega_{W_{\text{tot}}}^{\dagger}\right], \tag{5.3.64}$$

must be included in the squashing map. This inclusion effectively grants Eve control over the channel. Furthermore, the key rate must be calculated using the new POVM elements determined by the mode transformation T_d.

To conclude this section, we demonstrate that the constants $c_{\tilde{n}}$ are monotonically increasing with respect to $\tilde{n}$ for both lossy and lossless detection setups.

Corollary 5.3.5 (Monotonicity of $c_{\tilde{n}}$)**.** *The constants $c_{\tilde{n}}$ of Theorem 5.3.1 and Theorem 5.3.3 are monotonically increasing in $\tilde{n} \in \mathbb{N}_0$ and satisfy*

$$\min_{\tilde{n} \geq N_B+1} c_{\tilde{n}} = c_{N_B+1}, \qquad \min_{\tilde{n} \leq N_B} c_{\tilde{n}} = c_0. \tag{5.3.65}$$

Proof. We will show the statement for the perfect case of Theorem 5.3.1. The case of Theorem 5.3.3 including losses follows by replacing G with V_d.

Let $\mathcal{I} = \bigcup_{k+1}^{k_{\max}} \mathcal{I}_k$ be the partitioning achieving the maximum in Theorem 5.3.1 and for simplicity, define the block $\Lambda_k := \Lambda(\mathcal{I}_k)$ and μ_k as the largest singular value of Λ_k.

Since G is a semi-unitary ($G^\dagger G = \mathbb{I}$) and G is composed of the blocks $\Lambda_1, \ldots, \Lambda_{k_{\max}}$, it holds for all $k = 1, \ldots, k_{\max}$,

$$\mathbb{I} = G^\dagger G = \sum_{k=1}^{k_{\max}} \Lambda_k^\dagger \Lambda_k \succeq \Lambda_k^\dagger \Lambda_k \succeq 0. \tag{5.3.66}$$

Therefore, the maximum eigenvalue of $\Lambda_k^\dagger \Lambda_k$ satisfies

$$1 \geq \lambda_{\max}\left(\Lambda_k^\dagger \Lambda_k\right) \geq 0, \tag{5.3.67}$$

for all blocks k. The maximum eigenvalue of $\Lambda_k^\dagger \Lambda_k$ is equal to the square of the maximum singular value of Λ_k, i.e.

$$\mu_k^2 = \lambda_{\max}\left(\Lambda_k^\dagger \Lambda_k\right), \tag{5.3.68}$$

and since $\mu_k \in [0, 1]$, it follows

$$\left(\mu_k^2\right)^n \geq \left(\mu_k^2\right)^{n+1} \quad \forall k. \tag{5.3.69}$$

Furthermore, it holds for the constants $c_{\tilde{n}}$

$$c_{\tilde{n}} = 1 - \sum_k \mu_k^{2\tilde{n}} \leq 1 - \sum_k \mu_k^{2(\tilde{n}+1)} = c_{\tilde{n}+1}. \tag{5.3.70}$$

Hence, the constants $c_{\tilde{n}}$ are monotonically increasing in $\tilde{n}$ and they satisfy

$$\min_{\tilde{n} \geq N_B+1} c_{\tilde{n}} = c_{N_B+1}, \qquad \min_{\tilde{n} \leq N_B} c_{\tilde{n}} = c_0. \tag{5.3.71}$$

□

5.3.2 Including Dark Counts

To conclude this section, one would also like to include dark counts in the detection setup. Dark counts can be modelled by a classical post processing, see e.g. [66, Eq. (24)] and the discussion before it. Let $\{F_{\text{no}}, F^i_{\text{sc}}\} =: \{F_i\}_{i\in\{0,1,\dots,N_{\text{out}}\}}$ be the no-click and single-click POVM elements excluding dark counts. These could either originate from a perfect detection setup using G or from an imperfect setup using V_d due to Lemma 5.3.2. Furthermore, let $\{\tilde{F}_i\}_{0,1\dots,N_{\text{out}}}$ be the POVM elements including dark counts. The following relation holds

$$\tilde{F}_i = \sum_{j=0}^{N_{\text{out}}} \mathcal{P}_{i|j} F_j, \tag{5.3.72}$$

where $\mathcal{P}$ is the relevant part of the stochastic matrix implementing the classical post processing corresponding to dark counts. We note that here we omitted any parts of $\mathcal{P}$ which map no-click events to multi-click events. Thus, it holds $\sum_i \mathcal{P}_{i|j} = p_j \leq 1\ \forall j$ and not equality. We will use this fact in the following corollary. Furthermore, Corollary 5.3.6 elevates the previous lower bounds to one which can also be used for detectors with dark counts.

Corollary 5.3.6 (Lower bound including dark counts.)**.** *Let $G \in \mathbb{C}^{N_{\text{out}}\times N_{\text{in}}}$ be the submatrix describing a detection setup consisting only of passive linear optical elements. Furthermore, let $\mathcal{P}$ be the stochastic matrix implementing the post processing corresponding to dark counts as in Eq. (5.3.72). Then, the lower bounds on the subspaces of the previous Theorem 5.3.1 and Theorem 5.3.3 are still valid using observations including dark counts and the constants $c_{\tilde{n}}$ from the setups without dark counts.*

Proof. Assume we define M^{dark} equivalently as M^{mult} in Eq. (5.2.12), but using the POVM elements $\{\tilde{F}_j\}$ including the dark count post-processing and M^{mult} just as before in Eq. (5.2.12). With this construction, the $\tilde{n}$-photon component of M^{dark} is a linear combination of $\tilde{n}$-photon components of the POVM elements $\{F_i\}$ before the post-processing. Thus, it

follows

$$
\begin{aligned}
M_{\tilde{n}}^{\text{dark}} &= \Pi_{\tilde{n}} - \sum_{k,i=0}^{N_{\text{out}}} \mathcal{P}_{k|i} \frac{\left(b_i^\dagger\right)^{\tilde{n}}}{\tilde{n}!} |0\rangle\langle 0| \left(b_i\right)^{\tilde{n}} \\
&= \Pi_{\tilde{n}} - \sum_{i=0}^{N_{\text{out}}} \left(\sum_{k=0}^{N_{\text{out}}} \mathcal{P}_{k|i} \right) \frac{\left(b_i^\dagger\right)^{\tilde{n}}}{\tilde{n}!} |0\rangle\langle 0| \left(b_i\right)^{\tilde{n}} \\
&= \Pi_{\tilde{n}} - \sum_{i=0}^{N_{\text{out}}} p_i \frac{\left(b_i^\dagger\right)^{\tilde{n}}}{\tilde{n}!} |0\rangle\langle 0| \left(b_i\right)^{\tilde{n}} \geq M_{\tilde{n}},
\end{aligned}
\tag{5.3.73}
$$

where we used $\sum_i \mathcal{P}_{k|i} = p_i \leq 1 \; \forall j$. Hence, we can conclude

$$
\text{Tr}\left[M_{\tilde{n}}^{\text{dark}} \rho^{(\tilde{n})}\right] \geq \text{Tr}\left[M_{\tilde{n}}^{\text{mult}} \rho^{(\tilde{n})}\right] \geq c_{\tilde{n}}. \tag{5.3.74}
$$

Since this holds for all $\tilde{n} \in \mathbb{N}$ it follows for the observations of the POVM element with and without dark counts $m_{\text{obs}}^{\text{dark}} \geq m_{\text{obs}}$. Hence, we can write

$$
p_L(\tilde{n} \leq N_B|x) = 1 - \frac{m_{\text{mult}|x}^{\text{dark}}}{c_{\geq N_B+1}^{\text{dark}}} \geq 1 - \frac{m_{\text{mult}|x}^{\text{dark}}}{c_{\geq N_B+1}}, \tag{5.3.75}
$$

which concludes the proof. □

Again, let us point out that the bound resulting from Corollary 5.3.6 is strictly worse than Theorem 5.3.1 or Theorem 5.3.3, as the dark counts do not decrease the lower bounds on the $c_{\tilde{n}}$, but they increase the observation m_{obs}, thus decreasing the lower bound on $p(\tilde{n} \leq N_B|x)$.

5.4 Discussion of Previous Results

Now, it is instructive to compare our results with those reported in Ref. [42]. In that work, lower bounds were obtained via numerical experiments using a specific POVM and a truncated subspace dimension until $N_B = 20$. While numerical optimization can yield tight bounds within a finite subspace, it is fundamentally limited by the inability to represent the full, infinite-dimensional Hilbert space. Consequently, observing monotonic behaviour up to $N_B = 20$ provides insufficient guarantees for QKD security. Crucially, such a truncation leaves open the possibility that Eve could conceal her attack within the higher-dimensional subspace ($N_B > 20$) that remains unchecked by the numerical test.

On the other hand, our results here give a generic lower bound which is monotonically increasing in N_B, see Corollary 5.3.5, and do not require designing a specific POVM element for each detection setup. Thus, the results presented here can be used in a QKD security proof and are far more convenient to apply.

5.5 Detector Characterization for Flag-State Squasher

In this section, we answer the question of how one could determine the detector matrix G experimentally. Since there will always be some loss involved in any experimental setup, we only discuss the more general lossy case here.

At first, consider a coherent state with complex amplitude $\vec{\alpha}$ in the input modes entering the detection setup and it is transformed according to G via

$$\beta_i = \sum_{j=1}^{N_{\text{in}}} G_{ij}\alpha_j, \tag{5.5.1}$$

assuming there is no stray light entering the setup. For threshold detectors the probability of observing a click in detector i corresponding to mode b_i originating from a coherent laser pulse is given by

$$p_i = 1 - e^{-|\beta_i|^2} = 1 - e^{-|\sum_{j=1}^{N_{\text{in}}} G_{ij}\alpha_j|^2}. \tag{5.5.2}$$

Thus, with the knowledge of the incoming complex amplitude $\vec{\alpha}$ of a coherent state and the probability of a detection, one can solve for G, provided enough test states are sent. We do not have access to the exact probabilities, but we can use normalized counts in the detectors as an approximation for the detection probability. In the limit of many pulses, this will converge to the probability p_i by the law of large numbers. Furthermore, if one uses a bright laser or assumes low dark count rates, one can neglect recorded counts due to dark counts.

The only question that remains is how many different states need to be sent in order to characterize the detector matrix G sufficiently. Any detector matrix G can be split horizontally into a detection G_d and a loss part G_l, such that

$$G = \begin{pmatrix} G_d \\ G_l \end{pmatrix}. \tag{5.5.3}$$

Then, let us apply the singular value decomposition on G_l to find

$$G_l = U_l \Sigma_l R_l, \tag{5.5.4}$$

where Σ_l contains the singular values $\vec{\sigma}$ on the diagonal and in this case has the special form

$$\Sigma_l = \begin{pmatrix} \mathrm{diag}(\vec{\sigma}) \\ \hline 0_{(N_{\mathrm{out}}-N_{\mathrm{in}})\times N_{\mathrm{in}}} \end{pmatrix}. \tag{5.5.5}$$

We can combine R_l and Σ_l, thus let us define $Q_l := \mathrm{diag}(\vec{\sigma})R_l$. Therefore, the detector matrix G can equivalently be written as

$$G = \begin{pmatrix} G_d \\ \hline G_l \end{pmatrix} = \begin{pmatrix} \mathbb{I}_d & 0 \\ 0 & U_l \end{pmatrix} \cdot \begin{pmatrix} G_d \\ \hline Q_l \\ 0 \end{pmatrix} =: \begin{pmatrix} \mathbb{I}_d & 0 \\ 0 & U_l \end{pmatrix} \cdot G'. \tag{5.5.6}$$

Hence, after redefining only the loss modes as the ones after the unitary U_l has been applied, it suffices to characterize G'. This reduces the free parameters by quite a bit, since now only $2N_{\mathrm{in}} \cdot (N_{\mathrm{out}} + N_{\mathrm{in}})$ free real parameters need to be found.

In fact this can be further restricted, since we only need to characterize each row of G' up to a global phase and moreover we can use $(G')^\dagger G' = \mathbb{I}_{N_{\mathrm{in}}}$. This already gives us $N_{\mathrm{out}} + N_{\mathrm{in}} + N_{\mathrm{in}}^2$ constraints. Hence, we need to determine

$$n_{\mathrm{free}} = N_{\mathrm{in}}\,(2N_{\mathrm{out}} + N_{\mathrm{in}} - 1) - N_{\mathrm{out}} \tag{5.5.7}$$

free real parameters during testing. Due to Eq. (5.5.2), each states generates N_{out} constraints from observations. Therefore, one needs to probe a detection setup with

$$n_{\mathrm{states}} = \lceil \frac{N_{\mathrm{in}}}{N_{\mathrm{out}}}\,(2N_{\mathrm{out}} + N_{\mathrm{in}} - 1) - 1 \rceil \tag{5.5.8}$$

distinct states to find G' up to a global phase in each row. Then, Lemma 5.3.2 allows us to find a decomposition of G'. If we combine the resulting decomposition of G' with Eq. (5.5.6) we reach a valid decomposition of G, again up to global phases in each row. Therefore, n_{states}-many states are also required for characterizing G.

We can generalize this idea to a procedure to characterize a passive linear optical detection setup with losses and threshold detectors with the steps below.

1. Determine $\vec{\alpha}_1, \ldots \vec{\alpha}_{n_{\mathrm{states}}}$ different complex amplitudes.
2. Shine a coherent laser pulse with complex amplitude $\vec{\alpha}_j$ in the input modes into the detection setup and record the counts in the output ports corresponding to modes b_i, $i = 1, \ldots N_{\mathrm{out}}$.

3. Convert the counts to normalized counts, which approximate the detection probability and connect to G by

$$p_i = 1 - e^{-|\beta_i|^2} = 1 - e^{-|\sum_{j=1}^{N_{\text{in}}} G_{ij}\alpha_j|^2}.$$

4. Repeat steps 2. and 3. until sufficient statistics for a good approximation of the probabilities p_i is reached.

5. Repeat 2. – 4. for $j = 1 \dots n_{\text{states}}$ such that one has sufficient data to solve for G.

6. Then, use Lemma 5.3.2 to find V_d and Theorem 5.3.3 to calculate the lower bound minimizing over G within its uncertainties. Finally, use V_d to construct Bob's POVM elements.

One final remark needs to be made here, since V_d will be used to construct Bob's POVM and therefore for calculating the key rate, this allows us to incorporate any passive linear optical detection setup into the numerical frame work of [67].

5.6 Conclusion

In this chapter, we developed an analytical lower bound required for the subspace estimation in the flag-state squasher. Previous results from Refs. [42, 66, 68] either relied on sophisticated numerical optimizations or were highly specialized to specific detection setups. In contrast, our results presented here are applicable to arbitrary passive linear optical detection setups, including dark counts. Most importantly, our approach removes the previous numerical assumption of monotonicity in the lower bounds. Therefore, these bounds represent a significant step towards more realistic detection setups in QKD security proofs.

A few questions remain open. First, this framework should be extended to active detection setups. Secondly, and perhaps more importantly for practical implementations, it is necessary to incorporate detection setups that are not fully characterized. An initial step towards this was taken in Ref. [69], although it was again specialized to a passive BB84 detection setup.

Chapter 6

Asymptotic Key Rates with Decoy States

This chapter is based on Sections III and IV of our work in Ref. [1]. We begin by reviewing standard decoy-state methods before introducing our improved decoy-state methods. These methods were motivated by a distinct gap in the performance existing security analyses for decoy-state protocols.

Early decoy-state methods [11, 13, 14] used analytic techniques to bound the single-photon observations such as the error rate. Later, the decoy analysis was reformulated as a linear program in Refs. [70, 71] which is typically solved numerically, and tighter than the analytical bounds. Furthermore, the techniques of [70, 71] could incorporate more fine-grained data than simple error rates.

Conversely, the approach taken in Ref. [10] chose a different route to bound the adversary's influence on the single-photon signals. Since multi-photon signals are fully insecure, they were tagged and thus, announced to the adversary Eve. Then, minimizing the possible attacks an adversary could perform, led to an analytic key rate for a WCP source without any decoy states.

However, none of the decoy-state analyses could recover the results from Ref. [10] in the limit of no decoy intensities used. Therefore, none of these methods were tight in both the resulting key rates and the number of intensities used. Moreover, even the tighter decoy-state methods presented in [23], which use additional information on the vacuum error rate, did not recover the key rate of a single weak coherent pulse completely, although they are already an important step towards it.

The methods presented in this chapter successfully bridge this gap, recovering the results of Ref. [10]. While the techniques of this chapter have since been superseded by the more general frameworks in Section 10.6 and Section 10.7, they provide the foundational intuition for those later chapters. In Section 10.7.2, we discuss the tightness of those methods and compare them to the ones shown in this chapter.

6.1 Standard Decoy Methods

First, let us reintroduce the asymptotic key rate formula from Ref. [71] for decoy-state methods. For simplicity, let us assume that in key generation rounds Alice sends weak coherent pulses only with the signal intensity μ_s. Thus, the photon number distribution follows a Poisson distribution,

$$p_{\mu_s}(n) = \frac{\mu_s^n e^{-\mu_s}}{n!}, \tag{6.1.1}$$

with a mean corresponding to the mean-photon number μ_s. Then, using the source replacement scheme, see Section 3.3, and including a shield system A_s, see Section 3.4, one can write the signal state as

$$|\xi\rangle_{AA_sA'} = \sum_{x,n} \sqrt{p_x p_{\mu_s}(n)}\, |x\rangle_A |n\rangle_{A_s} \left|\psi^{(x,n)}\right\rangle_{A'}, \tag{6.1.2}$$

where A and A' represent Alice's systems and the system to be sent to Bob, respectively.

As shown in [71], phase randomization renders the transmitted state a mixture of Fock states, and Eve can determine the photon number via a quantum non-demolishing (QND) measurement. Thus, the final state shared by Alice (A) and Bob (B) is block diagonal in the photon number and given by

$$\rho_{AA_sB} = \sum_n p_n |n\rangle\langle n|_{A_s} \otimes \rho_{AB}^{(n)}. \tag{6.1.3}$$

Furthermore, let $\{\Gamma_{x,y}\}_{xy} = \{|x\rangle\langle x| \otimes F_y^B\}_{xy}$ be the POVM elements acting on systems A and B, corresponding Alice's and Bob's observations $\{\gamma_{x,y}\}$. Hence, following [66, App. B] the key rate is

$$R \geq \sum_{n=0}^{\infty} p_n \min_{\rho_{AB}^{(n)} \in S_n} f\left(\rho_{AB}^{(n)}\right) - p_{\text{pass}} \delta_{\text{leak}}, \tag{6.1.4}$$

where $f(\rho) := D\left(\mathcal{G}\left(\rho\right) || \mathcal{G}\left(\mathcal{Z}\left(\rho\right)\right)\right)$ and $D(.||.)$ is the relative entropy. The optimizations run over the feasible sets S_n conditioned on n photons sent by Alice, which are constructed

from the POVMs $\{\Gamma_{x,y}\}_{xy}$ and the observations $\{\gamma_{x,y}\}_{xy}$ and will be discussed in more detail below. To evaluate this optimization for given S_n and therefore the asymptotic key rate, we use the methods from [67].

If the sets S_n allowed for arbitrary states for each n, independent of the observations, the key rate would trivially be zero. Therefore, the most critical step for obtaining high key rates is finding a tight description of the sets S_n.

As a first example of this task, we introduce previous decoy methods [71], which have been phrased as linear programs. Define

$$\gamma^{\mu}_{y|x} := \Pr(y|x,\mu) \tag{6.1.5}$$

as the probability of observing outcome y given state x with intensity μ was sent by Alice. Then, define the n-photon yields

$$Y^{x,y}_{n} := \Pr(y|x,n), \tag{6.1.6}$$

as the conditional probability of observing outcome y given Alice sent state x with n photons. Since the yields $Y^{x,y}_n$ are independent of Alice's intensity choice, both quantities are connected by the law of total probability via

$$\gamma^{\mu}_{y|x} = \sum_{n=0}^{\infty} p_{\mu}(n) Y^{x,y}_{n}, \tag{6.1.7}$$

where $p_\mu(n)$ is again the Poisson distribution with mean μ. Using these definitions, we can describe the feasible sets S_n in Eq. (6.1.4) in more detail. If we had perfect knowledge of the n-photon yields the feasible set S_n for each photon number n would be given by

$$S_n = \{\rho^{(n)}_{AB} \in \mathrm{S}_{=}(\mathcal{H}_A \otimes \mathcal{H}_B) | \operatorname{Tr}\left[\Gamma_{x,y}\rho^{(n)}_{AB}\right] = p(x) Y^{x,y}_{n}\}, \tag{6.1.8}$$

where additional constraints like the normalization have been left out for simplicity. Here, again the operators $\{\Gamma_{x,y}\}_{xy}$ are Alice and Bob's POVM elements.

However, since the n-photon yield is not directly observable, we constrain it by deriving independent upper and lower bounds from the observations, as in Ref. [71]

$$\begin{aligned} S_n = \{&\rho^{(n)}_{AB} \in \mathrm{S}_{=}(\mathcal{H}_A \otimes \mathcal{H}_B) \mid \\ &\Pr(x) Y^{x,y}_{n,L} \leq \operatorname{Tr}\left[\Gamma_{x,y}\rho^{(n)}_{AB}\right] \leq \Pr(x) Y^{x,y}_{n,U}, \\ &\operatorname{Tr}_B(\rho_{AB}) = \sigma_A\}. \end{aligned} \tag{6.1.9}$$

where σ_A characterizes the reduced state constraints due to the source replacement scheme as shown in Section 3.3 used for P&M protocols. The upper and lower bounds $Y^{x,y}_{n,U/L}$ will be derived in the remaining part of this subsection. However, it is noteworthy at this point, that the bounds $Y^{x,y}_{n,U/L}$ we present are independent from each other for each combination of observations x, y, which is a simplification that will result in some looseness.

Thus, given the shape of the sets S_n and in order to calculate the asymptotic key rate, we are left with the problem of finding the upper and lower bounds of the n-photon yields. Typically, only the single-photon yields are considered, as due to the PNS attack all multi-photon signals are insecure and do not give a positive contribution to the key rate, and vacuum signals only contribute on the order of dark counts.

The n-photon yields and the observations $\gamma^{\mu}_{y|x}$ are related by

$$\gamma^{\mu}_{y|x} = \sum_{n=0}^{\infty} p_{\mu}(n) Y^{x,y}_{n}. \tag{6.1.10}$$

Since one cannot solve a system of equations with infinitely many variables, one instead uses the following upper and lower bounds given a photon number cutoff N_{ph},

$$\gamma^{\mu}_{y|x} \geq \sum_{n \leq N_{\mathrm{ph}}} p_{\mu}(n) Y^{x,y}_{n}, \tag{6.1.11}$$

$$\gamma^{\mu}_{y|x} \leq \sum_{n \leq N_{\mathrm{ph}}} p_{\mu}(n) Y^{x,y}_{n} + (1 - p_{\mathrm{tot}}(\mu)), \tag{6.1.12}$$

where $p_{\mathrm{tot}} = 1 - \sum_{n \leq N_{\mathrm{ph}}} p_{\mu}(n)$ was introduced.

Hence, the lower bound on the single photon yield of observation i, j can be found by the following minimization problem

$$\begin{aligned} Y^{i,j}_{n,L} := \min_{\mathbf{Y}_m} \; & Y^{i,j}_{n} \\ \text{s.t. } & \gamma^{\mu}_{j|i} \leq \sum_{m \leq N_{\mathrm{ph}}} p_{\mu}(m)\, Y^{i,j}_{m} + (1 - p_{tot}(\mu)), \\ & \gamma^{\mu}_{j|i} \geq \sum_{m \leq N_{\mathrm{ph}}} p_{\mu}(m)\, Y^{i,j}_{m}, \\ & \forall \mu \in \{\mu_1, \mu_2, \dots\}, \\ & 0 \leq Y^{i,j}_{m} \leq 1 \; \forall m \in \mathbb{N}_0, \end{aligned} \tag{6.1.13}$$

and a maximization problem with the same constraints has to be solved for the upper bound. This simple independent linear program achieves tight bounds for the sets S_n with

three and more intensities. Therefore, one also achieves tight asymptotic key rates in those cases.

However, this formulation yields no positive key rate for a single intensity; specifically, it fails to recover the no-decoy WCP rate of [10], as noted in the introduction of this chapter. Intuitively, a general decoy-state method should reduce to the no-decoy WCP limit when restricted to a single intensity setting. In Section 6.2, we present a method that explicitly recovers this limit.

The limitation of the linear program in Eq. (6.1.13) arises because each optimization problem processes the statistics for a given observable independently. Consequently, constraints linking different observables are ignored.

For instance, consider the scenario of uniform loss for an unbiased BB84 protocol. Because the loss is uniform (meaning Eve attacks each basis equally), the total detection probability in different bases should be the same. The standard linear program, however, processes these statistics independently and ignores this equivalence. Neglecting such cross-constraints discards available information, which results in a vanishing key rate in the no-decoy setting.

In contrast, the no-decoy approach of [10] utilizes the vacuum error rate and detection probability jointly. This introduces an interdependence between the bounds on the yields, distinguishing it from the independent constraints in Eq. (6.1.13).

6.2 Improved Decoy Methods

Now, we introduce our improved decoy methods. To reiterate, we rely on fully phase-randomized coherent states. To incorporate states which are not fully phase-randomized, the techniques of [68] can be used.

Since we consider fully phase-randomized coherent states, the signal states are block-diagonal in the photon number, as seen in Eq. (6.3.6). This property does not hold for partially phase-randomized states exhibiting phase imperfections, a scenario we explicitly analyze in Section 10.10. Given the ideal block-diagonal structure, Eve's channel can be recast as a direct sum acting on each block separately

$$\mathcal{E} = \bigoplus_{n=0}^{\infty} \mathcal{E}_n, \tag{6.2.1}$$

where each $\mathcal{E}_n : \mathcal{H}_{A'_n} \to \mathcal{H}_{B_M}$. This decomposition of Eve's channel holds because Eve could always apply a QND measurement to detect the photon number and then adjust her

attack based on the photon number [66]. Here, $\mathcal{H}_{B_M}$ is Bob's infinite dimensional state space.

To treat the infinite-dimensional full space $\mathcal{H}_{B_M}$ numerically, a squashing map (see Chapter 4) is required. This squashing map acts on $\mathcal{H}_{B_M}$ and maps it to a finite-dimensional space $\mathcal{H}_B$, such as a qubit space. Mathematically, the squashing map Λ is defined as

$$\Lambda : \mathrm{S}_{=}\left(\mathcal{H}_{B_M}\right) \to \mathrm{S}_{=}\left(\mathcal{H}_B\right), \tag{6.2.2}$$

where $\mathcal{H}_B$ is finite-dimensional.

Due to the block-diagonal structure of Eve's channel, the full channel (comprising Eve's attack and the squashing map) retains a block-diagonal form:

$$\Phi = \bigoplus_{n=0}^{\infty} \Phi_n = \bigoplus_{n=0}^{\infty} \Lambda \circ \mathcal{E}_n : \mathrm{S}_{=}\left(\mathcal{H}_{A'}\right) \to \mathrm{S}_{=}\left(\mathcal{H}_B\right). \tag{6.2.3}$$

By including the squashing map in the full channel, we effectively grant Eve control over it, thereby only overestimating their power. This allows us to use the Choi isomorphism for each channel Φ_n separately, representing its action via the Choi matrix J_n. Consequently, the states in each set S_n must be generated by the same channel J_n acting on the source states $\left|\psi^{(x,n)}\right\rangle_{A'}$ from Eq. (6.1.2).

Therefore, we can rewrite the n-photon yields for all x, y as

$$Y_n^{x,y} = \mathrm{Tr}\left[F_y^B \Phi_n\left(\rho_x^{(n)}\right)\right] = \mathrm{Tr}\left[J_n\left(F_y^B \otimes \left(\rho_x^{(n)}\right)^T\right)\right], \tag{6.2.4}$$

where F_y^B are Bob's POVM elements and $\rho_x^{(n)} = \left|\psi^{(x,n)}\right\rangle\!\left\langle\psi^{(x,n)}\right|$ is the density matrix of the state x sent by Alice with n photons.

Connecting the constraints on the yields with each other via the attack channel is the most crucial difference of our methods compared to previous ones. The channel Φ_n has to be the same across all observations. Hence, we can include all information about the channel in the optimization problem and any correlations between observations are incorporated automatically. The final minimization/maximization problem for n-photon

yield can be written as

$$
\begin{aligned}
Y_{n,L}^{i,j} := \min_{\mathbf{Y}_m, J_m} \; & Y_n^{i,j} \\
\text{s.t. } \gamma_{y|x}^{\mu} \leq & \sum_{m \leq N_{\text{ph}}} p_\mu(m)\, Y_m^{x,y} + (1 - p_{tot}(\mu)), \\
\gamma_{y|x}^{\mu} \geq & \sum_{m \leq N_{\text{ph}}} p_\mu(m)\, Y_m^{x,y}, \\
& Y_m^{x,y} = \text{Tr}\left[J_m \left(F_y^B \otimes \left(\rho_x^{(m)}\right)^T \right) \right], \\
& 0 \leq Y_m^{x,y} \leq 1, \\
& J_m \succeq 0, \; \text{Tr}_B\left[J_m\right] = \mathbb{1}_{A'}, \\
& \forall \mu \in \{\mu_1, \mu_2, \dots\}, \; \forall x, y, \\
& \forall\, 0 \leq m \leq N_{\text{ph}},
\end{aligned}
\tag{6.2.5}
$$

which is convex in the Choi matrices J_n. For the maximization problem, the minimum is simply replaced with a maximum. We note that the additional partial trace and positivity constraints on the Choi matrices result from the fact that we require the full channel Φ_n to be a CPTP map. These additional requirements on the channel are another difference to previous methods.

Moreover, note that in Eq. (6.2.5), the constraints encompass all observations x, y, in contrast to Eq. (6.1.13). The latter considers only the observations and constraints corresponding to a specific event i, j in isolation. This distinction highlights that the SDP formulation leverages significantly more information about the channel within each optimization problem.

6.3 Extension of Asymptotic Key Rate with Differing Intensities

Next, we address errors in signal preparation and detection. The analyses in [66, 71] are restricted to cases where intensities are identical across all signal states. Here, we derive a generalized asymptotic key rate expression that accommodates signal-dependent intensities, a modification motivated by common experimental imperfections. This derivation recovers the standard expression as a special case. We assume that while intensities may vary between signal states, they remain constant across the temporal rounds of the QKD protocol.

This generalization is motivated by practical hardware constraints. Ideally, Alice transmits all key-generation signals with the same intensity μ_s. However, experimental implementations often utilize distinct laser diodes for different states, leading to unavoidable intensity variations. Thus, the effective signal intensity is better described by a set $\{\mu_s^x\}_x$, where each entry is possibly different for each state.

To calculate the asymptotic key rates for this scenario, we start by defining the signal states within the source replacement scheme that includes a shield system, generalizing Section 6.1.

In the key generation rounds, we can write the state between Alice (A), her shield system (A_s) and the signal (A') as

$$|\xi\rangle_{AA_sA'} = \sum_{x,n} \sqrt{\Pr(x,n)}\, |x\rangle_A |n\rangle_{A_s} \left|\psi^{(x,n)}\right\rangle_{A'}, \tag{6.3.1}$$

where $\Pr(x,n)$ denotes the joint probability of Alice choosing setting x and sending n photons. Since the probability of sending n photons depends on the chosen intensity, this dependence is implicit in the notation. If the signal intensity is identical for all settings x (i.e., $\mu(x) = \mu_s$), then $\Pr(x,n)$ simplifies to the standard factorization $\Pr(x,n) = p_x p_{\mu_s}(n)$.

Next, we will explore the structure of the state shared by Alice and Bob and relate it to the standard case of equal intensities. For this purpose, Eve can still apply a QND measurement to determine the photon number of Alice's signal, which following [66] can be recast as an isometry $V : A' \to BE\tilde{E}$, where E and $\tilde{E}$ are Eve's side information and her additional register storing the photon number, respectively. Furthermore, Eve could always apply a different isometry depending on the photon number she observed, thus V can be written as a sum over the photon number,

$$V = \sum_m V_m \Pi_m \otimes |m\rangle_{\tilde{E}}, \tag{6.3.2}$$

where V_m represents the isometry conditioned on m photons and Π_m the projector onto m photons corresponding to Eve's QND measurement. Next, the state after Eve's action is

$$|\chi\rangle_{AA_sBE\tilde{E}} = \sum_{x,n,m} \sqrt{\Pr(x,n)}\, |x\rangle_A |n\rangle_{A_s} \otimes V_m \Pi_m \left|\psi^{(x,n)}\right\rangle_{A'} |m\rangle_{\tilde{E}}. \tag{6.3.3}$$

The projector Π_m acts on $\left|\psi^{(x,n)}\right\rangle_{A'}$ as $\Pi_m \left|\psi^{(x,n)}\right\rangle_{A'} = \delta_{nm} \left|\psi^{(x,n)}\right\rangle_{A'}$, and therefore we can simplify $|\chi\rangle$ to

$$|\chi\rangle_{AA_sBE\tilde{E}} = \sum_{x,n} \sqrt{\Pr(x,n)}\, |x\rangle_A |n\rangle_{A_s} \otimes V_n \left|\psi^{(x,n)}\right\rangle_{A'} |n\rangle_{\tilde{E}}. \tag{6.3.4}$$

After tracing out Eve's systems, the state between Alice, Alice's shield and Bob is

$$\rho_{A_sAB} = \sum_{n=0}^{\infty} \Pr(n) \left|n\right\rangle\!\left\langle n\right|_{A_s} \otimes \sum_{x,x'} \sqrt{\Pr(x|n)\Pr(x'|n)} \left|x\right\rangle\!\left\langle x'\right|_{A} \otimes \mathcal{E}_n \left(\left|\psi^{(x,n)}\right\rangle\!\left\langle\psi^{(x',n)}\right| \right), \tag{6.3.5}$$

where $\mathcal{E}_n(\rho) = \mathrm{Tr}_E \left[V_n \rho \left(V_n\right)^{\dagger} \right]$ is Eve's action transforming system A' to B as a channel conditioned on n photons sent by Alice.

Hence, the final state between Alice (both A and A_s) and Bob (B) can still be written as

$$\rho_{A_sAB} = \sum_{n} \Pr(n) \left|n\right\rangle\!\left\langle n\right|_{A_s} \otimes \rho_{AB}^{(n)}, \tag{6.3.6}$$

which coincides with the results from [66] if we have equal intensities. However, $\Pr(n)$ is now the marginal of the joint probability distribution $\Pr(x,n)$, and $\Pr(x|n)$ is found by Bayes' theorem,

$$\Pr(x|n) = \frac{\Pr(n|x)\Pr(x)}{\Pr(n)}. \tag{6.3.7}$$

Thus, following [66] the key rate is still lower bounded as

$$R \geq \sum_{n} \Pr(n) \min_{\rho_{AB}^{(n)} \in S_n} f\left(\rho_{AB}^{(n)}\right) - p_{\text{pass}} \delta_{\text{leak}}, \tag{6.3.8}$$

where $f(\rho) := D\left(\mathcal{G}\left(\rho\right) || \mathcal{Z} \circ \mathcal{G}\left(\rho\right)\right)$ and $D(\cdot||\cdot)$ is the relative entropy as in [67]. Furthermore, ss in Section 6.1, the feasible sets S_n are defined as

$$\begin{aligned} S_n = \{ & \rho_{AB}^{(n)} \in \mathrm{S}_{=}(\mathcal{H}_A \otimes \mathcal{H}_B) \mid \\ & \Pr(x|n) Y_{n,L}^{x,y} \leq \mathrm{Tr}\left[\Gamma_{x,y} \rho_{AB}^{(n)}\right] \leq \Pr(x|n) Y_{n,U}^{x,y}, \\ & \mathrm{Tr}_B(\rho_{AB}) = \sigma_A \}. \end{aligned} \tag{6.3.9}$$

As before, we are left with the problem of finding the upper and lower bounds on the single photon yields. These can be found by exactly the same SDP as in Eq. (6.2.5) because they are formulated as conditional probabilities (conditioned on Alice sending a particular state).

6.4 Key Rate Optimization Problem including Flag-state Squashing

In this section, we consider how the asymptotic key rate optimization problem must be adapted to incorporate the flag-state squasher discussed in Chapter 5. We assume an arbitrary bound on Bob's subspace weight, as in Eq. (5.1.2). Denoting this bound by $p_L(\tilde{n} \leq N_B|n=1,x)$, where $\tilde{n}$ represents the photon number arriving at Bob and N_B is the cutoff, we derive the following optimization problem for the single-photon contribution:

$$\begin{aligned} &\min_{\rho_{AB}^{(1)}} f\left(\rho_{AB}^{(1)}\right), \\ &\text{s.t. } \Pr(x|1)Y_{L,1}^{x,y} \leq \mathrm{Tr}\left[\Gamma_{x,y}\rho_{AB}^{(1)}\right] \leq \Pr(x|1)Y_{U,1}^{x,y}, \ \forall x,y, \\ &\quad p_L(\tilde{n} \leq N_B|n=1,x) \leq \mathrm{Tr}\left[\Pi_{\leq N_B}\rho_{AB}^{(n=1,x)}\right], \\ &\quad \mathrm{Tr}_B[\rho_{AB}^{(1)}] = \sigma_A \end{aligned} \tag{6.4.1}$$

where σ_A characterizes the constraints originating from the source replacement scheme and $\Pi_{\leq N_B}$ is the projector onto Bob's $\leq N_B$-photon subspace.

However, the bound $p_L(\tilde{n} \leq N_B|n=1,x)$ is not directly accessible from the results presented in Chapter 5. The reason for that is that the observations $m_{\mathrm{mult}|x,n}$ cannot be obtained since neither Alice nor Bob know which of Alice's signals contained exactly $n=1$ photons. Therefore, in the next section we derive how one can obtain the lower bound $p_L(\tilde{n} \leq N_B|n,x)$ purely based on the multi-click observations $m_{\mathrm{mult}|x}$, without conditioning on Alice's photon number. Such a bound is especially required for the decoy-state analysis itself.

6.5 Resulting Subspace Bounds for Decoy-State Methods

To formulate our decoy-state analysis and evaluate the key rate optimization Eq. (6.4.1) while including the flag-state squasher from Chapter 5, we need a lower bound on the weight in Bob's subspace conditioned on Alice sending n photons, i.e. $p(\tilde{n} \leq N_B|x,n)$, depending only on the observations $m_{\mathrm{mult}|x}$ instead of $m_{\mathrm{mult}|x,n}$, which cannot be observed. This bound can be reached in two separate ways. First, a simple version can be found by

considering

$$1 - \frac{m_{\text{mult}|x}}{c_{\geq N_B+1}} = p_L(\tilde{n} \leq N_B|x) \tag{6.5.1}$$

$$\leq p(\tilde{n} \leq N_B|x) = \sum_{m=0}^{\infty} p(m|x)p(\tilde{n} \leq N_B|x,m) \tag{6.5.2}$$

$$\leq p(n|x)p(\tilde{n} \leq N_B|x,n) + (1 - p(n|x)), \tag{6.5.3}$$

Here, the first equality is by the definition of the lower bounds from Theorem 5.3.1 or Theorem 5.3.3, the second equality is due to chain rules of probabilities, and the last inequality is by bounding the sum with one specific n. Then, solving for $p(\tilde{n} \leq N_B|x,n)$ yields

$$p_L(\tilde{n} \leq N_B|x,n) = 1 - \frac{m_{\text{mult}|x}}{p(n|x)\,(1 - c_{\geq N_B+1})}. \tag{6.5.4}$$

This is the lower bound on the subspace population we will use in the formulation of the decoy analysis.

However, one can also use the result of the decoy analysis to construct the lower bound conditioned on n photons sent by Alice and use that in the minimization of the secret key rate, Eq. (6.4.1). If we identify $m^U_{\text{obs}|x,n}$ with the upper bound on $m_{\text{obs}|x,n}$ resulting from the decoy analysis, a tighter bound on the subspace population is

$$p_L(\tilde{n} \leq N_B|x,n) = 1 - \frac{m^U_{\text{mult}|x,n}}{1 - c_{\geq N_B+1}}, \tag{6.5.5}$$

which we will use for the key rate optimization problem in Eq. (6.4.1).

6.6 Example: Biased Passive WCP 6-State

As an example we present the six-state protocol because here the squashing map from [62, 63] exists only without any bias in the basis choice. A bias is advantageous for the final secret key rate, but can also originate from beamsplitters having an imperfect splitting ratio. Therefore, we will use the flag-state squasher derived in the previous Chapter 5. This illustrates the modularity of our approach and allows us to use a bias in Bob's basis choices which is not possible with the simple squashing method from [62, 63]. Furthermore, we will show how to mitigate the problem of different intensities for each signal being sent.

For all results we optimize the signal intensity, which, if the true optimal value is not found, might lead to slight variations in the plots. However, this does not reduce the validity of the secret key rates presented.

6.6.1 Flag-State Squasher

For this specific example, we assume a perfect passive six-state analyser without any losses and use the POVM elements from [62, 63]. A schematic representation of a six-state receiver including our submatrix G can be seen in Fig. 6.1. If we order H, V, D, A, R and

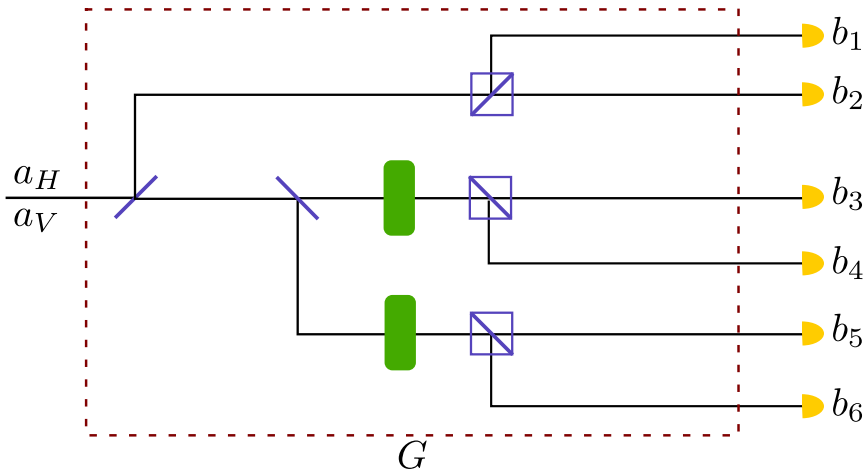


Figure 6.1: Example of a perfect passive detection setup for a six-state analyser for polarization encoding. We label beamsplitters with blue lines, polarizing beamsplitters with additional blue boxes, polarization rotators with green boxes and detectors with yellow paraboloids. The first two beamsplitters divert the incoming light such that with probabilities $p_{\mathrm{HV}}, p_{\mathrm{AD}}$ and p_{RL} the polarization is measured in the horizontal/vertical, anti-diagonal/diagonal and right-/left circular basis after appropriate rotation by the polarization rotators. Finally, the polarizing beamsplitters split the light with equal probability in each polarization basis, such that e.g. horizontal polarized light is measured in the b_1 mode. The submatrix G includes all elements in the red dashed box to a semi-unitary. While this example illustrates an ideal setup, imperfections such as biased splitting ratios or misalignment can be naturally captured within this framework simply by adjusting the corresponding matrix elements of the semi-unitary G. (Figure 6 from Ref. [1])

L as 1 to 6 we find the detector matrix G to be,

$$G = \begin{pmatrix} \sqrt{p_{\mathrm{HV}}} & 0 \\ 0 & \sqrt{p_{\mathrm{HV}}} \\ \sqrt{\frac{p_{\mathrm{AD}}}{2}} & \sqrt{\frac{p_{\mathrm{AD}}}{2}} \\ \sqrt{\frac{p_{\mathrm{AD}}}{2}} & -\sqrt{\frac{p_{\mathrm{AD}}}{2}} \\ \sqrt{\frac{p_{\mathrm{RL}}}{2}} & i\sqrt{\frac{p_{\mathrm{RL}}}{2}} \\ \sqrt{\frac{p_{\mathrm{RL}}}{2}} & -i\sqrt{\frac{p_{\mathrm{RL}}}{2}} \end{pmatrix}. \tag{6.6.1}$$

As we have already laid out in Section 4.1, one needs to find a lower bound for the weight in Bob's $\leq N_B$-photon subspace. Since we have a perfect detection setup, we can apply Theorem 5.3.1 and find

$$p_L(\tilde{n} \leq N_B|x) = 1 - \frac{m_{\mathrm{mult}|x}}{c_{\geq N_B+1}}, \tag{6.6.2}$$

where due to Corollary 5.3.5,

$$c_{\geq N_B+1} = 1 - \sum_{\alpha\in\{\mathrm{HV,AD,RL}\}} p_\alpha^{N_B+1}, \tag{6.6.3}$$

and p_α is Bob's probability of choosing basis α. As a side note it is worth mentioning that in this case as in the motivating example Section 5.2.1, the optimal partition of Theorem 5.3.1 of G is achieved by arranging all rows in G of the same detection basis together.

Furthermore, applying Eq. (6.5.4) for the weight inside the subspace, conditioned on state x sent by Alice with n photons, we find

$$p_L^{\text{6-state}}(\tilde{n} \leq N_B|x,n) = 1 - \frac{m_{\mathrm{mult}|x}}{p(n|x)\left(1 - \sum_\alpha p_\alpha^{N_B+1}\right)}. \tag{6.6.4}$$

Now, we have all necessary tools to formulate both our improved decoy methods and the final key rate optimization for this example.

6.6.2 Decoy Analysis

With our squashing map and the subspace estimation at hand, one can state the minimization problem for the n-photon yields for the 6-state receiver as

$$\begin{aligned}
Y_{n,L}^{i,j} := & \min_{\mathbf{Y}_m, J_0, J_1} \; Y_n^{i,j} \\
& \text{s.t. } \gamma_{y|x}^{\mu} \leq \sum_{m \leq N_{\text{ph}}} p_\mu(m)\, Y_m^{x,y} + (1 - p_{tot}), \\
& \qquad \gamma_{y|x}^{\mu} \geq \sum_{m \leq N_{\text{ph}}} p_\mu(m)\, Y_m^{x,y}, \\
& \qquad Y_m^{x,y} = \text{Tr}\left[J_m \left(F_y^B \otimes \left(\rho_x^{(m)}\right)^T \right) \right], \\
& \qquad 0 \leq Y_m^{x,y} \leq 1, \; \forall\, 0 \leq m \leq N_{\text{ph}}, \\
& \qquad 1 \geq \text{Tr}\left[J_m \left(\Pi_{\leq N_B} \otimes \left(\rho_x^{(m)}\right)^T \right) \right] \\
& \qquad\quad \geq p_L(\tilde{n} \leq N_B | x, m), \; m = 0, 1, \\
& \qquad J_0, \; J_1 \succeq 0, \text{Tr}_B\left[J_0\right] = 1, \; \text{Tr}_B\left[J_1\right] = \mathbb{1}_{A'}, \\
& \qquad \forall \mu \in \{\mu_1, \mu_2, \dots\}, \; \forall x, y,
\end{aligned} \tag{6.6.5}$$

where $\Pi_{\leq N_B}$ is again the projector onto Bob's subspace with $\tilde{n} \leq N_B$ photons and p_L is the lower bound from Eq. (6.6.4). Bob's POVM elements F_y^B acting on the $\leq N_B$-photon subspace are presented in Appendix C.1.2.

6.6.3 Final Key Rate Optimization

Next, using the bounds on single photon yields resulting from Eq. (6.6.5) and the bound for the weight inside the subspace, Eq. (6.5.5),we can formulate the final key rate optimization problem. Following Section 6.4 we find the final optimization to be

$$\begin{aligned}
& \min_{\rho_{AB}^{(1)}} f\left(\rho_{AB}^{(1)}\right), \\
& \text{s.t. } \Pr(x|1) Y_{L,1}^{x,y} \leq \text{Tr}\left[\Gamma_{x,y} \rho_{AB}^{(1)}\right] \leq \Pr(x|1) Y_{U,1}^{x,y}, \; \forall x, y \\
& \qquad \text{Tr}_B[\rho_{AB}^{(1)}] = \sigma_A, \\
& \qquad \text{Tr}\left[\Pi_{\leq N_B} \rho_{AB}^{(1)}\right] \geq p_L^{\text{6-state}}(\tilde{n} \leq N_B | n = 1).
\end{aligned} \tag{6.6.6}$$

6.6.4 Results with Equal Intensities

In this section, we compare our improved decoy-state methods with previous results and demonstrate the performance of the flag-state squashing methods from Chapter 5. For all results presented here, we assume a loss parameter of 0.2 dB/km.

First, Fig. 6.2 compares the flag-state squasher (with $N_B = 1$) against the qubit squashing map from Refs. [62, 63]. We employed uniform basis selection ($p_x = p_y = p_z = \frac{1}{3}$), as the applicability of the squashing map in Refs. [62, 63] is restricted to this symmetric regime. Furthermore, to isolate the effects of the squashing maps, we selected three intensities to minimize the influence of the decoy-state estimation. Both curves utilize the linear program for yield estimation from Eq. (6.1.13), and we assume error correction at the Shannon limit ($f_{\mathrm{EC}} = 1$). As expected, the results are virtually identical, indicating that the flag-state squasher recovers established results even for small subspace dimensions. Consequently, the flag-state squasher can effectively replace the qubit squashing map, offering greater flexibility for detection setups.

Next, Fig. 6.3 illustrates the improvements yielded by our improved decoy-state analysis from Eq. (6.2.5). Specifically, we compare the standard decoy-state method using the linear program from Eq. (6.1.13) with our improved method employing the SDP formulation in Eq. (6.6.5). Since the flag-state squasher supports biased detection setups, we selected $p_z = 0.8$ and $p_x = p_y = 0.1$, maintaining $f_{\mathrm{EC}} = 1$. It is evident that our improved method (orange stars) achieves the same key rate as the linear program from [71] (green circles) while using only one decoy intensity. When comparing the two methods for a fixed number of decoy intensities, our improved approach consistently yields higher key rates. However, as the total number of decoy intensities increases, the key rates from both approaches converge, becoming increasingly indistinguishable.

Finally, in Fig. 6.4, we examine the no-decoy case for the six-state protocol. With no decoy intensities and a flag dimension of $N_B = 1$, we do not strictly recover the analytical key rate of Ref. [10], as illustrated by the deviation between the blue solid curve and the red circles.

This discrepancy stems from the residual looseness in the flag-state squasher, rather than our improved decoy methods. As demonstrated in our work [1, Sec. V], we successfully recover the no-decoy key rates of Ref. [10] for the BB84 protocol when using the qubit squasher [62, 63]. For the six-state protocol, however, we can only approach the results of [10] by increasing N_B. Figure 6.4 shows that setting $N_B = 8$ (purple circles) already yields a significant improvement over $N_B = 1$. This confirms that the flag-state squasher is the source of the discrepancy. We expect to recover the analytical key rate in the limit

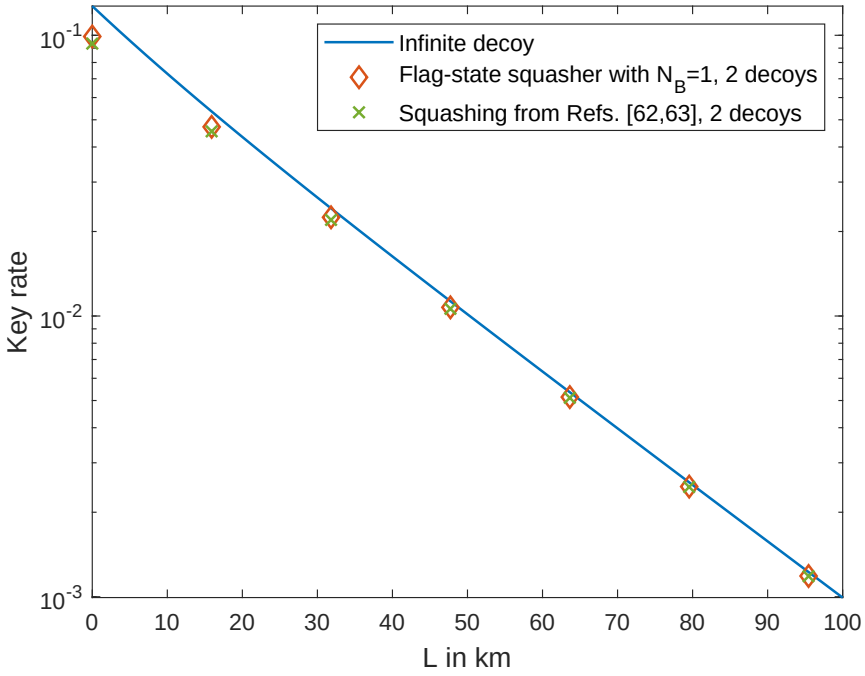


Figure 6.2: Comparison of secret key rates of the passive unbiased six-state protocol over the distance between numerical results for three intensities using the flag-state squasher with $N_B = 1$ presented in this work (orange diamonds) and results from [62, 63] (green crosses). (adapted Figure 7 from Ref. [1])

$N_B \to \infty$, as $p_L \overset{N_B \to \infty}{\longrightarrow} 1$. Importantly, this looseness becomes negligible once decoy states are introduced, as the bounds on the yields become significantly tighter. Furthermore, Fig. 6.4 displays results obtained by imposing only $e_0 = \frac{1}{2}$, following [23] (yellow triangles). Our methods demonstrate significantly better performance compared to this simplified approach, which fails to recover the analytical key rates even in the limit $N_B \to \infty$.

6.6.5 Results with Differing Intensities

As the final example, we analyze the impact of differing intensities on asymptotic key rates, applying the results from Section 6.3. Having different intensities for each state effectively modifies the conditional probability of transmitting a specific signal given a single-photon emission. Crucially, this introduces an effective bias in the bit choices, which may vary across bases. We discuss two strategies to address this issue.

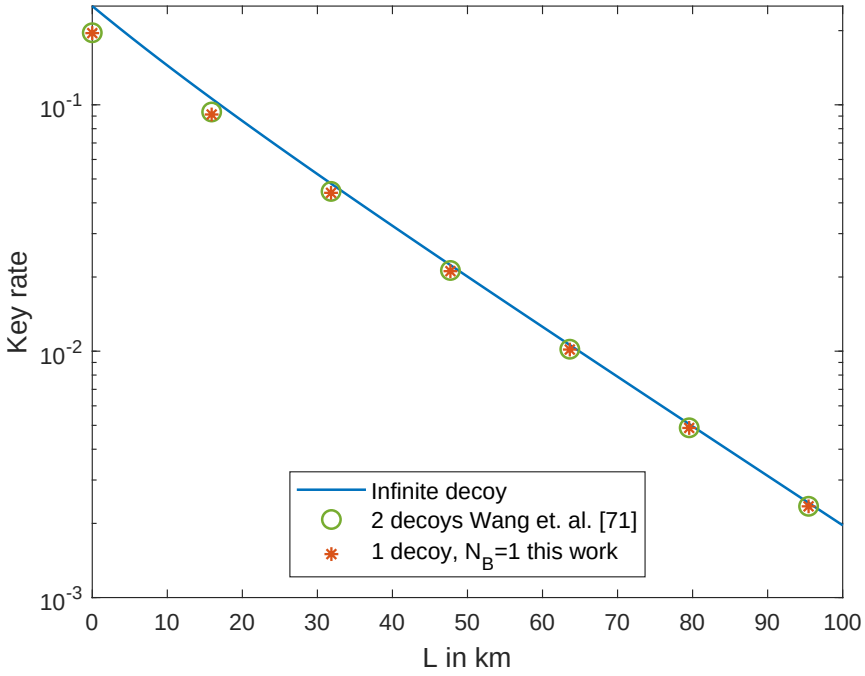


Figure 6.3: Comparison of secret key rates of the passive biased ($p_z = 0.8$, $p_x = p_y = 0.1$) six-state protocol over the distance between our numerical results for one decoy intensity (orange stars) with numerical results from Ref. [71] using two decoy intensities (green circles). (adapted Figure 8 from Ref. [1])

In general, there are two primary approaches, one can either retain the original protocol parameters and accept the altered probabilities, or actively adjust the bit bias to counteract this effect. In the latter case, one attempts to enforce the condition

$$\frac{1}{2} \stackrel{!}{=} \Pr(x|n) = \frac{\Pr(n|x)\Pr(x)}{\Pr(n)}, \tag{6.6.7}$$

which implies that signals with higher intensities must be selected less frequently. A distinct disadvantage of this approach is that the resulting probability of a single-photon event, $\Pr(1)$, may be reduced. Consequently, the prefactor $\Pr(1)$ in the key rate formula Eq. (6.3.8) decreases. While this maintains the relative entropy term, it effectively lowers the total key rate.

On the other hand, if we accept the effective bit bias resulting from the differing intensities, the prefactor $\Pr(1)$ remains unchanged. However, the amount of secret key generated per single photon decreases. This occurs because the effective bit bias reduces the relative

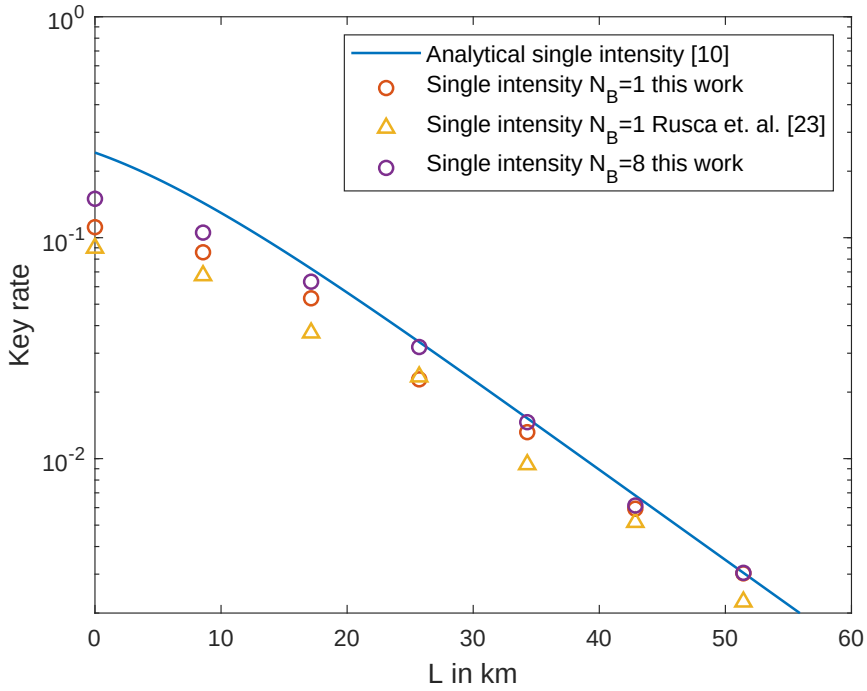


Figure 6.4: Comparison of secret key rates of the 6-state protocol without decoy intensities over distance between our numerical results with $N_B = 1$ (orange circles), $N_B = 8$ (purple circles), numerical results from [23] (yellow triangles) and analytical results from [10] (blue line), demonstrating that our numerical methods successfully recover the known analytical scaling. (adapted Figure 9 from Ref. [1])

entropy term in Eq. (6.3.8), representing the main drawback of applying no countermeasure.

Despite the presence of countervailing effects, the key rates are often very similar. To illustrate this trade-off, we calculated the asymptotic key rates for a signal intensity $\mu_{\text{sig}} = (0.95, 0.15, 0.1, 0.9, 0.95, 0.1)$, ordered according to the six signal states (H, V, D, A, R, L). The equivalent effective signal intensity, defined as the weighted average yielding the same single-photon probability, is $\bar{\mu}_{\text{sig}} = 0.3256$. For simplicity, we maintained a uniform second intensity $\mu_2 = 0.01$ for all states, as this parameter is primarily used for the yield estimation and has less impact on the bias.

Figure 6.5 displays the resulting secret key rates for both the compensated and uncompensated protocols. While both methods yield comparable results, neither achieves the performance of the equivalent ideal protocol with a uniform signal intensity of $\bar{\mu}_{\text{sig}} = 0.3256$.

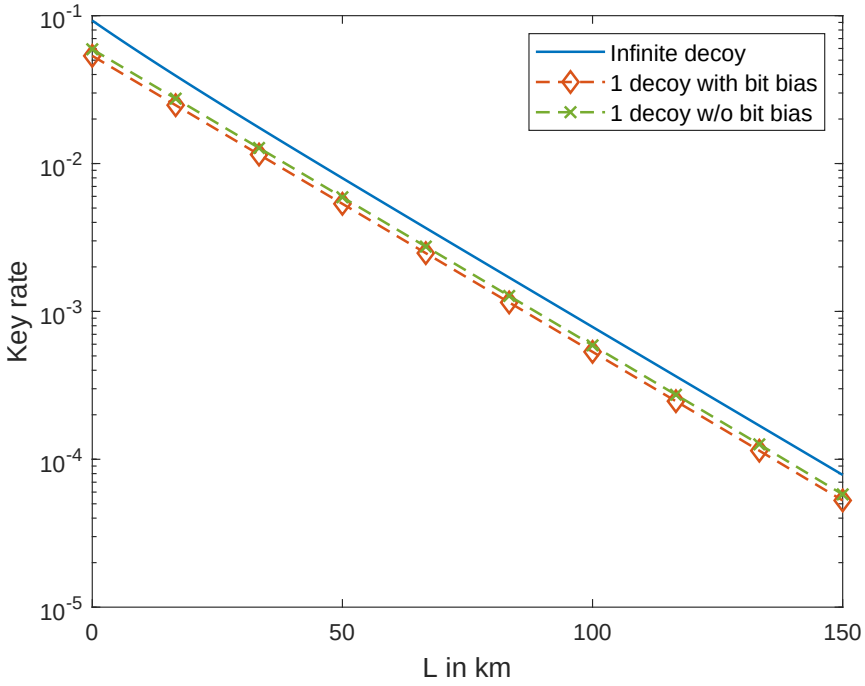


Figure 6.5: Comparison of secret key rates of the 6-state protocol with one decoy intensity over the distance between the rates using a bit bias (red diamonds) and no bit bias (green crosses). (Figure 10 from Ref. [1])

The approach without a bit bias compensation (green crosses) yields a lower rate due to the reduced entropy resulting from the effective bias. Conversely, the approach with a compensating bit bias (red diamonds) also falls short of the uniform-intensity benchmark; this is due to the reduction in the total single-photon probability (i.e., $\Pr(n=1) < p_{\bar{\mu}_{\text{sig}}}(1)$), as previously discussed.

Notably, if we analyze the key rate per single photon, defined as

$$\frac{R_\infty}{\Pr(1)}, \tag{6.6.8}$$

and shown in Fig. 6.6, we observe that the protocol compensating for the bit bias successfully recovers the same key rate per single photon as the equivalent uniform-intensity protocol. Ultimately, a trade-off exists between mitigating the bias caused by differing intensities and maximizing the single-photon emission probability $\Pr(n=1)$. While one could optimize Alice's state-selection probabilities to navigate this trade-off, our results suggest that the potential gain in the final key rate is marginal.

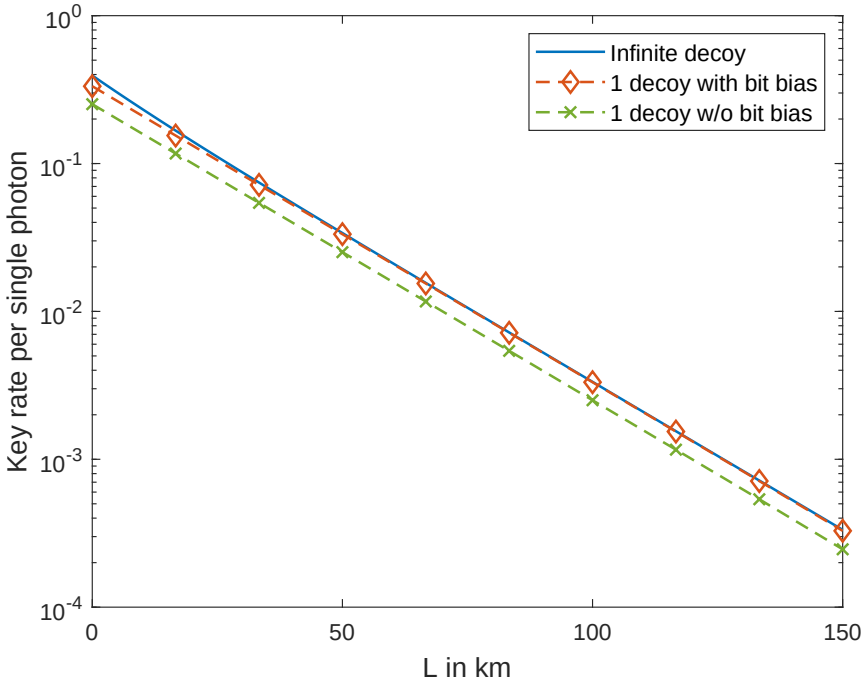


Figure 6.6: Comparison of secret key rates *per single photon* of the 6-state protocol with one decoy intensity over the distance between the rates using a bit bias (red diamonds) and no bit bias (green crosses). (Figure 11 from Ref. [1])

6.7 Conclusion

In this chapter, we presented improved decoy-state methods that tighten the achievable key rates in two significant ways. First, in contrast to the approach in Ref. [23], our methods recover the analytical no-decoy rates from Ref. [10]. Second, they improve upon the key rates obtained using the standard linear program approach of Ref. [71]. Consistent with the findings in Ref. [23], we observe that a single decoy intensity is sufficient for experimental implementations, while still achieving slightly better results than those reported in Ref. [23]. Furthermore, our methods yield higher key rates than previous results obtained with two decoy intensities [71], effectively providing a performance gain through more refined analysis rather than hardware upgrades. Finally, we demonstrated through an explicit example that the subspace estimation from Chapter 5 yields tight key rates for decoy-state protocols.

Chapter 7

Finite-Size Security for Prepare-and-Measure Protocols under IID Attacks

This chapter proves finite-size security of the *fixed-length* prepare-and-measure protocol Protocol 1 and its decoy-state version in the composable security framework of Ref. [54]. We assume an independently and identically distributed collective attack (in short IID). To highlight this IID nature of the attack, we denote all n-round registers with a vector symbol, for example, $\vec{E}$ for Eve's register. Finally, the content of this chapter is based on Sections III – VII and appendices A – C of our work in Ref. [3].

The steps for the security proof we will take are as follows. We first improve finite-size key rates for generic P&M protocols in Section 7.1 and later apply those results to decoy-state protocols in Section 7.2.

7.1 Improved Prepare-and-Measure Finite-size

Before proceeding to the security proof, we briefly define the terms *test rounds* and *generation rounds*, as they appear frequently in the following discussion. A *test round* refers to a round designated for estimating the influence of an adversary. These rounds typically involve the full disclosure of the sent signals and the observed measurement outcomes. A common example is the transmission of signals in a complementary basis (e.g., the X-basis) for parameter estimation.

In contrast, generation rounds are intended to establish the secret key. In these rounds, public announcements are strictly limited; for instance, signals might be sent in the Z-basis, with only the basis choice announced during sifting.

Crucially, the adversary cannot distinguish between test and generation rounds during the transmission phase. Our formal notion of test and generation rounds thus generalizes these specific examples.

7.1.1 Acceptance Test and Resulting Entropy Bounds

Again, this section uses the same protocol structure as presented in the background Section 3.2. We will work towards the security proof of Theorem 7.1.5 of a generic P&M protocol, which is proven in Section 7.1.2. However, first, we show the two improvements of the upcoming Lemma 7.1.3 and Theorem 7.1.1, which are a direct consequence of the acceptance test and its underlying acceptance and feasible sets, defined in this section.

For this purpose, let us briefly introduce some common notation used throughout this chapter, starting with the description of Alice's source. As discussed in Section 3.3, under the source replacement scheme [56, 57], the states Alice prepares can be rewritten in the following form

$$|\psi\rangle_{AA'} = \sum_{a=1}^{d_A} \sqrt{p(a)}\, |a\rangle_A\, |s_a\rangle_{A'}\,, \tag{7.1.1}$$

where $p(a) = p(\text{gen})p(a|\text{gen}) + p(\text{test})p(a|\text{test})$ is the total probability of Alice sending state $|s_a\rangle$. In this equivalent picture, Alice now acts with the POVM elements $\{|a\rangle\langle a|\}$ on her system A.

Furthermore, the single round state after Eve's channel has the form

$$\sigma_{ABE} = (\mathbb{I}_A \circ\ \mathcal{E}_{A'\to BE})\left[|\psi\rangle\langle\psi|_{AA'}\right]. \tag{7.1.2}$$

Under the assumption of an IID collective attack, the final n state can thus be written as

$$\rho_{\vec{A}\vec{B}\vec{E}} = \bigotimes_{i=1}^{n} \sigma_{ABE}, \tag{7.1.3}$$

where the vector is used to indicate that the register contains multiple rounds. We will use this notation extensively in the following. Due to the source replacement scheme, we only have to consider states which satisfy

$$\mathrm{Tr}_{A'}[|\psi\rangle\langle\psi|] = \sigma_A, \tag{7.1.4}$$

because Eve has no access to Alice's lab, as previously described below the security definition Eq. (3.3.7).

Now, we are in a position to discuss some important constructions such as the acceptance and feasible set, necessary for the security proof. Before we formally introduce them in Theorem 7.1.1, let us start by discussing their role in the security proof to gain some intuition. The acceptance set is used in the acceptance test of Protocol 1 to make the decision whether we accept the observations or abort the protocol. The acceptance sets we will consider in this work are always characterized by the expected frequency $\bar{F}$ and an allowed tolerance $\vec{t}$, i.e. we accept if the observed frequencies lie in some tolerance around the expected frequencies.

The feasible set $S_{\vec{\nu}}$ splits the set of all density matrices (with the marginal constraint originating from the source replacement scheme, Eq. (7.1.4)) into the states inside $S_{\vec{\nu}}$ and outside of $S_{\vec{\nu}}$. For the states outside of $S_{\vec{\nu}}$ we would want to show that they will be accepted with low probability, i.e. $\Pr[\Omega_{\text{acc}}] \leq \varepsilon_{\text{AT}}$ and that the protocol will abort. In Theorem 7.1.1, we will construct the acceptance set such that the probability of the protocol accepting any state outside of the feasible set $S_{\vec{\nu}}$ is small.

A one dimensional representation of the construction of the feasible set is depicted in Fig. 7.1. It is determined by the expected frequency $\bar{F}$, the tolerated fluctuation t and the parameter ν characterizing the uncertainty due to finite-size effects. As shown in Fig. 7.1 in $S_{\vec{\nu}}$ we need to consider all density matrices for which there exists an observed frequency which is at most ν away from the true, but unknown value $\text{Tr}[\Gamma\rho]$ (where Γ is the POVM element corresponding to this measurement), and still within the tolerated fluctuation from $\bar{F}$. All those density matrices make up the feasible set $S_{\vec{\nu}}$.

For the states in $S_{\vec{\nu}}$ we will later derive the secure key length in Theorem 7.1.5. Hence, the feasible set allows us to split the set of density matrices and thus, split the security proof by densities matrices in $S_{\vec{\nu}}$ and density matrices outside of $S_{\vec{\nu}}$.

Now, we will start with the security proof by proving secrecy for states outside the feasible set $S_{\vec{\nu}}$ in Theorem 7.1.1.

Theorem 7.1.1 (Accept Probability outside Feasible Set)**.** *Let $\Sigma = \Sigma_{test} \cup \{sift, \perp\}$ be the alphabet of observations from test and generation rounds and $\{\Gamma_k\}_{k\in\Sigma}$ the corresponding POVM elements. Furthermore, let the acceptance set S_{acc} and the feasible set $S_{\vec{\nu}}$ of*

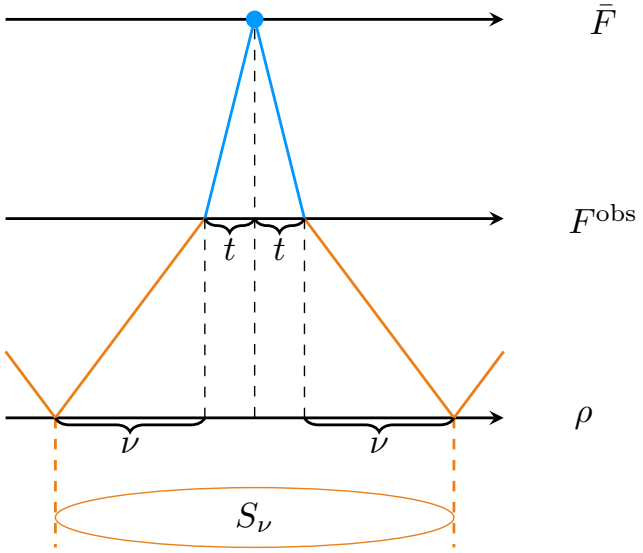


Figure 7.1: One dimensional representation of the construction of the feasible set from the expected frequency $\bar{F}$ with equal upper and lower ν. (Figure 1 from Ref. [3])

Protocol 1 be defined as

$$
\begin{aligned}
S_{\text{acc}} &:= \mathcal{Q}_1 \cap \mathcal{Q}_2, && (7.1.5)\\
\mathcal{Q}_1 &:= \{F^{obs} \in \mathcal{P}(\Sigma) \,|\, \left|F_k^{obs} - \bar{F}_k\right| \leq t_k \;\forall k \in \Sigma_{test}\} && (7.1.6)\\
\mathcal{Q}_2 &:= \{F^{obs} \in \mathcal{P}(\Sigma) \,|\, \left|F_{sift}^{obs} - \frac{n_{sift}}{n}\right| \leq t_{sift}\}, && (7.1.7)\\
S_{\vec{\nu}} &:= \{\rho \in \mathrm{S}_=(AB) \,|\exists F^{obs} \in S_{\text{acc}}, s.t.\; \forall k \in \Sigma \\
&\quad F_k^{obs} - \nu_k^L \leq \mathrm{Tr}[\Gamma_k \rho] \leq F_k^{obs} + \nu_k^U \,, && (7.1.8)\\
&\quad \rho_A = \mathrm{Tr}_{A'}[|\psi\rangle\langle\psi|]\}.
\end{aligned}
$$

Furthermore, define n_{sift} in $\mathcal{Q}_2$ as the total expected click count in the generation rounds surviving sifting. Then, the variational bounds $\nu_k^{L/U}$ can be chosen as

$$\nu_k^{U/L} = \min_{\nu'}\{\nu' \in [0,1] \mid C_k^{U/L}(\nu') = \varepsilon_{AT}\}, \tag{7.1.9}$$

such that for all $\rho_{\vec{A}\vec{B}\vec{E}} = \bigotimes_{i=1}^n \sigma_{ABE}$ and $\sigma \notin S_{\vec{\nu}}$,

$$\Pr[\Omega_{acc}] \leq \varepsilon_{AT}. \tag{7.1.10}$$

Furthermore, each C_k^L *and* C_k^U *is defined by*

$$C_k^L(\nu) := 1 - I_{1-\left(F_k^L-\nu\right)}\left(n+1-\lceil nF_k^L\rceil, \lceil nF_k^L\rceil\right), \tag{7.1.11}$$

$$C_k^U(\nu) := I_{1-\left(F_k^U+\nu\right)}\left(n-\lfloor nF_k^U\rfloor, \lfloor nF_k^U\rfloor+1\right), \tag{7.1.12}$$

where $F_k^{L/U} = \bar{F}_k \mp t_k$ *and* $I_x[a,b]$ *is the regularized incomplete beta function, defined by*

$$I_x(a,b) := \frac{B_x(a,b)}{B_1(a,b)}, \quad \textit{where} \tag{7.1.13}$$

$$B_x(a,b) := \int_0^x t^{a-1}(1-t)^{b-1}\,\mathrm{d}t, \tag{7.1.14}$$

see e.g. [72, §8.17(i)].

Remark 7.1.2. While concentration inequalities such as Chernoff-Hoeffding bounds are analytically simpler, utilizing the exact expression via the incomplete beta function yields significantly tighter variational bounds and thus higher key rates.

Proof. For simplicity define $p_k := \mathrm{Tr}[\Gamma_k\sigma]$ and let $X_1,\ldots,X_{|\Sigma|}$ be the random variables corresponding to the click statistics for each POVM element Γ_k, i.e. $\frac{X_k}{n} = F_k^{\mathrm{obs}}$.

We need to show that for all $\sigma \notin S_{\vec{\nu}}$, $\Pr[\Omega_{\mathrm{acc}}] \leq \varepsilon_{\mathrm{AT}}$. Since $\Omega_{\mathrm{acc}} = \Omega_{\mathrm{AT}} \wedge \Omega_{\mathrm{EV}}$, we find an upper bound as

$$\Pr[\Omega_{\mathrm{acc}}] \leq \Pr[\Omega_{\mathrm{AT}}] = \Pr\left[F^{\mathrm{obs}} \in S_{\mathrm{acc}}\right] \leq \Pr\left[\left|F_k^{\mathrm{obs}} - \bar{F}_k\right| \leq t_k\right] \forall k \in \Sigma, \tag{7.1.15}$$

which holds due to the union bound. For each k, this can be split in two cases as

$$\begin{aligned}&\Pr\left[\left|F_k^{\mathrm{obs}} - \bar{F}_k\right| \leq t_k\right]\\ &\leq \max\left\{\Pr\left[\left|F_k^{\mathrm{obs}} - \bar{F}_k\right| \leq t_k\right]\Big|_{p_k > \bar{F}_k + t_k + \nu_k^U}, \Pr\left[\left|F_k^{\mathrm{obs}} - \bar{F}_k\right| \leq t_k\right]\Big|_{p_k < \bar{F}_k - t_k - \nu_k^L}\right\}.\end{aligned} \tag{7.1.16}$$

As can be seen in Fig. 7.1 for $\sigma \notin S_{\vec{\nu}}$ only these two cases are possible, as otherwise the true value p_k could be generated from states inside the feasible set. Furthermore, since $\sigma \notin S_{\vec{\nu}}$, there must exist at least one k such that either $p_k > \bar{F}_k + t_k + \nu_k^U$ or $p_k < \bar{F}_k - t_k - \nu_k^L$.

First, consider $p_k > \bar{F}_k + t_k + \nu$, for some ν which we aim to find and will be denoted by ν_k^U. For a single index k we find

$$\Pr\left[\left|F_k^{\mathrm{obs}} - \bar{F}_k\right| \leq t_k\right]\Big|_{p_k > \bar{F}_k + t_k + \nu} \tag{7.1.17}$$

$$= \Pr\left[F_k^{\mathrm{obs}} \leq \bar{F}_k + t_k\right]\Big|_{p_k > \bar{F}_k + t_k + \nu} - \Pr\left[F_k^{\mathrm{obs}} < \bar{F}_k - t_k\right]\Big|_{p_k > \bar{F}_k + t_k + \nu} \tag{7.1.18}$$

$$\leq \Pr\left[F_k^{\mathrm{obs}} \leq \bar{F}_k + t_k\right]\Big|_{p_k > \bar{F}_k + t_k + \nu} \tag{7.1.19}$$

Under a collective IID attack, we know that each X_k is binomially distributed with parameter p_k and trials n, in short $X_k \sim \mathrm{Bin}_n(p_k)$. Then, identifying $F_k^{\mathrm{obs}} = \frac{X_k}{n}$, lets us write

$$F_k^{\mathrm{obs}} \leq \bar{F}_k + t \Longleftrightarrow X_k \leq \lfloor n\bar{F}_k + nt_k \rfloor. \tag{7.1.20}$$

We can use the fact that $X_k \sim \mathrm{Bin}_n(p_k)$ because observing outcome k alone is a n-times repeated Bernoulli trial. Then, note that the incomplete beta function $I_x(a,b)$ is the cumulative distribution function of the binomial distribution, which allows us to write $\Pr[X_k \leq m] = I_{1-p_k}(n-m, m+1)$. Thus,

$$\begin{aligned}
&\Pr\left[\left|F_k^{\mathrm{obs}} - \bar{F}_k\right| \leq t_k\right]\Big|_{p_k > \bar{F}_k + t + \nu} && (7.1.21)\\
\leq &I_{1-p_k}\left(n - \lfloor n\left(\bar{F}_k + t_k\right)\rfloor, \lfloor n\left(\bar{F}_k + t_k\right)\rfloor + 1\right) && (7.1.22)\\
\leq &I_{1-\left(F_k^U + \nu\right)}\left(n - \lfloor nF_k^U\rfloor, \lfloor nF_k^U\rfloor + 1\right) && (7.1.23)\\
= &C_k^U(\nu), && (7.1.24)
\end{aligned}$$

where for the second inequality, we used that $p_k > \bar{F}_k + t_k + \nu = F_k^U + \nu$ and that $I_x(a,b)$ is monotonically increasing in x. A similar consideration can be done for $p_k < \bar{F}_k - t_k - \nu = F_k^L - \nu$ and in summary one finds

$$C_k^U(\nu) \geq \Pr\left[\left|F_k^{\mathrm{obs}} - \bar{F}_k\right| \leq t_k\right]\Big|_{p_k > F_k^U + \nu}, \tag{7.1.25}$$

$$C_k^L(\nu) \geq \Pr\left[\left|F_k^{\mathrm{obs}} - \bar{F}_k\right| \leq t_k\right]\Big|_{p_k < F_k^L - \nu}. \tag{7.1.26}$$

For each k one can pick the smallest ν_k^L and ν_k^U such that $C_k^L(\nu_k^L) = \varepsilon_{AT}$ and $C_k^U(\nu_k^U) = \varepsilon_{AT}$, and it holds $\Pr\left[\left|F_k^{\mathrm{obs}} - \bar{F}_k\right| \leq t_k\right] \leq \varepsilon_{AT}$. Thus, we define each $\nu_k^{U/L}$ as

$$\nu_k^{U/L} = \min_{\nu'}\{\nu' \in [0,1] | C_k^{U/L}(\nu') = \varepsilon_{AT}\}, \tag{7.1.27}$$

as in the theorem statement. Therefore, we find for $\sigma \notin S_{\vec{\nu}}$,

$$\begin{aligned}
\Pr[\Omega_{\mathrm{acc}}] &\leq \max_k \Pr\left[\left|F_k^{\mathrm{obs}} - \bar{F}_k\right| \leq t_k\right]\\
&\leq \max_k\{C_k^L(\nu_k^L), C_k^U(\nu_k^U)\} \leq \varepsilon_{AT},
\end{aligned} \tag{7.1.28}$$

due to the definition of each $\nu_k^{U/L}$. □

Although it is not completely obvious at this point, our acceptance set S_{acc} uses observations from both test and generation rounds. The subset $\mathcal{Q}_1$ intuitively serves the purpose

of constraining Eve's attack by observing the outcomes of the test rounds. Even simpler, this could be viewed as if one had observed the error rate only and had made an accept decision based on that. On the other hand, the set $\mathcal{Q}_2$ serves the main purpose of reducing the finite-size corrections. This can be understood as follows, if in key generation rounds Alice and Bob announce which signals survived the sifting stage, one can use this additional information as well to achieve tighter finite-size bounds. Therefore, we added the events $\{\text{sift} , \perp\}$ to the alphabet of test rounds Σ_{test} to indicate that a signal in a key generation survived sifting or not, respectively. Later, in Lemma 7.1.3 we will utilize this additional information contained in the set $\mathcal{Q}_2$ for a better bound on the smooth min-entropy by reducing the finite-size corrections.

Finally, the appearance of the incomplete beta function might seem surprising, but under a collective IID attack, we know that each individual observation is binomial distributed and the incomplete beta function is the cumulative distribution function of the binomial distribution.

As stated previously, our acceptance set S_{acc} uses some information of the generation rounds. Now, we will utilize these additional constraints originating from $\mathcal{Q}_2$ to find a better lower bound on the smooth min-entropy, which will ultimately reduce the finite-size corrections.

In preparation for the next Lemma 7.1.3, let us split off the event $\Omega_{\text{sift}} = \{\mathbf{F}^{\text{obs}} \in \mathcal{Q}_2\} \subseteq \Omega_{\text{AT}}$ corresponding to having sufficient sifted signals during generation rounds. Furthermore, let us introduce a register $\vec{D}$ storing the information of a detection of a sifted signal. Here, we assign $D_i = 1$ if the signal i was part of a generation round, detected and sifted, otherwise we assign $D_i = 0$.

Lemma 7.1.3 (Bounds on the Smooth Min-Entropy for States in the Feasible Set)**.** *Let $\vec{S}, \vec{Y}, \vec{I}, \vec{E}, \vec{D}$ be the registers corresponding to Alice's and Bob's raw keys, public announcements, Eve's quantum side information and the locations of the sifted detections, respectively. We label $D_i = 1$ if a signal in a generation round passed sifting in round i and $D_i = 0$ otherwise. Thus, D_i is a deterministic function of I_i. For $\rho_{\vec{S}\vec{Y}\vec{I}\vec{E}\vec{D}}$ such that*

$$\rho_{\vec{S}\vec{Y}\vec{I}\vec{E}\vec{D}} = \bigotimes_{i=1}^{n} \sigma_{SYIED}, \tag{7.1.29}$$

and $\sigma \in S_{\vec{\nu}}$, with $S_{\vec{\nu}}$ and S_{acc} defined as in Theorem 7.1.1, it holds

$$H_{\min}^{\bar{\varepsilon}}(\vec{S}|\vec{I}\vec{E}\vec{D})_{\rho_{|\Omega_{sift}}} \geq H_{\min}^{\bar{\varepsilon}}(\vec{S}_{sift}|\vec{I}_{sift}\vec{E}_{sift})_{\tau^{sift}_{|\vec{D}=\vec{1}}}, \tag{7.1.30}$$

where

$$\tau^{sift}_{\vec{S}_{sift}\vec{Y}_{sift}\vec{I}_{sift}\vec{E}_{sift}|\vec{D}=\vec{1}} := \bigotimes_{i=1}^{\lfloor n_{sift}-nt_{sift}\rfloor} \sigma_{SYIE|D=1}. \tag{7.1.31}$$

Proof. The proof proceeds with the following steps, first, we will show that conditioning on a specific number of signals passing sifting maintains the IID structure. Then, we will find a lower bound on the smooth min-entropy by using the constraints from set $\mathcal{Q}_2$.

Let $\Omega_{\vec{d}} \subseteq \Omega_{\text{sift}} = \{\mathbf{F}^{\text{obs}} \in \mathcal{Q}_2\}$ be the event of a single string $\vec{d}$, i.e. a particular instance of sifted detection events. By Lemma 2.3.8, we can lower bound the smooth min-entropy by

$$H^{\bar{\varepsilon}}_{\min}(\vec{S}|\vec{I}\vec{E}\vec{D})_{\rho_{|\Omega_{\text{sift}}}} \geq \min_{\substack{\Omega_{\vec{d}}\in\Omega_{\text{sift}} \\ \text{s.t. } \Pr[\Omega_{\vec{d}}]\neq 0}} H^{\bar{\varepsilon}}_{\min}(\vec{S}|\vec{I}\vec{E})_{\rho_{|\Omega_{\vec{d}}}}. \tag{7.1.32}$$

Next, consider the collective IID assumption again, where we have

$$\rho_{\vec{S}\vec{Y}\vec{I}\vec{E}\vec{D}} = \bigotimes_{i=1}^{n} \sigma_{SYIED}. \tag{7.1.33}$$

Then, for all $\vec{d}$ and $\rho_{\vec{S}\vec{Y}\vec{I}\vec{E}\vec{D}|\Omega_{\vec{d}}}$ one can write

$$\rho_{\vec{S}\vec{Y}\vec{I}\vec{E}|\Omega_{\vec{d}}} = \sigma^{(1)}_{SYIE|D=d_1} \otimes \cdots \otimes \sigma^{(n)}_{SYIE|D=d_n}. \tag{7.1.34}$$

For all $\vec{d}$ with $\Omega_{\vec{d}} \subseteq \Omega_{\text{sift}}$, we know that at least $\lfloor n_{\text{sift}} - nt_{\text{sift}}\rfloor$ signals are detected and sifted. The remaining number of signals might contain sifted signals but we need to omit those in order to achieve a fixed length of the raw key register. (Note that this is only required in this proof due to subtleties of handling variable-length registers. In practice, we do not need to omit extra rounds in the raw key register due to [55, Section VII].)

If one discards a signal, no key is produced from this particular signal, and we assign $D_i = 0$. Therefore, we split $\vec{S}$, $\vec{I}$ and $\vec{E}$ in $\vec{S} = \vec{S}_{\text{sift}}\vec{S}_{\text{no}}$, $\vec{I} = \vec{I}_{\text{sift}}\vec{I}_{\text{no}}$ and $\vec{E} = \vec{E}_{\text{sift}}\vec{E}_{\text{no}}$, where all strings in $\vec{S}_{\text{sift}}$, $\vec{E}_{\text{sift}}$ have fixed length $\lfloor n_{\text{sift}} - nt_{\text{sift}}\rfloor$. Thus, we find up to reordering

$$\rho_{\vec{S}\vec{Y}\vec{I}\vec{E}|\Omega_{\vec{d}}} = \left(\bigotimes_{i=1}^{\lfloor n_{\text{sift}}-nt_{\text{sift}}\rfloor} \sigma_{SYIE|D=1}\right) \otimes \tau^{\text{no}}_{\vec{S}_{\text{no}}\vec{Y}_{\text{no}}\vec{I}_{\text{no}}\vec{E}_{\text{no}}|\vec{D}_{\text{no}}=\vec{d}_{\text{no}}}, \tag{7.1.35}$$

where τ^{no} contains all remaining rounds and $\vec{d}_{\text{no}}$ contains as many ones as signals were sifted past $\lfloor n_{\text{sift}} - nt_{\text{sift}}\rfloor$. We note that all registers in this expression have a fixed length.

For simplicity, let us also define the state

$$\tau^{\text{sift}}_{\vec{S}_{\text{sift}}\vec{Y}_{\text{sift}}\vec{I}_{\text{sift}}\vec{E}_{\text{sift}}|\vec{D}=\vec{1}} := \bigotimes_{i=1}^{\lfloor n_{\text{sift}} - n t_{\text{sift}} \rfloor} \sigma_{SYIE|D=1}. \tag{7.1.36}$$

Then, we can rewrite $\rho_{\vec{S}\vec{Y}\vec{I}\vec{E}|\Omega_{\vec{d}}}$ as

$$\rho_{\vec{S}\vec{Y}\vec{I}\vec{E}|\Omega_{\vec{d}}} = \rho_{\vec{S}_{\text{sift}}\vec{S}_{\text{no}}\vec{Y}\vec{I}\vec{E}_{\text{sift}}\vec{E}_{\text{no}}|\Omega_{\vec{d}}} = \tau^{\text{sift}}_{\vec{S}_{\text{sift}}\vec{Y}_{\text{sift}}\vec{I}_{\text{sift}}\vec{E}_{\text{sift}}|\vec{D}=\vec{1}} \otimes \tau^{\text{no}}_{\vec{S}_{\text{no}}\vec{Y}_{\text{no}}\vec{I}_{\text{no}}\vec{E}_{\text{no}}|\vec{D}_{\text{no}}=\vec{d}_{\text{no}}}. \tag{7.1.37}$$

Since both $\vec{S}_{\text{sift}}$ and $\vec{S}_{\text{no}}$ are classical, we can bound the smooth min-entropy regarding $\vec{S}_{\text{no}}$ by zero [47, Lemma 6.7] and thus,

$$H^{\bar{\varepsilon}}_{\min}(\vec{S}|\vec{I}\vec{E})_{\rho_{|\Omega_{\vec{d}}}} \geq H^{\bar{\varepsilon}}_{\min}(\vec{S}_{\text{sift}}|\vec{I}\vec{E})_{\rho_{|\Omega_{\vec{d}}}}. \tag{7.1.38}$$

Furthermore, since $\vec{I}_{\text{no}}\vec{E}_{\text{no}}$ are uncorrelated with $\vec{S}_{\text{sift}}$ one can apply the data processing inequality [47, Theorem 6.2] and remove them from the conditioning register. Hence, we find in summary,

$$H^{\bar{\varepsilon}}_{\min}(\vec{S}|\vec{I}\vec{E})_{\rho_{|\Omega_{\vec{d}}}} \geq H^{\bar{\varepsilon}}_{\min}(\vec{S}_{\text{sift}}|\vec{I}_{\text{sift}}\vec{E}_{\text{sift}})_{\tau^{\text{sift}}_{|\vec{D}=\vec{1}}}, \tag{7.1.39}$$

for all $\Omega_{\vec{d}} \subseteq \Omega_{\text{sift}}$. Since this holds for all $\Omega_{\vec{d}}$, it especially holds for the minimum in Eq. (7.1.32) and we reach the theorem statement

$$H^{\bar{\varepsilon}}_{\min}(\vec{S}|\vec{I}\vec{E}\vec{D})_{\rho_{|\Omega_{\text{sift}}}} \geq \min_{\substack{\Omega_{\vec{d}} \in \Omega_{\text{sift}} \\ \text{s.t. } \Pr[\Omega_{\vec{d}}] \neq 0}} H^{\bar{\varepsilon}}_{\min}(\vec{S}|\vec{I}\vec{E})_{\rho_{|\Omega_{\vec{d}}}} \tag{7.1.40}$$

$$\geq H^{\bar{\varepsilon}}_{\min}(\vec{S}_{\text{sift}}|\vec{I}_{\text{sift}}\vec{E}_{\text{sift}})_{\tau^{\text{sift}}_{|\vec{D}=\vec{1}}}. \tag{7.1.41}$$

□

To gain some insight into the usefulness of this lemma, consider the following situation. In most QKD protocols, one sifts certain events (and discards the remaining outcomes). Without applying Lemma 7.1.3, one would have to map these outcomes to e.g. 0 to keep the input to the privacy amplification procedure of constant length.

Now, one could apply the asymptotic equipartition property directly [28, Corollary 4.10], achieving a scaling of the following form

$$H^{\bar{\varepsilon}}_{\min}(\vec{S}|\vec{I}\vec{E})_{\rho^{\otimes n}_{SIE}} \geq nH(S|EI)_{\rho} - \sqrt{n}g(\bar{\varepsilon}), \tag{7.1.42}$$

where g is a correction term depending on $\bar{\varepsilon}$ and the dimension of S only. However, the von Neumann entropy still scales with the number of sifted signals, whereas the correction term is constant in the total number of signals sent. Hence, at some point the correction term will dominate; leading to zero key rate.

Our Lemma 7.1.3 improves on this situation by having a correction factor which also depends on the sifted signals. Thus, nonzero key rates in significantly higher loss-regimes are possible.

7.1.2 Security Proof against IID Attacks for Generic P&M Protocols

As we have seen in Lemma 7.1.3 for composite events such as $\Omega := \Omega_1 \wedge \Omega_2$, it is beneficial to retain some partial conditioning, particularly in the context of detected events. To address this, we present a version of the leftover-hashing lemma that allows partial conditioning to be preserved for composite events. The proof simply applies Theorem 3.5.2.

Corollary 7.1.4 (Leftover-Hashing with Partial Conditioning)**.** *Let $\sigma_{XYD} \in \mathrm{S}_{=}(XYD)$ be classical on X and Y $\varepsilon \in \left[0, \sqrt{\mathrm{Tr}[\sigma_{XYD}]}\right)$. Let $\mathcal{H}$ be a 2-universal family of hash functions from $\mathcal{X} = \{0,1\}^n$ to $\mathcal{K} = \{0,1\}^l$. Moreover, let $\rho_{S^H} = \frac{1}{|\mathcal{H}|}\sum_{h\in\mathcal{H}} |h\rangle\langle h|_{S^H}$. Similar to Theorem 3.5.2 let $\omega_{KS^HYD_{|\Omega}} = \mathrm{Tr}_X\left[\mathcal{R}\left(\sigma_{XYD_{|\Omega}} \otimes \rho_{S^H}\right)\right]$ and $\Omega := \Omega_1 \wedge \Omega_2 : \mathcal{X}\times\mathcal{Y} \to \{0,1\}$ be an event with $\Pr[\Omega]_\sigma > 0$. Then,*

$$\frac{1}{2}\Pr[\Omega]\left\|\omega_{KS^HYD_{|\Omega}} - \chi_K \otimes \omega_{S^HYD_{|\Omega}}\right\| \leq \frac{1}{2}2^{-\frac{1}{2}\left(H_{\min}^{\varepsilon}(X|YD)_{\sigma_{|\Omega_1}} - l\right)} + 2\varepsilon, \tag{7.1.43}$$

where $\chi_K = \frac{1}{2^l}\mathbb{I}_K$ is the fully mixed state.

Proof. First, observe that the probability of the event $\Omega := \Omega_1 \wedge \Omega_2$ is bounded by

$$\Pr[\Omega] = \Pr[\Omega_2|\Omega_1]\Pr[\Omega_1] \leq \Pr[\Omega_2|\Omega_1]. \tag{7.1.44}$$

Furthermore, we can write the state $\omega_{KS^HYD_{|\Omega}}$ conditioned on the event Ω as

$$\omega_{KS^HYD_{|\Omega}} = \frac{\left(\omega_{KS^HYD_{|\Omega_1}}\right)_{\wedge\Omega_2}}{\Pr[\Omega_2|\Omega_1]}. \tag{7.1.45}$$

Hence, inserting both observations into the left-hand side of (7.1.43) gives

$$\frac{1}{2}\Pr[\Omega]\left\|\omega_{KS^HYD_{|\Omega}}-\chi_K\otimes\omega_{S^HYD_{|\Omega}}\right\| \tag{7.1.46}$$

$$\leq\frac{1}{2}\left\|\left(\omega_{KS^HYD_{|\Omega_1}}\right)_{\wedge\Omega_2}-\chi_K\otimes\left(\omega_{S^HYD_{|\Omega_1}}\right)_{\wedge\Omega_2}\right\| \tag{7.1.47}$$

$$\leq\frac{1}{2}2^{-\frac{1}{2}\left(H_{\min}^{\varepsilon}(X|DY)_{\left(\sigma_{|\Omega_1}\right)_{\wedge\Omega_2}}-l\right)}+2\varepsilon\leq\frac{1}{2}2^{-\frac{1}{2}\left(H_{\min}^{\varepsilon}(X|DY)_{\sigma_{|\Omega_1}}-l\right)}+2\varepsilon, \tag{7.1.48}$$

where in the second step we used Theorem 3.5.2 and in the third step Lemma 2.3.7. □

Next, continuing with our security proof, in Theorem 7.1.5 we tie all of our previous theorems and lemmas together and prove the security of the prepare-and-measure Protocol 1 against IID attacks. In particular, we will make use the definition of the acceptance set from Theorem 7.1.1. Furthermore, we will apply the LHL with partial conditioning from above together with Lemma 7.1.3 to reduce the influence of second order corrections which achieves e.g. better scaling with loss.

Theorem 7.1.5 (Secure Key Rate of a Prepare-and-Measure QKD Protocol)**.** *Fix the parameters of Protocol 1 and define S_{acc} and $S_{\vec{\nu}}$ as in Theorem 7.1.1. Assuming a collective IID attack, the QKD protocol is $\varepsilon_{sec}=\varepsilon_{EV}+\max\{\varepsilon_{AT},\varepsilon_{PA}+2\bar{\varepsilon}\}$ secure, if the key length l, in the case of not aborting, is chosen to be*

$$\begin{aligned}l\leq\lfloor n_{sift}-nt_{sift}\rfloor\min_{\rho\in S_{\vec{\nu}}}\frac{H(S|EI)_{\rho}}{\Pr(sift)}-\lambda_{EC}-\log(2/\varepsilon_{EV})\\-\sqrt{\lfloor n_{sift}-nt_{sift}\rfloor}\Delta\left(\bar{\varepsilon}\right)-2\log\left(\frac{1}{2\varepsilon_{PA}}\right),\end{aligned} \tag{7.1.49}$$

where

$$\Delta(\bar{\varepsilon})=2\log(1+2\dim(S))\sqrt{\log\left(\frac{2}{\bar{\varepsilon}^2}\right)}. \tag{7.1.50}$$

Proof. As stated in Section 3.1, secrecy and correctness imply security. Therefore, we prove each part separately.

For correctness, let us consider Bob's guess $\vec{\bar{S}}$ of Alice's raw key $\vec{S}$. Then, correctness follows immediately as a property of hashing used in the error verification step since

$$\begin{aligned}&\Pr[K_A \neq K_B \wedge \Omega_{\text{acc}}] \leq \Pr[K_A \neq K_B \wedge \Omega_{\text{EV}}] \\ &\leq \Pr\left[(\vec{S} \neq \vec{\bar{S}}) \wedge \Omega_{\text{EV}}\right] \leq \Pr\left[\Omega_{\text{EV}} | (\vec{S} \neq \vec{\bar{S}})\right] \\ &\leq \varepsilon_{\text{EV}}.\end{aligned} \tag{7.1.51}$$

Turning our attention to secrecy two different cases have to be considered. First note, that under the collective IID assumption $\rho_{\vec{S}\vec{Y}\vec{E}\vec{I}\vec{D}}$ can be written as

$$\rho_{\vec{S}\vec{Y}\vec{E}\vec{I}\vec{D}} = \bigotimes_{i=1}^{n} \sigma_{SYEID}. \tag{7.1.52}$$

Thus, we split the proof in two cases $\sigma \in S_{\vec{\nu}}$ and $\sigma \notin S_{\vec{\nu}}$. In the case of $\sigma \notin S_{\vec{\nu}}$, we can apply Theorem 7.1.1 and it holds

$$\frac{1}{2}\Pr[\Omega_{\text{acc}}] \left\| \omega_{K_A G E_{|\Omega_{\text{acc}}}} - \chi_{K_A} \otimes \omega_{G E_{|\Omega_{\text{acc}}}} \right\| \leq \Pr[\Omega_{\text{AT}}] \leq \varepsilon_{\text{AT}}, \tag{7.1.53}$$

where G is the register storing the choice of hash function, and we applied the triangle inequality on the 1-norm in the first inequality.

Next, consider $\sigma \in S_{\vec{\nu}}$, where we will make use of Corollary 7.1.4. We split Ω_{AT} into two parts $\Omega_{\text{AT}} = \Omega_{\text{sift}} \wedge \Omega'_{\text{AT}}$, where $\Omega_{\text{sift}} = \{F_{\text{sift}}^{\text{obs}} \in \mathcal{Q}_2\}$, i.e. the number of detected signals in the generation rounds used for key generation varies at most nt_{sift} from n_{sift}. The event Ω'_{AT} corresponds to the remaining conditions in Ω_{AT}.

Furthermore, let $\vec{E}' = \vec{E}\vec{I}L\vec{D}$ be Eve's side information which contains Eve's quantum side information, public announcements, leakage during error correction and $\vec{D}$ corresponding to the information of a sifted and detected signals in the key generation rounds. As in Lemma 7.1.3 we assign $D_i = 1$ if a signal was detected and survived sifting in round i which was also a generation round, otherwise $D_i = 0$. Recall that Ω_{sift} is a necessary condition for the protocol to accept since it is a subset of Ω_{acc}. Then, still with G the register storing the choice of hash function, by Corollary 7.1.4 it follows

$$\frac{1}{2}\Pr[\Omega_{\text{acc}}] \left\| \omega_{K_A G \vec{E}'_{|\Omega_{\text{acc}}}} - \chi_K \otimes \omega_{S^H \vec{E}'_{|\Omega_{\text{acc}}}} \right\| \leq \frac{1}{2} 2^{-\frac{1}{2}\left(H_{\min}^{\bar{\varepsilon}}(\vec{S}|\vec{E}')_{\rho_{|\Omega_{\text{sift}}}} - l\right)} + 2\bar{\varepsilon}. \tag{7.1.54}$$

Thus, choose a value of l satisfying

$$l \leq H_{\min}^{\bar{\varepsilon}}(\vec{S}|\vec{E}\vec{I}L\vec{D})_{\rho_{|\Omega_{\text{sift}}}} - 2\log\left(\frac{1}{2\varepsilon_{\text{PA}}}\right) \tag{7.1.55}$$

is enough to guarantee $(\varepsilon_{\mathrm{PA}} + 2\bar{\varepsilon})$-secrecy. As usual, we split off the leakage L during the error correction and verification steps by using [47, Lemma 6.8]

$$H_{\min}^{\bar{\varepsilon}}(\vec{S}|\vec{E}\vec{I}L\vec{D})_{\rho_{|\Omega_{\mathrm{sift}}}} \geq H_{\min}^{\bar{\varepsilon}}(\vec{S}|\vec{E}\vec{I}\vec{D})_{\rho_{|\Omega_{\mathrm{sift}}}} - \lambda_{\mathrm{EC}} - \log(2/\varepsilon_{\mathrm{EV}}). \tag{7.1.56}$$

As the next step, we employ Lemma 7.1.3 and we find

$$H_{\min}^{\bar{\varepsilon}}(\vec{S}|\vec{E}\vec{I}\vec{D})_{\rho_{|\Omega_{\mathrm{sift}}}} \geq H_{\min}^{\bar{\varepsilon}}(\vec{S}_{\mathrm{sift}}|\vec{E}_{\mathrm{sift}}\vec{I}_{\mathrm{sift}})_{\tau^{\mathrm{sift}}_{|\vec{D}=\vec{1}}}. \tag{7.1.57}$$

Now, we invoke the asymptotic equipartition property [28, Corollary 4.10], which yields

$$\begin{aligned}
&H_{\min}^{\bar{\varepsilon}}(\vec{S}_{\mathrm{sift}}|\vec{E}_{\mathrm{sift}}\vec{I}_{\mathrm{sift}})_{\tau^{\mathrm{sift}}_{|\vec{D}=\vec{1}}} && (7.1.58)\\
\geq &\lfloor n_{\mathrm{sift}} - nt_{\mathrm{sift}} \rfloor H(S_{\mathrm{sift}}|E_{\mathrm{sift}}I_{\mathrm{sift}})_{\sigma_{|D=1}} && (7.1.59)\\
&- 2\sqrt{n_{\mathrm{sift}} - nt_{\mathrm{sift}}} \log(1 + 2\dim(S)) \sqrt{\log\left(\frac{2}{\bar{\varepsilon}^2}\right)},\\
= &\lfloor n_{\mathrm{sift}} - nt_{\mathrm{sift}} \rfloor H(S_{\mathrm{sift}}|E_{\mathrm{sift}}I_{\mathrm{sift}})_{\sigma_{|D=1}} - \sqrt{n_{\mathrm{sift}} - nt_{\mathrm{sift}}}\Delta(\bar{\varepsilon}). && (7.1.60)
\end{aligned}$$

where we inserted the definition of $\Delta(\bar{\varepsilon})$ from Eq. (7.1.50).

Let us simplify the expression of the entropy even further. For the single round expression one can drop the subscript sift of the registers S and E, if we condition on $D = 1$. Furthermore, we only generate key if $D = 1$, otherwise $H(S|EI; D = 0) = 0$. Hence, we find

$$H(S|EI; D = 1) = \frac{1}{\Pr(D = 1)} H(S|EI). \tag{7.1.61}$$

The event $D = 1$ is equivalent to the state surviving sifting, thus, $\Pr(D = 1) = \Pr(\mathrm{sift})$.

Finally, Eve could have picked any of the states in $S_{\vec{\nu}}$, thus we need to minimize over $S_{\vec{\nu}}$. Then, combining all previous equations together the theorem statement follows. □

We make a few remarks regarding this theorem. First, the label "sift" explicitly refers to signals that survive the sifting step during key generation rounds.

Additionally, in theoretical analyses, the leakage term λ_{EC} is typically estimated following [54] as

$$\lambda_{\mathrm{EC}} = n \Pr(\mathrm{sift}) f_{EC} H(X|Y; \mathrm{sift}). \tag{7.1.62}$$

We will adopt this method to estimate the error correction cost in our subsequent examples. However, in a practical implementation, λ_{EC} must be chosen to reflect the actual maximum

allowed amount of data exchanged during the error correction step. Importantly, the validity of the security proof remains intact regardless of the specific value assigned to the leakage term, as long as it is an upper bound on the actual data exchanged.

Finally, the presence of the probability term in the denominator of the objective function in Theorem 7.1.5 renders the optimization problem non-convex. To address this, we utilize the constraints implied by the acceptance set $\mathcal{Q}_2$ (see Eq. (7.1.7)) to derive a lower bound:

$$\min_{\rho \in S_{\vec{\nu}}} \frac{H(S|EI)_\rho}{\Pr(\text{sift})} \geq \frac{\min_{\rho \in S_{\vec{\nu}}} H(S|EI)_\rho}{\bar{F}_{\text{sift}} + t_{\text{sift}} + \nu^U_{\text{sift}}}. \tag{7.1.63}$$

We employ this bound in Section 7.2 and Section 7.5.

Determining the appropriate values for $\bar{F}$ in Theorem 7.1.5 generally requires a channel model to simulate the honest channel behaviour. This, in turn, necessitates accurate characterization of the transmitter and receiver.

A common pitfall is using the fluctuation parameters $\vec{t}$ to account for mismatches caused by inadequate characterization. This practice should be avoided, as it severely penalizes the key rate.

To remedy this, we present a variable-length security proof in Section 8.3 and Chapter 10. This approach relaxes the dependency on the channel model, however, one still requires Alice's and Bob's devices to operate as specified by their POVMs.

7.2 Decoy-State Methods in Finite-Size Regime

7.2.1 Source Replacement Scheme

Before we start with the finite-size analysis of decoy-state protocols, we recapitulate the relevant notation from Sections 3.3, 3.4 and 6.1. First, to purify Alice's signal states, we introduce a shield system [58] A_s which remains inaccessible to Alice and Bob during the protocol. Furthermore, we apply the source replacement scheme [56, 57], as in Eq. (7.1.1). Then, in each round, Alice prepares the following state,

$$|\xi\rangle_{AA_SA'} = \sum_{n=0}^{\infty} \sum_{a=1}^{d_A} \sqrt{\Pr(a,n)}\, |a\rangle_A |n\rangle_{A_S} |s^a_n\rangle_{A'}\,, \tag{7.2.1}$$

where $\Pr(a,n)$ is the joint probability distribution of Alice sending signal a and n photons. This distribution can be found as the marginal distribution of $\Pr(a,\mu,n)$, the joint

probability distribution of Alice sending signal a, n photons and using intensity μ, i.e.

$$\Pr(a,n) = \sum_{\mu} \Pr(a,n,\mu). \tag{7.2.2}$$

Additionally, we note that for fully phase-randomized weak coherent pulses, the signal states $|s_n^a\rangle$ only depend on a, thus no explicit μ dependence is stated in Eq. (7.2.1).

Since Eve could always perform a QND measurement we can assume ρ_{AA_SB} to be block-diagonal in the photon number n (sent by Alice) as shown in [66]. Under a collective IID attack, Eve's attack can be represented as the same channel $\mathcal{E}_{A'\to B}$ acting on system A' of $|\psi\rangle_{AA_SA'}$ each round. Furthermore, we still assume that Eve has no access to Alice's lab, which is represented as constraints of Alice's marginal state, see Eq. (7.2.6) below.

Hence, the shared state between Alice and Bob is

$$\rho_{AA_SB} = \sum_{n=0}^{\infty} \Pr(n)\, |n\rangle\langle n|_{A_S} \otimes \rho_{AB}^{(n)}, \tag{7.2.3}$$

where

$$\Pr(n) = \sum_{a,\mu} \Pr(a,\mu,n) = \sum_{a,\mu} \Pr(a)\Pr(\mu|a)\Pr(n|\mu,a), \tag{7.2.4}$$

and

$$\rho_{AB}^{(n)} = \sum_{a,a'} \sqrt{\Pr(a|n)\Pr(a'|n)}\, |a\rangle\langle a'|_A \otimes \mathcal{E}_n\left(\left|s_n^a\right\rangle\!\left\langle s_n^{a'}\right|_{A'}\right). \tag{7.2.5}$$

As shown in [66] or [1] this holds for intensities which are different for each signal state, i.e. Alice sends out a different intensity for horizontal polarized than for vertical polarized light. Again, due to the source replacement scheme, we only have to consider states which satisfy

$$\mathrm{Tr}_{A'}[|\xi\rangle\langle\xi|] = \sum_{n} \Pr(n)\, |n\rangle\langle n|_{A_S} \otimes \rho_A^{(n)}. \tag{7.2.6}$$

Furthermore, Bob receives the following state conditioned on Alice choosing signal a and sending n photons

$$\sigma_B^{(a,n)} = \mathcal{E}_{A'\to B}\left(|s_n^a\rangle\langle s_n^a|_{A'}\right). \tag{7.2.7}$$

Since $|s_n^a\rangle_{A'}$ is independent of the choice of the intensity μ, $\sigma_B^{(a,n)}$ is also independent of μ.

7.2.2 Asymptotic Scenario for Decoy-State Methods

Now, we will quickly summarize asymptotic decoy-state methods and introduce a common notation before turning our attention to the finite-size case. In the asymptotic regime, see Section 6.1, the yields are defined as the probability that Bob detects the outcome b given that Alice sent the state a with n photons and are denoted by $Y_n^{a,b}$. Importantly, for fully phase-randomized coherent states, the n-photon yields $Y_n^{a,b}$ are independent of the intensity μ, which does not hold for partially phase-randomized states exhibiting phase imperfections, a scenario we explicitly analyze in Section 10.10. Therefore, formally, we define the yields as $Y_n^{a,b} := \Pr(b|a,n)$, and due to the independence on μ they satisfy

$$Y_n^{a,b} := \Pr(b|a,n) = \Pr(b|a,n,\mu) = \operatorname{Tr}\left[M_b^B \sigma_B^{(a,n,\mu)}\right]. \tag{7.2.8}$$

Therefore, using simple chain rules for probabilities,

$$\begin{aligned} \Pr(b|a,\mu) &= \sum_{n=0}^{\infty} \Pr(n|a,\mu)\Pr(b|a,n,\mu) \\ &= \sum_{n=0}^{\infty} \Pr(n|a,\mu) Y_n^{a,b}, \end{aligned} \tag{7.2.9}$$

which holds for all intensities μ.

As described in Protocol 3, Alice chooses between a test and a generation round. Thus, if we knew the n-photon yields $Y_n^{a,b}$, we would find the following constraints in the test rounds

$$\Pr(\text{test}|a,n)\operatorname{Tr}\left[\Gamma_{a,b}\rho_{AB}^{(n)}\right] = \Pr(\text{test},a|n)Y_n^{a,b}. \tag{7.2.10}$$

Under an IID attack, Eve's attack (for each photon number) can still be described by probabilities, the yields, and the observations will give bounds on the probabilities.

Hence, the question is to find, or rather bound, the yields $Y_n^{a,b}$ in the finite-size regime, and in the next section we will lay out how the methods from Ref. [71] or Section 6.2 need to be altered in order to be valid in the finite-size regime.

7.2.3 Finite-Size Scenario for Decoy-State Methods

So far we assumed perfect knowledge of the observations $\Pr(b|a,\mu)$ and found a relation to the yields $Y_n^{a,b}$. However, in the finite-size regime, we do not have access to the probabilities (yields) $Y_n^{a,b} = \Pr(b|a,\mu)$ directly.

Thus, we turn our attention to estimating the n-photon yields $Y_n^{a,b}$ based on the *finite* information generated in the test rounds. After we found a relation of the yields to the frequencies observed in test rounds, we will formulate the key-length expression. Again, since we assume an IID attack, the yields will play the same role as in the asymptotic case, representing the probabilities of Bob observing an event given that Alice sent n photons, thus encompassing Eve's attack. However, the estimation of the yields is now much looser due to finite-size effects.

Let us start by relating the probability $\Pr(k) = \mathrm{Tr}[\Gamma_k \rho]$ of observing outcome $k = (a, b, \mu, \mathrm{test}) \in \Sigma_{\mathrm{test}}$ to the constraints imposed by the acceptance test. For any observed frequency $F^{\mathrm{obs}} \in S_{\mathrm{acc}}$ passing the acceptance test it holds for all $k = (a, b, \mu, \mathrm{test}) \in \Sigma_{\mathrm{test}}$,

$$\left| F_k^{\mathrm{obs}} - \bar{F}_k \right| \leq t_k, \tag{7.2.11}$$

or equivalently

$$\bar{F}_k - t_k \leq F_k^{\mathrm{obs}} \leq \bar{F}_k + t_k. \tag{7.2.12}$$

Furthermore, for $\rho_{AB} \in S_{\vec{\nu}}$, see Theorem 7.1.1, the observed frequency $F^{\mathrm{obs}} \in S_{\mathrm{acc}}$ satisfies,

$$F_k^{\mathrm{obs}} - \nu_k^L \leq \mathrm{Tr}[\Gamma_k \rho] \leq F_k^{\mathrm{obs}} + \nu_k^U, \tag{7.2.13}$$

for all $k = (a, b, \mu, \mathrm{test}) \in \Sigma_{\mathrm{test}}$. Therefore, inserting Eq. (7.2.12) gives

$$\bar{F}_k - t_k - \nu_k^L \leq \mathrm{Tr}[\Gamma_k \rho] \leq \bar{F}_k + t_k + \nu_k^U, \tag{7.2.14}$$

for all $k = (a, b, \mu, \mathrm{test}) \in \Sigma_{\mathrm{test}}$. Finally, we can identify $\mathrm{Tr}[\Gamma_k \rho]$ with

$$\begin{aligned} \mathrm{Tr}[\Gamma_k \rho] &= \Pr(\mu, \mathrm{test}|a, b)\, \mathrm{Tr}[\Gamma_{a,b} \rho] \\ &= \Pr(\mu, \mathrm{test}|a, b)\, \Pr(a, b) = \Pr(a, b, \mu, \mathrm{test}) \end{aligned} \tag{7.2.15}$$

and inserting this into Eq. (7.2.14) results in

$$\bar{F}_k - t_k - \nu_k^L \leq F_k^{\mathrm{obs}} - \nu_k^L \leq \Pr(a, b, \mu, \mathrm{test}) \leq F_k^{\mathrm{obs}} + \nu_k^U \leq \bar{F}_k + t_k + \nu_k^U, \tag{7.2.16}$$

which again holds for all $k = (a, b, \mu, \mathrm{test}) \in \Sigma_{\mathrm{test}}$.

Now, note that during the acceptance test, we accept on the frequencies generated by the joint probability distribution $\Pr(k) = \Pr(a, b, \mu, \mathrm{test})$. However, in the asymptotic regime we formulated the decoy analysis Eq. (7.2.9) in terms of conditional probabilities. Therefore, applying the same conditioning on (a, μ, test), we can write asymptotically

$$\begin{aligned} \frac{\Pr(a, b, \mu, \mathrm{test})}{\Pr(a, \mu, \mathrm{test})} &= \Pr(b|a, \mu, \mathrm{test}) \\ &= \sum_{n=0}^{\infty} \Pr(n|a, \mu, \mathrm{test}) Y_n^{a,b}, \end{aligned} \tag{7.2.17}$$

where we remark that in the finite-size regime, we have the additional conditioning on the "test" event.

Hence, inserting Eq. (7.2.16), we find the following relation between the yields and our acceptance set characterized by $\bar{F}$ and $\vec{t}$

$$\begin{aligned} \frac{\bar{F}_k + t_k + \nu_k^U}{\Pr(a,\mu,\text{test})} &\geq \sum_{n=0}^{\infty} \Pr(n|a,\mu,\text{test}) Y_n^{a,b}, \\ \frac{\bar{F}_k - t_k - \nu_k^L}{\Pr(a,\mu,\text{test})} &\leq \sum_{n=0}^{\infty} \Pr(n|a,\mu,\text{test}) Y_n^{a,b}, \end{aligned} \tag{7.2.18}$$

which holds for all $k = (a, b, \mu, \text{test}) \in \Sigma_{\text{test}}$. Again, as in the asymptotic case, the yields $Y_n^{a,b}$ satisfy

$$\Pr(\text{test}|a,n) \operatorname{Tr}\left[\Gamma_{a,b}\rho_{AB}^{(n)}\right] = \Pr(\text{test},a|n) Y_n^{a,b}, \tag{7.2.19}$$

which allows us to constraint the n-photon states.

The expressions in Eqs. (7.2.18) and (7.2.19) would already be enough to formulate the upcoming security statement. However, Eq. (7.2.18) contains an infinite number of unknown variables, the yields $Y_n^{a,b}$. Therefore, let us introduce the photon number cutoff N_{ph} which introduces a small tolerance accounted by

$$p_{\text{tot}}(\mu|a) := \sum_{n \leq N_{\text{ph}}} \Pr(n|a,\mu,\text{test}), \tag{7.2.20}$$

but allows us to reduce the infinite sum to a finite sum and formulate the finite linear program in Eq. (7.2.30) later.

Furthermore, to simplify the notation, we note the following. Typically, the constraints on the n-photon states will be of the form of Eq. (7.2.19) and are most naturally expressed in terms of the right-hand-side of Eq. (7.2.19), which we will call the joint yields. Therefore, let us define the joint yields as

$$\gamma_n^{a,b} := \Pr(\text{test},a|n) Y_n^{a,b}. \tag{7.2.21}$$

Whenever it will simplify the presentation, we will switch between the picture of the joint $\gamma_n^{a,b}$ and conditional yields $Y_n^{a,b}$ but one could equally well express the decoy analysis in terms of either yield definitions.

Finally, before we state Corollary 7.2.1, a remark is in order. Again, the entrywise bounding of $\Pr(k) = \Pr(a, b, \mu, \text{test})$ is in contrast to previous numerical finite-size methods presented in [43], where 1-norm constraints were used. With those constraints, one could still claim $\left|F_k^{\text{obs}} - \Pr(k)\right| \leq \nu$. However, this would be a "worst-case" upper bound because ν satisfies

$$\left\|\vec{F}^{\text{obs}} - \vec{p}\right\|_1 \leq \nu, \tag{7.2.22}$$

for $p_k := \Pr(k)$. Hence, we would allow each element to vary in the worst possible way. On the other hand, entrywise constraints are much more in line with previous numerical decoy-state methods of Ref. [71].

After this discussion, we can state Corollary 7.2.1, which lets us evaluate the secure key length for decoy-state protocols.

Corollary 7.2.1 (Security for Decoy-state Protocols (Optimized Yields))**.** *Fix the parameters of Protocol 3, and define S_{acc} and $S_{\vec{\nu}}$ as in Theorem 7.1.1. Assuming a collective IID attack, the QKD protocol is $\varepsilon_{sec} = \varepsilon_{EV} + \max\{\varepsilon_{AT}, \varepsilon_{PA} + 2\bar{\varepsilon}\}$ secure, if the key length l, in the case of not aborting, is chosen to be*

$$\begin{aligned} l \leq \frac{\lfloor n_{sift} - n t_{sift} \rfloor}{\bar{F}_{sift} + t_{sift} + \nu^U_{sift}} \sum_{n=0}^{\infty} \Pr(n) \min_{\substack{\rho^{(n)} \in S_n(\gamma_n) \\ \{\gamma_n\}_n \in G}} H(S|EI)_{\rho^{(n)}} \\ - \sqrt{\lfloor n_{sift} - n t_{sift} \rfloor}\, \Delta(\bar{\varepsilon}) - \lambda_{EC} - \log(2/\varepsilon_{EV}) - 2\log\left(\frac{1}{2\varepsilon_{PA}}\right), \end{aligned} \tag{7.2.23}$$

where $\Delta(\bar{\varepsilon})$ is defined as in Eq. (7.1.50), and the sets S_n correspond to the feasible n-photon subspaces and are defined by

$$\begin{aligned} S_n(\gamma_n) = \{\rho^{(n)} \in \mathrm{S}_=(AB) | \exists \{\gamma_n\} \in G \ s.t. \\ \forall (a, b, \mu, test) \in \Sigma_{test}, \\ \Pr(\text{test}|a, b, n) \operatorname{Tr}\left[\Gamma_{a,b} \rho^{(n)}\right] = \gamma_n^{a,b}, \\ \rho_A^{(n)} = \operatorname{Tr}_{A_s A'}[|n\rangle\langle n|_{A_s} |\xi\rangle\langle\xi|]\}, \end{aligned} \tag{7.2.24}$$

where

$$
\begin{aligned}
G := \{\{\gamma_m\} | \gamma_m^{a,b} = \Pr(\text{test}, a|m) Y_m^{a,b}, \\
&\frac{\bar{F}_k - t_k - \nu_k^L}{\Pr(a,\mu,\text{test})} \leq \sum_{m \leq n_{\text{ph}}} \Pr(m|a,\mu,\text{test}) Y_m^{a,b} \left(1 - p_{tot}(\mu|a)\right), \\
&\frac{\bar{F}_k + t_k + \nu_k^U}{\Pr(a,\mu,\text{test})} \geq \sum_{m \leq n_{\text{ph}}} \Pr(m|a,\mu,\text{test}) Y_m^{a,b}, \\
&0 \leq Y_m^{a,b} \leq 1, \ \forall m \in \mathbb{N}_0, \ \forall k = (a,b,\mu,\text{test}) \in \Sigma_{test}\}.
\end{aligned}
\tag{7.2.25}
$$

Proof. The correctness and secrecy for $\rho \notin S_{\vec{\nu}}$ follow from Theorem 7.1.5. For $\rho \in S_{\vec{\nu}}$ we continue from the key length expression in Theorem 7.1.5 and only focus on the term containing the optimization over $\rho \in S_{\vec{\nu}}$.

First, we bound the fraction by the same lower bound on the denominator Pr(sift) as in Eq. (7.1.63) to find

$$
\min_{\rho \in S_{\vec{\nu}}} \frac{H(S|EI)_\rho}{\Pr(\text{sift})} \geq \frac{\min_{\rho \in S_{\vec{\nu}}} H(S|EI)_\rho}{\bar{F}_{\text{sift}} + t_{\text{sift}} + \nu_{\text{sift}}^U}. \tag{7.2.26}
$$

Next, we will reformulate the entropy term in terms of contributions from different photon numbers. To do so, note that since we assume fully phase-randomized coherent states and as described in [66, App. B], Eve can perform a QND measurement on the photon number. Hence, Eve knows the photon number of the states sent and we split Eve's register into $E = E'M$, where M is classical and labels the photon number of Alice's sent states. Thus,

$$
\min_{\rho \in S_{\vec{\nu}}} H(S|EI)_\rho \geq \min_{\rho \in S_{\vec{\nu}}} \sum_{n=0}^{\infty} \Pr(M = n) H(S|E'I)_{\rho|M=n}. \tag{7.2.27}
$$

All states $\rho_{AB} \in S_{\vec{\nu}}$ must be of the form of Eq. (7.2.3) and furthermore all $\rho_{AB|M=n} = \rho_{AB}^{(n)}$ satisfy Eq. (7.2.5). Additionally, Eq. (7.2.10) establishes relations between the state $\rho_{AB}^{(n)}$ and the yields. With the definition of the joint yields γ_n in Eq. (7.2.21) and by applying the results of the acceptance test, we know that the yields need to satisfy the constraints

summarized in G, as laid out in the preceding section. Hence, we find

$$\min_{\rho\in S_{\vec{\nu}}} H(S|EI)_\rho \geq \sum_{n=0}^{\infty} \Pr(M=n) \min_{\rho\in S_{\vec{\nu}}} H(S|E'I;M=n)_{\rho_{|M=n}} \tag{7.2.28}$$

$$\geq \sum_{n=0}^{\infty} \Pr(M=n) \min_{\substack{\rho^{(n)}\in S_n(\gamma_n)\\ \{\gamma_n\}_n\in G}} H(S|E'I;n)_{\rho^{(n)}}, \tag{7.2.29}$$

where for the first inequality we swapped the minimization with the sum and for the second inequality we inserted the definitions of $S_n(\gamma_n)$ and G. Furthermore, we used that for $\rho \in S_{\vec{\nu}}$ it holds $\rho^{(n)} \in S_n(\gamma_n)$ and $\{\gamma_n\} \in G$.

Combining this with Eq. (7.2.26) the corollary follows. □

The double optimization of Corollary 7.2.1 can be numerically unstable and does not fit the framework for decoy-state protocols presented in Ref. [71] nor the general framework for QKD protocols shown in Ref. [67]. In Ref. [71], the decoy analysis was formulated as a linear program which outputs upper and lower bounds on the yields (or e.g. error rates) and then calculates the key rates based on these bounds.

Therefore, we will reformulate Corollary 7.2.1. However, the previous version will be tighter. As in the asymptotic case, we will formulate the key rate evaluation in a two-step process. First, find upper and lower bounds on the yields $Y_n^{a,b}$ based on the observations and, second, calculate the key rate with those bounds.

We can solve for upper and lower bounds of the n-photon yields Y_n with a linear program given as

$$\begin{aligned} Y_{n,L}^{a,b} := \min_{\{\mathbf{Y}_m\}_m}\ & Y_n^{a,b} \\ \text{s.t. } & \frac{\bar{F}_k - t_k - \nu_k^L}{\Pr(a,\mu,\text{test})} \leq \sum_{m\leq N_{\text{ph}}} \Pr(m|a,\mu,\text{test})\, Y_m^{a,b} + (1-p_{\text{tot}}(\mu|a))\,, \\ & \frac{\bar{F}_k + t_k + \nu_k^U}{\Pr(a,\mu,\text{test})} \geq \sum_{m\leq N_{\text{ph}}} \Pr(m|a,\mu,\text{test})\, Y_m^{a,b}, \\ & 0 \leq Y_m^{a,b} \leq 1\ \forall m\in\mathbb{N}_0,\ \forall k=(a,b,\mu,\text{test})\in\Sigma_{\text{test}}, \end{aligned} \tag{7.2.30}$$

where the yields $Y_m^{a,b}$ were rearranged into a vector $\mathbf{Y}_m$ for each $m \in \mathbb{N}_0$, and as before $p_{\text{tot}}(\mu|a) = \sum_{m\leq N_{ph}} \Pr(m|a,\mu,\text{test})$. Therefore, each $\mathbf{Y}_m$ is a vector containing the probabilities $Y_m^{a,b}$ as entries. The same optimization problem stated with a maximization gives rise to the upper bounds.

As before, we define upper and lower bounds on the joint yields $\gamma_{n,L}^{a,b}$ and $\gamma_{n,U}^{a,b}$ by

$$\begin{aligned}\gamma_{n,L}^{a,b} &= \Pr(a,\text{test}|n)Y_{n,L}^{a,b},\\ \gamma_{n,U}^{a,b} &= \Pr(a,\text{test}|n)Y_{n,U}^{a,b}.\end{aligned} \tag{7.2.31}$$

Thus, it follows

$$\gamma_{n,L}^{a,b} \leq \Pr(\text{test}|a,b,n)\,\mathrm{Tr}\left[\Gamma_{a,b}\rho_{AB}^{(n)}\right] \leq \gamma_{n,U}^{a,b}, \tag{7.2.32}$$

for all $n \in \mathbb{N}_0$ and $(a,b) \in \Sigma_{\text{test}}$, and we can state the following Corollary 7.2.2.

Corollary 7.2.2 (Security for Decoy-state Protocols (Upper and Lower Bounds on Yields))**.** *Fix the parameters of Protocol 3, and define S_{acc} and $S_{\vec{\nu}}$ as in Theorem 7.1.1. Assuming a collective IID attack, the QKD protocol is $\varepsilon_{sec} = \varepsilon_{EV} + \max\{\varepsilon_{AT}, \varepsilon_{PA} + 2\bar{\varepsilon}\}$ secure, if the key length l, in the case of not aborting, is chosen to be*

$$\begin{aligned}l \leq\ & \frac{\lfloor n_{sift} - nt_{sift}\rfloor}{\bar{F}_{sift} + t_{sift} + \nu_{sift}^{U}} \sum_{n=0}^{\infty} \Pr(n) \min_{\rho^{(n)} \in \tilde{S}_n} H(S|EI)_{\rho^{(n)}} \\ & - \sqrt{\lfloor n_{sift} - nt_{sift}\rfloor}\,\Delta\left(\bar{\varepsilon}\right) - \lambda_{EC} - \log(2/\varepsilon_{EV}) - 2\log(1/2\varepsilon_{PA}),\end{aligned} \tag{7.2.33}$$

where $\Delta\left(\bar{\varepsilon}\right)$ is defined as in Eq. (7.1.50), and the sets $\tilde{S}_n$ correspond to the feasible n-photon subspaces and are defined by

$$\begin{aligned}\tilde{S}_n := \{\rho^{(n)} \in \mathrm{S}_{=}(AB) | \forall (a,b,\mu,\text{test}) \in \Sigma_{test}\\ \gamma_{n,L}^{a,b} \leq \mathrm{Tr}\left[\Gamma_{a,b}\rho^{(n)}\right] \leq \gamma_{n,U}^{a,b},\\ \rho_A^{(n)} = \mathrm{Tr}_{A_s A'}[|n\rangle\langle n|_{A_s} |\xi\rangle\langle\xi|]\}.\end{aligned} \tag{7.2.34}$$

Proof. The correctness and secrecy for $\rho \notin S_{\vec{\nu}}$ again follow directly from Theorem 7.1.5. For $\rho \in S_{\vec{\nu}}$ we continue from the key length expression in Corollary 7.2.1 and only focus on the entropy term. For each $\{Y_n\} \in G$, see Eq. (7.2.25), it holds $S_n(Y_n) \subseteq \tilde{S}_n$, thus we can replace the lower bound on the conditional entropy from Corollary 7.2.1 with the weaker one,

$$\min_{\rho \in S_{\vec{\nu}}} H(S|EI)_\rho \geq \sum_{n=0}^{\infty} \Pr(M=n) \min_{\rho^{(n)} \in \tilde{S}_n} H(S|E'I;M=n)_{\rho^{(n)}}, \tag{7.2.35}$$

which already concludes the proof. □

Instead of Eq. (7.2.30) one could also use improved decoy methods [1] to achieve similar improvements in the finite-size regime, such as only requiring two total intensities. The optimization problem is

$$
\begin{aligned}
Y_{n,L}^{a,b} := \min_{\mathbf{Y}_m, J_m} \; & Y_n^{a,b} \\
\text{s.t.} \; & \frac{\bar{F}_k - t_k - \nu_k^L}{\Pr(a,\mu,\text{test})} \leq \sum_{m \leq N_{\text{ph}}} \Pr(m|a,\mu,\text{test})\, Y_m^{a,b} \\
& \qquad\qquad + (1 - p_{\text{tot}}(\mu|a))\,, \\
& \frac{\bar{F}_k + t_k + \nu_k^U}{\Pr(a,\mu,\text{test})} \geq \sum_{m \leq N_{\text{ph}}} \Pr(m|a,\mu,\text{test})\, Y_m^{a,b}, \\
& Y_m^{a,b} = \text{Tr}\left[J_m \left(M_b^B \otimes \left(\rho_a^{(m)}\right)^T \right) \right], \\
& 0 \leq Y_m^{a,b} \leq 1, \\
& J_m \succeq 0, \; \text{Tr}_B\left[J_m\right] = \mathbb{1}_{A'}, \\
& \forall \mu \in \{\mu_1, \mu_2, \dots\}, \; \forall a, b, \; \forall m \in N_0, \\
& \forall k = (a, b, \mu, \text{test}) \in \Sigma_{\text{test}},
\end{aligned}
\tag{7.2.36}
$$

which also allows the intensities to be different for each signal state. This part concludes the fixed-length security proofs and now we turn our attention to examples and optimizing finite-size key rates.

7.3 Optimizing Security Parameters

So far, in all security statements, Theorem 7.1.5, Corollary 7.2.1 and Corollary 7.2.2, one can arbitrarily choose the distribution of the individual ε-parameters to combine to the same final security parameter ε_{sec}. However, some contributions to ε_{sec} result in a higher key rate reduction than others. Therefore, in this section, we aim to heuristically optimize the total key length in terms of the security parameters.

The security parameter ε_{sec} in Theorem 7.1.5, Corollary 7.2.1 and Corollary 7.2.2 is given by

$$
\varepsilon_{\text{sec}} = \max\{\varepsilon_{\text{AT}}, \varepsilon_{\text{PA}} + 2\bar{\varepsilon}\}. \tag{7.3.1}
$$

Heuristically, ε_{AT} has the greatest influence, as a smaller ε_{AT} results in a larger variational bound ν and thus more states must be considered in minimization. To distribute

$\varepsilon_{\mathrm{AT}}$ ideally under our assumptions, one should choose $\varepsilon_{\mathrm{AT}} = \varepsilon_{\mathrm{PA}} + 2\bar{\varepsilon}$. Furthermore, if we made $\varepsilon_{\mathrm{AT}}$ bigger than $\varepsilon_{\mathrm{PA}} + 2\bar{\varepsilon}$, then one could increase the latter two parameters for free without any penalty in the security statement. Hence, $\varepsilon_{\mathrm{AT}} = \varepsilon_{\mathrm{PA}} + 2\bar{\varepsilon}$ is the optimal choice.

Therefore, as a starting point, let us assume that the maximum in the security parameter is achieved by

$$\max\{\varepsilon_{\mathrm{AT}}, \varepsilon_{\mathrm{PA}} + 2\bar{\varepsilon}\} = \varepsilon_{\mathrm{PA}} + 2\bar{\varepsilon}. \tag{7.3.2}$$

Thus, the total security parameter needs to satisfy

$$\varepsilon_{\mathrm{sec}} = \varepsilon_{\mathrm{EV}} + \varepsilon_{\mathrm{PA}} + 2\bar{\varepsilon}. \tag{7.3.3}$$

Furthermore, the reduction of the key rate from the error verification parameter $\varepsilon_{\mathrm{EV}}$ is smaller than from privacy amplification. Hence, we make another choice, by assuming that their correction terms should be equal, i.e.

$$\log(2/\varepsilon_{\mathrm{EV}}) = 2\log(1/2\varepsilon_{\mathrm{PA}}). \tag{7.3.4}$$

This choice results in

$$\varepsilon_{\mathrm{EV}} = 8\varepsilon_{\mathrm{PA}}^2. \tag{7.3.5}$$

Then, for a fixed $\varepsilon_{\mathrm{sec}}$, the smoothing parameter is given by

$$2\bar{\varepsilon} = \varepsilon_{\mathrm{sec}} - \varepsilon_{\mathrm{PA}}\left(8\varepsilon_{\mathrm{PA}} + 1\right). \tag{7.3.6}$$

Under these assumptions, the optimal choice for all ε-parameters is found by optimizing the key length l in terms of $\varepsilon_{\mathrm{PA}}$. Fix n_{sift}, n, the security parameter $\varepsilon_{\mathrm{sec}}$ and define the remaining ε parameters by Eq. (7.3.5) and Eq. (7.3.6). Thus, inserting the assumptions from above into the key length l of Eq. (7.1.49) and differentiating it with respect to $\varepsilon_{\mathrm{PA}}$, yields

$$\frac{\mathrm{d}l}{\mathrm{d}\varepsilon_{\mathrm{PA}}} = \frac{1}{\ln(2)\varepsilon_{\mathrm{PA}}} - \sqrt{\lfloor n_{\mathrm{sift}} - nt_{\mathrm{sift}} \rfloor} \cdot \frac{\log(1+\dim(S))\left(8\varepsilon_{\mathrm{PA}}+1\right)}{\ln(2)\bar{\varepsilon}\sqrt{\log\left(\frac{2}{\bar{\varepsilon}^2}\right)}}. \tag{7.3.7}$$

The optimal value $\varepsilon_{\mathrm{PA}}^*$ is the solution to $\frac{\mathrm{d}l}{\mathrm{d}\varepsilon_{\mathrm{PA}}} = 0$, which can be done numerically. Since both $\bar{\varepsilon}$ and $\varepsilon_{\mathrm{PA}}$ need to be positive and smaller than $\varepsilon_{\mathrm{sec}}$, the optimal value $\varepsilon_{\mathrm{PA}}^*$ needs to be in $\left[0, -\frac{1}{16} + \frac{1}{\sqrt{8}}\sqrt{\frac{1}{32} + \varepsilon_{\mathrm{sec}}}\right]$.

In summary, the optimal $\varepsilon_{\mathrm{AT}}$ is given by $\varepsilon_{\mathrm{AT}} = \varepsilon_{\mathrm{PA}}^* + 2\bar{\varepsilon}$. Thus, it makes up the majority of the available $\varepsilon_{\mathrm{sec}}$, since $\varepsilon_{\mathrm{EV}} = 8(\varepsilon_{\mathrm{PA}}^*)^2 \ll \varepsilon_{\mathrm{PA}}^*$.

7.4 Optimizing Expected Key Rates for Fixed Length Protocols

The acceptance set S_{acc} implicitly depends on the expected frequencies $\bar{F}$ and the tolerated fluctuations characterized in terms of $\vec{t}$ and t_{sift}. So far, the tolerance parameters $\vec{t}$ and t_{sift} in the acceptance set S_{acc} are free and can be chosen arbitrarily. In terms of pure key rates, they should be chosen to be zero, however, then the protocol would almost always abort since observing a single point has vanishing probability.

Hence, one would want to choose them such that under the honest behaviour we accept the observed frequencies with high probability while maintaining high secret key rates. However, both goals cannot be achieved simultaneously. A wider acceptance set will always decrease the key rate whereas a smaller acceptance set ultimately decreases the accept probability. Therefore, we introduce the expected key rate as a figure of merit because it captures the trade-off between both effects. In [55, 73] expected key rates have been calculated by simulating the channel numerically. Here we will instead proceed by finding tight bounds on the accept probability.

At first, let us define the expected key rate as

$$\bar{R}_{S_{\text{acc}}(\vec{t})} = \Pr\left[F^{\text{obs}} \in S_{\text{acc}}(\vec{t})\right] \cdot R_{\text{acc}} + \Pr\left[F^{\text{obs}} \notin S_{\text{acc}}(\vec{t})\right] \cdot R_{\text{abort}}, \tag{7.4.1}$$

where R_{acc} is the usual finite-size key rate after passing the acceptance test. In the case of aborting $R_{\text{abort}} = 0$ because then no key is generated.

In order to estimate the accept probability $\Pr[\Omega_{\text{acc}}]$ for an acceptance set S_{acc} characterized by $\vec{t}$ and $\bar{F}$, we assume a given channel model, which correctly describes the observations. This assumption means that the parameter $\bar{F}$ in the security proof is equal to the honest behaviour based on that channel model.

Additionally, we make certain assumptions about the accept probability, because it only influences the expected key rate, not the secret key rate upon acceptance. Specifically, Theorem 7.1.5 proves security for any width of the acceptance set, although the resulting key length might be zero.

Hence, we assume that the accept probability is mostly governed by $\Pr\left[F^{\text{obs}} \in S_{\text{acc}}\right]$, which is a reasonable assumption since under honest behaviour, a well-designed error correction procedure will succeed with high probability. Thus, let us assume

$$\Pr[\Omega_{\text{acc}}] \approx \Pr[\Omega_{\text{AT}}] = \Pr\left[F^{\text{obs}} \in S_{\text{acc}}\right] = \Pr\left[\left|F_k^{\text{obs}} - \bar{F}_k\right| \leq t,\ \forall k \in \Sigma\right]. \tag{7.4.2}$$

In order to bound this last equation, let us first define the events

$$A_k := \{F^{\text{obs}} \in \mathcal{P}(\Sigma) \text{ s.t. } \left|F_k^{\text{obs}} - \bar{F}_k\right| \leq t\} \tag{7.4.3}$$

for all $k \in \Sigma_{\text{test}} \cup \{\text{sift}\}$. Then, we can write

$$\Pr\left[F^{\text{obs}} \in S_{\text{acc}}\right] = \Pr\left[\bigcap_{k=1}^{|\Sigma_{\text{test}}|+1} A_k\right], \tag{7.4.4}$$

which allows us to apply the Frechet inequalities [74]. They state for events $A_1, \ldots, A_n$ (which can be dependent on each other), the probability of all events occurring simultaneously is bounded by

$$\Pr\left[\bigcap_{k=1}^{n} A_k\right] \leq \min_{1\leq k\leq n} \Pr[A_k], \tag{7.4.5}$$

$$\Pr\left[\bigcap_{k=1}^{n} A_k\right] \geq \max\left\{0, \sum_{k=1}^{n} \Pr[A_k] - (n-1)\right\}. \tag{7.4.6}$$

Thus, we can find upper and lower bounds on the accept probability given a channel model. Here it is important to note that we could not in general calculate the accept probability directly for a given channel model because the events A_k can depend arbitrarily on each other due to the underlying quantum state ρ_{AB}, justifying the estimation via the Frechet inequalities.

Hence, we can evaluate the performance of our acceptance set choice in terms of both the maximal and minimal achievable expected key rate. Furthermore, we can estimate the optimal choice for allowed fluctuations $\vec{t}$ and t_{sift} for a given channel model.

In Fig. 7.2 one can see the expected key rate of the qubit BB84 protocol plotted against the tolerated fluctuations t in the acceptance set for a loss-only channel model with 10 dB loss and a total number of signals of $n = 10^9$. For each data point, we optimized the testing probability. In order to present the allowed fluctuations on a one-dimensional line, we chose $t_k \equiv t$ for all $k \in \Sigma_{\text{test}} \cup \{\text{sift}\}$. The upper (blue circles) and lower (red circles) bounds use the Frechet inequalities of Eq. (7.4.5) and the independent bound assumes all A_k to be independent, i.e.

$$\Pr\left[\bigcap_{k=1}^{n} A_k\right] \approx \prod_{k=1}^{|\Sigma_{\text{test}}|+1} \Pr(A_k). \tag{7.4.7}$$

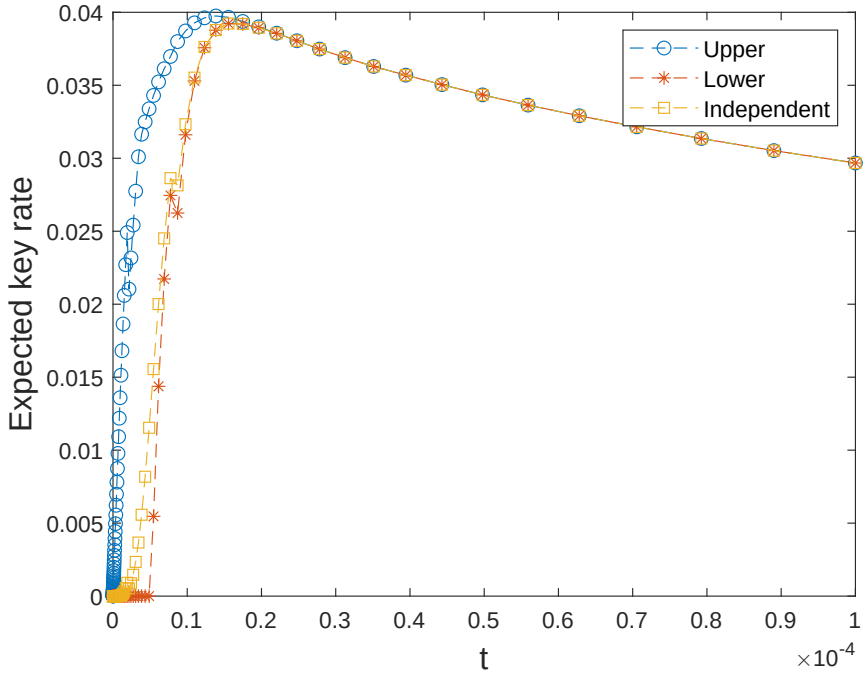


Figure 7.2: Plot of upper, lower and independent bounds on the expected key rate of the *fixed-length* qubit BB84 protocol for *IID attacks* with a channel loss of 10 dB loss and a total number of signals of $n = 10^9$ over the tolerated fluctuation parameter with $t_k = t_{\text{sift}} \equiv t$. (Figure 2 from Ref. [3])

One can clearly see that there is an optimal choice of t. The first increase in expected key rate originates from the accept probability increasing as we increase t. Later, while increasing t further we allow more fluctuations and thus more states in $S_{\vec{\nu}}$. This reduces the achievable key rate and therefore also reduces the expected key rate.

7.5 Examples for Collective Attacks

In this section, we only consider IID collective attacks and present key rates for fixed-length protocols with unique acceptance, i.e. $\vec{t} = 0$. In general, the leakage depends on the actual error correction procedure. However, for all examples presented here we use the upper bound

$$\lambda_{\text{EC}} = n\bar{F}_{\text{sift}} f_{EC} H(X|Y;\text{sift}), \tag{7.5.1}$$

which corresponds to the leaked number of bits during error-correction step for the expected observations. This expression for the leakage is not to be confused with a result from an optimization over the feasible set, it is just an upper bound on the error correction cost.

In each example we will optimize some free parameters such as the testing probability or the decoy intensities. Since this is different for each specific example we will discuss this in the corresponding subsections in detail.

For the plots presented in this work, we symmetrize the deviation bounds by defining a single parameter $\nu_k \coloneqq \max\{\nu_k^U, \nu_k^L\}$ and setting $\nu_k^U = \nu_k^L = \nu_k$. This simplification is justified because it strictly enlarges the feasible set, thereby maintaining the security guarantees provided by Theorem 7.1.1. The impact on the secret key rate is negligible for the finite block sizes considered here, as ν_k^U and ν_k^L are already nearly identical. Furthermore, this approximation becomes exact in the asymptotic limit, where the Binomial distribution converges to a Gaussian and the tail bounds become symmetric.

7.5.1 Qubit BB84 Protocol with Loss

As a simple example we present the BB84 protocol [35] with a perfect qubit source including loss in the channel. In this protocol, we assume that both the X and Z basis are used for key generation and with some probability each round is revealed as a test round. Thus, one acquires data from both bases to constraint the feasible set. We made the following parameter choices, Z-basis probability $p_z = 1/2$ for both Alice and Bob, misalignment parameter $\theta_{\text{misalign}} = 0.01$, error correction parameter $f_{\text{EC}} = 1.16$ and a security parameter $\varepsilon_{\text{sec}} = 10^{-8}$. We optimize the testing probability and all contributions to the final security parameter ε_{sec} for each data point.

The secret key rates are evaluated employing the numerical framework established by Ref. [67]. The associated Kraus operators for this protocol can be found in Appendix C.2.1. Figure 7.3 illustrates the resulting key rates as a function of total channel loss η (dB) for varying block sizes, $n = 10^4, \ldots, 10^9$. The convergence toward the asymptotic limit (indicated by the solid black line) is evident as the number of signals increases.

Additionally, in Fig. 7.4 we compare our results for the qubit BB84 protocol with the ones from Ref. [43]. The main difference in the techniques are the different constraint types. We use entrywise constraints, while in Ref. [43] 1-norm constraints were used.

Our entrywise constraints give a clear improvement. This is due to the much tighter bounds derived in Theorem 7.1.1 and the improvements on applying the AEP due to Lemma 7.1.3.

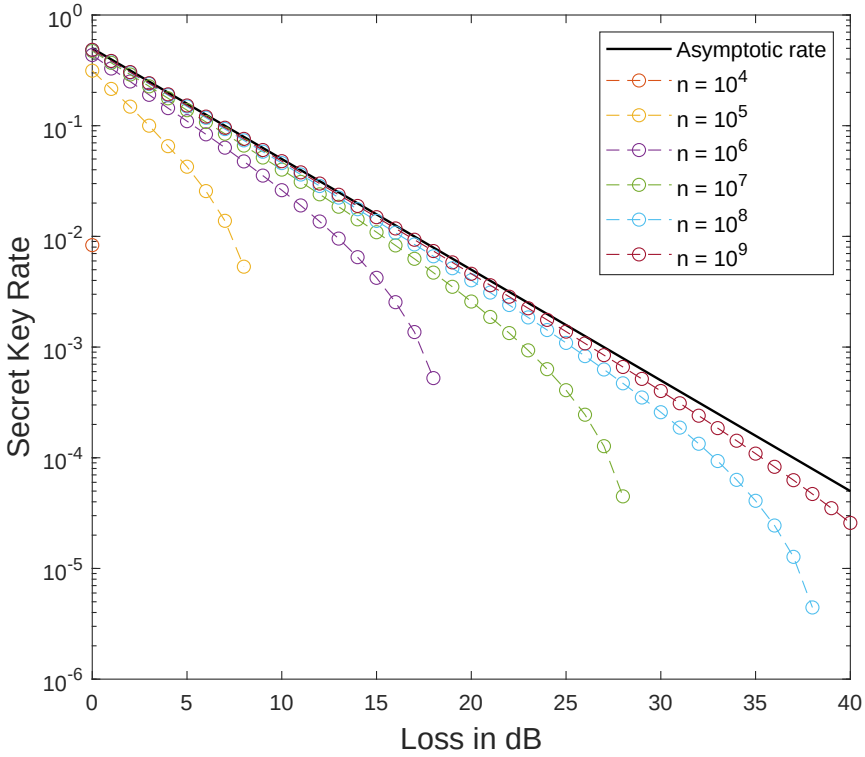


Figure 7.3: Secret key rate for *IID attacks* of the *fixed-length* qubit BB84 protocol against channel loss for varying number of total signals sent $n = 10^4, \ldots, 10^9$. (Figure 3 from Ref. [3])

With the results presented in this work, we essentially achieve asymptotic rates with $n = 10^{10}$ signals whereas 1-norm constraints only achieve positive key rates until approximately 25 dB in Fig. 7.4.

As an aside, in Ref. [43] the authors derived an improvement on the variational bound ν. This improvement uses multiple coarse grainings to effectively reduce the parameter ν. By incorporating all possible coarse grainings, they would effectively recover a vector $\vec{\nu}$ just as we presented. However, their bounds originating from a 1-norm, are inherently worse than the ones in Theorem 7.1.1. Therefore, even after including multiple coarse grainings, their results will remain worse than ours and our methods remain much simpler and easier to implement.

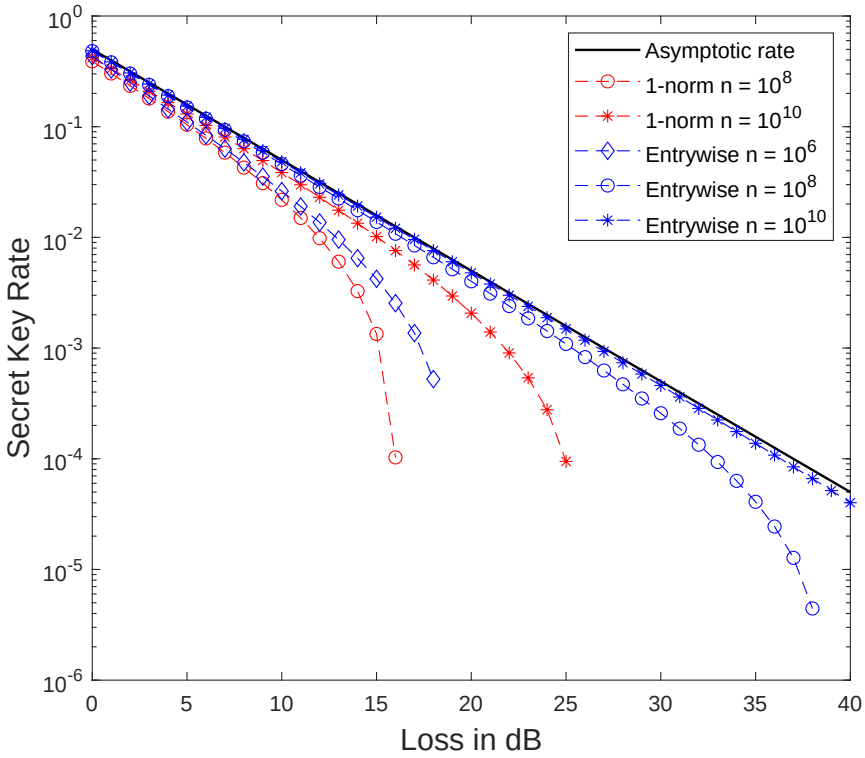


Figure 7.4: Comparison of the *fixed-length* qubit BB84 secret key rates for *IID attacks* between 1-norm constraints of [43] and our entrywise constraints of Theorem 7.1.1. The key rate is plotted against channel loss with a varying number of total signals sent n. (Figure 4 from Ref. [3])

7.5.2 Decoy BB84 Protocol

Next, we present a decoy-state version of the BB84 protocol. We assume a fully phase-randomized WCP source and a symmetric passive detection setup for Bob, i.e. Bob chooses the Z- and X-basis with probability $p_z = 1/2$. This allows us to use the squashing map from [62, 63]. We again use both bases for key generation, such that this protocol is equivalent to the qubit version apart from the additional overhead due to the decoy states.

In each round Alice decides with probability $p(\text{gen})$ if it is a test or generation round and always selects the signal intensity if it was a key generation round. In a test round Alice selects all intensities with equal probability, i.e. $p(\mu|\text{test}) = 1/3$. We use the same parameters as in Section 7.5.1 with the addition of two decoy intensities, which we set to be $\mu_2 = 0.02, \mu_3 = 0.001$, and a signal intensity μ_s as a free parameter. Now, in addition

to the testing fraction, we optimize the signal intensity μ_s for each data point as well.

Since Eve could always perform a QND measurement to measure the photon number and since vacuum signals only contribute on the order of dark counts, we restrict the sum in Corollary 7.2.2 to single photons sent by Alice. Then, we can again use the methods from [67] to evaluate the key rate. The required Kraus operators can be found in Appendix C.2.2.

In Fig. 7.5, we show the resulting key rates as a function of the total channel loss in dB for the total number of signals sent, ranging from $n = 10^6$ to $n = 10^{11}$. With $n = 10^{11}$, we approximately recover the asymptotic limit up to 40 dB. Additionally, in Fig. 7.6, we display the corresponding optimal signal intensities. For $n = 10^{11}$ signals sent, the signal intensity μ_s remains mostly constant at approximately 0.97. As the number of signals sent decreases, this constant region also shrinks. Eventually, across all values of n, the signal intensity sharply drops to zero as the key rate also falls to zero. Overall, all optimal intensities lie in the range of 0.6 to 1. These observations are in line with the expected scaling of the signal intensity in the asymptotic regime as shown in Ref. [11].

7.6 Variational Bounds for Small Security Parameters

Evaluating variational bounds for extremely small security parameters is a recurring requirement in finite-key security proofs. For example, when applying the postselection technique to reduce the analysis of coherent attacks to the IID case [26, 27], the resulting penalty forces the underlying IID security parameters to be vanishingly small.

However, this runs into numerical issues, as security parameters as small as $\varepsilon \approx 10^{-3000}$ are often needed. As we have seen before, for fixed-length protocols, the evaluation of the variational bounds $\vec{\nu}^{U/L}$ depends on the cumulative distribution function of the binomial distribution, which is given by the incomplete beta function. Standard numerical methods to calculate this function fail for such small parameter regimes due to machine precision.

To overcome this limitation, we determine the variational bounds $\vec{\nu}^{U/L}$ by bounding the binomial distribution with the normal distribution, following the approach in [75]. This approach requires calculating the inverse of the standard normal cumulative distribution function, Φ. Fortunately, the `MATLAB` function `erfcinv`, which implements the inverse of Φ, supports symbolic expressions, allowing us to achieve the required precision.

Lemma 7.6.1 (Gaussian Bounds on Variational bounds)**.** *Let the acceptance set S_{acc} and the feasible set $S_{\vec{\nu}}$ be defined as in Theorem 7.1.1. Furthermore, choose the variational*

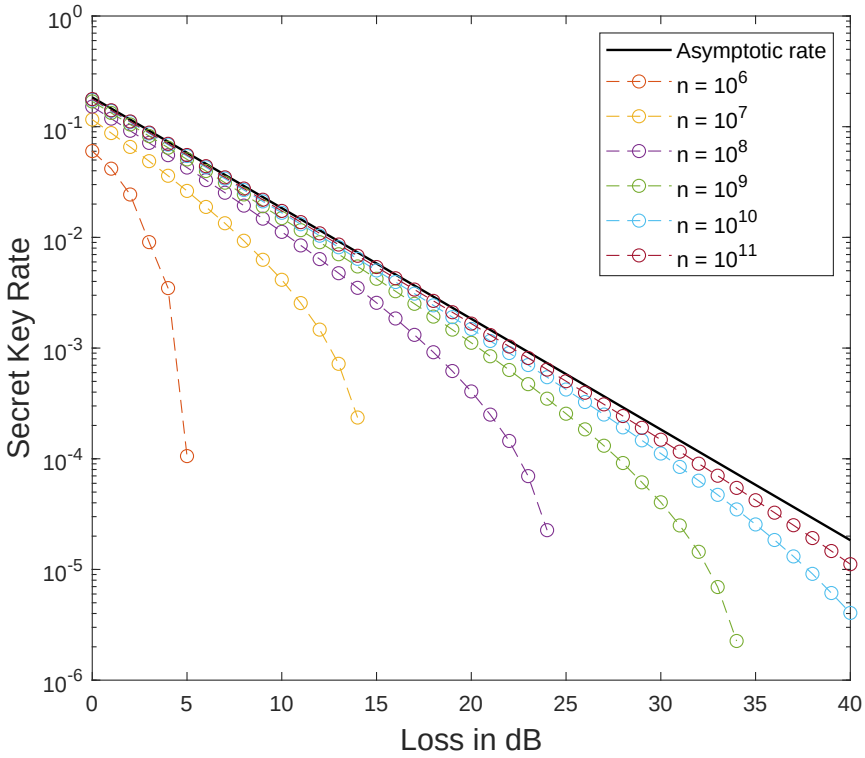


Figure 7.5: Secret key rate for *IID attacks* of the *fixed-length* decoy-state BB84 protocol using two decoy intensities against channel loss with a varying number of total signals sent $n = 10^6, \ldots, 10^{11}$. (Figure 5 from Ref. [3])

bounds $\nu_k^{L/U}$ *as the solutions*

$$\nu_k^L = \arg\min_{\nu'} \{\nu' \in [0,1] \, | 1 - C_{n,F_k^L - \nu'}(\lceil nF_k^L \rceil - 1) = \varepsilon_{AT}\}, \tag{7.6.1}$$

$$\nu_k^U = \arg\min_{\nu'} \{\nu' \in [0,1] \, | C_{n,F_k^U + \nu'}(\lfloor nF_k^U \rfloor + 1) = \varepsilon_{AT}\}, \tag{7.6.2}$$

where $F_k^{L/U} = \bar{F}_k \mp t_k$ *as before. Here, the functions* $C_{n,p}$ *are defined as in [75] by*

$$C_{n,p}(k) = \Phi\left(\text{sign}\left(\frac{k}{n} - p\right)\sqrt{2nH\left(\frac{k}{n}, p\right)}\right), \tag{7.6.3}$$

where $\Phi(x)$ *is the cumulative-distribution function of the standard normal distribution and*

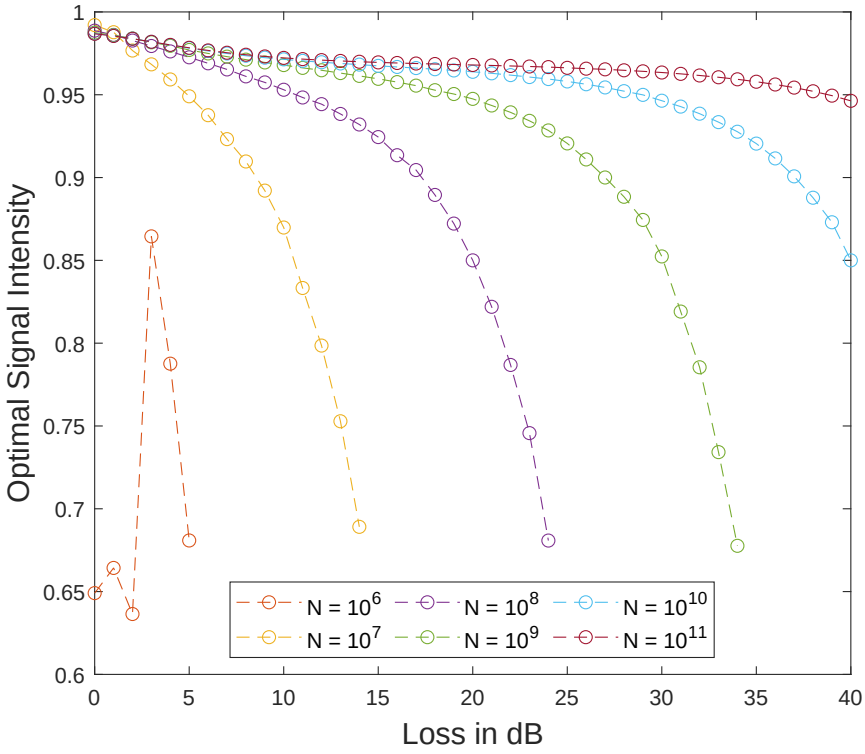


Figure 7.6: Optimal signal intensities for *IID attacks* of the *fixed-length* decoy-state BB84 protocol using two decoy intensities against channel loss with a varying number of total signals sent $n = 10^6, \ldots, 10^{11}$.

$H(x,p)$ is given by

$$H(x,p) = x \log\left(\frac{x}{p}\right) + (1-x)\log\left(\frac{1-x}{1-p}\right). \tag{7.6.4}$$

Then, it holds for all $\rho_{\vec{A}\vec{B}\vec{E}} = \bigotimes_{i=1}^{n} \sigma_{ABE}$ and $\sigma \notin S_{\vec{\nu}}$ that,

$$\Pr[\Omega_{acc}] \leq \varepsilon_{AT}. \tag{7.6.5}$$

Proof. Following the same steps as in Theorem 7.1.1 and again defining $p_k := \mathrm{Tr}\left[\Gamma_k \sigma\right]$ one can show that for $p_k > \bar{F}_k + t_k + \nu$

$$\Pr\left[\left|F_k^{\mathrm{obs}} - \bar{F}_k\right| \leq t_k\right]\Big|_{p_k > F_k^U + \nu} \leq \Pr\left[X_k \leq \lfloor n F_k^U \rfloor\right]\Big|_{p_k > F_k^U + \nu}, \tag{7.6.6}$$

where we again note that X_k is binomial distributed. Now, one can apply the bounds from [75] and note that it is monotonically decreasing in p to find

$$\Pr\left[\left|F_k^{\text{obs}} - \bar{F}_k\right| \leq t_k\right]\Big|_{p_k > F_k^U + \nu} \leq C_{n, F_k^U + \nu'}(\lfloor nF_k^U \rfloor + 1). \tag{7.6.7}$$

Similarly, one finds

$$\Pr\left[\left|F_k^{\text{obs}} - \bar{F}_k\right| \leq t_k\right]\Big|_{p_k < F_k^L - \nu} \leq 1 - C_{n, F_k^L - \nu'}(\lceil nF_k^L \rceil - 1). \tag{7.6.8}$$

Thus, if $\nu_k^{U/L}$ are chosen as in the theorem statement we find

$$\Pr\left[\left|F_k^{\text{obs}} - \bar{F}_k\right| \leq t_k\right] \tag{7.6.9}$$

$$\leq \max\left\{\Pr\left[\left|F_k^{\text{obs}} - \bar{F}_k\right| \leq t_k\right]\Big|_{p_k > F_k^U + \nu_k^U}, \Pr\left[\left|F_k^{\text{obs}} - \bar{F}_k\right| \leq t_k\right]\Big|_{p_k < F_k^L + \nu_k^L}\right\} \tag{7.6.10}$$

$$\leq \varepsilon_{\text{AT}}, \tag{7.6.11}$$

for all $\sigma \notin S_\nu$. □

7.7 Conclusion

In this chapter, we established a finite-size security proof against IID collective attacks for generic prepare-and-measure (P&M) protocols with probabilistic testing. Our approach advances previous methods by incorporating Lemma 7.1.3 and Theorem 7.1.1, resulting in significantly improved finite-size scaling. These gains stem from two key factors: scaling the finite-size correction terms with the number of detected signals rather than the total transmitted signals, and tightening the feasible set via sharper concentration inequalities.

Additionally, we applied the IID security proofs from Theorem 7.1.5 and Corollary 7.2.2 to qubit-based and decoy-state protocols, respectively. Notably, unlike methods based on the entropic uncertainty relation (EUR), such as those in [22, 24, 37], our approach is versatile and applicable to a broader range of setups, particularly those with passive detection setups. As a concrete example, we presented a decoy-state BB84 protocol with a symmetric passive receiver. However, we note that recently security proofs for a BB84 protocol with a passive receiver using the phase-error correction technique or the EUR have been derived in Refs. [69, 76]. Furthermore, for collective attacks, we estimated the expected key rate, enabling a balance between achieving high acceptance probabilities and optimizing key rates.

Despite these improvements, the analysis in this chapter remains restricted to collective attacks. Although the published version of this work in Ref. [3] extends these methods to coherent attacks and the variable-length regime via the postselection technique, the resulting key rates are highly conservative.

As demonstrated in Ref. [27], reducing a coherent attack to an IID attack via the postselection technique introduces two dimension-dependent correction terms. The first term is subtracted directly from the key length, causing an immediate reduction. The second term indirectly impacts the key rate by scaling up the security parameters. Consequently, to maintain the target security, the parameters within the IID proof must be chosen to be significantly smaller. This inherently reduces the key rate, as the feasible set $S_{\vec{\nu}}$ expands when security parameters decrease.

To quantify the severity of these correction terms, Chapter 10 compares the postselection technique against both the entropic uncertainty relation (EUR) and the full Rényi approach for a qubit BB84 (Fig. 10.3) and an active decoy-state BB84 (Fig. 10.7) protocol.

Therefore, rather than relying on the postselection technique to establish security against coherent attacks, this thesis adopts the full Rényi approach presented in Chapter 10. This method yields significantly higher finite-key performance and is robust against experimental imperfections.

Chapter 8

Entropy Accumulation Theorem-based QKD Proofs

In this chapter, we derive security proofs for both fixed- and variable-length QKD protocols using the marginal constrained entropy accumulation theorem (MEAT) [41] and the generalized entropy accumulation theorem (GEAT) [30]. We construct a formal model of a QKD protocol that satisfies the requirements of MEAT, allowing us to prove finite-size security against coherent attacks. A schematic representation of the MEAT model we will describe in this section is shown in Fig. 8.1 below.

Additionally, we show, how key rates based on the GEAT model are also valid lower bound on the key rate under the MEAT model. However, these lower bounds based on the GEAT will be strict lower bounds on what the MEAT could achieve.

The content of this chapter is mostly based on Section IV of our work of Ref. [4].

8.1 Background on EAT-based Proofs

As already mentioned in the introduction, one typically uses the LHL for smooth min-entropies, see Theorem 3.5.2, to prove secrecy for QKD protocols. Then, the task of finding a secure lower bound on the secret key rate reduces to finding a lower bound on the n-round smooth min-entropy.

The GEAT [30], then showed that the n-round entropy of a state ρ^n generated by a process of the form

$$\rho^n = \mathcal{M}_n \circ \cdots \circ \mathcal{M}_1(\rho^0) \tag{8.1.1}$$

can be informally bounded as

$$H_{\min}^{\varepsilon}(S_1^n|I_1^nE)_{\rho^n} \geq nh_{\text{vN}} - \mathcal{O}(\sqrt{n}), \tag{8.1.2}$$

where as in our protocol description S_i, I_i are the i-th secret and announcement register and E, is Eve's side information, respectively. The constant h_{vN} is a lower bound on the single-round von-Neumann entropy given the accept condition of the QKD protocol. Under this setting, each channel $\mathcal{M}_i$ has to obey a non-signalling condition, for more details see [30, Theorem II.1].

The GEAT already generalized the entropy accumulation theorem(EAT) [29], however for a QKD specific application, it still had a few restrictions. For example, the original EAT [28] could not be applied to P&M protocols. The GEAT was able to incorporate P&M protocols but only with the restriction of Eve interacting with *only one* single signal at a time.

Later on, the generalized Rényi entropy accumulation theorem (GREAT) [77] showed a lower bound on the n-round Rényi entropy under the same model as the GEAT, i.e.

$$\widetilde{H}_{\alpha}^{\uparrow}(S_1^n|I_1^nE)_{\rho^n} \geq nh_{\alpha} - \mathcal{O}(\sqrt{n}), \tag{8.1.3}$$

where h_α is a lower bound on the single-round Rényi entropy given the accept condition of the QKD protocol. This analysis significantly improved on the GEAT in terms of performance.

Then, recently in Refs. [41, 78] (originally developed in [79]) the marginal constrained entropy accumulation theorem (MEAT) proved similar lower bounds for processes creating the final state ρ^n as

$$\rho^n = \mathcal{M}_n \circ \cdots \circ \mathcal{M}_1(\rho^0) \quad \text{or} \quad \rho^n = \mathcal{M}_n \otimes \cdots \otimes \mathcal{M}_1(\rho^0). \tag{8.1.4}$$

A schematic representation of the MEAT channel structure with tensor products is shown in Fig. 8.1. Most importantly, the MEAT can directly incorporate a marginal constraint and thus, is readily applicable to P&M protocols via the source replacement scheme, see Section 3.3. Notably, the channel structure of the MEAT with concatenated channel is still different from the GEAT. However, and most importantly for our application, both channel structures the MEAT can accommodate, allow one to prove security against coherent attacks for QKD protocols.

Therefore, in Section 8.2 we will formalize the QKD protocol of Protocol 1 in the MEAT framework. Afterwards, we will show a theorem leading to a lower bound like h_α. Finally, we will also draw a connection between the MEAT and GEAT key length expressions in Section 8.4, showing that the GEAT rates are strict lower bounds under the MEAT model.

8.2 Formalizing QKD protocols in MEAT

In this section we formalize a QKD protocol such that the methods of Refs. [41, 78, 79] can be used. Intuitively, in these works the action of each round of a QKD protocol is specified by a channel $\mathcal{M}_i$ which creates the secret data and all public announcements.

Furthermore, as described above, the channel structure of Refs. [41, 78, 79] requires a set of channels $\{\mathcal{M}_i\}$ which generate the final state of the protocol by acting either as a tensor product of channels, i.e., $\bigotimes_{i=1}^{n} \mathcal{M}_i$, or as a concatenation of channels $\mathcal{M}_n \circ \cdots \circ \mathcal{M}_1$.

For simplicity, we chose a tensor product structure and note that by the same argument as in Ref. [30] the same key rates can be found under the model of concatenated channels. Again, in Fig. 8.1 we show a schematic representation of a channel acting in this tensor product form. As represented there, starting from a big n-round state of which Eve holds the purification $\tilde{E}$, the channels $\mathcal{M}_i$ create the secret data S_i and the public announcements $\widehat{C}_i I_i$ for each round $i = 1 \dots n$. For simplicity of the description, we assume each channel $\mathcal{M}$ is equivalent to the same channel $\mathcal{M}$ in each round, but in general, the channels could be different. We will make use of this degree of freedom later on in Section 10.8.

As a starting point, in this section, we define a channel $\mathcal{M}_{\text{QKD}}$ suitable for describing QKD protocols in this tensor-product manner. We will use a similar approach as in Ref. [67] and Ref. [80, App A] in the asymptotic regime.

8.2.1 Source Replacement Scheme

Under the source replacement scheme, see Section 3.3, Alice's state preparation process in each round can be viewed as having her first prepare a *pure* state,

$$|\psi\rangle_{AA'} = \sum_{k=1}^{d_A} \sqrt{p(k)}\, |k\rangle_A\, |s_k\rangle_{A'}\,, \tag{8.2.1}$$

where $p(k)$ is the total probability of Alice sending state $|s_k\rangle_{A'}$ to Bob, and then performing a measurement on A described by the following POVM elements:

$$M_k^A = |k\rangle\langle k|\,, \tag{8.2.2}$$

for all $k = 1 \dots d_A$. Furthermore, let us write $\tau_A := \text{Tr}_{A'}[|\psi\rangle\langle\psi|]$ to denote the marginal of the single-round state Alice prepares in the source-replaced picture. If Alice sends mixed states, for example in the case of weak coherent pulse (WCP) sources, one first has to purify the states with a shield system, as described in Section 3.4.

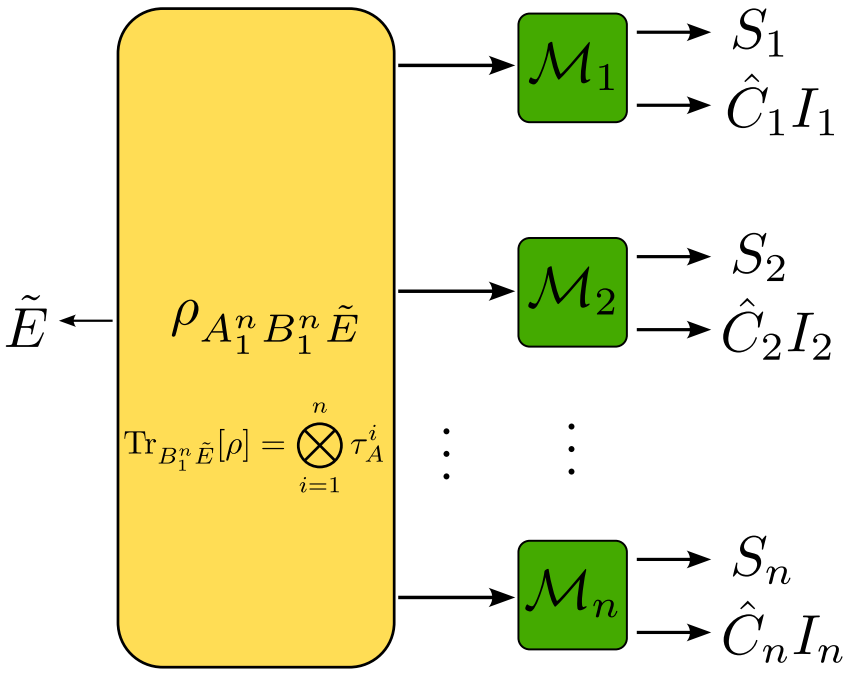


Figure 8.1: Representation of the tensor product structure of the MEAT theorem [41, Corollary 4.1]. (Figure 1 from Ref. [4])

Again following the justification in Section 3.3, we can equivalently let Eve prepare the state $\rho_{A_1^n B_1^n \tilde{E}}$, where $\tilde{E}$ is a purifying system. Eve then sends all systems A_1^n to Alice and systems B_1^n to Bob. Since in the original prepare-and-measure protocol Eve has no access to the A_1^n registers, the marginal state on these registers will be unchanged under Eve's attack. This is again reflected by requiring that the state $\rho_{A_1^n B_1^n \tilde{E}}$ satisfies a marginal constraint $\rho_{A_1^n} = \tau_A^{\otimes n}$.

In this equivalent picture, Alice and Bob then measure the registers $A_1^n B_1^n$ to generate the raw data strings $X_1^n Y_1^n$, which are then processed further as described in the protocol description. Note that for protocols of the form we described, the process of generating the registers $S_1^n I_1^n \widehat{C}_1^n$ from the registers $A_1^n B_1^n$ can be written as a channel of the form $\mathcal{M}_{\mathrm{QKD}}^{\otimes n}$, for a single channel $\mathcal{M}_{\mathrm{QKD}}$ we describe in Section 8.2.3 below.

Finally, a quick note on the numerical implementation. It might be preferred to incorporate a Schmidt decomposition to remove unnecessary dimensions of the source replacement scheme, however, for the discussion presented here, we assume that each of Alice's POVM

elements is a projector as given by the generic source replacement scheme.

8.2.2 Announcements and Conditioning on Test and Generation rounds

Next, we need to take care of Alice's and Bob's announcements, such as basis announcements. Given the distinction between test and generation rounds, we split Alice's set of announcements and measurement outcomes into test and generation subsets, i.e.

$$\mathcal{C}^A = \mathcal{C}^{A,\mathtt{test}} \cup \mathcal{C}^{A,\mathtt{gen}}, \tag{8.2.3}$$

$$\mathcal{X} = \mathcal{X}^{\mathtt{test}} \cup \mathcal{X}^{\mathtt{gen}}. \tag{8.2.4}$$

This split allows us to partition Alice's measurement outcomes similarly to [80, App. A],

$$\mathcal{X}^{\mathtt{test}} = \bigcup_{\alpha \in \mathcal{C}^{A,\mathtt{test}}} \mathcal{X}_\alpha^{\mathtt{test}}, \quad \mathcal{X}^{\mathtt{gen}} = \bigcup_{\alpha \in \mathcal{C}^{A,\mathtt{gen}}} \mathcal{X}_\alpha^{\mathtt{gen}}, \tag{8.2.5}$$

and equivalently for Bob

$$\mathcal{Y} = \bigcup_{\beta \in \mathcal{C}^B} \mathcal{Y}_\beta. \tag{8.2.6}$$

As stated in the protocol description Protocol 1, we will perform different processing depending on being in a test or generation round. Therefore, let us define the partition operators for test and generation rounds respectively as

$$\Pi^{\mathtt{test}} = \sum_{\substack{\alpha \in \mathcal{C}^{A,\mathtt{test}}, \\ x \in \mathcal{X}_\alpha^{\mathtt{test}}}} M_{\alpha,x}^A, \quad \Pi^{\mathtt{gen}} = \sum_{\substack{\alpha \in \mathcal{C}^{A,\mathtt{gen}}, \\ x \in \mathcal{X}_\alpha^{\mathtt{gen}}}} M_{\alpha,x}^A. \tag{8.2.7}$$

Furthermore, in the same spirit to the announcements, we aim to partition Alice's POVM elements into "`test`" and "`gen`" subsets and it will turn out to be useful to have a notion of POVM elements conditioned on these events. Under the basic source replacement scheme, we can set these POVM elements to be equal to the original POVM elements, though the test subset will then form a POVM on the subspace $\mathrm{supp}(\Pi^{\mathtt{test}})$ rather the original space, and analogously for the generation subset.

However, we note that when the analysis is simplified using a Schmidt decomposition of $|\psi\rangle_{AA'}$, this might not be the case and again we discuss this situation in Appendix C.3.1. To clearly indicate the required conditioning, especially for the general case, we add $|\mathtt{test}$

or $|\mathtt{gen}$ where appropriate, and we will denote the resulting POVMs as $\{M_k^{A|\mathtt{test}}\}$ and $\{M_k^{A|\mathtt{gen}}\}$, respectively.

After partitioning them appropriately, Alice's sets of POVM elements can be written as

$$\{M_{\alpha,x}^{A|\mathtt{test}}\}_{\alpha\in\mathcal{C}^{A,\mathtt{test}},x\in\mathcal{X}_\alpha^{\mathtt{test}}} = \{M_k^{A|\mathtt{test}}\}_k, \tag{8.2.8}$$

$$\{M_{\alpha,x}^{A|\mathtt{gen}}\}_{\alpha\in\mathcal{C}^{A,\mathtt{gen}},x\in\mathcal{X}_\alpha^{\mathtt{gen}}} = \{M_k^{A|\mathtt{gen}}\}_k. \tag{8.2.9}$$

On Bob's side following Ref.[67] we also partition his POVM elements according to the announcements such that

$$\{M_{\beta,y}^{B}\}_{\beta\in\mathcal{C}^{B},y\in\mathcal{Y}_\beta} = \{M_k^{B}\}_{k=1\ldots d_B}. \tag{8.2.10}$$

Finally, for brevity, let us define

$$\mathcal{C}_{AB}^{\mathrm{test}} :=\{(\alpha,x,\beta,y)|\alpha\in\mathcal{C}^{A,\mathtt{test}},x\in\mathcal{X}_\alpha^{\mathtt{test}},\beta\in\mathcal{C}^{B},y\in\mathcal{Y}_\beta\}, \tag{8.2.11}$$

as the set containing all announcements during test rounds. In addition, we assume that the alphabet $\mathcal{I}$ of registers I_i contains $\mathcal{C}_{AB}^{\mathrm{test}}$.

8.2.3 Defining the QKD channel

After these preliminary definitions, we have all tools available to define the channel $\mathcal{M}_{\mathrm{QKD}}$ characterizing a QKD protocol acting as a tensor product in each round as shown in Fig. 8.1.

Before we formally state the definition, let us discuss the constituents of the definition of $\mathcal{M}_{\mathrm{QKD}}$ on an intuitive level. We define $\mathcal{M}_{\mathrm{QKD}}$ based on four isometries. We will discuss them in the order in which they will be applied. A circuit diagram representing the action of all isometries can be found in Fig. 8.2.

First, the isometry $V_{\mathtt{test}}$ applies the split into test and generation rounds because different processing takes place based on each branch. Next, the isometries $V_{\mathcal{G}}$ and V_{meas} act together as a single block applying operations conditioned on the flag in register F. Informally, $V_{\mathcal{G}}$ creates an intermediate version of the secret register, S_Q, and Alice's and Bob's announcements and private data in generation rounds. The isometry V_{meas} applies the measurements in test rounds and creates the corresponding announcements in register I.

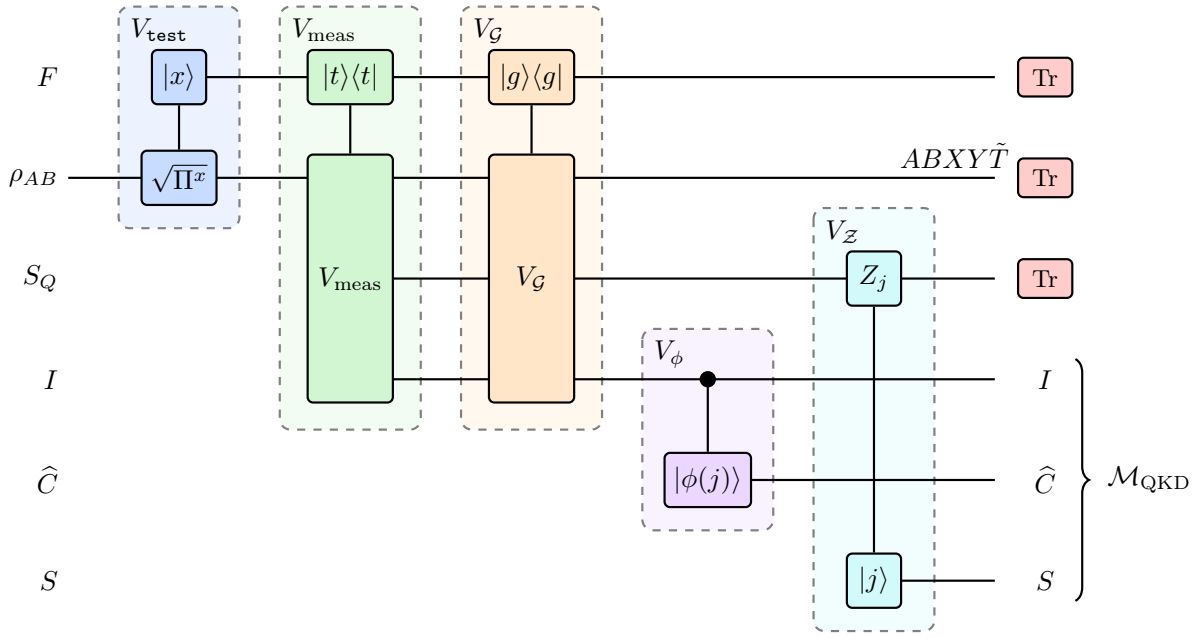


Figure 8.2: Circuit representation of the concatenated isometries from Definition 8.2.1 defining the QKD channel $\mathcal{M}_{\text{QKD}}$. The channel maps the bipartite input state ρ_{AB} to the final classical registers I, $\widehat{C}$, and S by applying the isometries $V_{\mathtt{test}}$, V_{meas}, $V_{\mathcal{G}}$, V_ϕ, and $V_{\mathcal{Z}}$ and tracing out the complementary systems $(ABXY\tilde{T}FS_Q)$. The isometry $V_{\mathtt{test}}$ applies the projectors $\sqrt{\Pi^x}$ and creates the value x in register F. Note that the isometries $V_{\text{meas}} \otimes |t\rangle\langle t|_F$ and $V_{\mathcal{G}} \otimes |g\rangle\langle g|_F$ are depicted sequentially for clarity only. They act as a single block but commute due to their mutually orthogonal support on the flag register F.

Afterwards, the isometry V_ϕ performs a post-processing that applies the deterministic function ϕ to register I and stores the outcomes in register $\widehat{C}$. Thus, it effectively creates the announcements that are used for the acceptance test or the variable-length decision.

Finally, the isometry $V_{\mathcal{Z}}$ effectively measures the intermediate secret register S_Q and creates the final classical secret register S.

Definition 8.2.1. Let $\hat{\mathcal{G}} : AB \to ABXY\hat{S}_QI$ be the CPTNI map as described in [80, App A] of a QKD protocol defined with Alice's POVM elements conditioned on a generation round, i.e. $\{M_{\alpha,x}^{A|\mathtt{gen}}\}$[1]. We extend $\hat{\mathcal{G}}$ to a CPTP map $\mathcal{G}$ through the following construction.

[1]Ref. [80] derives the $\hat{\mathcal{G}}$ for reverse reconciliation, but the replacements for direct reconciliation are straightforward and already partly described in Ref. [80].

Let $\hat{\mathcal{S}}$ be the alphabet of $\hat{S}_Q$ and $\mathcal{S} = \hat{\mathcal{S}} \cup \{\perp\}$ be the alphabet of the extended register S_Q and define $\mathcal{G}$ as

$$\mathcal{G} : AB \to ABXYS_QI, \tag{8.2.12}$$

$$\mathcal{G}(\rho) \coloneqq \hat{\mathcal{G}}(\rho) + \left(1 - \mathrm{Tr}\Big(\hat{\mathcal{G}}(\rho)\Big)\right) |\perp\rangle\langle\perp| \,, \tag{8.2.13}$$

for all $\rho \in \mathrm{S}_=(AB)$, and where $|\perp\rangle$ is orthogonal to $\hat{\mathcal{G}}(\rho)$ for any $\rho \in \mathrm{S}_=(AB)$. Additionally, let $V_{\mathcal{G}} : AB \to ABXYS_Q I\tilde{T}$ be the Stinespring dilation of $\mathcal{G}$[2]. Furthermore, we define the following isometries

$$V_{\mathtt{test}} \coloneqq \sqrt{\Pi^{\mathtt{test}}} \otimes |t\rangle_F + \sqrt{\Pi^{\mathtt{gen}}} \otimes |g\rangle_F \,, \tag{8.2.14}$$

$$V_{\mathcal{Z}} \coloneqq \sum_{j \in \mathcal{S}} |j\rangle_S \otimes Z_j, \tag{8.2.15}$$

$$V_{\mathrm{meas}} \coloneqq \sum_{(\alpha,x,\beta,y) \in \mathcal{C}_{AB}^{\mathtt{test}}} \sqrt{M_{\alpha,x}^{A|\mathtt{test}}} \otimes \sqrt{M_{\beta,y}^{B}} \otimes |\alpha, x, \beta, y\rangle_I \otimes |\perp\rangle_{S_QXY} \otimes |\mathtt{test}\rangle_{\tilde{T}} \,, \tag{8.2.16}$$

$$V_{\phi} \coloneqq \sum_{j \in \mathcal{I}} |j\rangle\langle j|_I \otimes |\phi(j)\rangle_{\widehat{C}} \,. \tag{8.2.17}$$

where the orthonormal projectors Z_j satisfy $\sum_{j\in\mathcal{S}} Z_j = \mathbb{I}_{S_Q}$ [3]. Finally, let us define the concatenation of all isometries as

$$W \coloneqq V_{\mathcal{Z}} V_{\phi} \big(V_{\mathrm{meas}} \otimes |t\rangle\langle t|_F + V_{\mathcal{G}} \otimes |g\rangle\langle g|_F \big) V_{\mathtt{test}}. \tag{8.2.18}$$

Then, we define the channel $\mathcal{M}_{\mathrm{QKD}}$ of a QKD protocol for any ρ_{AB} as

$$\mathcal{M}_{\mathrm{QKD}}[\rho_{AB}] \coloneqq \mathrm{Tr}_{\mathrm{all}\setminus SI\widehat{C}} \left[W \rho_{AB} W^{\dagger} \right], \tag{8.2.19}$$

where with $\mathrm{Tr}_{\mathrm{all}\setminus SI\widehat{C}}$ we indicate tracing out all systems apart from $SI\widehat{C}$. Finally, we also define the map creating the statistics in test rounds as

$$\mathbf{\Phi}[\sigma] \coloneqq \sum_{\hat{c} \in \widehat{\mathcal{C}}_{\setminus \mathrm{gen}}} \sum_{\substack{(\alpha,x,\beta,y) \\ \in \phi^{-1}(\hat{c})}} \mathrm{Tr}\left[\left(M_{\alpha,x}^{A|\mathtt{test}} \otimes M_{\beta,y}^{B} \right) \sigma \right] \hat{e}_{\hat{c}}, \tag{8.2.20}$$

where $\hat{e}_{\hat{c}}$ is a unit vector with a one at the $\hat{c}$-th position and ϕ^{-1} stand for the preimage.

[2] Due to the properties of $\hat{\mathcal{G}}$, its Kraus operators are indexed by the announcements contained in register I, which in generation rounds is equal to T. Therefore, the environment in a Stinespring dilation of $\mathcal{G}$ only contains a copy of T which is exactly Eve's copy $\tilde{T}$.

[3] We assume $\perp$ to be appended to any register if required.

Remark 8.2.2. The partition operators $\Pi^{\mathtt{test}}$ and $\Pi^{\mathtt{gen}}$ can be chosen as orthogonal projectors. However, when a Schmidt decomposition is used to reduce the dimensions in the source-replacement scheme, they will not necessarily be projectors anymore.

8.2.4 Fixed-Length Security

Now, we will use this construction of the channel $\mathcal{M}_{\mathrm{QKD}}$ and apply the results of [41] to find a lower bound on the Rényi entropy of a QKD protocol.

Theorem 8.2.3 (MEAT Lower bound on Conditional Entropy)**.** *Let $\mathcal{M}_{QKD} : AB \to SI\widehat{C}$ be the channel of a QKD protocol as in Definition 8.2.1. Then, for $\omega_{S_1^n I_1^n \widehat{C}_1^n \widetilde{E}} = \mathcal{M}_{QKD}^{\otimes n}[\rho_{A_1^n B_1^n \widetilde{E}}]$ with $\mathrm{Tr}_{B_1^n \widetilde{E}}[\rho_{A_1^n B_1^n \widetilde{E}}] = \tau_A^{\otimes n}$ and $\rho_{A_1^n B_1^n \widetilde{E}}$ a purification of $\rho_{A_1^n B_1^n}$, it holds for any $\alpha \in (1, \infty]$*

$$\widetilde{H}_\alpha^\uparrow(S_1^n|I_1^n\widetilde{E})_{\omega_{|\Omega_{\mathrm{AT}}}} = \widetilde{H}_\alpha^\uparrow(S_1^n|I_1^n\widehat{C}_1^n\widetilde{E})_{\omega_{|\Omega_{\mathrm{AT}}}} \geq n h_\alpha^{\uparrow,\mathrm{QKD}} - \frac{\alpha}{\alpha-1}\log\frac{1}{\Pr[\Omega_{\mathrm{AT}}]}. \tag{8.2.21}$$

Here the quantity $h_\alpha^{\uparrow,\mathrm{QKD}}$ is given by

$$h_\alpha^{\uparrow,\mathrm{QKD}} = \inf_{\mathbf{q}\in S_{\mathrm{acc}}} \inf_{\substack{\rho\in \mathrm{S}_=(AB) \\ s.t.\ \mathrm{Tr}_B[\rho]=\tau_A}} \left(\frac{\alpha}{\alpha-1} D\left(\mathbf{q}\middle\|\boldsymbol{\nu}_{\widehat{C}}^\rho\right) + q(\mathtt{gen})\widetilde{H}_\alpha^\uparrow(S|\tilde{T}E)_{\nu_{|\mathtt{gen}}^\rho} \right), \tag{8.2.22}$$

where $D(.\|.)$ is the relative entropy and

$$\nu_{S\tilde{T}\widehat{C}E}^\rho = \mathcal{M}_{QKD}(\rho_{ABE}), \tag{8.2.23}$$

$$\boldsymbol{\nu}_{\widehat{C}}^\rho = (\gamma\boldsymbol{\nu}_{\widehat{C}|\mathtt{test}}^\rho, 1-\gamma)^T, \tag{8.2.24}$$

$$\boldsymbol{\nu}_{\widehat{C}|\mathtt{test}}^\rho = \boldsymbol{\Phi}[\rho_{AB|\mathtt{test}}], \tag{8.2.25}$$

and E is a purifying system such that ρ_{ABE} is pure.

Proof. First, note that since each $\widehat{C}_i$ can be computed from I_i with the deterministic function ϕ, one finds

$$\widetilde{H}_\alpha^\uparrow(S_1^n|I_1^n\widetilde{E})_{\omega_{|\Omega_{\mathrm{AT}}}} = \widetilde{H}_\alpha^\uparrow(S_1^n|I_1^n\widehat{C}_1^n\widetilde{E})_{\omega_{|\Omega_{\mathrm{AT}}}}. \tag{8.2.26}$$

Next, we will apply [41, Corollary 4.2], which gives us a lower bound on $\widetilde{H}_\alpha^\uparrow(S_1^n|I_1^n\widehat{C}_1^n\widetilde{E})_{\omega_{|\Omega_{\mathrm{AT}}}}$ in terms of an expression equivalent to $h_\alpha^{\uparrow,\mathrm{QKD}}$. Therefore, observe that the channel $\mathcal{M}_{\mathrm{QKD}}$

satisfies the assumptions of [41, Corollary 4.2]. Hence, making the following identifications between the register and variable names in this work and in [41, Corollary 4.2],

$$\begin{aligned} I &\mapsto I, \widehat{C} \mapsto \widehat{C}, \bar{C} = \emptyset, A_j \mapsto A_{j-1}, \\ B_j &\mapsto B_{j-1}, \widetilde{E} \mapsto \widehat{E}, E \mapsto \widetilde{E}, \\ \omega &\mapsto \rho, \rho \mapsto \omega, \tau \mapsto \sigma, \\ S_{\Omega_{\mathrm{AT}}} &\mapsto S_{\Omega}. \end{aligned} \tag{8.2.27}$$

we can apply [41, Corollary 4.2], and $h_\alpha^{\uparrow,\mathrm{QKD}}$ has the following intermediate form

$$h_\alpha^{\uparrow,\mathrm{QKD}} = \inf_{\mathbf{q}\in S_{\mathrm{acc}}} \inf_{\substack{\rho\in \mathrm{S}_=(AB) \\ \text{s.t. } \mathrm{Tr}_B[\rho]=\tau_A}} \left(\frac{\alpha}{\alpha-1} D\left(\mathbf{q}\middle\|\boldsymbol{\nu}^\rho_{\widehat{C}}\right) + q(\mathtt{gen})\widetilde{H}_\alpha^{\uparrow}(S|IE)_{\nu^\rho_{|\mathtt{gen}}} \right). \tag{8.2.28}$$

However, this is equal to the formula claimed in the theorem statement, since conditioned on a generation round, we have $T = I$. Thus, also Eve's copy satisfies $\tilde{T} = I$ in generation rounds.

Moreover, the equality between the probability distribution $\boldsymbol{\nu}^\rho_{\widehat{C}|\mathtt{test}}$ on $\widehat{C}$ induced by $\mathcal{M}_{\mathrm{QKD}}[\rho_{ABE}]$ and $\mathbf{\Phi}[\rho_{AB|\mathtt{test}}]$ follows from a direct calculation. □

We have not yet laid out how to calculate $h_\alpha^{\uparrow,\mathrm{QKD}}$ for a QKD protocol. One requires a significant amount of reformulations to reach an expression which is easily implementable with numerical methods. Therefore, we conclude this section with the security proof assuming that one has access to a lower bound $\underline{h}^{\mathrm{QKD}}$ on $h_\alpha^{\uparrow,\mathrm{QKD}}$.

Theorem 8.2.4 (MEAT Fixed-Length Security)**.** *For any $\alpha \in (1,2)$, $\varepsilon_{\mathrm{PA}}, \varepsilon_{\mathrm{EV}} \in (0,1]$, the fixed-length version of Protocol 1 is $\varepsilon_{\mathrm{PA}}$-secret, $\varepsilon_{\mathrm{EV}}$-correct, and hence $(\varepsilon_{\mathrm{PA}} + \varepsilon_{\mathrm{EV}})$-secure, when the length l of the final key satisfies*

$$l \leq n\underline{h}^{QKD} - \lambda_{EC} - \left\lceil \log \frac{1}{\varepsilon_{\mathrm{EV}}} \right\rceil - \frac{\alpha}{\alpha-1} \log \frac{1}{\varepsilon_{\mathrm{PA}}} + 1, \tag{8.2.29}$$

where $\underline{h}^{QKD}$ is a lower bound on $h_\alpha^{\uparrow,\mathrm{QKD}}$, which is defined as in Theorem 8.2.3 and λ_{EC} is the length of the error-correction string.

Proof. As stated in Remark 3.1.4 below Definition 3.1.2, secrecy and correctness imply security. Therefore, we prove each part separately.

Starting with correctness; the protocol is $\varepsilon_{\mathrm{EV}}$-correct since,

$$\Pr[K_A \neq K_B \wedge \Omega_{\mathrm{acc}}] \leq \Pr[K_A \neq K_B \wedge \Omega_{\mathrm{EV}}] \tag{8.2.30}$$

$$\leq \Pr\left[S_1^n \neq \overline{S}_1^n \wedge \Omega_{\mathrm{EV}}\right] \leq \Pr\left[\Omega_{\mathrm{EV}} | S_1^n \neq \overline{S}_1^n\right] \tag{8.2.31}$$

$$\leq 2^{-\left\lceil \log \frac{1}{\varepsilon_{\mathrm{EV}}} \right\rceil} \leq \varepsilon_{\mathrm{EV}}, \tag{8.2.32}$$

by the property of 2-universal hashing.

For secrecy, let us start with the secrecy definition and apply the LHL for Rényi entropies, Corollary 3.5.6, yielding

$$\frac{1}{2}\Pr[\Omega_{\mathrm{acc}}] \left\| \sigma_{K_A I_1^n E L G_{|\Omega_{\mathrm{acc}}}} - \mathbb{U}_{K_A}^l \otimes \sigma_{I_1^n E L G_{|\Omega_{\mathrm{acc}}}} \right\|_1 \tag{8.2.33}$$

$$\leq \Pr[\Omega_{\mathrm{acc}}] 2^{\frac{2(1-\alpha)}{\alpha}} 2^{\frac{1-\alpha}{\alpha}\left(\widetilde{H}_\alpha^\uparrow(S_1^n | I_1^n E L)_{\omega_{|\Omega_{\mathrm{acc}}}} - l\right)} \tag{8.2.34}$$

$$= \Pr[\Omega_{\mathrm{EV}} | \Omega_{\mathrm{AT}}] \Pr[\Omega_{\mathrm{AT}}] 2^{\frac{1-\alpha}{\alpha}\left(\widetilde{H}_\alpha^\uparrow(S_1^n | I_1^n E L)_{\omega_{|\Omega_{\mathrm{acc}}}} - l + 1\right)}. \tag{8.2.35}$$

Continuing with the Rényi entropy itself, we apply [28, Lemma B.5] together with $\Omega_{\mathrm{acc}} = \Omega_{\mathrm{AT}} \wedge \Omega_{\mathrm{EV}}$ and find

$$\widetilde{H}_\alpha^\uparrow(S_1^n | I_1^n E L)_{\omega_{|\Omega_{\mathrm{AT}} \wedge \Omega_{\mathrm{EV}}}} \geq \widetilde{H}_\alpha^\uparrow(S_1^n | I_1^n E L)_{\omega_{|\Omega_{\mathrm{AT}}}} - \frac{\alpha}{\alpha - 1} \log \frac{1}{\Pr[\Omega_{\mathrm{EV}} | \Omega_{\mathrm{AT}}]}. \tag{8.2.36}$$

Next, we split off the error correction data contained in register L by applying the chain rule of [47, Eq. (5.94)], which combines [47, Lemmas 5.3 & 5.4], to reach

$$\begin{aligned} &\widetilde{H}_\alpha^\uparrow(S_1^n | I_1^n E L)_{\omega_{|\Omega_{\mathrm{AT}} \wedge \Omega_{\mathrm{EV}}}} \\ &\geq \widetilde{H}_\alpha^\uparrow(S_1^n | I_1^n E)_{\omega_{|\Omega_{\mathrm{AT}}}} - \lambda_{\mathrm{EC}} - \left\lceil \log \frac{1}{\varepsilon_{\mathrm{EV}}} \right\rceil - \frac{\alpha}{\alpha - 1} \log \frac{1}{\Pr[\Omega_{\mathrm{EV}} | \Omega_{\mathrm{AT}}]}. \end{aligned} \tag{8.2.37}$$

Due to the source replacement scheme, we only need to consider states with a marginal $\mathrm{Tr}_{B_1^n \widetilde{E}}[\rho_{A_1^n B_1^n \widetilde{E}}] = \tau_A^{\otimes n}$ defined by Alice's sending probabilities and signal states, which we can include in Theorem 8.2.3. Then, by Theorem 8.2.3 the Rényi entropy conditioned on the event Ω_{AT} is bounded by

$$\widetilde{H}_\alpha^\uparrow(S_1^n | I_1^n E)_{\omega_{|\Omega_{\mathrm{AT}}}} \geq n h_\alpha^{\uparrow,\mathrm{QKD}} - \frac{\alpha}{\alpha - 1} \log \frac{1}{\Pr[\Omega_{\mathrm{AT}}]}. \tag{8.2.38}$$

By assumption, $\underline{h}^{\mathrm{QKD}}$ is a lower bound on $h_\alpha^{\uparrow,\mathrm{QKD}}$, therefore

$$\begin{aligned}&\widetilde{H}_\alpha^\uparrow(S_1^n|I_1^n EL)_{\omega_{|\Omega_{\mathrm{AT}}\wedge\Omega_{\mathrm{EV}}}}\\&\geq n\underline{h}^{\mathrm{QKD}} - \frac{\alpha}{\alpha-1}\log\frac{1}{\Pr[\Omega_{\mathrm{AT}}]} - \lambda_{\mathrm{EC}} - \left\lceil\log\frac{1}{\varepsilon_{\mathrm{EV}}}\right\rceil - \frac{\alpha}{\alpha-1}\log\frac{1}{\Pr[\Omega_{\mathrm{EV}}|\Omega_{\mathrm{AT}}]}.\end{aligned} \tag{8.2.39}$$

Now, if we insert Eq. (8.2.39) and the key length expression from Eq. (8.2.29) into Eq. (8.2.35), we find

$$\frac{\Pr[\Omega_{\mathrm{acc}}]}{2}\left\|\sigma_{K_A I_1^n EL_{|\Omega_{\mathrm{acc}}}} - \mathbb{U}_{K_A}^l \otimes \sigma_{I_1^n EL_{|\Omega_{\mathrm{acc}}}}\right\|_1 \leq \varepsilon_{\mathrm{PA}}. \tag{8.2.40}$$

Hence, the protocol is $\varepsilon_{\mathrm{PA}}$-secret and $\varepsilon_{\mathrm{EV}}$-correct, and thus $\varepsilon_{\mathrm{PA}}+\varepsilon_{\mathrm{EV}}$-secure. □

8.3 Variable-Length Protocols in MEAT

For any security proof presented in this thesis so far, we have only considered fixed-length QKD protocols. Now, in this section we will extend the analysis to variable-length protocols and prove their security. In preparation for this, we need to establish a few more definitions as presented in Ref. [77].

Definition 8.3.1 (f-weighted Rényi entropies (partially restated from [77, 79]))**.** Let $\rho \in \mathrm{S}_=(\widehat{C}QQ')$ be a state where $\widehat{C}$ is classical with alphabet $\widehat{\mathcal{C}}$. A *tradeoff function*[4] *on* $\widehat{C}$ is a function $f : \widehat{\mathcal{C}} \to \mathbb{R}$; equivalently, we may denote it as a real-valued tuple $\mathbf{f} \in \mathbb{R}^{|\widehat{\mathcal{C}}|}$ where each term in the tuple specifies the value $f(\hat{c})$. Given a tradeoff function f and a value $\alpha \in (0,1)\cup(1,\infty)$, we define two versions of an *f-weighted Rényi entropy* of order α for ρ, as follows:

$$\widetilde{H}_\alpha^{\uparrow,f}(Q|\widehat{C}Q')_\rho := \frac{\alpha}{1-\alpha}\log\left(\sum_{\hat{c}}\rho(\hat{c})\,2^{\frac{1-\alpha}{\alpha}\left(-f(\hat{c})+\widetilde{H}_\alpha^\uparrow(Q|Q')_{\rho_{|\hat{c}}}\right)}\right), \tag{8.3.1}$$

$$\widetilde{H}_\alpha^{\downarrow,f}(Q|\widehat{C}Q')_\rho := \frac{1}{1-\alpha}\log\left(\sum_{\hat{c}}\rho(\hat{c})\,2^{(1-\alpha)\left(-f(\hat{c})+\widetilde{H}_\alpha^\downarrow(Q|Q')_{\rho_{|\hat{c}}}\right)}\right), \tag{8.3.2}$$

where the sums run over all $\hat{c}$ values such that $\rho(\hat{c}) > 0$. We extend both definitions to $\alpha = \infty$ by taking the $\alpha \to \infty$ limit.

[4] Ref. [77] instead referred to this as a "quantum estimation score-system" (QES).

Definition 8.3.2 (Normalized tradeoff functions)**.** Let f be a tradeoff function on a register $\widehat{C}$ as in Definition 8.3.1. Given any set of states $\mathcal{S} \subseteq \mathrm{S}_{=}(\widehat{C}QQ')$, we define the $\widetilde{H}_\alpha^{\uparrow,f}$-*normalization constant* and $\widetilde{H}_\alpha^{\downarrow,f}$-*normalization constant* for that set to be, respectively,

$$\kappa^{\uparrow} \coloneqq \inf_{\rho \in \mathcal{S}} \widetilde{H}_\alpha^{\uparrow,f}(Q|\widehat{C}Q')_\rho, \tag{8.3.3}$$

$$\kappa^{\downarrow} \coloneqq \inf_{\rho \in \mathcal{S}} \widetilde{H}_\alpha^{\downarrow,f}(Q|\widehat{C}Q')_\rho. \tag{8.3.4}$$

Given either of the above values, we then define a corresponding $\widetilde{H}_\alpha^{\uparrow,f}$-*normalized* or $\widetilde{H}_\alpha^{\downarrow,f}$-*normalized* tradeoff function f respectively, via

$$\hat{f}(\hat{c}) \coloneqq f(\hat{c}) + \kappa^{\uparrow}, \tag{8.3.5}$$

$$\hat{f}(\hat{c}) \coloneqq f(\hat{c}) + \kappa^{\downarrow}. \tag{8.3.6}$$

With these definitions, we are already able to prove security of variable-length QKD protocols.

Theorem 8.3.3 (Variable-Length Security)**.** *Consider Protocol 1, and take any $\alpha \in (1,2)$, $\varepsilon_{\mathrm{PA}}, \varepsilon_{\mathrm{EV}} \in (0,1]$, and any tradeoff function $f : \widehat{\mathcal{C}} \to \mathbb{R}$. Let $\kappa^{\uparrow}$ be the $\widetilde{H}_\alpha^{\uparrow,f}$-normalization constant (Definition 8.3.2) corresponding to the set of all states that can be produced in a single round, i.e.*

$$\kappa^{\uparrow} \coloneqq \inf_{\substack{\rho_{AB} \in \mathrm{S}_{=}(AB) \\ s.t.\ \mathrm{Tr}_B[\rho_{AB}] = \tau_A}} \widetilde{H}_\alpha^{\uparrow,f}(S|\widehat{C}IE)_\nu, \tag{8.3.7}$$

where the set of possible output states $\nu_{S\widehat{C}IE}$ is parameterized via the input state ρ_{AB} on Alice and Bob's registers as described in the previous sections. We then define the following tradeoff function on $\widehat{C}_1^n$:

$$\hat{f}_{\mathrm{full}}(\hat{c}_1^n) \coloneqq \sum_i (f(\hat{c}_i) + \kappa^{\uparrow}). \tag{8.3.8}$$

Then, the variable-length version of Protocol 1 is $\varepsilon_{\mathrm{PA}}$-secret, $\varepsilon_{\mathrm{EV}}$-correct, and hence $(\varepsilon_{\mathrm{PA}} + \varepsilon_{\mathrm{EV}})$-secure, whenever the length l_{var} of the final key (when error verification does not abort) is chosen as some function of $\hat{c}_1^n$ satisfying

$$l_{\mathrm{var}} \leq \max\left\{0, \hat{f}_{\mathrm{full}}(\hat{c}_1^n) - \lambda_{EC}(\hat{c}_1^n) - \left\lceil \log \frac{1}{\varepsilon_{\mathrm{EV}}} \right\rceil - \frac{\alpha}{\alpha - 1} \log \frac{1}{\varepsilon_{\mathrm{PA}}} + 1\right\}, \tag{8.3.9}$$

where $\lambda_{EC}(\hat{c}_1^n)$ is the length of error correction data used. In particular, this holds if we define the function

$$\hat{f}_{\text{prot}}(\hat{c}_1^n) \coloneqq \sum_i (f(\hat{c}_i) + \underline{\kappa}^{\uparrow}), \tag{8.3.10}$$

where $\underline{\kappa}^{\uparrow}$ is any lower bound on $\kappa^{\uparrow}$, and choose

$$l_{\text{var}} = \max\left\{0, \left\lfloor \hat{f}_{\text{prot}}(\hat{c}_1^n) - \lambda_{EC}(\hat{c}_1^n) - \left\lceil \log \frac{1}{\varepsilon_{\text{EV}}} \right\rceil - \frac{\alpha}{\alpha - 1} \log \frac{1}{\varepsilon_{\text{PA}}} + 1 \right\rfloor \right\}. \tag{8.3.11}$$

Remark 8.3.4. If we implement a variable-length protocol by setting l_{var} as specified in Eq. (8.3.11), then we need to compute a lower bound $\underline{\kappa}^{\uparrow}$ on the $\widetilde{H}_{\alpha}^{\uparrow,f}$-normalization constant $\kappa^{\uparrow}$. We can use any method that securely computes such a lower bound. Even more, since by [77, Lemma 6.2], the two versions of f-weighted Rényi entropies are related by $\widetilde{H}_{\alpha}^{\uparrow,f}(Q|\widehat{C}Q')_{\rho} \geq \widetilde{H}_{\alpha}^{\downarrow,f}(Q|\widehat{C}Q')_{\rho}$, one can pick either of these normalization constants as $\underline{\kappa}^{\uparrow}$ and achieve secure key rates.

The main claim in Theorem 8.3.3 was proven in Ref. [79]. However, at the current time of writing, the manuscript for that work is still in preparation. Thus for the sake of completeness, we re-state here the essential steps of their analysis, but we emphasize that this is simply an exposition of their proof, and that work should be cited as the source whenever possible.

Proof. As stated in Definition 3.1.2, secrecy and correctness imply security. Therefore, we prove each part separately. As a reminder, in line with the protocol description, we use the following register names, K_A for Alice's final key, L for error correction and verification data, G storing the choice of hash function, and $I, \widehat{C}$ to indicate the different public announcements.

Starting with correctness; the protocol is ε_{EV}-correct since,

$$\Pr[K_A \neq K_B \wedge \Omega_{\text{acc}}] \leq \Pr[K_A \neq K_B \wedge \Omega_{\text{EV}}] \tag{8.3.12}$$

$$\leq \Pr\left[S_1^n \neq \overline{S}_1^n \wedge \Omega_{\text{EV}}\right] \leq \Pr\left[\Omega_{\text{EV}} | S_1^n \neq \overline{S}_1^n\right] \tag{8.3.13}$$

$$\leq 2^{-\left\lceil \log \frac{1}{\varepsilon_{\text{EV}}} \right\rceil} \leq \varepsilon_{\text{EV}}. \tag{8.3.14}$$

Next, to show secrecy, we observe the following critical bound: by [41, Theorem 4.1a and Corollary 4.2] (also proven in [78, 79]), the state ω at the end of the public announcement

step satisfies

$$\widetilde{H}_{\alpha}^{\uparrow,\hat{f}_{\text{full}}}(S_1^n|\widehat{C}_1^n I_1^n \tilde{E})_{\omega} \geq 0. \tag{8.3.15}$$

With this, we can prove that the variable-length definition of secrecy in Definition 3.1.2 holds. We again highlight that the main steps shown here were first presented in [79]. We find:

$$\frac{1}{2}\sum_{m=0}^{\infty} \Pr[\Omega_{\text{len}=m}] \left\| \sigma_{K_A I_1^n LEG|\Omega_{\text{len}=m}} - \mathbb{U}_{K_A}^m \otimes \sigma_{I_1^n LEG|\Omega_{\text{len}=m}} \right\|_1 \tag{8.3.16}$$

$$\leq \frac{1}{2}\sum_{\hat{c}_1^n} \Pr\left[\widehat{C}_1^n = \hat{c}_1^n \wedge \Omega_{\text{EV}}\right] \left\| \sigma_{K_A I_1^n LEG|\widehat{C}_1^n=\hat{c}_1^n \wedge \Omega_{\text{EV}}} - \mathbb{U}_{K_A}^m \otimes \sigma_{I_1^n LEG|\widehat{C}_1^n=\hat{c}_1^n \wedge \Omega_{\text{EV}}} \right\|_1 \tag{8.3.17}$$

$$\leq \sum_{\hat{c}_1^n \text{ s.t. } l_{\text{var}}>0} \Pr\left[\widehat{C}_1^n = \hat{c}_1^n \wedge \Omega_{\text{EV}}\right] 2^{\frac{1-\alpha}{\alpha}\left(\widetilde{H}_{\alpha}^{\uparrow}(S_1^n|I_1^n LE)_{\omega|\widehat{C}_1^n=\hat{c}_1^n \wedge \Omega_{\text{EV}}} - l_{\text{var}} + 1\right)} \tag{8.3.18}$$

$$\leq \sum_{\hat{c}_1^n \text{ s.t. } l_{\text{var}}>0} \Pr\left[\widehat{C}_1^n = \hat{c}_1^n \wedge \Omega_{\text{EV}}\right] 2^{\frac{1-\alpha}{\alpha}\left(\widetilde{H}_{\alpha}^{\uparrow}(S_1^n|I_1^n LE)_{\omega|\widehat{C}_1^n=\hat{c}_1^n} - \frac{\alpha}{\alpha-1}\log\frac{1}{\Pr\left[\Omega_{\text{EV}}|\widehat{C}_1^n=\hat{c}_1^n\right]} - l_{\text{var}} + 1\right)} \tag{8.3.19}$$

$$= \sum_{\hat{c}_1^n \text{ s.t. } l_{\text{var}}>0} \Pr\left[\widehat{C}_1^n = \hat{c}_1^n\right] 2^{\frac{1-\alpha}{\alpha}\left(\widetilde{H}_{\alpha}^{\uparrow}(S_1^n|I_1^n LE)_{\omega|\widehat{C}_1^n=\hat{c}_1^n} - l_{\text{var}} + 1\right)} \tag{8.3.20}$$

$$\leq \sum_{\hat{c}_1^n \text{ s.t. } l_{\text{var}}>0} \Pr\left[\widehat{C}_1^n = \hat{c}_1^n\right] 2^{\frac{1-\alpha}{\alpha}\left(\widetilde{H}_{\alpha}^{\uparrow}(S_1^n|I_1^n LE)_{\omega|\widehat{C}_1^n=\hat{c}_1^n} - \hat{f}_{\text{full}}(\hat{c}_1^n) + \lambda_{\text{EC}}(\hat{c}_1^n) + \left\lceil \log\frac{1}{\varepsilon_{\text{EV}}}\right\rceil + \frac{\alpha}{\alpha-1}\log\frac{1}{\varepsilon_{\text{PA}}}\right)} \tag{8.3.21}$$

$$\leq \sum_{\hat{c}_1^n \text{ s.t. } l_{\text{var}}>0} \Pr\left[\widehat{C}_1^n = \hat{c}_1^n\right] 2^{\frac{1-\alpha}{\alpha}\left(\widetilde{H}_{\alpha}^{\uparrow}(S_1^n|I_1^n E)_{\omega|\widehat{C}_1^n=\hat{c}_1^n} - \hat{f}_{\text{full}}(\hat{c}_1^n) + \frac{\alpha}{\alpha-1}\log\frac{1}{\varepsilon_{\text{PA}}}\right)} \tag{8.3.22}$$

$$= \varepsilon_{\text{PA}} \sum_{\hat{c}_1^n \text{ s.t. } l_{\text{var}}>0} \Pr\left[\widehat{C}_1^n = \hat{c}_1^n\right] 2^{\frac{1-\alpha}{\alpha}\left(\widetilde{H}_{\alpha}^{\uparrow}(S_1^n|I_1^n E)_{\omega|\widehat{C}_1^n=\hat{c}_1^n} - \hat{f}_{\text{full}}(\hat{c}_1^n)\right)} \tag{8.3.23}$$

$$\leq \varepsilon_{\text{PA}} \sum_{\hat{c}_1^n} \Pr\left[\widehat{C}_1^n = \hat{c}_1^n\right] 2^{\frac{1-\alpha}{\alpha}\left(\widetilde{H}_{\alpha}^{\uparrow}(S_1^n|I_1^n E)_{\omega|\widehat{C}_1^n=\hat{c}_1^n} - \hat{f}_{\text{full}}(\hat{c}_1^n)\right)} \tag{8.3.24}$$

$$= \varepsilon_{\text{PA}} 2^{\frac{1-\alpha}{\alpha}\left(\widetilde{H}_{\alpha}^{\uparrow,\hat{f}_{\text{full}}}(S_1^n|I_1^n \widehat{C}_1^n E)_{\omega}\right)} \leq \varepsilon_{\text{PA}}. \tag{8.3.25}$$

The second line (Eq. (8.3.17)) is due to the key length m depending only on the announcements $\hat{c}_1^n$ and error verification succeeding. Furthermore, for any $\hat{c}_1^n$ leading to the same

key length m, we used strong convexity to bound the trace distance. Thus, one can rewrite the sum in terms of $\hat{c}_1^n$ as in the second line. The third line (Eq. (8.3.18)) follows by applying the Renyi LHL (Corollary 3.5.6) together with the fact that the trace-distance term is zero whenever $l_{\text{var}} = 0$ (note that here we view l_{var} as the function of $\hat{c}_1^n$ satisfying Eq. (8.3.9)), observing that in all the trace-distance terms in the preceding expression, we have conditioned on error verification accepting and thus the key length would indeed be given by that function). The fourth line (Eq. (8.3.19)) is due to [28, Lemma B.5]. Line six (Eq. (8.3.21)) simply inserts Eq. (8.3.9) (recalling the sum is restricted only to terms with $l_{\text{var}} > 0$) and line seven (Eq. (8.3.22)) follows by using the chain rule of [47, Eq. (5.94)], which combines [47, Lemmas 5.3 & 5.4]. The ninth line (Eq. (8.3.24)) holds simply because we have extended the sum with only non-negative terms, and the tenth line (Eq. (8.3.25)) follows by the definition of f-weighted Rényi entropies, Definition 8.3.1. Finally, the last inequality follows from the fundamental bound noted in Eq. (8.3.15) above. □

8.3.1 Properties of Variable-Length Key Rates

We now discuss an important connection between our variable-length key lengths and the expected key length. As we already discussed in Section 7.4, expected key rates capture the tradeoff between high achievable key rates and accepting the observations with high probability. Variable-length protocols are particularly well suited for this purpose as they intuitively adapt the resulting key length based on the quality of the observations.

Therefore, recall the definition of the observed frequency vector $\mathbf{F}^{\text{obs}}$ defined for each component as

$$F^{\text{obs}}(\hat{c}) = \frac{\text{number of occurrences of } \hat{c} \text{ in } \hat{c}_1^n}{n}. \tag{8.3.26}$$

For any $\underline{\kappa}^{\uparrow}$ as described in Theorem 8.3.3, the resulting function $\hat{f}_{\text{prot}}$ applied to any string $\hat{c}_1^n$ can be written in terms of $\mathbf{F}^{\text{obs}}$ as follows:

$$\hat{f}_{\text{prot}}(\hat{c}_1^n) = n\left(\mathbf{f} \cdot \mathbf{F}^{\text{obs}} + \underline{\kappa}^{\uparrow}\right), \tag{8.3.27}$$

where $\mathbf{f}$ denotes the tradeoff function $f : \widehat{\mathcal{C}} \to \mathbb{R}$ viewed as a vector $\mathbf{f} \in \mathbb{R}^{|\widehat{\mathcal{C}}|}$. Since the expectation value is linear it follows

$$\mathbb{E}[\hat{f}_{\text{prot}}(\hat{c}_1^n)] = n\left(\mathbf{f} \cdot \mathbb{E}[\mathbf{F}^{\text{obs}}] + \underline{\kappa}^{\uparrow}\right). \tag{8.3.28}$$

Let us write $\mathbf{q}^{\text{hon}}$ to denote the probability vector of the honest IID channel behaviour, i.e. the channel were Eve does not interact with the signals at all. Each component $q^{\text{hon}}(\hat{c})$ of

the vector $\mathbf{q}^{\text{hon}}$ thus gives the probability of observing outcome $\hat{c}$. Then the expected value of the frequency vector given by this behaviour is connected to $\mathbf{q}^{\text{hon}}$ by

$$\mathbb{E}[\mathbf{F}^{\text{obs}}] = \mathbf{q}^{\text{hon}}. \tag{8.3.29}$$

Therefore, the expectation value of the function $\hat{f}_{\text{prot}}$ satisfies

$$\mathbb{E}[\hat{f}_{\text{prot}}(\hat{c}_1^n)] = n\left(\mathbf{f}\cdot\mathbf{q}^{\text{hon}} + \underline{\kappa}^{\uparrow}\right). \tag{8.3.30}$$

Let us additionally assume that the error correction procedure is chosen such that it satisfies $\mathbb{E}[\lambda_{\text{EC}}(\hat{c}_1^n)] \leq \lambda_{\text{EC}}(\mathbf{q}^{\text{hon}})$, and has negligible probability of failing for the honest behaviour, which has been made previously as well see [55, 73]. Then when the key length is chosen according to Eq. (8.3.11), and ignoring the floor function, its expected value would satisfy

$$\begin{aligned}\mathbb{E}[l_{\text{var}}] \geq & n\left(\mathbf{f}\cdot\mathbf{q}^{\text{hon}} + \underline{\kappa}^{\uparrow}\right) - \lambda_{\text{EC}}(\mathbf{q}^{\text{hon}}) - \left\lceil \log\frac{1}{\varepsilon_{\text{EV}}}\right\rceil \\ & - \frac{\alpha}{\alpha-1}\log\frac{1}{\varepsilon_{\text{PA}}} + 1,\end{aligned} \tag{8.3.31}$$

since the maximum on the right-hand-side of Eq. (8.3.11) is lower bounded by its second argument.

This expression has one immediate important property. It is simply equal to the key length in Eq. (8.3.11) evaluated at the expected frequency distribution of the honest channel. In the subsequent sections, when plotting the expected key rates in the variable-length case, we will be taking the expression in Eq. (8.3.31) as the formula for the expected key length.

8.3.2 Optimality of the Tradeoff Function

Note that while Theorem 8.3.3 is valid for any choice of the (single-round) tradeoff function f, it does not by itself indicate how to choose f in a manner that produces the best finite-size performance. However, in [41, Lemma 4.12], a method was presented to address this which gives a numerical technique to find the optimal f in

$$\sup_f\left(\mathbf{f}\cdot\mathbf{q}^{\text{hon}} + \kappa^{\uparrow}\right) = \sup_f \inf_{\nu\in\Sigma}\left(\widetilde{H}_\alpha^{\uparrow,f}(S|\widehat{C}IE)_\nu + \mathbf{f}\cdot\mathbf{q}^{\text{hon}}\right). \tag{8.3.32}$$

In [41, Lemma 4.12] it was shown that this optimization problem is equal to $h_\alpha^{\uparrow,\mathrm{QKD}}$ adding an auxiliary variable $\boldsymbol{\lambda} \in \mathbb{P}(\widehat{\mathcal{C}})$ that is a probability vector on register $\widehat{C}$. Hence, one can find the optimal tradeoff function by solving the modified optimization problem

$$\inf_{\substack{\rho\in \mathrm{S}_=(AB)\\ \text{s.t. } \mathrm{Tr}_B[\rho]=\tau_A}} \inf_{\boldsymbol{\lambda}\in\mathbb{P}(\widehat{\mathcal{C}})} \left(\frac{\alpha}{\alpha-1} D\left(\boldsymbol{\lambda}\middle\| \boldsymbol{\nu}^\rho_{\widehat{C}}\right) + q(\mathtt{gen}) \widetilde{H}^\uparrow_\alpha(S|\tilde{T}E)_{\nu^\rho_{|\mathtt{gen}}} \right) \tag{8.3.33}$$

$$\text{s.t. } \mathrm{Tr}_B[\rho] = \tau_A,$$

$$\boldsymbol{\nu}^\rho_{\widehat{C}} = (\gamma \boldsymbol{\nu}^\rho_{\widehat{C}|\mathtt{test}}, 1-\gamma)^T, \tag{8.3.34}$$

$$\boldsymbol{\nu}^\rho_{\widehat{C}|\mathtt{test}} = \boldsymbol{\Phi}[\rho_{AB|\mathtt{test}}], \tag{8.3.35}$$

$$\mathbf{q}^{\mathrm{hon}} - \boldsymbol{\lambda} = 0, \tag{solve for dual}$$

where $\mathbf{q}^{\mathrm{hon}}$ is fixed and Ref. [41] also showed that this is a convex optimization problem. If $\mathbf{q}^{\mathrm{hon}}$ is not fully known, one can alternatively add another minimization over $\mathbf{q}$. Most importantly, the dual variables with respect to the constraint Eq. (solve for dual) define each component of the tradeoff function $f(\hat{c})$.

Moreover, observe that by picking this tradeoff function, we maximize the expected key length formula in Eq. (8.3.31) for $\underline{\kappa}^\uparrow = \kappa^\uparrow$. Hence, later on we will use a slightly modified version of this method to choose f, as it would maximize that formula for the $\underline{\kappa}^\uparrow = \kappa^\uparrow$ case, and any suboptimality in the expected key rate caused by using a smaller $\underline{\kappa}^\uparrow$ would then be at most the difference $\kappa^\uparrow - \underline{\kappa}^\uparrow$. The required modification are swapping the optimized sandwiched Rényi entropy $\widetilde{H}^\uparrow_\alpha$ with the unoptimized version $\widetilde{H}^\downarrow_\alpha$ as this is yields much simpler numerical methods. We will lay out the numerical methods in Chapter 10.

Remark 8.3.5. We stress that while numerical methods for finding the dual solution might not find the exact optimal choice of f, this does not affect the security of our results. This is because Theorem 8.3.3 holds for *any* choice of f, and hence using a suboptimal f would only mean that the resulting key rates we obtain might not be optimal. The main requirement for security is to ensure that we compute a rigorous lower bound $\underline{\kappa}^\uparrow$ on $\kappa^\uparrow$, which we now discuss.

8.4 Prepare-and-Measure Protocols in GEAT

Finally, let us also consider fixed-length security proofs under the generalized entropy accumulation theorem (GEAT) [30, 31, 33]. The channel structure of the GEAT is a concatenation of channels, i.e.

$$\mathcal{M}_n \circ \cdots \circ \mathcal{M}_1. \tag{8.4.1}$$

In principle, the GEAT allows for finite-size security proofs of prepare-and-measure protocols. However, to rigorously incorporate a prepare-and-measure QKD protocol in the GEAT, each channel $\mathcal{M}_i$ needs to perform Alice's state generation, Eve's attack, Bob's measurements and all announcements, see [33]. This implies, that Eve could only interact with one signal in each round, thus restricting the repetition rate of Alice's source.

However, one can alternatively analyze the same protocol under the MEAT model of a tensor product channel as shown in Fig. 8.1. As described in Refs. [41, 77] one can then achieve the *exact same* key rates as in [2] against arbitrary attacks especially without any restrictions on Eve's attack or repetition rates.

We demonstrate that the lower bound on the n-round entropy derived in [2, Theorem 1] using the GEAT framework also serves as a valid lower bound for $h_\alpha^{\uparrow,\mathrm{QKD}}$ (defined in Eq. (8.2.22)) under the tensor product MEAT model. Consequently, by invoking Theorem 8.2.4, we conclude that the key length expression established in [2, Theorem 3] ensures security against coherent attacks, again under the MEAT model.

We can establish an even stronger result regarding the GEAT key rates. To ensure numerical tractability in Chapter 10, we will define h_α^{QKD} by replacing the conditional sandwiched Rényi entropy $\widetilde{H}_\alpha^{\uparrow}$ in Eq. (8.2.22) with $\widetilde{H}_\alpha^{\downarrow}$. This yields:

$$h_\alpha^{\mathrm{QKD}} = \inf_{\mathbf{q}\in S_{\mathrm{acc}}} \inf_{\substack{\rho\in \mathrm{S}_=(AB) \\ \text{s.t. } \mathrm{Tr}_B[\rho]=\tau_A}} \left(\frac{\alpha}{\alpha-1} D\left(\mathbf{q}\middle\|\boldsymbol{\nu}^\rho_{\widehat{C}}\right) + q(\mathtt{gen})\widetilde{H}_\alpha^{\downarrow}(S|\tilde{T}E)_{\nu^\rho|\mathtt{gen}} \right), \tag{8.4.2}$$

where $D(.\|.)$ denotes the relative entropy and, as defined previously:

$$\nu^\rho_{S\tilde{T}\widehat{C}E} = \mathcal{M}_{\mathrm{QKD}}(\rho ABE), \tag{8.4.3}$$

$$\boldsymbol{\nu}^\rho_{\widehat{C}} = (\gamma \boldsymbol{\nu}^\rho_{\widehat{C}|\mathtt{test}}, 1-\gamma)^T, \tag{8.4.4}$$

$$\boldsymbol{\nu}^\rho_{\widehat{C}|\mathtt{test}} = \boldsymbol{\Phi}[\rho_{AB|\mathtt{test}}]. \tag{8.4.5}$$

By definition, $h_\alpha^{\uparrow,\mathrm{QKD}} \geq h_\alpha^{\mathrm{QKD}}$, and crucially, we demonstrate below that the GEAT key rates strictly lower bound this new quantity h_α^{QKD}.

To establish this result, we require additional concepts from the GEAT framework. Specifically, the derivations in Ref. [2] relied on *min-tradeoff* functions, defined as follows:

Definition 8.4.1 (Min-tradeoff function)**.** Let $\mathcal{M} : Q \to S\widehat{IC}$ be an arbitrary channel with classical $\widehat{C}$ and let Σ denote the set of all states of the form $\mathcal{M}[\omega_{QE}]$ for some pure

state $\omega \in \mathrm{S}_=(QE)$, where E is a purifying register of Q. A *min-tradeoff function* $f_{\min}$ for the channel $\mathcal{M}$ is an affine function of probability distributions on $\widehat{C}$, with the property

$$H(S|\widehat{C}IE)_\nu \geq f_{\min}(\nu_{\widehat{C}}) \tag{8.4.6}$$

for all $\nu \in \Sigma$.

We note that in contrast to the tradeoff functions defined previously in Section 8.3, a min-tradeoff function is a true lower bound on the von-Neumann entropy, not just in the log-mean-exponential sense.

With this definition at hand, we can already show the following theorem, making our claim from above rigorous.

Theorem 8.4.2 (GEAT bounds as lower bounds on MEAT bounds)**.** *Let $\mathcal{M}_{QKD} : AB \to SI\widehat{C}$ be the channel of a QKD protocol as in Definition 8.2.1. Additionally, let h_{vN} be a value satisfying $f_{\min}(\mathrm{freq}_{\hat{c}_1^n}) \geq h_{\mathrm{vN}}$ for all $\hat{c}_1^n$. Then, for $\omega_{S_1^n I_1^n \widehat{C}_1^n \widetilde{E}} = \mathcal{M}_{QKD}^{\otimes n}[\rho_{A_1^n B_1^n \widetilde{E}}]$ with $\mathrm{Tr}_{B_1^n \widetilde{E}}[\rho_{A_1^n B_1^n \widetilde{E}}] = \tau_A^{\otimes n}$, $\rho_{A_1^n B_1^n \widetilde{E}}$ a purification of $\rho_{A_1^n B_1^n}$ and $\alpha \in (1,2)$, it holds*

$$\widetilde{H}_\alpha^\uparrow(S_1^n|I_1^n\widetilde{E})_{\omega_{|\Omega_{\mathrm{AT}}}} = \widetilde{H}_\alpha^\uparrow(S_1^n|I_1^n\widehat{C}_1^n\widetilde{E})_{\omega_{|\Omega_{\mathrm{AT}}}} \tag{8.4.7}$$

$$\geq nh_{\mathrm{vN}} + nT_\alpha(f_{\min}) - n\left(\alpha-1\right)^2 K_\alpha(f_{\min}) - \frac{\alpha}{\alpha-1}\log\frac{1}{\Pr[\Omega_{\mathrm{AT}}]} \tag{8.4.8}$$

$$\geq nh_{\mathrm{vN}} + nT_{\widehat{\alpha}}(f_{\min}) - n\left(\widehat{\alpha}-1\right)^2 K_{\widehat{\alpha}}(f_{\min}) - \frac{\alpha}{\alpha-1}\log\frac{1}{\Pr[\Omega_{\mathrm{AT}}]}, \tag{8.4.9}$$

and we especially have

$$h_\alpha^{\mathrm{QKD}} \geq nh_{\mathrm{vN}} + nT_\alpha(f_{\min}) - n\left(\alpha-1\right)^2 K_\alpha(f_{\min}) \tag{8.4.10}$$

$$\geq nh_{\mathrm{vN}} + nT_{\widehat{\alpha}}(f_{\min}) - n\left(\widehat{\alpha}-1\right)^2 K_{\widehat{\alpha}}(f_{\min}), \tag{8.4.11}$$

where $\widehat{\alpha} = 1/(2-\alpha)$. The functions T_α and K_α are defined as:

$$T_\alpha(f_{\min}) = \inf_{\nu\in\Sigma}\left(H(S|I\widehat{C}E)_\nu - f_{\min}(\nu_{\widehat{C}}) - (\alpha-1)\,V(\nu_{\widehat{C}}, f_{\min})\right)$$

$$\mathrm{Var}(\mathbf{p}, f_{\min}) \coloneqq \sum_{\hat{c}} p(\hat{c}) f_{\min}(\delta_{\hat{c}})^2 - \left(\sum_{\hat{c}} p(\hat{c}) f_{\min}(\delta_{\hat{c}})\right)^2$$

$$V(\mathbf{p}, f_{\min}) \coloneqq \frac{\ln(2)}{2}\left(\log(1+2d_S^\kappa) + \sqrt{2+\mathrm{Var}(\mathbf{p}, f_{\min})}\right)^2$$

$$K_\alpha(f_{\min}) \coloneqq \frac{2^{(\alpha-1)(\kappa\log d_S + \max(f_{\min}) - \min(f_{\min}))}}{6(2-\alpha)^3\ln 2}\ln^3\left(2^{\kappa\log d_S + \max(f_{\min}) - \min(f_{\min})} + e^2\right),$$

$$\max(f_{\min}) \coloneqq \sup_{\mathbf{p}\in\mathbb{P}_{\widehat{C}}} f_{\min}(\mathbf{p}), \quad \min(f_{\min}) \coloneqq \inf_{\nu\in\Sigma} f_{\min}(\nu_{\widehat{C}}), \tag{8.4.12}$$

with $\kappa = 1$ if the S_i systems are classical and $\kappa = 2$ if they are quantum, $d_S = \dim(S)$ and $\delta_{\hat{c}}$ being the distribution with all weight on element $\hat{c}$.

Proof. The proof proceeds by proving the first bound in Eq. (8.4.10) through various properties of the entropies involved and then the rest of the statement follows from Theorem 8.2.3.

We start by bounding h_α^{QKD} from below and note that for all $\alpha \in (1,2)$ it holds $\widehat{\alpha} \geq \alpha$. Thus,

$$\frac{\alpha}{\alpha - 1} \geq \frac{1}{\widehat{\alpha} - 1} \quad \text{and} \quad \widetilde{H}_\alpha^{\downarrow}(S|\tilde{T}E)_{\nu^\rho_{|\mathtt{gen}}} \geq \widetilde{H}_{\widehat{\alpha}}^{\downarrow}(S|\tilde{T}E)_{\nu^\rho_{|\mathtt{gen}}}, \tag{8.4.13}$$

and h_α^{QKD} is bounded from below by

$$h_\alpha^{\mathrm{QKD}} \geq \inf_{\mathbf{q} \in S_{\mathrm{acc}}} \inf_{\substack{\rho \in \mathrm{S}_=(AB) \\ \text{s.t. } \mathrm{Tr}_B[\rho] = \tau_A}} \left(\frac{1}{\widehat{\alpha} - 1} D\left(\mathbf{q} \middle\| \boldsymbol{\nu}^\rho_{\widehat{C}}\right) + q(\mathtt{gen}) \widetilde{H}_{\widehat{\alpha}}^{\downarrow}(S|\tilde{T}E)_{\nu^\rho_{|\mathtt{gen}}} \right) =: h_{\widehat{\alpha}}, \tag{8.4.14}$$

where we defined $h_{\widehat{\alpha}}$ for brevity. Next, we express $h_{\widehat{\alpha}}$ in terms of its dual, see [77, Lemma 5.4],

$$h_{\widehat{\alpha}} = \sup_{\mathbf{f}} \inf_{\mathbf{q} \in S_{\mathrm{acc}}} \inf_{\nu \in \Sigma} \left(\widetilde{H}_{\widehat{\alpha}}^{\downarrow, f}(S|I\widehat{C}E)_\nu + \mathbf{f} \cdot \mathbf{q} \right), \tag{8.4.15}$$

where each component of the vector $\mathbf{f}$ is defined by $f(\hat{c})$. This is again lower bounded by dropping the supremum and using the min-tradeoff function $f_{\min}$ instead,

$$h_{\widehat{\alpha}} \geq \inf_{\mathbf{q} \in S_{\mathrm{acc}}} \inf_{\nu \in \Sigma} \left(\widetilde{H}_{\widehat{\alpha}}^{\uparrow, f_{\min}}(S|I\widehat{C}E)_\nu + \mathbf{f}_{\min} \cdot \mathbf{q} \right). \tag{8.4.16}$$

Now, consider the read-and-prepare channel $\mathcal{D}$ of [81, Supplement, Theorem 2] (also see [77, Definition 2.4]). Furthermore, for $\mathcal{D}$ consider a constant $M \in \mathbb{R}$ big enough such that

$$\widetilde{H}_\alpha^{\downarrow, f_{\min}}(S|I\widehat{C}E)_\nu = \widetilde{H}_\alpha^{\downarrow}(DS|I\widehat{C}E)_\nu - M, \tag{8.4.17}$$

and

$$H(D)_{\nu|\hat{c}} \geq \widetilde{H}_\alpha^{\downarrow}(D)_{\nu|\hat{c}} = M - f_{\min}(\hat{c}), \tag{8.4.18}$$

which can always be found due to [77, Lemma 4.1]. Then, by [82, Corollary 4] together with [81, Supplement, Eq. (15)] we can bound the conditional Rényi entropy $\widetilde{H}_{\widehat{\alpha}}^{\downarrow}(DS|I\widehat{C}E)_\nu$ as

$$\widetilde{H}_{\widehat{\alpha}}^{\downarrow}(DS|I\widehat{C}E)_\nu \geq \widetilde{H}_{2 - \frac{1}{\widehat{\alpha}}}^{\uparrow}(DS|I\widehat{C}E)_\nu = \widetilde{H}_\alpha^{\uparrow}(DS|I\widehat{C}E)_\nu \tag{8.4.19}$$

$$\geq H(DS|I\widehat{C}E)_\nu - (\alpha - 1)\, V(\nu_{\widehat{C}}, f_{\min}) - (\alpha - 1)^2\, K_\alpha(f_{\min}). \tag{8.4.20}$$

Next, we rewrite the von Neumann entropy by applying chain rules, see e.g. [50, Eq. 7.2.111]

$$H(DS|I\widehat{C}E)_\nu = H(S|I\widehat{C}E)_\nu + H(D|SI\widehat{C}E)_\nu = H(S|I\widehat{C}E)_\nu + H(D|\widehat{C})_\nu, \tag{8.4.21}$$

where the last equality holds due to [77, Eq. (152)].

Now, let us combine these intermediate results and we find with the constant $M \in \mathbb{R}$ as defined above

$$\begin{aligned} h_\alpha^{\mathrm{QKD}} \geq h_{\widehat{\alpha}} \geq \inf_{\mathbf{q}\in S_{\mathrm{acc}}} \inf_{\nu\in\Sigma} \big(& H(S|I\widehat{C}E)_\nu + H(D|\widehat{C})_\nu - M + \mathbf{f}_{\min}\cdot\mathbf{q} \\ & - (\alpha-1)\, V(\nu_{\widehat{C}}, f_{\min}) - (\alpha-1)^2\, K_\alpha(f_{\min})\big). \end{aligned} \tag{8.4.22}$$

Next, since $\widehat{C}$ is classical, we can write

$$H(D|\widehat{C})_\nu = \sum_{\hat{c}} \nu(\hat{c}) H(D)_{\nu|\hat{c}}, \tag{8.4.23}$$

and using $H(D)_{\nu|\hat{c}} \geq M - f_{\min}(\hat{c})$, we can simplify this to

$$H(D|\widehat{C})_\nu \geq M - \sum_{\hat{c}} f_{\min}(\hat{c})\nu(\hat{c}) = M - f_{\min}(\nu_{\widehat{C}}). \tag{8.4.24}$$

Finally, we have $\inf_{\mathbf{q}\in S_{\mathrm{acc}}} \mathbf{f}_{\min}\cdot\mathbf{q} \geq h_{\mathrm{vN}}$ by assumption and we reach

$$\begin{aligned} h_\alpha^{\mathrm{QKD}} \geq\ & h_{\mathrm{vN}} - (\alpha-1)^2\, K_\alpha(f_{\min}) \\ & + \inf_{\nu\in\Sigma} \big(H(S|I\widehat{C}E)_\nu + M - f_{\min}(\nu_{\widehat{C}}) - M - (\alpha-1)\, V(\nu_{\widehat{C}}, f_{\min})\big), \end{aligned} \tag{8.4.25}$$

which is independent of M as the terms cancel. Finally, after inserting the definition of T_α we reach the first inequality in Eq. (8.4.10). The second inequality follows by noting that $\widehat{\alpha} \geq \alpha$ for all $\alpha \geq 1$ and that each correction term is increasing in α.

Lastly, the statement regarding the conditional Rényi entropies follows from the inequality Eq. (8.4.10) immediately since $h_\alpha^{\uparrow,\mathrm{QKD}}$ already satisfied this property and thus, any lower bound on $h_\alpha^{\uparrow,\mathrm{QKD}}$ inherits it. □

Remark 8.4.3. Since it holds $h_\alpha^{\uparrow,\mathrm{QKD}} \geq h_\alpha^{\mathrm{QKD}}$, it also immediately follows that the GEAT key rates are a lower bound on $h_\alpha^{\uparrow,\mathrm{QKD}}$.

Using the theorem above, we can conclude the security against coherent attacks of the key rates derived form the GEAT formula in [2, Theorem 3] under the MEAT model.

Corollary 8.4.4 (GEAT Security)**.** *For any $\alpha \in (1,2)$, $\varepsilon_{\mathrm{PA}}, \varepsilon_{\mathrm{EV}} \in (0,1]$, the fixed-length version of Protocol 1 is $\varepsilon_{\mathrm{PA}}$-secret, $\varepsilon_{\mathrm{EV}}$-correct, and hence $(\varepsilon_{\mathrm{PA}} + \varepsilon_{\mathrm{EV}})$-secure, when the length l of the final key satisfies*

$$l \leq nh_{\mathrm{vN}} + nT_{\widehat{\alpha}}(f_{\min}) - n\left(\widehat{\alpha}-1\right)^2 K_{\widehat{\alpha}}(f_{\min}) - \lambda_{EC} - \left\lceil \log \frac{1}{\varepsilon_{\mathrm{EV}}} \right\rceil - \frac{\alpha}{\alpha-1} \log \frac{1}{\varepsilon_{\mathrm{PA}}} + 1, \tag{8.4.26}$$

where $\widehat{\alpha} = 1/(2-\alpha)$ and T_α, K_α are defined as in Eq. (8.4.12) of Theorem 8.4.2 and λ_{EC} is the length of the error-correction string.

Proof. The statement follows immediately from Eq. (8.4.10) together with Theorem 8.2.4. □

Chapter 9

Security of Prepare-and-Measure Protocols via GEAT

In this chapter we will present the results from our work in Ref. [2], mainly focusing on Sections VI and VII. Due to Theorem 8.4.2, the key rates from Ref. [2] are strict lower bounds on the full Rényi approach using the MEAT, which we will present in Chapter 10. Therefore, we omit any technical details and only showcase the results.

As we have seen in Corollary 8.4.4, the finite-size key length of protocol 1 can be calculated by the GEAT formula

$$l \leq nh_{\mathrm{vN}} + nT_{\widehat{\alpha}}(f_{\min}) - n\left(\widehat{\alpha}-1\right)^2 K_{\widehat{\alpha}}(f_{\min}) - \lambda_{\mathrm{EC}} - \left\lceil \log \frac{1}{\varepsilon_{\mathrm{EV}}} \right\rceil - \frac{\alpha}{\alpha-1}\log\frac{1}{\varepsilon_{\mathrm{PA}}} + 1, \tag{9.0.1}$$

and most importantly, each term in this expression only depends on a single round of the protocol.

To achieve the best key rates, one would like to maximize the key length with respect to the min-tradeoff function $f_{\min}$, i.e. taking the supremum over $f_{\min}$. However, the terms involving the min-tradeoff function are very involved and closed-form solution such as in Eq. (10.2.1) proved elusive. Thus, in [2], we developed a framework to find a min-tradeoff function that approximately maximizes the key length expression Eq. (9.0.1).

Additionally, we derived certain improvements to the decoy-state analysis in [2], but these techniques will also be explained in Section 10.7.

9.1 Qubit BB84 with Loss

In this section we consider a version of the BB84 protocol [6] using a perfect qubit source. We take the following model for the honest behaviour, described by two parameters: a depolarizing-noise parameter $p_{\text{hon}}^{\text{depol}} \in [0,1]$, and a loss-in-decibels parameter $\zeta_{\text{hon}} \in [0,\infty)$. Namely, after Alice uses the source to prepare perfect qubit states (in any specified basis), they are first subject to a depolarizing channel

$$\mathcal{E}_{\text{depol}}[\rho] \coloneqq (1-p_{\text{hon}}^{\text{depol}})\rho + p_{\text{hon}}^{\text{depol}}\frac{\mathbb{I}}{2}, \tag{9.1.1}$$

and then, with probability $1-10^{-\frac{\zeta_{\text{hon}}}{10\text{dB}}}$ the qubit is lost (in which case Bob always obtains a no-detection outcome), and otherwise Bob performs an ideal projective measurement on the qubit, in some specified basis.

For a conceptual overview of these protocol steps, we refer the reader to the detailed description provided in Protocol 1 and here we add some explicit details. In each round Alice decides with probability γ whether it is a test or generation round. If it is a generation round, Alice always sends one of the states $|0\rangle, |1\rangle$ in the Z-basis. Otherwise, she always sends $|+\rangle, |-\rangle$ in the X-basis. Alice records her choice of basis and signal state in a classical register X with alphabet $\{(\texttt{Z},0),(\texttt{Z},1),(\texttt{X},0),(\texttt{X},1)\}$. Upon arrival, Bob measures the incoming states with probability γ in the X-basis and with $1-\gamma$ in the Z-basis, recording the basis choice and outcome in a classical register Y with alphabet $\{(\texttt{Z},0),(\texttt{Z},1),(\texttt{Z},\texttt{no-det}),(\texttt{X},0),(\texttt{X},1),(\texttt{X},\texttt{no-det})\}$.

Then Alice announces a classical register C^A that is set to `gen` if it was a generation round, and otherwise set to the value of the X register. Bob then announces a classical register C^B as follows: if it was a generation round (i.e. $C^A = \texttt{gen}$) *and* he measured in the Z basis, then he sets C^B to either $(\texttt{Z},\texttt{det})$ or $(\texttt{Z},\texttt{no-det})$ depending on whether there was a detection. Otherwise he sets $C^B = Y$. These form all the single-round public announcements in this protocol, i.e. we have $I = (C^A, C^B)$.

First, in Fig. 9.1, we present the finite-size key rates for the qubit BB84 protocol described above plotted against loss in dB. In order to represent a realistic setup, we chose a depolarization of $p_{\text{hon}}^{\text{depol}} = 0.01$ for all signals and a security parameter of $\varepsilon_{\text{sec}} = 10^{-8}$. For simplicity we first only show the unique-acceptance scenario, i.e. a single-point acceptance set $S_{\text{acc}} = \{\mathbf{p}^{\text{hon}}\}$. The testing probability γ and the Rényi parameter have been optimized for each data point.

We compare our results with the key rates resulting from the unique-acceptance IID analysis as presented in Chapter 7. Our key rates compared to that scenario are worse,

but this is to be expected since here we cover a much more general attack, compared to the IID attack. We see that in the low-loss regime, we in fact obtain basically the same key rates as the IID scenario for all n values. However, our key rates drop off significantly faster as the loss increases.

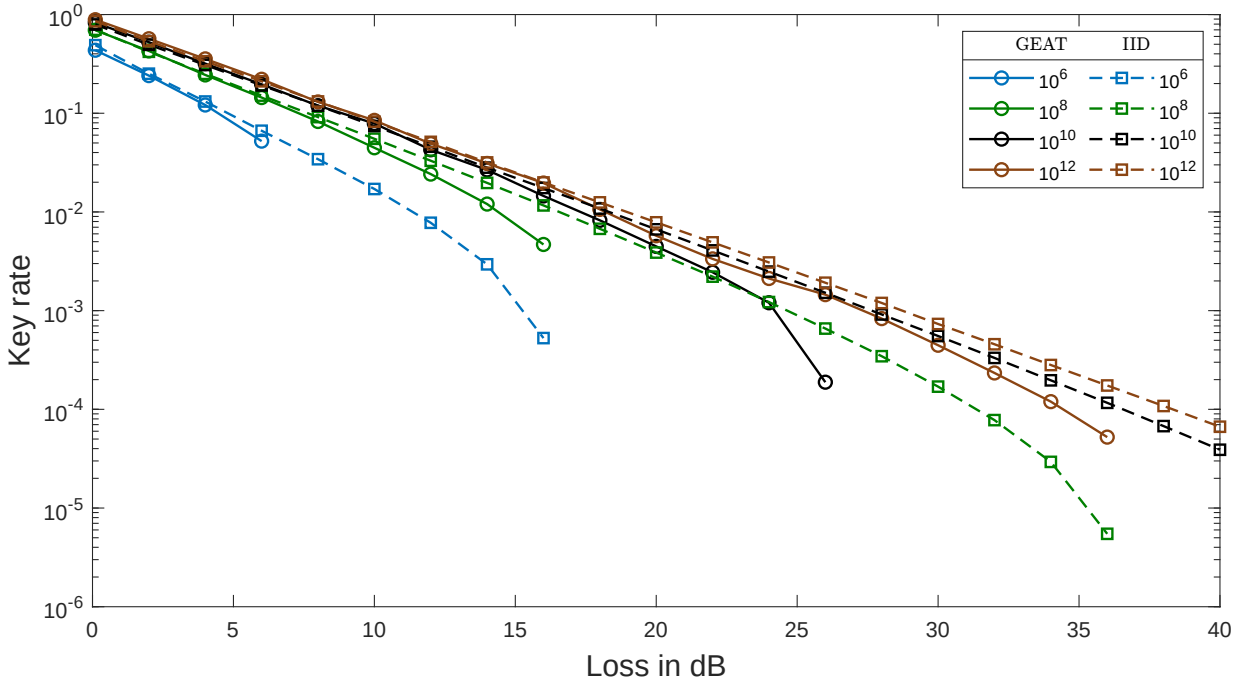


Figure 9.1: Key rate comparison for the qubit BB84 protocol with loss, between the GEAT analysis (solid lines), and the IID setting (dashed lines). In all cases, the signals are depolarized with $p_{\text{hon}}^{\text{depol}} = 0.01$, the security parameter is $\varepsilon_{\text{sec}} = 10^{-8}$, and for simplicity we show only the results for unique acceptance. The testing probability γ and the Rényi parameter α are optimized for the EAT analysis, and the former was also optimized in the IID analysis. (Figure 1 from Ref. [2])

Next, in Fig. 9.2, we show a comparison of our results for the realistic acceptance set with the key rates obtained from a EUR-based security proof, combining results from recent works [24, 83]. We find that the latter generally performs better than our approach for this protocol, though the difference becomes smaller at large n.

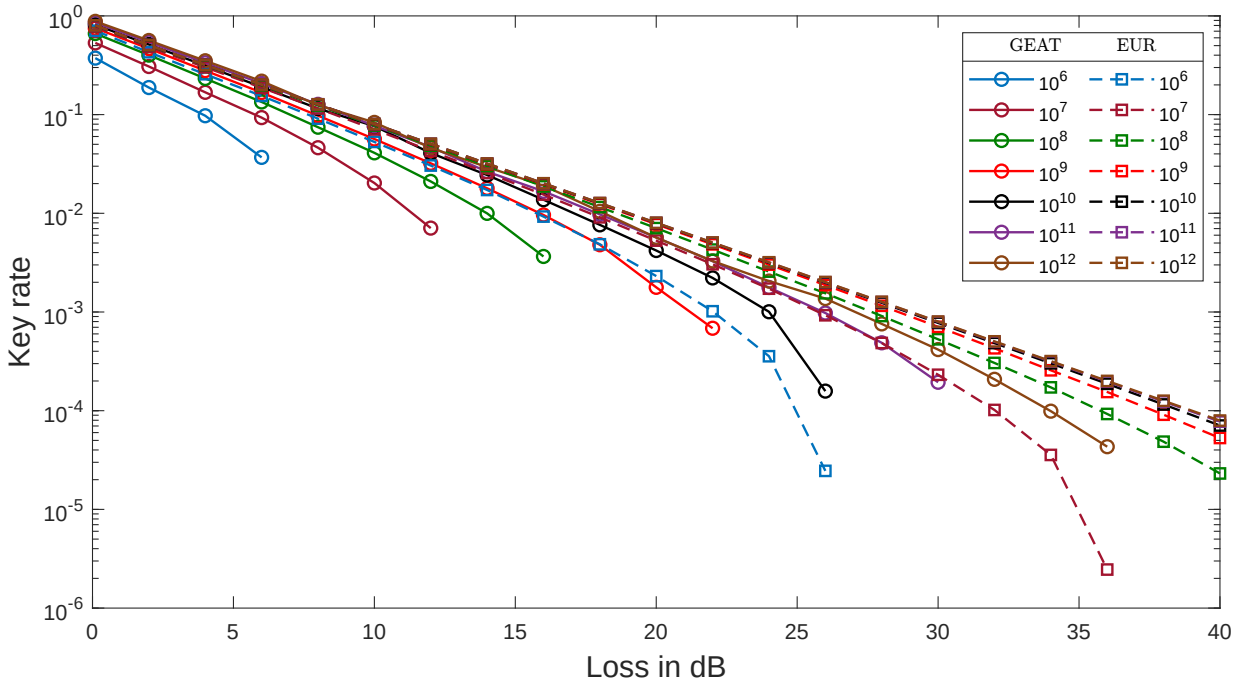


Figure 9.2: Key rate comparison for the qubit BB84 protocol with loss for a realistic acceptance set, between the GEAT analysis (solid lines), and the EUR-based analysis in [24] incorporating improvements from [83] (dashed lines). In all cases, the signals are depolarized with $p_{\text{hon}}^{\text{depol}} = 0.01$ and the security and completeness parameters are $\varepsilon_{\text{sec}} = 10^{-8}$ and $\varepsilon^{\text{com}} = 10^{-3}$, respectively. The testing probability γ and the Rényi parameter α are again optimized for the EAT analysis, and the former was also optimized in the EUR analysis. The dashed lines based on the EUR results from Ref. [24] show the expected key rates. (Bottom of Figure 2 from Ref. [2])

9.2 Decoy BB84 Protocol with Improved Analysis

In this section we present our numerical results for the decoy-state BB84 protocol using the GEAT analysis. For this decoy-state protocol, Alice can choose between fully phase-randomized weak coherent pulses (WCP) with intensities $\{\mu_{\text{sig}}, \mu_2, \mu_3\}$ to send her signals. The intensity μ_{sig} is the so-called signal intensity and will be used predominantly for key generation. Furthermore, we assume that the information is encoded in the polarization degree of freedom, i.e. Alice will send states from the set $\{\rho_H, \rho_V, \rho_D, \rho_A\}$, which are mixed because fully phase-randomized states are a classical mixture of photon-number states.

In each round Alice decides with probability γ whether it is a test or generation round.

If it is a test round, Alice now selects an intensity $\mu_i \in \{\mu_{\text{sig}}, \mu_2, \mu_3\}$ with some specified probability $p(\mu_i|t)$, and sends a uniformly random choice out of ρ_D, ρ_A in the X-basis with that intensity. Otherwise, she uses the signal intensity μ_{sig} to send a uniformly random choice out of the states ρ_H, ρ_V in the Z-basis. In either case, she uses a classical register X to record her choice of intensity, basis, and signal state.

We consider an active measurement setup where Bob has control over a polarization-rotator, which then determines the basis to be measured in. This choice is just out of convenience for our optimization problem, to keep the dimensions small. There are no inherent issues with using a passive detection setup, unlike methods relying on the entropic uncertainty relation, e.g. [22].

Applying the active detection setup, Bob still measures the incoming states with probability γ in the X-basis and with $1-\gamma$ in the Z-basis, recording his basis choice and outcome. Bob then applies the post-processing of [62, 63] for active BB84 detection setups to his outcomes. This converts his outcomes to an equivalent qubit detection scheme. Finally, Bob stores his post-processed outcomes in a classical register Y. Then, the remaining public announcements are analogous to the previous qubit BB84 protocol.

In Fig. 9.3 we show the secret key rates for the decoy-state BB84 protocol, plotted against loss in dB. We suppose that the honest implementation is subject to misalignment with an angle of $\theta_{\text{hon}}^{\text{misalign}} = \text{asin}(0.1) \approx 0.1002$ under the model described in [71]. We chose the intensities as $\{\mu_{\text{sig}} = 0.9, \mu_2 = 2 \times 10^{-2}, \mu_3 = 10^{-3}\}$ and the security parameter as $\varepsilon_{\text{sec}} = 10^{-8}$. For each data point we optimized the testing probability γ and the Rényi parameter.

As in the case of the qubit BB84 protocol, the key rates resulting from our proof technique are mostly lower than the IID key rates from Chapter 7 for the same protocol. Again, since we prove security against a wider class of attacks, one could expect such a behaviour. We see that to reach positive key rates for losses up to 25 dB, we require about 10^{12} signals sent.

Finally, in Fig. 9.4, we show a comparison of our results for the realistic acceptance set with the key rates obtained from a EUR-based security proof, again combining results from Refs. [24, 83]. For this protocol, we find that the former appears to perform better in low-loss regimes, while the latter performs better in high-loss regimes. However, as EUR-based proof techniques are quite different from the GEAT-based approach we used, it is difficult to draw broad conclusions from this about general protocols, and the question of which approach performs better for a given protocol would have to be addressed on a case-by-case basis.

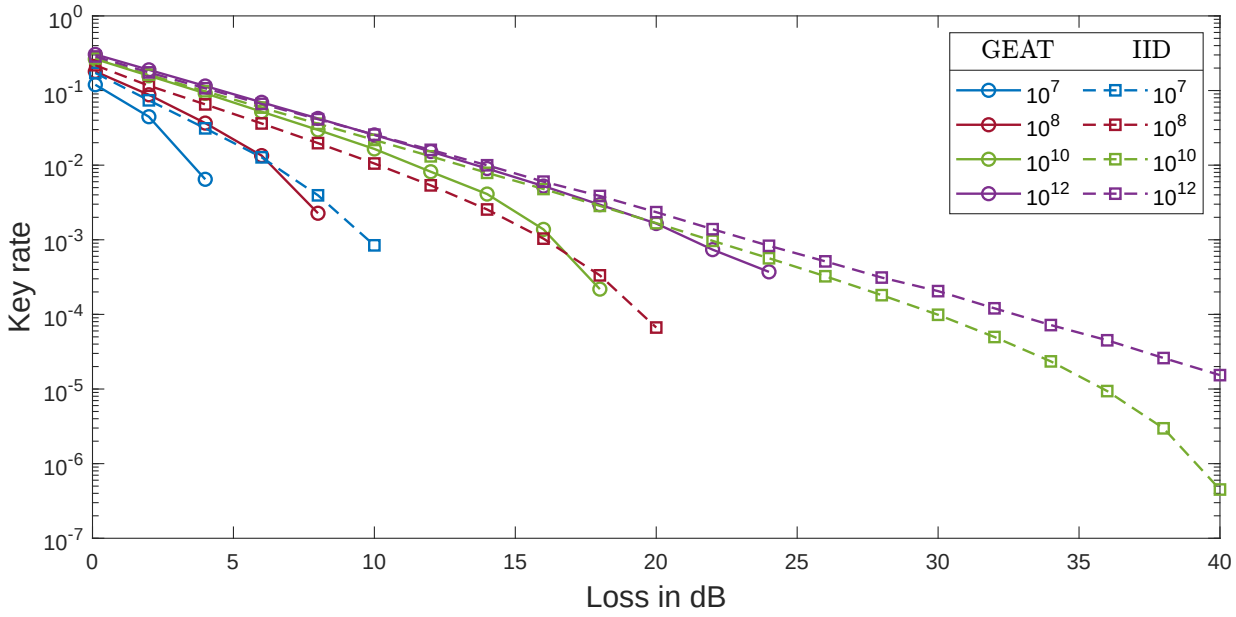


Figure 9.3: Key rate comparison for the decoy-state protocol, between the GEAT analysis in this work (solid lines), and the IID setting (dashed lines). In all cases, fully phase-randomized weak coherent pulses with intensities $\{\mu_{\text{sig}} = 0.9, \mu_2 = 2 \times 10^{-2}, \mu_3 = 10^{-3}\}$ were used. Signals are misaligned with $\theta_{\text{hon}}^{\text{misalign}} = \text{asin}(0.1)$, the photon number cutoff is $N_{\text{ph}} = 10$, the security parameter is $\varepsilon_{\text{sec}} = 10^{-8}$, and for simplicity we show only the results for unique acceptance. The testing probability γ and the Rényi parameter α are optimized for the EAT analysis, and the former was also optimized for the IID analysis. (Figure 3 from Ref. [2])

9.3 Conclusion

In this chapter, we presented the resulting finite-size key rates of the GEAT under the MEAT model. As such, the key rates presented here are valid against coherent attacks. Furthermore, we demonstrated that positive key rates can be achieved for reasonable numbers of signals sent and realistic protocols. Additionally, our results for the decoy-state BB84 protocol outperform the EUR-based results from Ref. [24] in the low-loss regime. This shows the potential of the methods developed here but also highlights that the key rates based on the GEAT formula Eq. (9.0.1) are significantly lower than what the MEAT framework could potentially achieve. An indicator that the MEAT key rates should actually be higher than those resulting from the EUR is the data presented in Ref. [77, Figure 1].

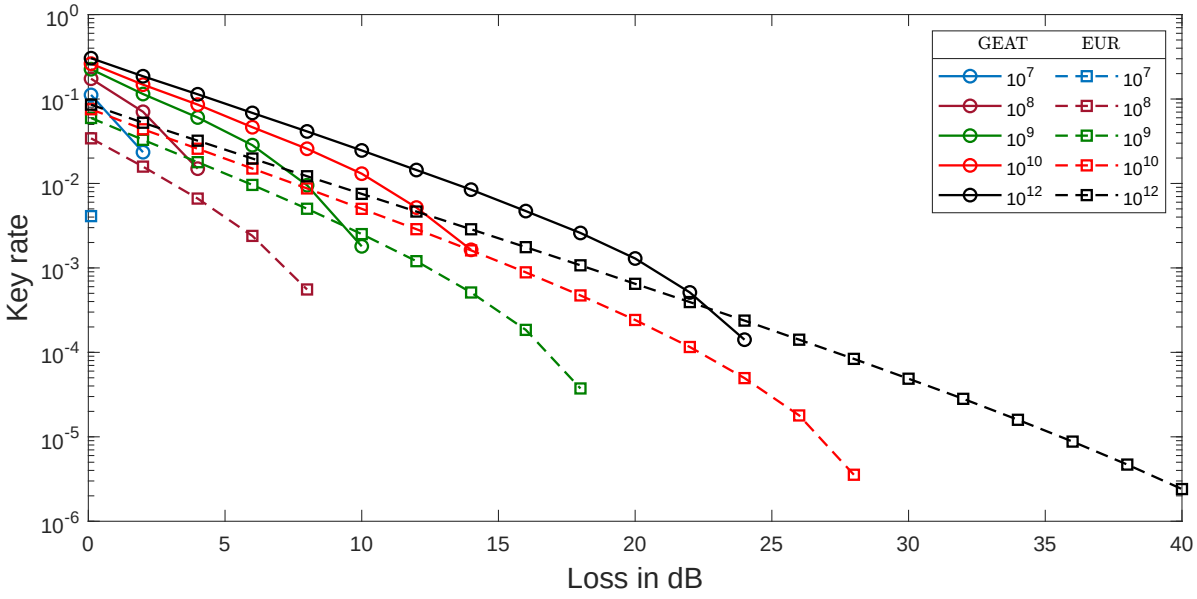


Figure 9.4: Key rate comparison for the decoy-state protocol for a realistic acceptance set, between the GEAT analysis (solid lines), and the EUR-based analysis in [24] incorporating improvements from [83] (dashed lines). In all cases, two decoy signals are used with intensities $\mu_2 = 2 \times 10^{-2}$ and $\mu_3 = 10^{-3}$, while the actual signal has the intensity of $\mu_{\text{sig}} = 0.9$, and the signals are chosen uniformly at random in each round. Signals are misaligned with $\theta_{\text{hon}}^{\text{misalign}} = \text{asin}(0.1)$, the photon number cutoff is $N_{\text{ph}} = 10$, the security parameter is chosen to be $\varepsilon_{\text{sec}} = 10^{-8}$, and the completeness parameter for the realistic tolerance intervals is $\varepsilon^{\text{com}} = 10^{-3}$. The testing probability γ and the Rényi parameter α are optimized for each data point. We find that for this specific protocol, our approach appears to perform better at low loss values, while the EUR-based approach appears to perform better at high loss values. The dashed lines based on the EUR results from Ref. [24] show the expected key rates. (Bottom of Figure 4 from Ref. [2])

However, especially in the high-loss regime, the results of this chapter still fall short of the EUR-based proofs from Ref. [24]. This is due to two main factors; the presence of larger correction terms, $T_{\widehat{\alpha}}(f_{\min})$ and $K_{\widehat{\alpha}}(f_{\min})$, and the sub-optimal choice of the min-tradeoff function produced by the algorithm in Ref. [2]. Together, these lead to lower key rates, an effect that becomes particularly pronounced at high loss.

Chapter 10

Rényi Security Framework

In this chapter, we will apply the fixed-length security proof for QKD protocols derived under the MEAT model in Theorem 8.2.4 in Section 8.2 and derive a numerical method to evaluate fixed-length key rates. Additionally, we will extend these numerical methods to the variable-length framework of Section 8.3 especially Theorem 8.3.3 and use them to calculate variable-length key rates for several examples such as a qubit and decoy-state BB84 protocol. We finish this section by generalizing the framework to QKD protocols with independent source imperfections, exemplified by fluctuating intensities and faulty phase-randomization. Crucially, we analyze the combination of realistic detection setups and source imperfections, which addresses the specific requirements highlighted by recent experimental collaborations, including the Reference Frame Independent Quantum Key Distribution (ReFQ) project, which is part of the Quantum EncrYption and Science Satellite (QEYSSat) mission [39]. The content of this chapter is based on our work in [4] excluding Sections I–III.

10.1 Calculating Versions of $h_\alpha^{\uparrow,\mathrm{QKD}}$

In general, numerical minimizations of Rényi entropies or divergences for generic values of α can be a challenging task. For example, for $\alpha = 1/2$, the sandwiched divergence can be reformulated as a semidefinite program; however, for $\alpha > 1/2$ a similar reformulation remains unknown [84, Problem 1]. However, there has been a considerable effort for interior point solvers of sandwiched Rényi divergences, see e.g. [85].

In this chapter, we address this by following a similar approach as in Ref. [67], where numerical methods for calculating asymptotic QKD key rates have been presented. Their

numerical approach relied on the reformulation of the required conditional von Neumann entropy in terms of the relative entropy presented in [86, Theorem 1], which then was minimized with the Frank-Wolfe algorithm [53] or see Algorithm 1, an algorithm for constrained convex optimization problems.

The reformulation was required because the conditional von Neumann entropy $H(S|E)$ is concave with respect to the state on SE, but when evaluated in terms of a purification ρ_{ABE} of the pre-measurement state ρ_{AB}, it is convex with respect to ρ_{AB}. This allows one to address it using convex optimization methods.

Motivated by the same reasons, we derive a similar reformulation for the Rényi entropy framework. To address the computational complexity associated with evaluating different Rényi entropies, in this section, we outline the path to a computable lower bound for $h_\alpha^{\uparrow,\mathrm{QKD}}$, which is formally derived in Theorem 8.2.3.

We start by proving a similar duality relation as in [86, Theorem 1], but for Rényi entropies. In [87, Proposition 17] a special case of the relations in the following lemma was found. Somewhat different formulations for $\widetilde{H}_\alpha^\uparrow$ and $\widetilde{H}_\alpha^\downarrow$ were also obtained in [88, Lemma A.2], but it appears unclear whether those formulas have the required convexity properties for our numerical methods. Finally, recent and concurrent work [89] implemented Frank-Wolfe methods to minimize some forms of sandwiched Rényi divergence; in particular, they used this to compute lower bounds on $\widetilde{H}_\alpha^\uparrow$ in terms of $\overline{H}_\alpha^\uparrow$ by deriving a relation similar to those we present here.

Theorem 10.1.1. *Let $\rho_{QE} \in \mathrm{S}_=(QE)$ be pure and let $\{Z_j\}_{1\dots d_S}$ be a set of orthogonal projections on Q such that $\sum_{j=1}^{d_S} Z_j = \mathbb{I}_Q$. Furthermore, define the isometry*

$$V_{\mathcal{Z}} \coloneqq \sum_j |j\rangle_S \otimes Z_j, \tag{10.1.1}$$

and the state σ_{SQE} after applying the isometry $V_{\mathcal{Z}}$ onto ρ_{QE} as

$$\sigma_{SQE} \coloneqq V_{\mathcal{Z}} \rho_{QE} V_{\mathcal{Z}}^\dagger. \tag{10.1.2}$$

Then, it holds

$$\widetilde{H}_\alpha^\downarrow(S|E)_{\sigma_{SE}} = \frac{1}{1-\alpha} \log \mathrm{Tr}\left[\left(\mathcal{Z}\left(\rho_Q^{\frac{1}{\alpha}}\right)\right)^\alpha\right], \tag{10.1.3}$$

$$\overline{H}_\alpha^\downarrow(S|E)_{\sigma_{SE}} = \overline{D}_{2-\alpha}\left(\rho_Q || \mathcal{Z}(\rho_Q)\right) \tag{10.1.4}$$

where

$$\mathcal{Z}(\tau) \coloneqq \sum_{j=1}^{d_S} Z_j \tau Z_j^\dagger, \tag{10.1.5}$$

for all $\tau \in \mathrm{S}_=(Q)$.

Proof. We first note that $\sigma_Q = \mathcal{Z}(\rho_Q)$ since

$$\sigma_Q = \mathrm{Tr}_{SE}\left[\sigma_{SQE}\right] = \mathrm{Tr}_{SE}\left[V_{\mathcal{Z}} \rho_{QE} V_{\mathcal{Z}}^\dagger\right] = \mathcal{Z}(\rho_Q). \tag{10.1.6}$$

Furthermore, it holds

$$\mathrm{Tr}_S\left[\sigma_{SQ}^\beta\right] = \mathrm{Tr}_S\left[V_{\mathcal{Z}} \rho_Q^\beta V_{\mathcal{Z}}^\dagger\right] = \mathcal{Z}\left(\rho_Q^\beta\right) \ \forall \beta \geq 0. \tag{10.1.7}$$

Using both properties, one can prove both theorem statements. Due to [82, Theorem 2], it holds,

$$\widetilde{H}_\alpha^\downarrow(S|E)_{\sigma_{SE}} = -\overline{H}_{\frac{1}{\alpha}}^\uparrow(S|Q)_{\sigma_{SQ}}. \tag{10.1.8}$$

Furthermore, we can use [82, Lemma 1] to rewrite the Petz Rényi entropy $\overline{H}_{\frac{1}{\alpha}}^\uparrow(S|Q)_{\sigma_{SQ}}$, which yields

$$\widetilde{H}_\alpha^\downarrow(S|E)_{\sigma_{SE}} = \frac{1}{1-\alpha} \log \mathrm{Tr}\left[\left(\mathrm{Tr}_S\left(\sigma_{SQ}^{\frac{1}{\alpha}}\right)\right)^\alpha\right], \tag{10.1.9}$$

and inserting Eq. (10.1.7) already gives Eq. (10.1.3).

Similarly, by applying the duality relation for $\overline{H}_\alpha^\downarrow(S|E)_{\sigma_{SE}}$ presented in [90, Lemma 6], we obtain

$$\overline{H}_\alpha^\downarrow(S|E)_{\sigma_{SE}} = -\overline{H}_{2-\alpha}^\downarrow(S|Q)_{\rho_{SQ}} = \overline{D}_{2-\alpha}\left(\rho_{SQ} || \mathbb{I}_S \otimes \rho_Q\right), \tag{10.1.10}$$

after which Eq. (10.1.4) follows by further simplifying $\overline{D}_{2-\alpha}\left(\rho_{SQ} || \mathbb{I}_S \otimes \rho_Q\right)$ using the definition of $\overline{D}_\alpha$ and the properties in Eqs. (10.1.6) and (10.1.7). □

For this work, we choose to focus only on the $\widetilde{H}_\alpha^\downarrow$ case presented in the above lemma, because it provides a reasonable balance between tightness of the bounds and ease of numerical work. In Fig. 10.1, we represent the relationships between the different definitions of Rényi entropies of Definition 2.3.11. Specifically, as indicated in this figure, among the

conditional entropies in Definition 2.3.11, $\overline{H}_\alpha^\downarrow$ is smaller than all the others [47, Fig. 5.1] and hence we choose not to use it. Out of the remaining options, the expressions we obtain for $\widetilde{H}_\alpha^\uparrow$ and $\overline{H}_\alpha^\uparrow$ involve the sandwiched divergence, while the expression we obtain for $\widetilde{H}_\alpha^\downarrow$ involves the Petz divergence. The sandwiched divergences have more complicated gradient expressions (required to implement the Frank-Wolfe algorithm or an interior point method) than the Petz divergences, and hence we choose to focus on $\widetilde{H}_\alpha^\downarrow$.

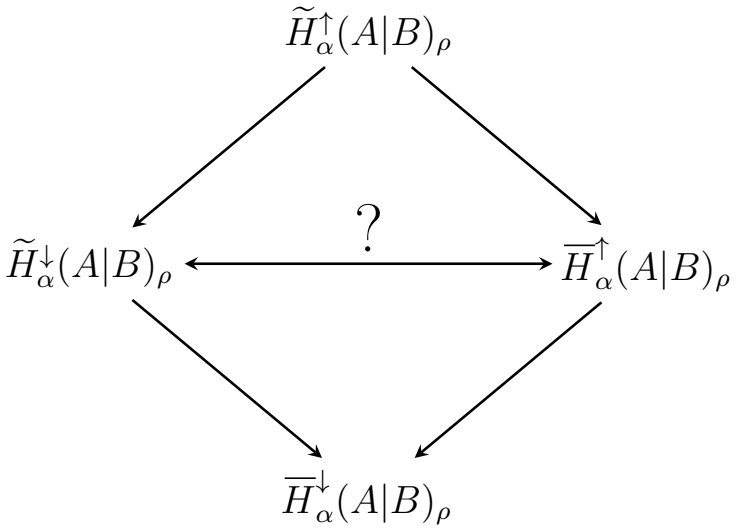


Figure 10.1: Schematic representation of the Rényi entropy hierarchy. An arrow indicates that the source entropy upper-bounds the target. The question mark signifies an indeterminate ordering.

Most importantly, in contrast to the work in [89], we note that in general it is unknown which of $\widetilde{H}_\alpha^\downarrow$ or $\overline{H}_\alpha^\uparrow$ provides tighter bounds. One only knows that $\widetilde{H}_\alpha^\uparrow$ would in principle give the tightest possible bounds. Additionally, for our later approaches involving variable-length key rates we require a formalism to find rates for either $\widetilde{H}_\alpha^\uparrow$ or $\widetilde{H}_\alpha^\downarrow$. Again, the above result suffices to lower bound both, simply because $\widetilde{H}_\alpha^\uparrow$ is lower bounded by $\widetilde{H}_\alpha^\downarrow$, which is covered by the above result.

In any case, for typical values of α used in this work (which approach 1 as n increases), all of the entropies in Definition 2.3.11 would only differ by small amounts, due to explicit converse bounds derived in Ref. [82, Corollary 4] that constrain how much they can differ. With this in mind, we will lower bound the quantity $h_\alpha^{\uparrow,\mathrm{QKD}}$ by replacing the Rényi entropy $\widetilde{H}_\alpha^\uparrow$ with its counterpart $\widetilde{H}_\alpha^\downarrow$ (expressed in the form in Theorem 10.1.1), with the under-

standing that this will make little difference whenever α is close to 1. We note also that it has been empirically observed in Ref. [88] that in the context of randomness generation, this relaxation only makes a difference for very small sample sizes below $n = 10^4$ signals.

Therefore, we formulate Corollary 10.1.2 in terms of the sandwiched Rényi entropy $\widetilde{H}^{\downarrow}_{\alpha}$, which proves security of a generic P&M QKD protocol against coherent attacks.

Corollary 10.1.2 (Fixed-Length security)**.** *For any $\alpha \in (1,2)$, $\varepsilon_{\mathrm{PA}}, \varepsilon_{\mathrm{EV}} \in (0,1]$, the fixed-length version of Protocol 1 is $\varepsilon_{\mathrm{PA}}$-secret, $\varepsilon_{\mathrm{EV}}$-correct, and hence $(\varepsilon_{\mathrm{PA}} + \varepsilon_{\mathrm{EV}})$-secure, when the length l of the final key satisfies*

$$l \leq n h_{\alpha}^{\mathrm{QKD}} - \lambda_{EC} - \left\lceil \log \frac{1}{\varepsilon_{\mathrm{EV}}} \right\rceil - \frac{\alpha}{\alpha - 1} \log \frac{1}{\varepsilon_{\mathrm{PA}}} + 1, \tag{10.1.11}$$

where λ_{EC} is the length of the error-correction string and $h_{\alpha}^{\mathrm{QKD}}$ is defined by

$$h_{\alpha}^{\mathrm{QKD}} := \inf_{\mathbf{q} \in S_{\mathrm{acc}}} \inf_{\substack{\rho \in \mathrm{S}_{=}(AB), \\ s.t.\ \mathrm{Tr}_B[\rho] = \tau_A}} \left(\frac{\alpha}{\alpha - 1} D\left(\mathbf{q} \middle\| \boldsymbol{\nu}^{\rho}_{\widehat{C}}\right) + q(\mathtt{gen}) \widetilde{H}^{\downarrow}_{\alpha}(S|\tilde{T}E)_{\nu^{\rho}_{|\mathtt{gen}}} \right), \tag{10.1.12}$$

where

$$\nu^{\rho}_{S\tilde{T}\widehat{C}E} = \mathcal{M}_{QKD}(\rho_{ABE}), \tag{10.1.13}$$

$$\boldsymbol{\nu}^{\rho}_{\widehat{C}} = (\gamma \boldsymbol{\nu}^{\rho}_{\widehat{C}|\mathtt{test}}, 1 - \gamma)^T, \tag{10.1.14}$$

$$\boldsymbol{\nu}^{\rho}_{\widehat{C}|\mathtt{test}} = \boldsymbol{\Phi}[\rho_{AB|\mathtt{test}}]. \tag{10.1.15}$$

The Rényi entropy $\widetilde{H}^{\downarrow}_{\alpha}(S|\tilde{T}E)_{\nu^{\rho}_{|\mathtt{gen}}}$ can further be simplified to

$$\widetilde{H}^{\downarrow}_{\alpha}(S|\tilde{T}E)_{\nu_{|\mathtt{gen}}} = \frac{1}{1 - \alpha} \log \mathrm{Tr}\left[\left(\hat{\mathcal{Z}} \left(\hat{\mathcal{G}}(\rho_{AB|\mathtt{gen}})^{\frac{1}{\alpha}} \right) \right)^{\alpha} + 1 - \mathrm{Tr}\left[\hat{\mathcal{G}}(\rho_{AB|\mathtt{gen}}) \right] \right], \tag{10.1.16}$$

where $\hat{\mathcal{G}}$ is the CPTNI map defined in Definition 8.2.1 omitting the discard symbol. The CPTNI map $\hat{\mathcal{Z}}$ similarly omits the discard symbol.

Before we state the proof, we note that in [77, Theorem 5.1] it was shown that the optimization problem for $h_{\alpha}^{\mathrm{QKD}}$ is convex in its arguments. Hence, it could be evaluated with convex optimization methods, as claimed earlier.

Proof. Correctness follows immediately from Theorem 8.2.4 and secrecy also follows from Theorem 8.2.4 if $h_\alpha^{\mathrm{QKD}} \leq h_\alpha^{\uparrow,\mathrm{QKD}}$. This is true since $\widetilde{H}_\alpha^{\downarrow}(S|IE)_{\nu_{|\mathtt{gen}}^{\rho}} \leq \widetilde{H}_\alpha^{\uparrow}(S|IE)_{\nu_{|\mathtt{gen}}^{\rho}}$ for any ν and thus the infimum in h_α^{QKD} is also smaller.

Hence, it only remains to show that $\widetilde{H}_\alpha^{\downarrow}(S|\tilde{T}E)_{\nu_{|\mathtt{gen}}^{\rho}}$ can be simplified to the expression stated in the corollary. Applying the definition of $\mathcal{M}_{\mathrm{QKD}}$ in Definition 8.2.1, yields that $\nu_{|\mathtt{gen}}$ is given by

$$\nu_{S\tilde{T}ES_QIABXY|\mathtt{gen}} = V_{\mathcal{Z}} V_{\mathcal{G}} \rho_{ABE|\mathtt{gen}} V_{\mathcal{G}}^\dagger V_{\mathcal{Z}}^\dagger. \tag{10.1.17}$$

Now, by identifying $\tilde{T}E \mapsto E$ and $S_QIABXY \mapsto Q$ in Eq. (10.1.3) of Theorem 10.1.1, one finds

$$\widetilde{H}_\alpha^{\downarrow}(S|\tilde{T}E)_{\nu_{S\tilde{T}E|\mathtt{gen}}} = \frac{1}{1-\alpha} \log \mathrm{Tr}\left[\left(\mathcal{Z}\left(\mathrm{Tr}_{\tilde{T}}\left[V_{\mathcal{G}}\rho_{AB|\mathtt{gen}}V_{\mathcal{G}}^\dagger\right]^{\frac{1}{\alpha}}\right)\right)^{\alpha}\right]. \tag{10.1.18}$$

Finally, we can simplify the trace over $\tilde{T}$ to find

$$\mathrm{Tr}_{\tilde{T}}\left[V_{\mathcal{G}}\rho_{AB|\mathtt{gen}}V_{\mathcal{G}}^\dagger\right] = \mathcal{G}(\rho_{AB|\mathtt{gen}}), \tag{10.1.19}$$

which leaves us with

$$\widetilde{H}_\alpha^{\downarrow}(S|\tilde{T}E)_{\nu^{\rho}_{S\tilde{T}E|\mathtt{gen}}} = \frac{1}{1-\alpha} \log \mathrm{Tr}\left[\left(\mathcal{Z}\left(\mathcal{G}(\rho_{AB|\mathtt{gen}})^{\frac{1}{\alpha}}\right)\right)^{\alpha}\right]. \tag{10.1.20}$$

By inserting the definitions of $\mathcal{Z}$ and $\mathcal{G}$

$$\mathcal{Z}(\rho) = \hat{\mathcal{Z}}(\rho) + |\perp\rangle\langle\perp|\,\rho\,|\perp\rangle\langle\perp|\,, \tag{10.1.21}$$

$$\mathcal{G}(\rho) = \hat{\mathcal{G}}(\rho) + \left(1 - \mathrm{Tr}\left(\hat{\mathcal{G}}(\rho)\right)\right)|\perp\rangle\langle\perp|\,, \tag{10.1.22}$$

and using the orthogonality of $|\perp\rangle\langle\perp|$, the corollary statement follows. □

We will spend the remainder of this section to simplify h_α^{QKD} even further and bring it into a form that should be more numerically stable.

Hence, let us consider the following situation of small testing fractions. In general, the components of the state ρ_{AB} resulting in test constraints in h_α^{QKD} will be proportional to the testing fraction γ. However, due to this proportionality, when choosing a sufficiently small testing fraction, these constraints can quickly be obscured by numerical noise.

To address this, we first make the trivial observation that we can always "undo" the source replacement argument on single rounds, to say that the set of all states $\mathcal{M}_{\mathrm{QKD}}(\rho_{ABE})$

satisfying $\rho_A = \tau_A$ can equivalently be generated by Alice preparing the classical-quantum state $\sigma_{XA'}$ in the original P&M description, Eve applying some arbitrary attack channel $A' \to BE$, and Bob measuring B, followed by the relevant processing of Alice and Bob's classical values. Furthermore, the simplification in Eq. (10.1.16) tells us that in fact we do not need to consider Eve's full attack channel $A' \to BE$, but rather only its restriction $\mathcal{E} : A' \to B$ obtained by tracing out E from the output.

Now, the state $\sigma_{XA'}$ is a mixture of the test and generation cases according to the protocol structure. However, the critical observation is that the channel $\mathcal{E}$ must still act in the same way on both the state conditioned on a test round and conditioned on a generation round. Numerically, it would be much more desirable to optimize over the channel instead of the state ρ_{AB}. When γ is small, the state ρ_{AB} contains components scaled by γ, which can lead to numerical instability in the solver. By reparameterizing the problem in terms of Eve's channel we avoid these scaling issues. Hence, similar to the approach in Ref. [2], we will now rephrase the optimization problem of h_α^{QKD} in terms of the Choi state J of Eve's channel $\mathcal{E}$ creating the states conditioned on test and generation rounds.

To define the optimization problem in terms of Eve's channel, we now apply the source replacement argument again, but *separately* on the test and generation components. That is to say, we view Alice sending some signal state to Bob and storing her choice as instead preparing some pure states $|\xi^t\rangle_{AA'}$ and $|\xi^g\rangle_{AA'}$ with probabilities γ and $1-\gamma$ respectively, and then measuring A with some suitable measurement (which can differ in the two cases). To describe this, we define these pure states:

$$|\xi^t\rangle_{AA'} := \sqrt{\Pi^{\mathtt{test}}} \frac{|\psi\rangle_{AA'}}{\sqrt{\gamma}}, \tag{10.1.23}$$

$$|\xi^g\rangle_{AA'} := \sqrt{\Pi^{\mathtt{gen}}} \frac{|\psi\rangle_{AA'}}{\sqrt{1-\gamma}}. \tag{10.1.24}$$

as the states Alice prepares conditioned on test and generation rounds, respectively. Then we can write

$$\begin{aligned} \rho_{AB|\mathtt{gen}} &= (\mathbb{I}_A \otimes \mathcal{E}) \left(|\xi^g\rangle\langle\xi^g|_{AA'}\right), \\ \rho_{AB|\mathtt{test}} &= (\mathbb{I}_A \otimes \mathcal{E}) \left(|\xi^t\rangle\langle\xi^t|_{AA'}\right), \end{aligned} \tag{10.1.25}$$

where again $\mathcal{E}$ defines Eve's channel. Thus, we could immediately write $\rho_{AB|\mathtt{gen}}$ required for Eq. (10.1.16) in terms of Eve's channel. Moreover, $\boldsymbol{\nu}^{\rho}_{\widehat{C}|\mathtt{test}} = \boldsymbol{\Phi}[\rho_{AB|\mathtt{test}}]$ could also straightforwardly be written as Eve's channel acting on the state $|\xi^t\rangle\langle\xi^t|_{AA'}$.

Next, we rephrase Eve's channel in terms of its Choi state. This allows for a treatment with convex optimization methods because Choi states of CPTP maps need to be positive semidefinite and satisfy a partial trace constraint.

To simplify this formulation, we define the CPTNI maps $\chi_g : A' \to A$ and $\chi_t : A' \to A$ as

$$\begin{aligned} \chi_g[X] &\coloneqq \mathrm{Tr}_{A'}\left[\left(\mathbb{I}_A \otimes X^{T_{A'}}\right)|\xi^g\rangle\langle\xi^g|\right], \\ \chi_t[X] &\coloneqq \mathrm{Tr}_{A'}\left[\left(\mathbb{I}_A \otimes X^{T_{A'}}\right)|\xi^t\rangle\langle\xi^t|\right], \end{aligned} \tag{10.1.26}$$

which let us write

$$\rho_{AB|\mathtt{gen}/\mathtt{test}} = \chi_{g/t}(J), \tag{10.1.27}$$

i.e. a linear map acting directly on the Choi state. Additionally, to achieve the same for the statistics in test rounds, let us define the map $\boldsymbol{\Psi}$ generating the statistics in test rounds directly from a Choi state by

$$\boldsymbol{\Psi}[J] \coloneqq \boldsymbol{\Phi} \circ \chi_t[J]. \tag{10.1.28}$$

Finally, for a condensed notation of the Rényi entropy evaluated on the state $\nu_{|\mathtt{gen}}$, let us define the function $g_\alpha : \mathrm{Pos}(A'B) \to \mathbb{R}_{\geq 0}$ as

$$g_\alpha(J) \coloneqq \widetilde{H}^{\downarrow}_\alpha(S|\tilde{T}E)_{\mathcal{M}_{\mathrm{QKD}}\circ\chi_g(J)}, \tag{10.1.29}$$

for all Choi states $J \in \mathrm{Pos}(A'B)$, such that $\mathrm{Tr}_B[J] = \mathbb{I}_{A'}$.

Combining all these definitions we can recast the optimization problem of h^{QKD}_α, originally defined in terms of the state ρ_{AB}, as an optimization problem over the Choi state J of Eve's channel. The resulting optimization problem is given by

$$\begin{aligned} h^{\mathrm{QKD}}_\alpha = \inf_{\substack{\mathbf{q}\in S_{\mathrm{acc}},\\ J\in\mathrm{Pos}(A'B)}} & \left(\frac{\alpha D\left(\mathbf{q}\middle\|\boldsymbol{\nu}_{\widehat{C}}\right)}{\alpha-1} + q(\mathtt{gen})g_\alpha(J)\right) \\ \text{s.t. } & \mathrm{Tr}_B[J] = \mathbb{I}_{A'}, \\ & \boldsymbol{\nu}_{\widehat{C}} = (\gamma\boldsymbol{\nu}_{\widehat{C}|\mathtt{test}}, 1-\gamma)^T, \\ & \boldsymbol{\nu}_{\widehat{C}|\mathtt{test}} = \boldsymbol{\Psi}[J]. \end{aligned} \tag{10.1.30}$$

In Corollary A.1.9, in the appendix, we show that this reformulation of h^{QKD}_α is indeed convex. Therefore, similar to the algorithm of Ref. [67], we employ a Frank-Wolfe method [53] to find the minimum. Hence, we require the gradient of h^{QKD}_α with respect to $\mathbf{q}$ and J. Again similar to the case of the relative entropy in [67], we need some form of perturbation for the gradient to exist everywhere on its domain. We leave the details of the numerical algorithm out for brevity. The details can be found in Appendix G of Ref. [4].

10.2 Computing the Normalization Constant κ

Now, we turn our attention to the $\widetilde{H}_\alpha^{\uparrow,f}$-normalization constant $\kappa^\uparrow$ required for variable-length key rates as derived in Theorem 8.3.3. Here, given a choice of f, we need to find a lower bound $\underline{\kappa}^\uparrow$ on the $\widetilde{H}_\alpha^{\uparrow,f}$-normalization constant $\kappa^\uparrow$.

As we explained in the previous section, to avoid minimizing the optimized conditional Rényi entropy $\widetilde{H}_\alpha^\uparrow$, we replaced it with its unoptimized counterpart $\widetilde{H}_\alpha^\downarrow$ in Corollary 10.1.2. For the same reasons we will make the same replacement in Eq. (8.3.33) which gives the following optimization problem for finding a tradeoff function

$$\inf_{\substack{\rho\in \mathrm{S}_=(AB)\\ \text{s.t. } \mathrm{Tr}_B[\rho]=\tau_A}} \inf_{\boldsymbol{\lambda}\in\mathbb{P}(\widehat{\mathcal{C}})} \left(\frac{\alpha}{\alpha-1} D\left(\boldsymbol{\lambda}\middle\|\boldsymbol{\nu}^\rho_{\widehat{C}}\right) + q(\mathtt{gen})\widetilde{H}_\alpha^\downarrow(S|\tilde{T}E)_{\nu^\rho_{|\mathtt{gen}}} \right) \tag{10.2.1}$$

$$\text{s.t. } \boldsymbol{\nu}^\rho_{\widehat{C}} = (\gamma \boldsymbol{\nu}^\rho_{\widehat{C}|\mathtt{test}}, 1-\gamma)^T, \tag{10.2.2}$$

$$\boldsymbol{\nu}^\rho_{\widehat{C}|\mathtt{test}} = \boldsymbol{\Phi}[\rho_{AB|\mathtt{test}}],$$

$$\mathbf{q}^{\mathrm{hon}} - \boldsymbol{\lambda} = 0. \tag{solve for dual}$$

The components $f(\hat{c})$ of the tradeoff function are then given by the dual variables of the constraint Eq. (solve for dual).

In line with the reformulations of the previous section for h_α^{QKD}, one can equivalently write this optimization problem in terms of Eve's channel as

$$\inf_{\substack{\mathbf{q}\in S_{\mathrm{acc}},\\ J\in \mathrm{Pos}(A'B)}} \inf_{\boldsymbol{\lambda}\in\mathbb{P}(\widehat{\mathcal{C}})} \left(\frac{\alpha}{\alpha-1} D\left(\boldsymbol{\lambda}\middle\|\boldsymbol{\nu}^\rho_{\widehat{C}}\right) + q(\mathtt{gen}) g_\alpha(J) \right) \tag{10.2.3}$$

$$\text{s.t. } \mathrm{Tr}_B[J] = \mathbb{I}_{A'},$$

$$\boldsymbol{\nu}^\rho_{\widehat{C}} = (\gamma \boldsymbol{\nu}^\rho_{\widehat{C}|\mathtt{test}}, 1-\gamma)^T,$$

$$\boldsymbol{\nu}^\rho_{\widehat{C}|\mathtt{test}} = \boldsymbol{\Psi}[\rho_{AB|\mathtt{test}}],$$

$$\mathbf{q}^{\mathrm{hon}} - \boldsymbol{\lambda} = 0. \tag{solve for dual}$$

However, the tradeoff function resulting from either version of the optimization problem in Eq. (10.2.1) or Eq. (10.2.3), will be slightly suboptimal if we evaluate either the $\widetilde{H}_\alpha^{\downarrow,f}$-normalization constant or the $\widetilde{H}_\alpha^{\uparrow,f}$-normalization constant. That is because the correct dual of the optimization problem above will contain a different, mixed f-weighted entropy, in between $\widetilde{H}_\alpha^{\uparrow,f}$ and $\widetilde{H}_\alpha^{\downarrow,f}$, which we will define below.

Definition 10.2.1 (Mixed f-weighted Rényi entropies)**.** Let $\rho \in \mathrm{S}_=(\widehat{C}QQ')$ be a state where $\widehat{C}$ is classical with alphabet $\widehat{\mathcal{C}}$. A *tradeoff function on* $\widehat{C}$ is a function $f : \widehat{\mathcal{C}} \to \mathbb{R}$, equivalently denoted by a real-valued tuple $\mathbf{f} \in \mathbb{R}^{|\widehat{\mathcal{C}}|}$ where each term in the tuple specifies the value $f(\hat{c})$. Given a tradeoff function f and a value $\alpha \in (0,1) \cup (1,\infty)$, we define two versions of a *mixed f-weighted Rényi entropy* of order α for ρ, as follows:

$$\widetilde{H}_\alpha^{\wr,f}(Q|\widehat{C}Q')_\rho := \frac{\alpha}{1-\alpha} \log \left(\sum_{\hat{c}} \rho(\hat{c})\, 2^{\frac{1-\alpha}{\alpha}\left(-f(\hat{c})+\widetilde{H}_\alpha^{\downarrow}(Q|Q')_{\rho_{|\hat{c}}}\right)} \right), \tag{10.2.4}$$

$$\widetilde{H}_\alpha^{\wr,f}(Q|\widehat{C}Q')_\rho := \frac{1}{1-\alpha} \log \left(\sum_{\hat{c}} \rho(\hat{c})\, 2^{(1-\alpha)\left(-f(\hat{c})+\widetilde{H}_\alpha^{\uparrow}(Q|Q')_{\rho_{|\hat{c}}}\right)} \right), \tag{10.2.5}$$

where the sums run over all $\hat{c}$ values such that $\rho(\hat{c}) > 0$. We extend both definitions to $\alpha = \infty$ by taking the $\alpha \to \infty$ limit.

The choice of $\wr$ ($\wr$) is meant to indicate that $\widetilde{H}^{\wr,f}$ ($\widetilde{H}^{\wr,f}$) is created from $\widetilde{H}^{\uparrow,f}$ ($\widetilde{H}^{\downarrow,f}$) upon replacing the conditional Rényi entropy in the original definition Definition 8.3.1 with its lower bound $\widetilde{H}^{\downarrow}$ (upper bound $\widetilde{H}^{\uparrow}$).

Both mixed f-weighted entropies satisfy the data-processing inequality, which we will make use of for e.g. squashing maps, see Section 10.3.

Lemma 10.2.2 (Data-processing mixed f-weighted entropies)**.** *Let $\rho \in \mathrm{S}_=(\widehat{C}QQ')$ be classical on $\widehat{C}$, let f be a tradeoff function on $\widehat{C}$ and $\alpha \in [\frac{1}{2},1) \cup (1,\infty]$. Then, for any channel $\mathcal{E} \in \mathrm{CPTP}(Q',Q'')$,*

$$\widetilde{H}_\alpha^{\wr,f}(Q|\widehat{C}Q'')_{\mathcal{E}(\rho)} \geq \widetilde{H}_\alpha^{\wr,f}(Q|\widehat{C}Q')_\rho, \tag{10.2.6}$$

$$\widetilde{H}_\alpha^{\wr,f}(Q|\widehat{C}Q'')_{\mathcal{E}(\rho)} \geq \widetilde{H}_\alpha^{\wr,f}(Q|\widehat{C}Q')_\rho. \tag{10.2.7}$$

Proof. Both properties follow from the data-processing inequality of the sandwiched Rényi entropies, see [91, Theorem 1] (or [47, 92–94]). □

Similarly to the $\widetilde{H}_\alpha^{\uparrow,f}$-normalization constant, one can define a $\widetilde{H}_\alpha^{\wr,f}$-normalization constant by

$$\kappa^{\wr} := \inf_{\substack{\rho_{AB} \\ \text{s.t. } \mathrm{Tr}_B[\rho_{AB}]=\tau_A}} \widetilde{H}_\alpha^{\wr,f}(S|\widehat{C}IE)_\nu. \tag{10.2.8}$$

Then, the resulting tradeoff function from the dual variables in Eq. (10.2.1) is the optimal tradeoff function for

$$\sup_{\mathbf{f}} \left(\mathbf{f} \cdot \mathbf{q} + \kappa^{\wr} \right), \tag{10.2.9}$$

which we will show in the following theorem.

Theorem 10.2.3 (Convex Conjugate of Mixed f-weighted Entropy)**.** *Let $\alpha \in (1, \infty)$ and $\rho \in \mathrm{S}_=(\widehat{C}QQ')$, where $\widehat{C}$ is classical and has the alphabet $\widehat{\mathcal{C}}$. Let $\mathbb{P}_\alpha$ denote the probability distributions on $\widehat{C}$ and $\boldsymbol{\rho}_{\widehat{C}}$ denote the probability on $\widehat{C}$ induced by ρ. Furthermore, define*

$$G_{\alpha,\rho}(\mathbf{f}) \coloneqq -\widetilde{H}_\alpha^{\wedge,f}(Q|\widehat{C}Q')_\rho. \tag{10.2.10}$$

Then, the convex conjugate G^* of G, defined as

$$G^*_{\alpha,\rho}(\boldsymbol{\lambda}) \coloneqq \sup_{\mathbf{f}\in\mathbb{R}^{|\widehat{\mathcal{C}}|}} \left(\mathbf{f}\cdot\boldsymbol{\lambda} - G_{\alpha,\rho}(\mathbf{f})\right) \tag{10.2.11}$$

satisfies

$$G^*_{\alpha,\rho}(\boldsymbol{\lambda}) = \begin{cases} \frac{\alpha}{\alpha-1} D\left(\boldsymbol{\lambda}\middle\|\boldsymbol{\nu}^\rho_{\widehat{C}}\right) + \sum_{\hat{c}} \lambda(\hat{c}) \widetilde{H}^\downarrow_\alpha(S|\tilde{T}E)_{\nu^\rho_{|\hat{c}}}, & \boldsymbol{\lambda} \in \mathbb{P}_{\widehat{C}} \\ \infty, & \textit{otherwise.} \end{cases} \tag{10.2.12}$$

Proof. The proof follows immediately from [41, Lemma 4.11] upon replacing $\widetilde{H}^\uparrow(Q|Q')_{\rho|\hat{c}}$ with $\widetilde{H}^\downarrow(Q|Q')_{\rho|\hat{c}}$ at the very first step. □

From this theorem we can conclude that the optimal tradeoff function for Eq. (10.2.9) is indeed given by solving for the dual variables in Eq. (solve for dual) of Eq. (10.2.3), summarized in the corollary below.

Corollary 10.2.4. *Let $\mathcal{M}_{QKD} : AB \to SI\widehat{C}$ be the channel from Definition 8.2.1 and S_{acc} be the convex acceptance set. For any $\rho_{AB} \in \mathrm{S}_=(AB)$ we define $\nu = \mathcal{M}_{QKD}[\rho_{AB}]$. Then, the optimization problem*

$$\sup_{\mathbf{f}} \inf_{\substack{\rho\in\mathrm{S}_=(AB),\\ s.t.\ \mathrm{Tr}_B[\rho]=\tau_A}} \left(\mathbf{f}\cdot\mathbf{q} + \widetilde{H}_\alpha^{\wedge,f}(S|\widehat{C}IE)_\nu\right), \tag{10.2.13}$$

is the Lagrange dual problem of

$$\inf_{\mathbf{q}\in S_{\mathrm{acc}}} \inf_{\substack{\rho\in\mathrm{S}_=(AB)\\ s.t.\ \mathrm{Tr}_B[\rho]=\tau_A}} \inf_{\boldsymbol{\lambda}\in\mathbb{P}(\widehat{\mathcal{C}})} \left(\frac{\alpha}{\alpha-1} D\left(\boldsymbol{\lambda}\middle\|\boldsymbol{\nu}^\rho_{\widehat{C}}\right) + \lambda(\mathtt{gen})\widetilde{H}^\downarrow_\alpha(S|\tilde{T}E)_{\nu^\rho_{|\mathtt{gen}}}\right) \tag{10.2.14}$$

$$s.t.\ \boldsymbol{\nu}^\rho_{\widehat{C}} = (\gamma\boldsymbol{\nu}^\rho_{\widehat{C}|\mathtt{test}}, 1-\gamma)^T, \tag{10.2.15}$$

$$\boldsymbol{\nu}^\rho_{\widehat{C}|\mathtt{test}} = \boldsymbol{\Phi}[\rho_{AB|\mathtt{test}}],$$

$$\mathbf{q} - \boldsymbol{\lambda} = 0,$$

and the optimal value is equal to h^{QKD}_α.

Proof. The proof follows from [41, Lemma 4.12] by replacing $\widetilde{H}_\alpha^{\uparrow,f}$ with $\widetilde{H}_\alpha^{\wr,f}$ and inserting the function G from Theorem 10.2.3. □

In addition to the corollary above, it holds that the standard f-weighted entropies are bounded by the mixed f-weighted entropies as $\widetilde{H}_\alpha^{\uparrow,f}(Q|\widehat{C}Q')_\rho \geq \widetilde{H}_\alpha^{\wr,f}(Q|\widehat{C}Q')_\rho$ and $\widetilde{H}_\alpha^{\wr,f}(Q|\widehat{C}Q')_\rho \geq \widetilde{H}_\alpha^{\downarrow,f}(Q|\widehat{C}Q')_\rho$ for all $\rho \in \mathrm{S}_=(\widehat{C}QQ')$. However, one cannot compare the two mixed f-weighted entropies as this would depend on how much the entropy $\widetilde{H}_\alpha^{\uparrow}(Q|Q')_{\rho|\hat{c}}$ differs from $\widetilde{H}_\alpha^{\downarrow}(Q|Q')_{\rho|\hat{c}}$ for each $\hat{c}$.

Lemma 10.2.5. *Let $\rho \in \mathrm{S}_=(\widehat{C}QQ')$ be arbitrary. Then, it holds for $\alpha \in (1,\infty)$*

$$\widetilde{H}_\alpha^{\uparrow,f}(Q|\widehat{C}Q')_\rho \geq \widetilde{H}_\alpha^{\wr,f}(Q|\widehat{C}Q')_\rho \geq \widetilde{H}_\alpha^{\downarrow,f}(Q|\widehat{C}Q')_\rho, \tag{10.2.16}$$

$$\widetilde{H}_\alpha^{\uparrow,f}(Q|\widehat{C}Q')_\rho \geq \widetilde{H}_\alpha^{\wr,f}(Q|\widehat{C}Q')_\rho \geq \widetilde{H}_\alpha^{\downarrow,f}(Q|\widehat{C}Q')_\rho. \tag{10.2.17}$$

Proof. We only show the proof for $\widetilde{H}_\alpha^{\wr,f}(Q|\widehat{C}Q')_\rho$ since the statement for $\widetilde{H}_\alpha^{\wr,f}(Q|\widehat{C}Q')_\rho$ follows analogously. For every $\hat{c}$ it holds

$$\rho(\hat{c})\, 2^{\frac{1-\alpha}{\alpha}\left(-f(\hat{c})+\widetilde{H}_\alpha^{\uparrow}(Q|Q')_{\rho|\hat{c}}\right)} \leq \rho(\hat{c})\, 2^{\frac{1-\alpha}{\alpha}\left(-f(\hat{c})+\widetilde{H}_\alpha^{\downarrow}(Q|Q')_{\rho|\hat{c}}\right)}. \tag{10.2.18}$$

Therefore, it it follows for the sum over $\hat{c}$

$$\sum_{\hat{c}} \rho(\hat{c})\, 2^{\frac{1-\alpha}{\alpha}\left(-f(\hat{c})+\widetilde{H}_\alpha^{\uparrow}(Q|Q')_{\rho|\hat{c}}\right)} \leq \sum_{\hat{c}} \rho(\hat{c})\, 2^{\frac{1-\alpha}{\alpha}\left(-f(\hat{c})+\widetilde{H}_\alpha^{\downarrow}(Q|Q')_{\rho|\hat{c}}\right)}, \tag{10.2.19}$$

and by taking the logarithm with the prefactor $-\frac{\alpha}{\alpha-1}$ we find

$$\widetilde{H}_\alpha^{\uparrow,f}(Q|\widehat{C}Q')_\rho \geq \widetilde{H}_\alpha^{\wr,f}(Q|\widehat{C}Q')_\rho. \tag{10.2.20}$$

The second inequality involving $\widetilde{H}_\alpha^{\wr,f}(Q|\widehat{C}Q')_\rho \geq \widetilde{H}_\alpha^{\downarrow,f}(Q|\widehat{C}Q')_\rho$ follows from Lemma 21 in Appendix E of Ref. [4] which states that the following log-mean exponential quantity is non-decreasing in ζ

$$\mathrm{lme}_\zeta[\mathbf{p},\mathbf{v}] := \frac{1}{\zeta}\log\left(\sum_{\hat{c}} p(\hat{c}) 2^{\zeta v(\hat{c})}\right). \tag{10.2.21}$$

By choosing $\zeta = -\frac{\alpha-1}{\alpha}$ and identifying

$$p(\hat{c}) = \rho(\hat{c}) \quad v(\hat{c}) = -f(\hat{c}) + \widetilde{H}_\alpha^{\downarrow}(Q|Q')_{\rho|\hat{c}}, \tag{10.2.22}$$

for all $\hat{c}$, we can recast the f-weighted entropy $\widetilde{H}_\alpha^{\wp,f}(Q|Q'\widehat{C})_\rho$ as

$$\widetilde{H}_\alpha^{\wp,f}(Q|Q'\widehat{C})_\rho = \mathrm{lme}_\zeta[\mathbf{p},\mathbf{v}]. \tag{10.2.23}$$

Then, if we pick $\zeta' = -\frac{1}{\alpha}$ it holds

$$\widetilde{H}_\alpha^{\downarrow,f}(Q|Q'\widehat{C})_\rho = \mathrm{lme}_{\zeta'}[\mathbf{p},\mathbf{v}], \tag{10.2.24}$$

and because $\alpha > 1$ we find $\zeta' = -\frac{1}{\alpha} \leq -\frac{\alpha-1}{\alpha} = \zeta$. Since [4, Lemma 21 in App. E] states that $\mathrm{lme}_\zeta[\mathbf{p},\mathbf{v}]$ is non-decreasing in ζ, it follows

$$\widetilde{H}_\alpha^{\downarrow,f}(Q|Q'\widehat{C})_\rho = \mathrm{lme}_{\zeta'}[\mathbf{p},\mathbf{v}] \leq \mathrm{lme}_\zeta[\mathbf{p},\mathbf{v}] = \widetilde{H}_\alpha^{\wp,f}(Q|Q'\widehat{C})_\rho, \tag{10.2.25}$$

which concludes the proof. The inequalities for the other mixed f-weighted entropy follow similarly. □

In light of the discussion above and the methods presented in Section 10.1 it is therefore most natural to bound $\kappa^\uparrow$ with either the $\widetilde{H}_\alpha^{\downarrow,f}$-normalization constant $\kappa^\downarrow$ or the $\widetilde{H}_\alpha^{\wp,f}$-normalization constant $\kappa^\wp$. Additionally, from Lemma 10.2.5 we can conclude that

$$\kappa^\wp \geq \kappa^\downarrow, \tag{10.2.26}$$

and with the methods presented here the $\widetilde{H}_\alpha^{\wp,f}$-normalization constant $\kappa^\wp$ will result in the highest key rates.

10.2.1 Reformulating the Optimization Problem for κ

Now, similar to Eq. (10.1.30) we will simplify the optimization problem for

$$\kappa^\wp = \inf_{\substack{\rho_{AB}\in \mathrm{S}_=(AB) \\ \text{s.t. } \mathrm{Tr}_B[\rho_{AB}]=\tau_A}} \widetilde{H}_\alpha^{\wp,f}(S|\widehat{C}IE)_\nu, \tag{10.2.27}$$

and rewrite it in terms of the Choi state J of Eve's channel. As a first step, let us expand the definition of the mixed f-weighted entropy

$$\widetilde{H}_\alpha^{\wp,f}(S|I\widehat{C}E)_\nu = \frac{\alpha}{1-\alpha}\log\left(\sum_{\hat{c}} \nu(\hat{c})\, 2^{\frac{1-\alpha}{\alpha}\left(-f(\hat{c})+\widetilde{H}_\alpha^{\downarrow}(S|IE)_{\nu|\hat{c}}\right)}\right), \tag{10.2.28}$$

where $\nu(\hat{c}) = \Pr(\hat{c})$ indicates the probability of outcome $\hat{c}$. Next, we note that $\widetilde{H}^{\downarrow}(S|IE)_{\nu|\hat{c}} = 0$ for all $\hat{c} \neq \texttt{gen}$ since in test rounds all data is announced. Additionally, as in the proof of Theorem 8.2.4 it holds

$$\widetilde{H}^{\downarrow}_{\alpha}(S|IE)_{\nu^{\rho}_{|\texttt{gen}}} = \widetilde{H}^{\downarrow}_{\alpha}(S|\tilde{T}E)_{\nu^{\rho}_{|\texttt{gen}}} = g_{\alpha}(J). \tag{10.2.29}$$

Then, we can rewrite the mixed f-weighted entropy as

$$\begin{aligned}\widetilde{H}^{\supset,f}_{\alpha}(S|I\widehat{C}E)_{\nu} = \frac{\alpha}{1-\alpha}\log\Big(\gamma \sum_{\hat{c}\in\widehat{\mathcal{C}}_{\setminus\texttt{gen}}} \nu(\hat{c}|\texttt{test})\, 2^{-\frac{\alpha-1}{\alpha}f(\hat{c})} \\ + (1-\gamma)\, 2^{-\frac{\alpha-1}{\alpha}\left(\widetilde{H}^{\downarrow}_{\alpha}(S|\tilde{T}E)_{\nu|\texttt{gen}} - f(\texttt{gen})\right)}\Big).\end{aligned} \tag{10.2.30}$$

Finally, we can make use of our definition of $\boldsymbol{\Psi}$ of Eq. (10.1.28) to relate the probabilities of outcomes in test rounds by

$$\nu(\hat{c}|\texttt{test}) = \Pr(\hat{c}|\texttt{test}) = \boldsymbol{\Psi}[J]. \tag{10.2.31}$$

Thus, we can recast the optimization problem of the $\widetilde{H}^{\supset,f}_{\alpha}$-normalization constant $\kappa^{\supset}$ in terms of the state as

$$\begin{aligned}\kappa^{\supset} = \inf_{\substack{\rho\in \mathrm{S}_{=}(AB),\\ \boldsymbol{\nu}\in\mathbb{P}(\widehat{\mathcal{C}}_{\setminus\texttt{gen}})}} \Bigg(\frac{\alpha}{1-\alpha}\log\Big(\gamma \sum_{\hat{c}\in\widehat{\mathcal{C}}_{\setminus\texttt{gen}}} \boldsymbol{\nu}(\hat{c}) 2^{\frac{\alpha-1}{\alpha}f(\hat{c})} \\ + (1-\gamma)\, 2^{-\frac{\alpha-1}{\alpha}\left(\widetilde{H}^{\downarrow}_{\alpha}(S|\tilde{T}E)_{\nu|\texttt{gen}} - f(\texttt{gen})\right)}\Big)\Bigg) \\ \text{s.t. } \boldsymbol{\nu} = \boldsymbol{\Phi}[\rho_{AB|\texttt{test}}], \\ \mathrm{Tr}_B[\rho_{AB}] = \tau_A,\end{aligned} \tag{10.2.32}$$

where we defined the vector $\nu(\hat{c}) \coloneqq \Pr(\hat{c}|\texttt{test})$ and assigned $f(\hat{c}) = \mathrm{f}_{\hat{c}}$. Alternatively, while again identifying $\widetilde{H}^{\downarrow}_{\alpha}(S|\tilde{T}E)_{\nu|\texttt{gen}} = g_{\alpha}(J)$, we can equivalently write it in terms of the Choi state of Eve's attack channel as

$$\begin{aligned}\kappa^{\supset} = \inf_{\substack{J\in\mathrm{Pos}(A'B),\\ \boldsymbol{\nu}\in\mathbb{P}(\widehat{\mathcal{C}}_{\setminus\texttt{gen}})}} \Bigg(\frac{\alpha}{1-\alpha}\log\Big(\gamma \sum_{\hat{c}\in\widehat{\mathcal{C}}_{\setminus\texttt{gen}}} \boldsymbol{\nu}(\hat{c}) 2^{\frac{\alpha-1}{\alpha}f(\hat{c})} \\ + (1-\gamma)\, 2^{-\frac{\alpha-1}{\alpha}(g_{\alpha}(J) - f(\texttt{gen}))}\Big)\Bigg) \\ \text{s.t. } \mathrm{Tr}_B[J] = \mathbb{I}_{A'}, \\ \boldsymbol{\nu} = \boldsymbol{\Psi}[J].\end{aligned} \tag{10.2.33}$$

Similarly, one can write the optimization problem of the $\widetilde{H}_\alpha^{\downarrow,f}$-normalization constant $\kappa^\downarrow$ in terms of the channel as

$$\kappa^\downarrow = \inf_{\substack{J\in \mathrm{Pos}(A'B),\\ \boldsymbol{\nu}\in\mathbb{P}(\widehat{\mathcal{C}}_{\setminus \mathrm{gen}})}} \left(\frac{1}{1-\alpha} \log \left(\gamma \sum_{\hat{c}\in\widehat{\mathcal{C}}_{\setminus \mathrm{gen}}} \boldsymbol{\nu}(\hat{c}) 2^{(\alpha-1)f(\hat{c})} + (1-\gamma)\, 2^{-(\alpha-1)(g_\alpha(J)-f(\mathtt{gen}))} \right) \right) \tag{10.2.34}$$

$$\text{s.t. } \mathrm{Tr}_B[J] = \mathbb{I}_{A'},$$
$$\boldsymbol{\nu} = \boldsymbol{\Psi}[J],$$

These optimization problems are again convex problems that can be solved by iterative methods such as the Frank-Wolfe algorithm. In the appendix, in Theorem A.1.11, we prove the joint convexity in J and $\boldsymbol{\nu}$ of the optimization problem for computing $\kappa^\downarrow$.

Remark 10.2.6. We emphasize that for an experimental implementation of a QKD protocol utilizing our Rényi framework, one is only required to find the gradient and the normalization constant of the tradeoff function $\hat{f}_{\mathrm{prot}}$ *once* before running the protocol. Moreover, since $\hat{f}_{\mathrm{prot}}$ is an affine function determined by its gradient $\mathbf{f}$ and the $\widetilde{H}_\alpha^{\downarrow,f}$-normalization constant $\kappa^\downarrow$, it can be implemented very efficiently. This is a huge advantage compared to other proof techniques such as the post-selection technique [27] or EUR- and phase error correction-based techniques [24, 95].

We have now established the necessary framework for calculating both fixed- and variable-length key rates within the context of finite-dimensional protocols, specifically qubit-based implementations. Fixed-length key rates are derivable from Eq. (10.1.30), which in turn facilitates the direct determination of the tradeoff function via Eq. (10.2.1). Then, variable-length key rates can be obtained by solving Eq. (10.2.34).

10.3 Squashing Maps

Before, we turn our attention to more sophisticated protocols such as decoy-state protocols, we present how squashing maps can be included in this framework. Specifically, here we show that the flag-state squasher of [42] and the qubit squasher from [62, 63] can be incorporated into the optimization problems for h_α^{QKD}, see Eq. (10.1.12), required in fixed-length security and into the optimization problem for the $\widetilde{H}_\alpha^{\supset,f}$-normalization constant

Eq. (10.2.32) required for variable-length security. In the following theorem the CPTP map Λ should be thought of as a squashing map, satisfying the properties in Eqs. (4.1.2) and (4.1.3).

Theorem 10.3.1 (Squashing Maps in h_α^{QKD} and $h_\alpha^{\uparrow,\mathrm{QKD}}$)**.** *Let $\Lambda : B \to B'$ be a squashing map satisfying Eq. (4.1.3), and for any CPTP map $\mathcal{E} : A' \to B$, let $\mathcal{E}_{\mathrm{sq}}$ denote the combined CPTP map $\mathcal{E}_{\mathrm{sq}} := \Lambda \circ \mathcal{E} : A' \to B'$ and let $V_\Lambda : B' \to B'E_{sq}$, $V_\mathcal{E} : A' \to B$, and $V_{\mathcal{E}_{\mathrm{sq}}} : A' \to B'EE_{sq}$ be the respective Stinespring dilations. By $\mathcal{M}'_{QKD}$ we denote the channel from Definition 8.2.1 with Bob's squashed POVM elements on B' defined by*

$$M'^{B'}_{\beta,y} := \Lambda^\dagger[M^B_{\beta,y}] \tag{10.3.1}$$

in place of the original POVM elements on B. Moreover, suppose there exist constants θ_1, θ_2 and a subset $\mathcal{F}_t \subseteq \widehat{\mathcal{C}}$ such that

$$\sum_{\hat{c}\in\mathcal{F}_t} \Phi_{\hat{c}}[\rho] \geq \theta_1 \left(\theta_2 - \mathrm{Tr}\left[\left(M^A \otimes \Pi^B\right)\rho\right]\right) \ \forall \rho \in \mathrm{S}_=(AB), \tag{10.3.2}$$

where Π^B is a projector onto a subspace in B which is invariant under Λ, and M^A is one of Alice's POVM elements.

Then, there exists a map $\mathbf{\Phi}' : AB' \to \widehat{C}$ such that

$$\mathbf{\Phi}'[\sigma_{AB'}] = \mathbf{\Phi}' \circ \Lambda[\rho_{AB}] = \mathbf{\Phi}[\rho_{AB}], \tag{10.3.3}$$

for all states ρ_{AB} and h_α^{QKD} can be bounded from below as

$$h_\alpha^{\mathrm{QKD}} \geq \inf_{\mathbf{q}\in S_{\mathrm{acc}}} \inf_{\substack{V_{\mathcal{E}_{\mathrm{sq}}},\\ V^\dagger_{\mathcal{E}_{\mathrm{sq}}} V_{\mathcal{E}_{\mathrm{sq}}} = \mathbb{I}_{A'}}} \left(\frac{\alpha}{\alpha-1} D\big(\mathbf{q}\big\|\boldsymbol{\nu}_{\widehat{C}}\big) + q(\mathtt{gen})\widetilde{H}^\downarrow_\alpha(S|\tilde{T}EE_{\mathrm{sq}})_{\nu_{|\mathtt{gen}}}\right), \tag{10.3.4}$$

$$\begin{aligned}
s.t.\ & \nu_{S\tilde{T}\widehat{C}EE_{\mathrm{sq}}|\mathtt{gen}} = \mathcal{M}'_{QKD}(V_{\mathcal{E}_{\mathrm{sq}}}|\xi^g\rangle\langle\xi^g|_{AA'} V^\dagger_{\mathcal{E}_{\mathrm{sq}}}),\\
& \boldsymbol{\nu}_{\widehat{C}} = (\gamma \boldsymbol{\nu}_{\widehat{C}|\mathtt{test}}, 1-\gamma)^T,\\
& \boldsymbol{\nu}_{\widehat{C}|\mathtt{test}} = \mathbf{\Phi}'[\sigma_{AB'|\mathtt{test}}],\\
& \sum_{\hat{c}\in\mathcal{F}_t} \Phi'_{\hat{c}}[\sigma_{AB'|\mathtt{test}}] \geq \theta_1 \left(\theta_2 - \mathrm{Tr}\left[\left(M^A \otimes \Pi^{B'}\right)\sigma_{AB'|\mathtt{test}}\right]\right),\\
& \sigma_{AB'|\mathtt{test}} = \mathrm{Tr}_{EE_{\mathrm{sq}}}\left[V_{\mathcal{E}_{\mathrm{sq}}}|\xi^t\rangle\langle\xi^t|_{AA'} V^\dagger_{\mathcal{E}_{\mathrm{sq}}}\right],
\end{aligned}$$

where $\Pi^{B'} = \Lambda[\Pi^B]$*, i.e., it projects onto the image of* Π^B *under* Λ*. Furthermore, for the* $\widetilde{H}_\alpha^{\downarrow,f}$*-normalization constant* $\kappa^{\downarrow}$ *it holds*

$$\begin{aligned}
\kappa^{\downarrow} \geq \inf_{\substack{V_{\mathcal{E}_{\mathrm{sq}}}, \\ V_{\mathcal{E}_{\mathrm{sq}}}^\dagger V_{\mathcal{E}_{\mathrm{sq}}} = \mathbb{I}_{A'}}} \inf_{\boldsymbol{\nu} \in \mathbb{P}(\widehat{\mathcal{C}}_{\setminus \mathrm{gen}})} \Bigg(& \frac{\alpha}{1-\alpha} \log \Big(\gamma \sum_{\hat{c} \in \widehat{\mathcal{C}}_{\setminus \mathrm{gen}}} \boldsymbol{\nu}(\hat{c}) 2^{\frac{\alpha-1}{\alpha} f(\hat{c})} \\
& + (1-\gamma)\, 2^{-\frac{\alpha-1}{\alpha}\left(\widetilde{H}_\alpha^{\downarrow}(S|\tilde{T}EE_{\mathrm{sq}})_{\nu|\mathtt{gen}} - f(\mathtt{gen})\right)} \Big) \Bigg) \\
\text{s.t. } & \nu_{S\tilde{T}\widehat{C}EE_{\mathrm{sq}}|\mathtt{gen}} = \mathcal{M}'_{QKD}(V_{\mathcal{E}_{\mathrm{sq}}} |\xi^g\rangle\langle\xi^g|_{AA'} V_{\mathcal{E}_{\mathrm{sq}}}^\dagger), \\
& \boldsymbol{\nu} = \boldsymbol{\Phi}'[\sigma_{AB'|\mathtt{test}}], \\
& \sum_{\hat{c} \in \mathcal{F}_t} \Phi'_{\hat{c}}[\sigma_{AB'|\mathtt{test}}] \geq \theta_1 \left(\theta_2 - \mathrm{Tr}\left[\left(M^A \otimes \Pi^{B'} \right) \sigma_{AB'|\mathtt{test}} \right] \right), \\
& \sigma_{AB'|\mathtt{test}} = \mathrm{Tr}_{EE_{\mathrm{sq}}} \left[V_{\mathcal{E}_{\mathrm{sq}}} \left|\xi^t\right\rangle\!\left\langle\xi^t\right|_{AA'} V_{\mathcal{E}_{\mathrm{sq}}}^\dagger \right].
\end{aligned} \tag{10.3.5}$$

Proof. First, we note that since Λ is a squashing map, there exists a map $\boldsymbol{\Phi}' : AB' \to \widehat{C}$ such that

$$\boldsymbol{\Phi}'[\sigma_{AB'}] = \boldsymbol{\Phi}' \circ \Lambda[\rho_{AB}] = \boldsymbol{\Phi}[\rho_{AB}], \tag{10.3.6}$$

for all states ρ_{AB}. Since Π^B projects onto a subspace that is invariant under Λ, we can write

$$\mathrm{Tr}\left[\left(M^A \otimes \Pi^B \right) \rho_{AB} \right] = \mathrm{Tr}\left[\left(M^A \otimes \Pi^{B'} \right) \Lambda(\rho_{AB}) \right], \tag{10.3.7}$$

for all $\rho_{AB} \in \mathrm{S}_=(AB)$. Furthermore, it holds for the Rényi entropy

$$\widetilde{H}_\alpha^{\downarrow}(S|\tilde{T}EE_{\mathrm{sq}})_{\mathcal{M}'_{\mathrm{QKD}}\left(V_\Lambda V_{\mathcal{E}} |\xi^g\rangle\langle\xi^g|_{AA'} V_{\mathcal{E}}^\dagger V_\Lambda^\dagger\right)} \tag{10.3.8}$$

$$\leq \widetilde{H}_\alpha^{\downarrow}(S|\tilde{T}E)_{\mathcal{M}'_{\mathrm{QKD}}\left(\mathrm{Tr}_{E_{\mathrm{sq}}}\left[V_\Lambda V_{\mathcal{E}} |\xi^g\rangle\langle\xi^g|_{AA'} V_{\mathcal{E}}^\dagger V_\Lambda^\dagger\right]\right)} \tag{10.3.9}$$

$$= \widetilde{H}_\alpha^{\downarrow}(S|\tilde{T}E)_{\mathcal{M}_{\mathrm{QKD}}\left(V_{\mathcal{E}} |\xi^g\rangle\langle\xi^g|_{AA'} V_{\mathcal{E}}^\dagger\right)}, \tag{10.3.10}$$

where the first inequality follows from the data-processing inequality [92, Theorem 1]. The second equality follows from the definition of the squashed POVM elements on B' and that the state on $S\tilde{T}E$ remains the same independent of whether the squashing map is applied [64, Theorem 3.3.1].

Combining both observations, we can bound h_α^{QKD} by

$$h_\alpha^{\mathrm{QKD}} \geq \inf_{\mathbf{q}\in S_{\mathrm{acc}}} \inf_{\substack{V_{\mathcal{E}},\\ V_{\mathcal{E}}^\dagger V_{\mathcal{E}}=\mathbb{I}_{A'}}} \left(\frac{\alpha}{\alpha-1} D\big(\mathbf{q}\big\|\boldsymbol{\nu}_{\widehat{C}}\big) + q(\mathtt{gen})\widetilde{H}_\alpha^{\downarrow}(S|\tilde{T}EE_{\mathrm{sq}})_{\nu_{|\mathtt{gen}}} \right), \tag{10.3.11}$$

$$\begin{aligned}
\text{s.t. } & \nu_{S\tilde{T}\widehat{C}EE_{\mathrm{sq}}|\mathtt{gen}} = \mathcal{M}'_{\mathrm{QKD}}(V_\Lambda V_{\mathcal{E}} \left|\xi^g\right\rangle\!\left\langle\xi^g\right|_{AA'} V_{\mathcal{E}}^\dagger V_\Lambda^\dagger),\\
& \boldsymbol{\nu}_{\widehat{C}} = (\gamma \boldsymbol{\nu}_{\widehat{C}|\mathtt{test}}, 1-\gamma)^T,\\
& \boldsymbol{\nu}_{\widehat{C}|\mathtt{test}} = \boldsymbol{\Phi}'[\sigma_{AB'|\mathtt{test}}],\\
& \sum_{\hat{c}\in\mathcal{F}_t} \Phi'_{\hat{c}}[\sigma_{AB'|\mathtt{test}}] \geq \theta_1 \left(\theta_2 - \mathrm{Tr}\left[\left(M^A \otimes \Pi^{B'}\right)\sigma_{AB'|\mathtt{test}}\right]\right),\\
& \sigma_{AB'|\mathtt{test}} = \mathrm{Tr}_{EE_{\mathrm{sq}}}\left[V_\Lambda V_{\mathcal{E}} \left|\xi^t\right\rangle\!\left\langle\xi^t\right|_{AA'} V_{\mathcal{E}}^\dagger V_\Lambda^\dagger\right],
\end{aligned}$$

since by assumption, every feasible point of Eq. (10.1.30) is a feasible point. Finally, by minimizing over $V_{\mathcal{E}_{\mathrm{sq}}} \coloneqq V_\Lambda V_{\mathcal{E}}$ instead of only $V_{\mathcal{E}}$, the theorem statement follows because the minimization runs over a bigger set that only decreases the infimum further.

Repeating the same steps for the $\widetilde{H}_\alpha^{\wp,f}$-normalization constant yields the similar bound for $\kappa^{\wp}$. □

Remark 10.3.2. The same statement holds true for the $\widetilde{H}_\alpha^{\downarrow,f}$-normalization constant.

Remark 10.3.3. For notational simplicity, we did not include constraints from the states in generation rounds, however one can include those straightforwardly following the same proof. The construction would go as follows. One would define a similar subset $\mathcal{F}_g \subseteq \mathcal{C}$ for constraints on generation rounds and a map acting similar to $\boldsymbol{\Phi}$ but with POVM elements conditioned on generation rounds. Then, the argument follows in the way as above.

Before we conclude this subsection, let us discuss which squashing maps can be included in this construction and how one can find the constants θ_1 and θ_2. Here, we follow the same lines as in [2, Appendix B].

First, if we consider the squashing maps from [62, 63], no additional constraints as in Eq. (10.3.2) are required, which could equivalently be written as $\theta_1 = 0$, enforcing only trivial constraints. However, much more importantly, this implies that one can use these squashing maps in our framework immediately.

Alternatively, if the flag-state squasher from [42] is employed, constraints of the form as in Eq. (10.3.2) are strictly necessary. Let us consider the bounds on Bob's $\leq N_B$-photon

subspace derived in Chapter 5. These state that the weight $p(\leq N_B|x)$ inside Bob's $\leq N_B$-photon subspace is bounded as

$$p(\leq N_B|x) \geq 1 - \frac{m_{\text{mult}|x}}{c_{\geq N_B+1}}, \tag{10.3.12}$$

where the multi-click probability $m_{\text{mult}|x}$ given that Alice sent state x is given by

$$m_{\text{mult}|x} = \text{Tr}\left[\left(M_x^A \otimes M_{\text{mult}}^B\right) \rho_{|x}\right] \tag{10.3.13}$$

and $c_{\geq N_B+1}$ is a constant determined by the optical detection setup, see Theorem 5.3.3.

Next, we note that the weight inside Bob's $\leq N_B$-photon subspace given that Alice sent state x can be equivalently written as

$$p(\leq N_B|x) = \text{Tr}\left[\left(M_x^A \otimes \Pi_{\leq N_B}^B\right) \rho_{|x}\right]. \tag{10.3.14}$$

Hence, by multiplying with $p(x)$ and removing the conditioning of ρ, the bound in Eq. (10.3.12) can be recast as

$$\text{Tr}\left[\left(M_x^A \otimes M_{\text{mult}}^B\right) \rho\right] \geq p(x) c_{\geq N_B+1} - \text{Tr}\left[\left(M_x^A \otimes \Pi_{\leq N_B}^B\right) \rho\right] \; \forall \rho \in \text{S}_=(AB). \tag{10.3.15}$$

Therefore, the constants required are $\theta_1 = c_{\geq N_B+1}$ and $\theta_2 = p(x)$, and the set $\mathcal{F}_t$ contains Alice's signal choice x and all of Bob's POVM elements included in the multi-click POVM element M_{mult}^B. Additionally, one can rewrite this expression in the more convenient formulation

$$\text{Tr}\left[\left(M_x^A \otimes \left(M_{\text{mult}}^B + c_{\geq N_B+1} \Pi_{\leq N_B}^B\right)\right) \rho\right] \geq p(x) c_{\geq N_B+1} \; \forall \rho \in \text{S}_=(AB). \tag{10.3.16}$$

In summary, we have shown that both the qubit squasher and the flag-state squasher can be included to calculate fixed- and variable-length key rates.

10.4 Source Maps

For many imperfections, it is favourable to use so-called source maps [68]. In [27] a proof-technique independent proof was given, however here we explicitly prove that source maps can indeed be incorporated into our framework.

Since the proof for the $\widetilde{H}_\alpha^{\downarrow,f}$-normalization constant is simpler and already lays out the general idea, we start with it and then present the proof for h_α^{QKD}.

Theorem 10.4.1 (Source Maps in f-weighted Entropies)**.** *Let $\sigma_{AA'}$ and $\tau_{AA''}$ be the states emitted by the real and virtual source, respectively. Furthermore, let $\Xi : A'' \to A'$ be a source map, i.e. a channel that satisfies $\sigma_{AA'} = \Xi(\tau_{AA''})$, and $V_\Xi : A'' \to A'E_{\mathrm{sou}}$ its Stinespring dilation. Then, for the QKD channel $\mathcal{M}_{QKD}$ and the $\widetilde{H}_\alpha^{\downarrow,f}$ entropy it holds*

$$\begin{aligned}&\inf_{V_{\mathcal{E}_{\mathrm{vir}}}} \widetilde{H}_\alpha^{\downarrow,f}(S|\widehat{C}IEE_{\mathrm{sou}})_{\mathcal{M}_{QKD}(V_{\mathcal{E}_{\mathrm{vir}}}\tau_{AA''}V_{\mathcal{E}_{\mathrm{vir}}}^\dagger)}\\ &\leq \inf_{V_{\mathcal{E}_{\mathrm{real}}}} \widetilde{H}_\alpha^{\downarrow,f}(S|\widehat{C}IE)_{\mathcal{M}_{QKD}(V_{\mathcal{E}_{\mathrm{real}}}\rho_{AA''}V_{\mathcal{E}_{\mathrm{real}}}^\dagger)},\end{aligned} \tag{10.4.1}$$

where $V_{\mathcal{E}_{\mathrm{vir}}} : A'' \to BEE_{\mathrm{sou}}$ and $V_{\mathcal{E}_{\mathrm{real}}} : A' \to BE$ are the Stinespring dilations of Eve's attack channels for the virtual and real source, respectively.

The same statement holds upon replacing the entropy $\widetilde{H}_\alpha^{\downarrow,f}$ with $\widetilde{H}_\alpha^{\uparrow,f}$ or $\widetilde{H}_\alpha^{\diamond,f}$.

Proof. The proof is based on a simple calculation.

$$\inf_{V_{\mathcal{E}_{\mathrm{vir}}}} \widetilde{H}_\alpha^{\downarrow,f}(S|\widehat{C}IEE_{\mathrm{sou}})_{\mathcal{M}_{\mathrm{QKD}}(V_{\mathcal{E}_{\mathrm{vir}}}\tau_{AA''}V_{\mathcal{E}_{\mathrm{vir}}}^\dagger)} \tag{10.4.2}$$

$$\leq \inf_{V_{\mathcal{E}_{\mathrm{real}}}} \widetilde{H}_\alpha^{\downarrow,f}(S|\widehat{C}IEE_{\mathrm{sou}})_{\mathcal{M}_{\mathrm{QKD}}(V_{\mathcal{E}_{\mathrm{real}}}V_\Xi\tau_{AA''}V_\Xi^\dagger V_{\mathcal{E}_{\mathrm{real}}}^\dagger)} \tag{10.4.3}$$

$$\leq \inf_{V_{\mathcal{E}_{\mathrm{real}}}} \widetilde{H}_\alpha^{\downarrow,f}(S|\widehat{C}IE)_{\mathcal{M}_{\mathrm{QKD}}(V_{\mathcal{E}_{\mathrm{real}}}\,\mathrm{Tr}_{E_{\mathrm{sou}}}[V_\Xi\tau_{AA''}V_\Xi^\dagger]V_{\mathcal{E}_{\mathrm{real}}}^\dagger)} \tag{10.4.4}$$

$$= \inf_{V_{\mathcal{E}_{\mathrm{real}}}} \widetilde{H}_\alpha^{\downarrow,f}(S|\widehat{C}IE)_{\mathcal{M}_{\mathrm{QKD}}(V_{\mathcal{E}_{\mathrm{real}}}\sigma_{AA'}V_{\mathcal{E}_{\mathrm{real}}}^\dagger)}. \tag{10.4.5}$$

The steps are as follows. The first inequality is due to optimizing over a smaller set, which can only increase the minimum. The second inequality is due to the data-processing inequality for $\widetilde{H}_\alpha^{\downarrow,f}$ entropies [77, Lemma 4.3]. The final equality holds because Ξ is a source map, i.e.

$$\mathrm{Tr}_{E_{\mathrm{sou}}}[V_\Xi\tau_{AA''}V_\Xi^\dagger] = \Xi(\tau_{AA''}) = \sigma_{AA'}. \tag{10.4.6}$$

The proof for the other f-weighted entropies follows by the same arguments upon replacing the data-processing inequality for $\widetilde{H}_\alpha^{\downarrow,f}$ entropies with the data-processing inequality for $\widetilde{H}_\alpha^{\uparrow,f}$ entropy [41, Lemma 4.4] or Lemma 10.2.2. □

Remark 10.4.2. The same statements holds for th mixed f-weighted entropy and the proof follows the same steps applying the data-processing inequality from Lemma 10.2.2.

This theorem, covers source maps in the $\widetilde{H}_\alpha^{\downarrow,f}$-normalization constant $\kappa^\downarrow$ and hence, next we will show that we can do the same for h_α^{QKD}.

Theorem 10.4.3 (Source maps in h_α^{QKD} and $h_\alpha^{\uparrow,\mathrm{QKD}}$)**.** *Let $\sigma_{AA'}$ and $\tau_{AA''}$ be the states emitted by the real and virtual source, respectively. Furthermore, let $\Xi : A'' \to A'$ be a source map, i.e. a channel that satisfies $\sigma_{AA'} = \Xi(\tau_{AA''})$, and $V_\Xi : A'' \to A'E_{\mathrm{sou}}$ its Stinespring dilation. Then, h_α^{QKD} can be bounded from below as*

$$h_\alpha^{\mathrm{QKD}} \geq \inf_{\mathbf{q}\in S_{\mathrm{acc}}} \inf_{\substack{V_{\mathcal{E}_{\mathrm{vir}}},\\ V_{\mathcal{E}_{\mathrm{vir}}}^\dagger V_{\mathcal{E}_{\mathrm{vir}}} = \mathbb{I}_{A''}}} \left(\frac{\alpha}{\alpha-1} D\big(\mathbf{q}\big\|\boldsymbol{\nu}_{\widehat{C}}\big) + q(\mathtt{gen}) \widetilde{H}_\alpha^{\downarrow}(S|\tilde{T}EE_{\mathrm{sou}})_{\nu_{|\mathtt{gen}}} \right), \tag{10.4.7}$$

where

$$\nu_{S\tilde{T}\widehat{C}EE_{\mathrm{sou}}|\mathtt{gen}} = \mathcal{M}_{QKD}(V_{\mathcal{E}_{\mathrm{vir}}} \tau_{AA''} V_{\mathcal{E}_{\mathrm{vir}}}^\dagger)_{|\mathtt{gen}}, \tag{10.4.8}$$

$$\boldsymbol{\nu}_{\widehat{C}} = (\gamma \boldsymbol{\nu}_{\widehat{C}|\mathtt{test}}, 1-\gamma)^T, \tag{10.4.9}$$

$$\boldsymbol{\nu}_{\widehat{C}|\mathtt{test}} = \boldsymbol{\Phi}[\mathrm{Tr}_{EE_{\mathrm{sou}}}[V_{\mathcal{E}_{\mathrm{vir}}} \tau_{AA'} V_{\mathcal{E}_{\mathrm{vir}}}^\dagger]]. \tag{10.4.10}$$

Here we again defined $V_{\mathcal{E}_{\mathrm{vir}}} : A'' \to BEE_{\mathrm{sou}}$ as the Stinespring dilation of Eve's attack channel for the virtual source. The same statement holds for $h_\alpha^{\uparrow,\mathrm{QKD}}$, i.e. replacing $\widetilde{H}_\alpha^{\downarrow}(S|\tilde{T}EE_{\mathrm{sou}})_{\nu_{|\mathtt{gen}}}$ with $\widetilde{H}_\alpha^{\uparrow}(S|\tilde{T}EE_{\mathrm{sou}})_{\nu_{|\mathtt{gen}}}$.

Proof. First, let us rewrite h_α^{QKD} from Corollary 10.1.2 (using the real source) as an optimization problem over Eve's attack channel characterized by its Stinespring dilation $V_{\mathcal{E}_{\mathrm{real}}} : A' \to BE$. This can be written as

$$h_\alpha^{\mathrm{QKD}} = \inf_{\mathbf{q}\in S_{\mathrm{acc}}} \inf_{\substack{V_{\mathcal{E}_{\mathrm{real}}},\\ V_{\mathcal{E}_{\mathrm{real}}}^\dagger V_{\mathcal{E}_{\mathrm{real}}} = \mathbb{I}_{A'}}} \left(\frac{\alpha}{\alpha-1} D\big(\mathbf{q}\big\|\boldsymbol{\nu}_{\widehat{C}}\big) + q(\mathtt{gen}) \widetilde{H}_\alpha^{\downarrow}(S|\tilde{T}E)_{\nu_{|\mathtt{gen}}} \right), \tag{10.4.11}$$

where

$$\nu_{S\tilde{T}\widehat{C}E|\mathtt{gen}} = \mathcal{M}_{\mathrm{QKD}}(V_{\mathcal{E}_{\mathrm{real}}} \sigma_{AA'} V_{\mathcal{E}_{\mathrm{real}}}^\dagger)_{|\mathtt{gen}}, \tag{10.4.12}$$

$$\boldsymbol{\nu}_{\widehat{C}} = (\gamma \boldsymbol{\nu}_{\widehat{C}|\mathtt{test}}, 1-\gamma)^T, \tag{10.4.13}$$

$$\boldsymbol{\nu}_{\widehat{C}|\mathtt{test}} = \boldsymbol{\Phi}[\mathrm{Tr}_E[V_{\mathcal{E}_{\mathrm{real}}} \sigma_{AA'} V_{\mathcal{E}_{\mathrm{real}}}^\dagger]]. \tag{10.4.14}$$

Since Ξ is a source map it satisfies

$$\mathrm{Tr}_{E_{\mathrm{sou}}}[V_\Xi \tau_{AA''} V_\Xi^\dagger] = \Xi(\tau_{AA''}) = \sigma_{AA'}, \tag{10.4.15}$$

and we can rewrite

$$\nu_{S\tilde{T}\widehat{C}E|\mathtt{gen}} = \mathcal{M}_{\mathrm{QKD}}(V_{\mathcal{E}_{\mathrm{real}}} \operatorname{Tr}_{E_{\mathrm{sou}}}[V_{\Xi}\tau_{AA''}V_{\Xi}^{\dagger}]V_{\mathcal{E}_{\mathrm{real}}}^{\dagger})_{|\mathtt{gen}}, \tag{10.4.16}$$

$$\boldsymbol{\nu}_{\widehat{C}|\mathtt{test}} = \boldsymbol{\Phi}[\operatorname{Tr}_{EE_{\mathrm{sou}}}[V_{\mathcal{E}_{\mathrm{real}}}V_{\Xi}\tau_{AA'}V_{\Xi}^{\dagger}V_{\mathcal{E}_{\mathrm{real}}}^{\dagger}]]. \tag{10.4.17}$$

Next, due to the data-processing inequality for Rényi entropies [91, Theorem 1] (or [47, 92–94]), we find

$$\begin{aligned} &\widetilde{H}_{\alpha}^{\downarrow}(S|\widehat{C}\tilde{T}EE_{\mathrm{sou}})_{\mathcal{M}_{\mathrm{QKD}}(V_{\mathcal{E}_{\mathrm{real}}}V_{\Xi}\tau_{AA''}V_{\Xi}^{\dagger}V_{\mathcal{E}_{\mathrm{real}}}^{\dagger})_{|\mathtt{gen}}} \\ &\leq \widetilde{H}_{\alpha}^{\downarrow}(S|\widehat{C}\tilde{T}E)_{\mathcal{M}_{\mathrm{QKD}}(V_{\mathcal{E}_{\mathrm{real}}}\operatorname{Tr}_{E_{\mathrm{sou}}}[V_{\Xi}\tau_{AA''}V_{\Xi}^{\dagger}]V_{\mathcal{E}_{\mathrm{real}}}^{\dagger})_{|\mathtt{gen}}}. \end{aligned} \tag{10.4.18}$$

Combining these relations, we find a lower bound on $h_{\alpha}^{\mathrm{QKD}}$ as

$$h_{\alpha}^{\mathrm{QKD}} \geq \inf_{\mathbf{q}\in S_{\mathrm{acc}}} \inf_{\substack{V_{\mathcal{E}_{\mathrm{real}}}, \\ V_{\mathcal{E}_{\mathrm{real}}}^{\dagger}V_{\mathcal{E}_{\mathrm{real}}}=\mathbb{I}_{A'}}} \left(\frac{\alpha}{\alpha-1} D\big(\mathbf{q}\big\|\boldsymbol{\nu}_{\widehat{C}}\big) + q(\mathtt{gen})\widetilde{H}_{\alpha}^{\downarrow}(S|\tilde{T}EE_{\mathrm{sou}})_{\nu_{|\mathtt{gen}}} \right), \tag{10.4.19}$$

where

$$\nu_{S\tilde{T}\widehat{C}EE_{\mathrm{sou}}|\mathtt{gen}} = \mathcal{M}_{\mathrm{QKD}}(V_{\mathcal{E}_{\mathrm{real}}}V_{\Xi}\tau_{AA''}V_{\Xi}^{\dagger}V_{\mathcal{E}_{\mathrm{real}}}^{\dagger})_{|\mathtt{gen}}, \tag{10.4.20}$$

$$\boldsymbol{\nu}_{\widehat{C}} = (\gamma\boldsymbol{\nu}_{\widehat{C}|\mathtt{test}}, 1-\gamma)^{T}, \tag{10.4.21}$$

$$\boldsymbol{\nu}_{\widehat{C}|\mathtt{test}} = \boldsymbol{\Phi}[\operatorname{Tr}_{EE_{\mathrm{sou}}}[V_{\mathcal{E}_{\mathrm{real}}}V_{\Xi}\tau_{AA'}V_{\Xi}^{\dagger}V_{\mathcal{E}_{\mathrm{real}}}^{\dagger}]]. \tag{10.4.22}$$

Finally, by optimizing over $V_{\mathcal{E}_{\mathrm{vir}}} \coloneqq V_{\mathcal{E}_{\mathrm{real}}}V_{\Xi}$ instead of only $V_{\mathcal{E}_{\mathrm{real}}}$, the theorem statement follows because the minimization runs over a bigger set that only decreases the infimum further.

The proof for $h_{\alpha}^{\uparrow,\mathrm{QKD}}$ follows exactly the same steps replacing $\widetilde{H}_{\alpha}^{\downarrow}$ with $\widetilde{H}_{\alpha}^{\uparrow}$ appropriately, while noting that the entropy $\widetilde{H}_{\alpha}^{\uparrow}$ still satisfies the required data-processing inequality. □

10.5 Qubit BB84 and Comparisons with other Proof Techniques

Before we turn our focus to more advanced protocols, in this section we present a comparison between our variable-length key rates for the qubit BB84 protocol and variable-length key rates obtained in the EUR framework. The variable-length key rates we obtain from

our techniques use the $\widetilde{H}_\alpha^{\downarrow,f}$-normalization constant $\kappa^\downarrow$. Thus, technically the results shown here could still be slightly improved.

For the comparison, we present EUR-based variable-length key rates from Ref. [24] incorporating the improvements from [83], which at the time of this writing should amount to the best known variable-length key rates for this protocol. Other works using the phase error correction approach, such as e.g. [95], would give similar results.

Our protocol choices are the following. The Z-basis is always used for key generation rounds and the X-basis is always used for test rounds. Hence, with probability $1-\gamma = p_z$ Alice sends a state in the Z-basis and with $\gamma = p_x$ she sends a state in the X basis. Bob in turn measures his incoming states with the same probabilities in the Z- and X-basis, respectively.

For the error correction cost we assume

$$\lambda_{\mathrm{EC}} = n p_{\mathrm{sift}}^{\mathrm{hon}} f_{\mathrm{EC}} H(X|Y;\mathrm{sift})_{\mathrm{hon}}, \tag{10.5.1}$$

with $f_{\mathrm{EC}} = 1.1$ and $p_{\mathrm{sift}}^{\mathrm{hon}} = (1-\eta)(1-\gamma)^2$ being the probability of a signal surviving the sifting procedure. Furthermore, we chose the security parameters $\varepsilon_{\mathrm{PA}} = \varepsilon_{\mathrm{EV}} = \frac{1}{2}10^{-80}$ for a total security of $\varepsilon_{\mathrm{sec}} = 10^{-80}$. This choice is motivated by the Advanced Encryption Standard with a 256-bit key, in short AES256. The probability of randomly guessing such a key used for AES256 is $2^{-256} \leq 10^{-80}$.

We model channel loss with $\eta = 10^{-\frac{\zeta_{\mathrm{hon}}}{10\,\mathrm{dB}}}$ and depolarization with the channel

$$\mathcal{E}_{\mathrm{depol}}[\rho] \coloneqq \left(1 - p_{\mathrm{hon}}^{\mathrm{depol}}\right)\rho + p_{\mathrm{hon}}^{\mathrm{depol}} \frac{\mathbb{I}}{2}, \tag{10.5.2}$$

where for this example we chose $p_{\mathrm{hon}}^{\mathrm{depol}} = 0.03$. Finally, to calculate the key rates of our work, we use the Kraus operators for the maps $\mathcal{G}$, $\mathcal{Z}$ and states $|\xi^t\rangle$, $|\xi^t\rangle$ as stated in Appendix C.3.2.

The resulting key rates for both this work and the comparison with EUR-based security proofs can be seen in Fig. 10.2. For both approaches we optimized the testing probability γ and for our work additionally the Rényi parameter $\alpha \in (1,2)$.

Especially for small block sizes our approach significantly outperforms the EUR key rates. Although not shown explicitly, we are even able to achieve a positive key rate with $n = 10^4$ signals, but only at zero loss, hence we omitted this data point for clarity of the plot. With larger block sizes, the differences become smaller and smaller and both techniques converge to more or less the same key rates.

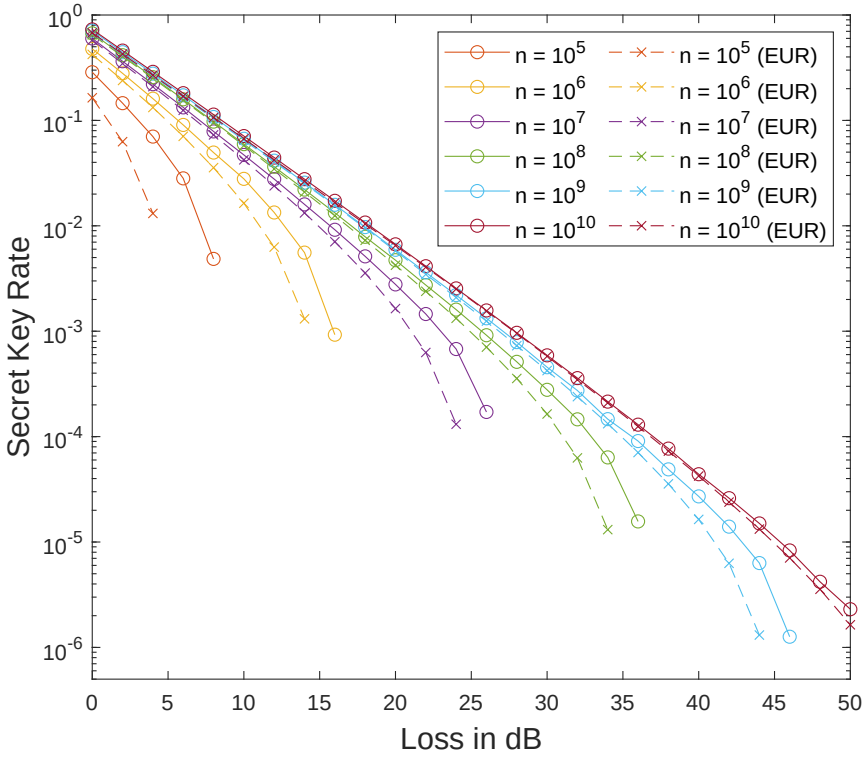


Figure 10.2: Results for variable-length secret key rates of the qubit BB84 protocol plotted against the channel loss in dB and a constant depolarization with $p_{\text{hon}}^{\text{depol}} = 0.03$ with a varying number of total signals sent $n = 10^5, \ldots, 10^{10}$. The security parameters, and error correction efficiency were chosen as $\varepsilon_{\text{PA}} = \varepsilon_{\text{EV}} = \frac{1}{2}10^{-80}$ and $f_{\text{EC}} = 1.1$. As a comparison, we showcase results using the EUR approach of [24] incorporating the improvements from [83]. For both approaches, we optimize the probability γ of using the X-basis in each round, and for our work we additionally optimize the Rényi parameter α. (Figure 2 from Ref. [4])

Additionally, Fig. 10.3 presents a comparison between the results of this chapter, EUR-based key rates, and the postselection technique for the same qubit BB84 protocol as described above. We plot the minimum number of signals required to achieve a non-zero key rate, as well as the key rate for $n = 10^{10}$ signals. Notably, our results yield non-zero key rates with two orders of magnitude fewer signals than the postselection technique and one order of magnitude fewer than the EUR-based technique.

Furthermore, consistent with Fig. 10.2, our method outperforms both reference techniques at $n = 10^{10}$ signals. The advantage over the postselection technique is even more

pronounced than that over the EUR-based rates.

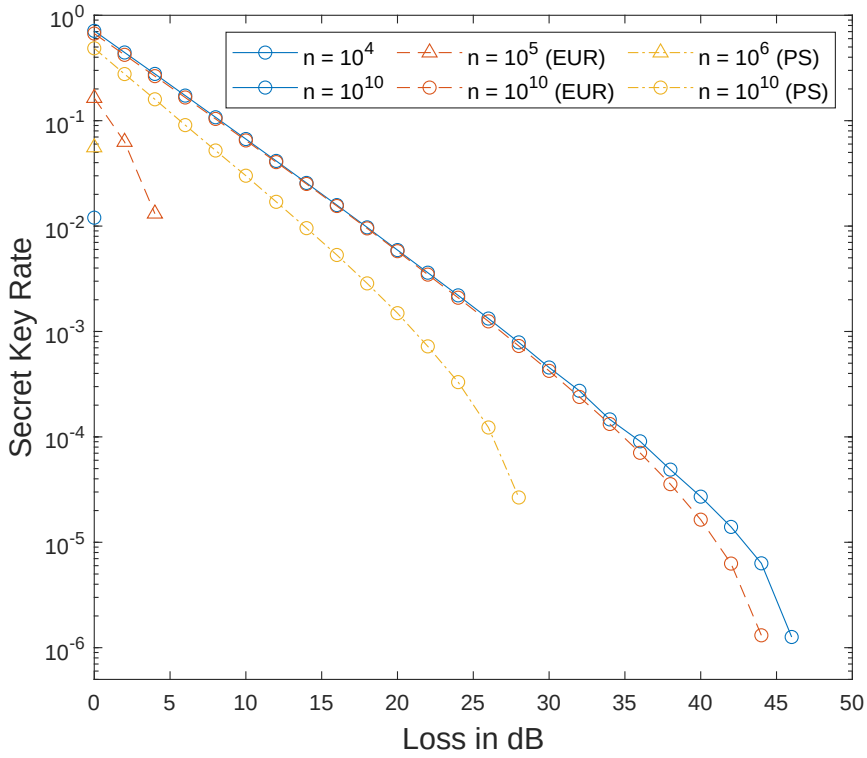


Figure 10.3: Comparison of the variable-length secret key rates for the qubit BB84 protocol. We compare our approach against the EUR [24] and postselection technique [27], plotted against channel loss in dB for a constant depolarization of $p_{\text{hon}}^{\text{depol}} = 0.03$. The lines with circles indicate the key rates for $n = 10^{10}$ signals sent, while the lines with triangles show the minimum total signals sent required to achieve a non-zero key rate. As before, the security parameters and error correction efficiency were chosen as $\varepsilon_{\text{PA}} = \varepsilon_{\text{EV}} = \frac{1}{2}10^{-80}$ and $f_{\text{EC}} = 1.1$. For all approaches, we optimize the testing probability γ of using the X-basis in each round; for our work, we additionally optimize the Rényi parameter α.

In the context of qubit protocols, both comparisons substantiate our claim of achieving the best key rates to date. Moreover, in contrast to EUR- or complementarity-based security proofs, our framework allows for the incorporation of arbitrary detection setups.

10.6 Extension to Block-Diagonal States

In this section, we derive the application of Corollary 10.1.2 and Theorem 8.3.3 to e.g. decoy-state protocols. Theorem 8.2.3 as it is stated can already be validly applied to decoy-state protocols; however, the system A' at least would have an infinite-dimensional Hilbert space, albeit a separable one. WCP sources exhibit a special block-diagonal structure in the photon number. Hence, to solve the issue of infinite dimensions and simplify calculations, we exploit this block-diagonal structure exhibited in the signal states. The approach chosen here is similar to other works; see e.g. [2] in the GEAT framework, and [66] in the asymptotic limit.

Therefore, let us assume:

(1) The signal states $\{\sigma_i\}_{i=1\ldots d_A}$ are simultaneously block-diagonal in some basis with block numbers $m = 0, 1\ldots$, such that we can write $|\xi^g\rangle$ and $|\xi^t\rangle$ as

$$|\xi^g\rangle_{A_S AA'} = \sum_{m=0}^{\infty} \sqrt{p(m|\texttt{gen})}\, |m\rangle_{A_S} \left|\xi^g_{|m}\right\rangle_{AA'}, \tag{10.6.1}$$

$$\left|\xi^t\right\rangle_{A_S AA'} = \sum_{m=0}^{\infty} \sqrt{p(m|\texttt{test})}\, |m\rangle_{A_S} \left|\xi^t_{|m}\right\rangle_{AA'}. \tag{10.6.2}$$

(2) In addition to the system A, the system A_S (often called a shield system) also remains with Alice and is inaccessible to Eve.

(3) Eve can perform a quantum non-demolishing (QND) measurement of system A' and learn the block-diagonal structure without disturbing the state.

Note that assumption 3 is made without loss of generality. This can be seen with the following construction based on a decomposition for WCP sources in [66, App. B], which is also valid for any block-diagonal state where Eve could perform a QND measurement.

Following Ref. [66], for a single round, Eve's attack in Stinespring form, including a QND measurement, can be written as

$$V_{A'\to BE\tilde{E}} = \sum_m V^m_{A'\to BE} \Pi^m_{A'} \otimes |m\rangle_{\tilde{E}}, \tag{10.6.3}$$

where Π^m is a projector for the QND measurement and register $\tilde{E}$ stores the block number m. On the other hand, if Eve does not attempt to measure the block number, her attack acts the same on each block, i.e. $V^m_{A'\to BE} = V_{A'\to BE}$ for all m.

Hence, for states satisfying assumption 1, any feasible point in Eq. (10.1.30), characterizing Eve's channel applied to system A', is also a feasible point in an optimization over direct-sum channels. Thus, we only further reduce the infimum when the optimization is performed over direct-sum channels, and assumption 3 does not restrict Eve's attack.

In summary, the assumptions 1–3 are naturally satisfied for WCP sources, and have been exploited in many works, e.g. [11–13], to formulate the first analytical decoy-state methods and have also been extended to numerical methods [70, 71]. We will discuss how these assumptions are met in more detail in Section 10.7.

Additionally, let us also quickly discuss how one can satisfy assumption 1 if the states are not block-diagonal in the same basis and, thus Eve cannot perform a QND measurement. In this case, one can apply a source-map [68] that announces the block number m to Eve, which intuitively only reduces the infimum. In Section 10.4, we formally show that such source maps can be incorporated and we make use of a source-map for this exact purpose in Section 10.10.

When there are shield systems, we slightly modify the definition of the single-round channel $\mathcal{M}_{\mathrm{QKD}}$ so that it acts on $A_S AB$ rather than just AB. In most scenarios, we take $\mathcal{M}_{\mathrm{QKD}}$ to just apply the operations already described in Definition 8.2.1 on AB, while simply tracing out A_S. However, for the sake of generality we note that it can also be allowed to perform joint operations across all the registers $A_S AB$ if necessary. Then when applying the MEAT, we take the marginal constraints to be imposed on all the $A_S A$ registers, rather than just the A registers — this is valid due to assumption 2, i.e. Eve cannot act on the $A_S A$ registers. Finally, when performing the single-round analysis, we again make the observation that we can "undo" the source replacement argument, viewing a minimization over states $\rho_{A_S ABE}$ satisfying a marginal constraint on $A_S A$ as being equivalent to minimizing over Eve's channel acting on A'.

Now, we will incorporate the block-diagonal structure into the optimization problem for h_α^{QKD} and the $\widetilde{H}_\alpha^{\downarrow,f}$-normalization constant $\kappa^{\downarrow}$. As a consequence of assumption 3 and the discussion above, Eve's channel $\mathcal{E}$ can be written as a direct sum acting on each block m separately,

$$\mathcal{E} = \bigoplus_{m=0}^{\infty} \mathcal{E}_m, \tag{10.6.4}$$

or equivalently, in terms of Choi states

$$J = \bigoplus_{m=0}^{\infty} J_m, \tag{10.6.5}$$

see for example [2, Sec. 7.2]. Moreover, assumptions 1–3 imply that the state $\rho_{A_S AB|\mathtt{test}}$ has the form

$$\rho_{A_S AB|\mathtt{test}} = \sum_{m=0}^{\infty} p(m|\mathtt{test})\, |m\rangle\langle m|_{A_S} \otimes \rho_{AB|\mathtt{test},m}, \tag{10.6.6}$$

and similarly for $\rho_{A_S AB|\mathtt{gen}}$. An equivalent decomposition was also found in [66, App. B].

Next, we will simplify the expression for h_α^{QKD} in Eq. (10.1.30) and the $\widetilde{H}_\alpha^{\downarrow,f}$-normalization constant $\kappa^\downarrow$ such that only finite-dimensional spaces and finite sums are involved. We will start with the entropy and later simplify the constraints.

10.6.1 Bounding the Entropy

Since, by assumption 3, Eve can always perform a QND measurement, we can assume her quantum side-information to include a classical register M determining the block-number m. Thus, by exploiting the properties of Rényi entropies when conditioned on classical registers [92, Sec. III.B.2 and Prop. 9] (or [47, Prop. 5.1]), we find for the entropy in a generation round

$$\widetilde{H}_\alpha^\downarrow(S|\tilde{T}EM)_{\nu^\rho_{|\mathtt{gen}}} = \frac{1}{1-\alpha} \log \left(\sum_{m=0}^{\infty} p(m|\mathtt{gen}) 2^{(1-\alpha)\widetilde{H}_\alpha^\downarrow(S|\tilde{T}E)_{\nu^\rho_{|m,\mathtt{gen}}}} \right). \tag{10.6.7}$$

Because $\nu_{S\tilde{T}E|m,\mathtt{gen}}$ is classical on S (and thus is a separable state across S and $\tilde{T}E$), we have $\widetilde{H}_\alpha^\downarrow(S|\tilde{T}E)_{\nu^\rho_{|m,\mathtt{gen}}} \geq 0$ by [47, Lemma 5.2]. Hence, picking a cutoff $N_{\mathrm{ent}} \in \mathbb{N}_0$ allows us to bound the entropy by

$$\begin{aligned} \widetilde{H}_\alpha^\downarrow(S|\tilde{T}EM)_{\nu^\rho_{|\mathtt{gen}}} \geq \frac{1}{1-\alpha} \log \Bigg\{ & \sum_{m \leq N_{\mathrm{ent}}} p(m|\mathtt{gen}) 2^{(1-\alpha)\widetilde{H}_\alpha^\downarrow(S|\tilde{T}E)_{\nu^\rho_{|m,\mathtt{gen}}}} \\ & + 1 - \sum_{m \leq N_{\mathrm{ent}}} p(m|\mathtt{gen}) \Bigg\}, \end{aligned} \tag{10.6.8}$$

where the inequality can be made arbitrarily tight by increasing N_{ent}.

Finally, in line with Eq. (10.1.29), we aim to define $g_\alpha^{N_{\mathrm{ent}}}$, a version of g_α incorporating the cutoff. Therefore, similarly to Eq. (10.1.26), let us define the CPTNI maps $\chi_{g/t}^m : A' \to A$,

$$\begin{aligned} \chi_g^m[X] &:= \mathrm{Tr}_{A'} \left[\left(\mathbb{I}_A \otimes X^{T_{A'}} \right) \left| \xi^g_{|m} \right\rangle\!\left\langle \xi^g_{|m} \right| \right], \\ \chi_t^m[X] &:= \mathrm{Tr}_{A'} \left[\left(\mathbb{I}_A \otimes X^{T_{A'}} \right) \left| \xi^t_{|m} \right\rangle\!\left\langle \xi^t_{|m} \right| \right], \end{aligned} \tag{10.6.9}$$

which are now based on the states conditioned on block m. This lets us write

$$\nu_{S\tilde{T}\widehat{C}E|m,\mathtt{gen}} = \mathcal{M}_{\mathrm{QKD}} \circ \chi_g^m[J_m], \tag{10.6.10}$$

for all $m \in \mathbb{N}_0$. Thus, we define $g_\alpha^{N_{\mathrm{ent}}}$ incorporating the cutoff N_{ent} by

$$g_\alpha^{N_{\mathrm{ent}}}(J_0, \ldots, J_{N_{\mathrm{ent}}}) \coloneqq \frac{1}{1-\alpha} \log \left\{ 1 - \sum_{m \leq N_{\mathrm{ent}}} p(m|\mathtt{gen}) \left(1 - 2^{(1-\alpha)\widetilde{H}_\alpha^\downarrow(S|\tilde{T}E)_{\nu(J_m)|m,\mathtt{gen}}} \right) \right\}, \tag{10.6.11}$$

where each $\nu(J_m)_{|m,\mathtt{gen}}$ depends on J_m as stated above in Eq. (10.6.10).

10.6.2 Bounding the Constraints

Next, we turn our attention to the constraints. Here, the only constraint in Eq. (10.1.30) depending on Eve's channel is $\boldsymbol{\nu}_{\widehat{C}|\mathtt{test}} = \boldsymbol{\Psi}[J]$. In other words, we need to reformulate $\boldsymbol{\Psi}$ using the direct-sum structure in Eve's channel.

Similarly to the discussion preceding the definition of the QKD channel $\mathcal{M}_{\mathrm{QKD}}$, see Eq. (8.2.7) and below, we require the POVM elements conditioned on m and `test` (or `gen`). Again, under the generic source replacement scheme these conditional POVM elements remain the same as the unconditional ones, as they are projectors. However, under a Schmidt decomposition the POVM elements not conditioned on block m may contain factors proportional to the probability of sending block m. Then, one needs to appropriately incorporate the conditioning as described in Appendix C.3.1.

For the discussion presented here, we again assume Alice's POVM elements are projectors, however, for clarity we indicate the conditioning on m by

$$\{M_{\alpha,x}^{A|\mathtt{test},m}\}_{\alpha \in \mathcal{C}^{A,\mathtt{test}}, x \in \mathcal{X}_\alpha^{\mathtt{test}}}, \quad \text{and} \quad \{M_{\alpha,x}^{A|\mathtt{gen},m}\}_{\alpha \in \mathcal{C}^{A,\mathtt{gen}}, x \in \mathcal{X}_\alpha^{\mathtt{gen}}}. \tag{10.6.12}$$

Therefore, let us define the maps $\boldsymbol{\Phi}_m$ generating the statistics conditioned on block m,

$$\boldsymbol{\Phi}_m[\sigma] \coloneqq \sum_{\hat{c} \in \widehat{\mathcal{C}}_{\setminus \mathrm{gen}}} \sum_{\substack{(\alpha,x,\beta,y) \\ \in \phi^{-1}(\hat{c})}} \mathrm{Tr}\left[\left(M_{\alpha,x}^{A|\mathtt{test},m} \otimes M_{\beta,y}^B \right) \sigma \right] \hat{e}_{\hat{c}}, \tag{10.6.13}$$

and their concatenation with the maps χ_t^m as

$$\boldsymbol{\Psi}_m[J_m] \coloneqq \boldsymbol{\Phi}_m \circ \chi_t^m[J_m]. \tag{10.6.14}$$

Then, a straightforward calculation using Eq. (10.6.6) yields

$$\boldsymbol{\nu}_{\widehat{C}|\mathtt{test}} = \boldsymbol{\Psi}[J] = \sum_{m=0}^{\infty} p(m|\mathtt{test})\boldsymbol{\Psi}_m[J_m]. \tag{10.6.15}$$

To consider only finitely many parts of the sum in the constraints, we pick another cutoff $N_{\text{cons}} \in \mathbb{N}_0$. Each entry of the vector $\boldsymbol{\Psi}_m[J_m]$ must be greater than zero, thus

$$\boldsymbol{\nu}_{\widehat{C}|\mathtt{test}} \geq \sum_{m \leq N_{\text{cons}}} p(m|\mathtt{test})\boldsymbol{\Psi}_m[J_m]. \tag{10.6.16}$$

On the other hand, each entry in $\boldsymbol{\Psi}_m[J_m]$ must be smaller than the probability of Alice sending a particular state determined by (α, x) (bit value and announcement) conditioned on block m and $\mathtt{test}$. Hence, the outcomes in test rounds $\boldsymbol{\nu}_{\widehat{C}|\mathtt{test}}$ can be upper bounded by

$$\boldsymbol{\nu}_{\widehat{C}|\mathtt{test}} \leq \sum_{m \leq N_{\text{cons}}} p(m|\mathtt{test})\boldsymbol{\Psi}_m[J_m] + 1 - \sum_{\substack{\hat{c} \in \widehat{\mathcal{C}}_{\setminus \text{gen}}, \\ (\alpha,x) \in \phi^{-1}(\hat{c})}} \sum_{m \leq N_{\text{cons}}} p(\alpha, x, m|\mathtt{test}) \hat{e}_{\hat{c}}, \tag{10.6.17}$$

and to simplify the notation, we define the vector

$$\mathbf{p}_{\text{tot}} \coloneqq \sum_{\hat{c} \in \widehat{\mathcal{C}}_{\setminus \text{gen}}} \sum_{\substack{(\alpha,x) \\ \in \phi^{-1}(\hat{c})}} \sum_{m \leq N_{\text{cons}}} p(\alpha, x, m|\mathtt{test})\ \hat{e}_{\hat{c}}. \tag{10.6.18}$$

In summary, we can bound $\boldsymbol{\nu}_{\widehat{C}|\mathtt{test}}$ by

$$\boldsymbol{\nu}_{\widehat{C}|\mathtt{test}} \leq \sum_{m \leq N_{\text{cons}}} p(m|\mathtt{test})\boldsymbol{\Psi}_m[J_m] + 1 - \mathbf{p}_{\text{tot}}, \tag{10.6.19}$$

$$\boldsymbol{\nu}_{\widehat{C}|\mathtt{test}} \geq \sum_{m \leq N_{\text{cons}}} p(m|\mathtt{test})\boldsymbol{\Psi}_m[J_m], \tag{10.6.20}$$

and also these bounds can be made arbitrarily tight by increasing N_{cons}.

10.6.3 Reformulated Optimization Problems

Now, we combine both Section 10.6.1 and Section 10.6.2, such that we can bound h_α^{QKD} by

$$
\begin{aligned}
h_\alpha^{\mathrm{QKD}} \geq \inf_{\substack{\mathbf{q}\in S_{\mathrm{acc}}, \boldsymbol{\nu}_{\widehat{C}}\in\mathbb{P}(\widehat{\mathcal{C}}) \\ J_m\in\mathrm{Pos}(A'B)}} & \left(\frac{\alpha D\big(\mathbf{q}\big\|\boldsymbol{\nu}_{\widehat{C}}\big)}{\alpha-1} + q(\mathtt{gen}) g_\alpha^{N_{\mathrm{ent}}}(J_0 \dots J_{N_{\mathrm{ent}}}) \right) \\
\text{s.t. } & \mathrm{Tr}_B[J_m] = \mathbb{I}_{A'}, \\
& \boldsymbol{\nu}_{\widehat{C}} = (\gamma \boldsymbol{\nu}_{\widehat{C}|\mathtt{test}}, 1-\gamma)^T, \\
& \boldsymbol{\nu}_{\widehat{C}|\mathtt{test}} \leq \sum_{m\leq N_{\mathrm{cons}}} p(m|\mathtt{test}) \boldsymbol{\Psi}_m[J_m] + 1 - \mathbf{p}_{\mathrm{tot}}, \\
& \boldsymbol{\nu}_{\widehat{C}|\mathtt{test}} \geq \sum_{m\leq N_{\mathrm{cons}}} p(m|\mathtt{test}) \boldsymbol{\Psi}_m[J_m].
\end{aligned}
\tag{10.6.21}
$$

Computing variable-length key rates requires a tradeoff function f and the $\widetilde{H}_\alpha^{\downarrow,f}$-normalization constant $\kappa^\downarrow$. Regarding the tradeoff function, similar to Eq. (10.2.3), one can add a constraint and an auxiliary variable $\boldsymbol{\lambda}$ and solve for its dual solution to find the tuple $\mathbf{f}$ defining the best choice of tradeoff function f. For the $\widetilde{H}_\alpha^{\downarrow,f}$-normalization constant $\kappa^\downarrow$, one can follow exactly the same steps as laid out above for h_α^{QKD} and finds

$$
\begin{aligned}
\kappa^\downarrow \geq \inf_{\substack{J_m\in\mathrm{Pos}(A'B), \\ \boldsymbol{\nu}\in\mathbb{P}(\widehat{\mathcal{C}}_{\setminus\mathrm{gen}})}} & \left(\frac{1}{1-\alpha} \log \Big(\gamma \sum_{c\in\widehat{\mathcal{C}}_{\setminus\mathrm{gen}}} \boldsymbol{\nu}(\hat{c}) 2^{(\alpha-1)f(\hat{c})} \right. \\
& \left. \left. + (1-\gamma)\, 2^{-(\alpha-1)\left(g_\alpha^{N_{\mathrm{ent}}}(J_0,\dots J_{N_{\mathrm{ent}}}) - f(\mathtt{gen})\right)} \Big) \right) \\
\text{s.t. } & \mathrm{Tr}_B[J_m] = \mathbb{I}_{A'}, \\
& \boldsymbol{\nu} \leq \sum_{m\leq N_{\mathrm{cons}}} p(m|\mathtt{test}) \boldsymbol{\Psi}_m[J_m] + 1 - \mathbf{p}_{\mathrm{tot}}, \\
& \boldsymbol{\nu} \geq \sum_{m\leq N_{\mathrm{cons}}} p(m|\mathtt{test}) \boldsymbol{\Psi}_m[J_m].
\end{aligned}
\tag{10.6.22}
$$

A similar expression can be found for the normalization constant in terms of the mixed f-weighted entropy $\widetilde{H}_\alpha^{\supset,f}$.

Most importantly, we emphasize that the inequalities in both Eq. (10.6.21) and Eq. (10.6.22) can be made arbitrarily tight, by increasing the cutoffs N_{ent} and N_{cons}. Thus, with enough such terms, the fixed-length and variable-length key rates resulting from these lower bounds will converge to the best possible values available under this general proof approach.

10.7 Decoy-state Protocols

After having laid out the general structure in the previous section, we now focus on decoy-state protocols. We note that decoy-state protocols possess even more structure than the generic format required to reach Eq. (10.6.21) and one can exploit this to further simplify the formulations of Eq. (10.6.21) and Eq. (10.6.22). Thus, the goal of this section is to verify the assumptions 1–3 and simplify the optimization problems for decoy-state protocols.

We start with the assumptions 1–3; in particular the block-diagonal structure of the signal states in Eq. (10.6.1). As before, we include a shield system [58], and apply the source replacement scheme [56, 57]. Furthermore, we note that for weak coherent pulses the states in system A' sent to Bob with m photons do not depend on the intensity. Therefore, the states Alice prepares in generation rounds can be written as

$$|\phi^g\rangle_{A\bar{A}A_SA'} = \sum_{m=0}^{\infty} \sum_{a,\mu} \sqrt{p(a,\mu,m|\texttt{gen})}\, |a\rangle_A |\mu\rangle_{\bar{A}} \otimes |m\rangle_{A_S} \left|s_m^{(a)}\right\rangle_{A'}, \tag{10.7.1}$$

where we abbreviated Alice's choice (α, x) as a. A similar expression upon appropriate replacements holds for `test` rounds. Again, we note that for weak coherent pulses the signal states $\left|s_m^{(a)}\right\rangle_{A'}$ only depend on a, and therefore, we stated no explicit dependence on the intensity μ. Hence, assumption 1 is already satisfied.

The fact that the signal states $\left|s_m^{(a)}\right\rangle_{A'}$ do not depend on the intensity, is one crucial difference from the generic block-diagonal framework presented above and we will exploit this fact in the following simplifications. For this purpose, we also kept the intensity choice in a separate register $\bar{A}$. In the context of defining the channels $\mathcal{M}_{\text{QKD}}$ and applying the MEAT, we define $\mathcal{M}_{\text{QKD}}$ to act on $A\bar{A}A_SB$, and apply the marginal constraint to the $A\bar{A}A_S$ registers.

Next, assumption 2 is trivially satisfied through the construction of the shield system. Finally, due to the photon number splitting attack [8, 9] we can assume without loss of generality that for WCP sources Eve always performs a QND measurement of the photon number first and then applies an attack based on the photon number. Thus, assumption 3 is satisfied as well, and by [66, App. B] the channel can immediately be decomposed in the block-diagonal form.

10.7.1 Simplifications of the Optimization Problems

Now, we turn our attention to the simplifications specific to decoy-state protocols. The goal of the reformulations here is to find an equivalent form of the constraints in Eq. (10.6.21) that instead uses the statistics conditioned on each intensity choice μ_j and has a similar form to commonly used decoy-state methods, e.g. [71]. We again highlight that this reformulation heavily relies on fully phase-randomized weak coherent pulses and is in general not possible.

First, we note that the vector $\boldsymbol{\Psi}_m[J_m]$, is ordered by the announcements μ and $a = (\alpha, x)$ for Alice and $b = (\beta, y)$ for Bob. Hence, we can write for each component of $\boldsymbol{\Psi}_m[J_m]$

$$\begin{aligned}\left(\Psi_m[J_m]\right)_{\mu,a,b} &= \mathrm{Tr}\left[\left(M_a^{A|\mathtt{test},m} \otimes |\mu\rangle\langle\mu| \otimes M_b^B\right) \chi_m^t[J_m]\right] \\ &= \mathrm{Tr}\left[\left(M_a^{A|\mathtt{test},m} \otimes M_b^B\right) \langle\mu| \chi_m^t[J_m] |\mu\rangle\right]. \end{aligned} \tag{10.7.2}$$

To simplify the inner product $\langle\mu| \chi_m^t[J_m] |\mu\rangle$ further, let us define the states conditioned on Alice sending intensity μ,

$$\left|\xi^t_{|m,\mu}\right\rangle_{AA'} \coloneqq \sum_a \sqrt{p(a|m,\mu,\mathtt{test})}\, |a\rangle_A \left|s_m^{(a)}\right\rangle_{A'}, \tag{10.7.3}$$

where we again omitted any dependence on the intensity μ of the states sent to Bob due to the properties of WCP sources. With these states we define the CPTNI map $\chi_t^{m,\mu}$ as

$$\chi_t^{m,\mu}[X] \coloneqq \mathrm{Tr}_{A'}\left[\left(\mathbb{I}_A \otimes X^{T_{A'}}\right) \left|\xi^t_{|m,\mu}\right\rangle\!\left\langle\xi^t_{|m,\mu}\right|\right], \tag{10.7.4}$$

which allows us to write

$$\langle\mu| \chi_m^t[J_m] |\mu\rangle = p(\mu|m,\mathtt{test})\chi_t^{m,\mu}[J_m]. \tag{10.7.5}$$

Furthermore, we define the map generating the statistics conditioned on an intensity μ and photon number m equivalently to before, see Eq. (10.1.28),

$$\boldsymbol{\Psi}_{m,\mu}[J_m] \coloneqq \Phi_m \circ \chi_t^{m,\mu}[J_m]. \tag{10.7.6}$$

Then, we find for the μ component of each $\boldsymbol{\Psi}_m[J_m]$

$$\begin{aligned} p(m|\mathtt{test})(\boldsymbol{\Psi}_m[J_m])_\mu &= p(m|\mathtt{test})p(\mu|m,\mathtt{test})\boldsymbol{\Psi}_{m,\mu}[J_m] \\ &= p(\mu|\mathtt{test})p(m|\mu,\mathtt{test})\boldsymbol{\Psi}_{m,\mu}[J_m]. \end{aligned} \tag{10.7.7}$$

Thus, we can rewrite the constraints of Eq. (10.6.19) additionally conditioned on the intensity μ_j as

$$\begin{aligned}\boldsymbol{\nu}_{\widehat{C}|\mathtt{test},\mu_j} &\leq \sum_{m\leq N_{\text{cons}}} p(m|\mathtt{test},\mu_j)\boldsymbol{\Psi}_{m,\mu_j}[J_m] + 1 - \mathbf{p}_{\text{tot}|\mu_j},\\ \boldsymbol{\nu}_{\widehat{C}|\mathtt{test},\mu_j} &\geq \sum_{m\leq N_{\text{cons}}} p(m|\mathtt{test},\mu_j)\boldsymbol{\Psi}_{m,\mu_j}[J_m],\end{aligned} \tag{10.7.8}$$

where $\mathbf{p}_{\text{tot}|\mu_j}$ is defined as in Eq. (10.6.18) with an additional conditioning on the intensity μ_j.

Finally, for typical decoy-state protocols, signals with the same encoding $a = (\alpha, x)$ but different intensities are usually mapped to the same key. This is especially true if only the signal intensity is used in key generation rounds. In this case, we define

$$|\xi^g\rangle_{A_S AA'} = \sum_{m=0}^{\infty} \sqrt{p(m|\mathtt{gen})}\,|m\rangle_{A_S}\left|\xi^g_{|m}\right\rangle_{AA'}, \tag{10.7.9}$$

as Alice's signal state in generation rounds, where

$$\left|\xi^g_{|m}\right\rangle = \sum_a \sqrt{p(a|m,\mathtt{gen})}\,|a\rangle_A\left|s^{(a)}_m\right\rangle_{A'}. \tag{10.7.10}$$

Crucially, here $p(m|\mathtt{gen}) = \sum_{\mu_j} p(m,\mu_j|\mathtt{gen})$ is the total probability of sending m photons in a key generation round. Then, the objective function $g_\alpha^{N_{\text{ent}}}$ is defined in terms of the states $|\xi^g\rangle$. If this assumption is not satisfied, we simply use $|\phi^g\rangle$.

Incorporating these reformulations into Eq. (10.6.21) we find for h_α^{QKD}

$$\begin{aligned}h_\alpha^{\text{QKD}} \geq \inf_{\substack{\mathbf{q}\in S_{\text{acc}},\boldsymbol{\nu}_{\widehat{C}}\in\mathbb{P}(\widehat{\mathcal{C}})\\ J_m\in\text{Pos}(A'B)}} &\left(\frac{\alpha D(\mathbf{q}\|\boldsymbol{\nu}_{\widehat{C}})}{\alpha-1} + q(\mathtt{gen})g_\alpha^{N_{\text{ent}}}(J_0\ldots J_{N_{\text{ent}}})\right)\\ \text{s.t. }\ &\text{Tr}_B[J_m] = \mathbb{I}_{A'},\\ &\boldsymbol{\nu}_{\widehat{C}} = (\gamma p(\mu_s|\mathtt{test})\boldsymbol{\nu}_{\widehat{C}|\mathtt{test},\mu_s},\ldots,1-\gamma)^T,\\ &\boldsymbol{\nu}_{\widehat{C}|\mathtt{test},\mu_j} \leq \sum_{m\leq N_{\text{cons}}} p(m|\mathtt{test},\mu_j)\boldsymbol{\Psi}_{m,\mu_j}[J_m] + 1 - \mathbf{p}_{\text{tot}|\mu_j},\\ &\boldsymbol{\nu}_{\widehat{C}|\mathtt{test},\mu_j} \geq \sum_{m\leq N_{\text{cons}}} p(m|\mathtt{test},\mu_j)\boldsymbol{\Psi}_{m,\mu_j}[J_m].\end{aligned} \tag{10.7.11}$$

This formulation of h_α^{QKD} already improves the numerical precision because the probability vector $\boldsymbol{\nu}_{\widehat{C}|\mathtt{test},\mu_j}$ is conditioned on the intensity μ_j. However, this optimization problem would still require solving for N_{cons} Choi states that typically only have a very small influence on the constraints for higher photon numbers.

Alternatively, one could replace the optimization over the Choi state of high photon numbers with an optimization over the probabilities or so-called yields, similar to [2, Sec. 7]. This version is also much closer to common numerical decoy-state methods such as those in Refs. [1, 68, 71] or as presented in Chapter 6.

Hence, to reduce the overhead from optimizing many Choi states, we introduce the yields

$$Y_m^{a,b} \coloneqq \Pr(b|a,\mu,m,\mathtt{test}). \tag{10.7.12}$$

These yields are actually independent of the intensity μ, because the state sent to Bob does not depend on the intensity. Therefore, we can relate the components (a,b) of each vector $\boldsymbol{\Psi}_{m,\mu_j}[J_m]$ to the yields by

$$\begin{aligned}(\Psi_{m,\mu}[J_m])_{a,b} &= p(a,b|m,\mu,\mathtt{test}) \\ &= p(a|m,\mu,\mathtt{test})p(b|a,\mu,m,\mathtt{test}) \\ &= p(a|m,\mu,\mathtt{test})Y_m^{a,b}.\end{aligned} \tag{10.7.13}$$

Moreover, to simplify the notation, we appropriately stack all yields $Y_m^{a,b}$ of the same photon number m into a vector $\mathbf{Y}_m$, such that it holds component-wise

$$\boldsymbol{\Psi}_{m,\mu_j}[J_m] = p(a|m,\mu_j,\mathtt{test})\mathbf{Y}_m, \tag{10.7.14}$$

or equivalently

$$p(m|\mathtt{test},\mu_j)\boldsymbol{\Psi}_{m,\mu_j}[J_m] = p(a,m|\mu_j,\mathtt{test})\mathbf{Y}_m. \tag{10.7.15}$$

Using this relation and inserting it into the previous version of the constraints Eq. (10.7.8),

one finds the fully simplified version of h_α^{QKD} for decoy-state protocols

$$
\begin{aligned}
h_\alpha^{\mathrm{QKD}} \geq \inf_{\substack{\mathbf{q}\in S_{\mathrm{acc}}, \boldsymbol{\nu}_{\widehat{C}}\in\mathbb{P}(\widehat{\mathcal{C}}) \\ J_m\in\mathrm{Pos}(A'B), \\ \mathbf{Y}_m}} & \left(\frac{\alpha D\big(\mathbf{q}\big\|\boldsymbol{\nu}_{\widehat{C}}\big)}{\alpha-1} + q(\mathtt{gen}) g_\alpha^{N_{\mathrm{ent}}}(J_0 \ldots J_{N_{\mathrm{ent}}}) \right) \\
\text{s.t. } & \mathrm{Tr}_B[J_m] = \mathbb{I}_{A'}, \\
& \boldsymbol{\nu}_{\widehat{C}} = (\gamma p(\mu_s|\mathtt{test}) \boldsymbol{\nu}_{\widehat{C}|\mathtt{test},\mu_s}, \ldots, 1-\gamma)^T, \\
& \boldsymbol{\nu}_{\widehat{C}|\mathtt{test},\mu_j} \leq \sum_{m\leq N_{\mathrm{ph}}} p(a,m|\mathtt{test},\mu_j)\mathbf{Y}_m + 1 - \mathbf{p}_{\mathrm{tot}|\mu_j}, \\
& \boldsymbol{\nu}_{\widehat{C}|\mathtt{test},\mu_j} \geq \sum_{m\leq N_{\mathrm{ph}}} p(a,m|\mathtt{test},\mu_j)\mathbf{Y}_m, \\
& p(a|m,\mu_j,\mathtt{test})\mathbf{Y}_m = \boldsymbol{\Psi}_{m,\mu_j}[J_m] \; \forall m \leq N_{\mathrm{ent}}.
\end{aligned}
\tag{10.7.16}
$$

Repeating the same arguments as before for the $\widetilde{H}_\alpha^{\downarrow,f}$-normalization constant $\kappa^{\downarrow}$, one reaches the decoy reformulation:

$$
\begin{aligned}
\kappa^{\downarrow} \geq \inf_{\substack{J_m\in\mathrm{Pos}(A'B), \\ \boldsymbol{\nu}\in\mathbb{P}(\widehat{\mathcal{C}}_{\setminus\mathrm{gen}}), \\ \mathbf{Y}_m}} & \left(\frac{1}{1-\alpha} \log\Big(\gamma \sum_{c\in\widehat{\mathcal{C}}_{\setminus\mathrm{gen}}} \boldsymbol{\nu}(\hat{c}) 2^{(\alpha-1)f(\hat{c})} \right. \\
& \left. \left. + (1-\gamma)\, 2^{-(\alpha-1)\left(g_\alpha^{N_{\mathrm{ent}}}(J_0,\ldots J_{N_{\mathrm{ent}}}) - f(\mathtt{gen})\right)} \Big) \right) \\
\text{s.t. } & \mathrm{Tr}_B[J_m] = \mathbb{I}_{A'}, \\
& \boldsymbol{\nu} = (p(\mu_s|\mathtt{test})\boldsymbol{\nu}_{|\mu_s}, \ldots)^T, \\
& \boldsymbol{\nu}_{|\mu_j} \leq \sum_{m\leq N_{\mathrm{ph}}} p(a,m|\mathtt{test},\mu_j)\mathbf{Y}_m + 1 - \mathbf{p}_{\mathrm{tot}|\mu_j}, \\
& \boldsymbol{\nu}_{|\mu_j} \geq \sum_{m\leq N_{\mathrm{ph}}} p(a,m|\mathtt{test},\mu_j)\mathbf{Y}_m, \\
& p(a|m,\mu_j,\mathtt{test})\mathbf{Y}_m = \boldsymbol{\Psi}_{m,\mu_j}[J_m] \; \forall m \leq N_{\mathrm{ent}}.
\end{aligned}
\tag{10.7.17}
$$

Again, the same steps can be taken to simplify the $\widetilde{H}_\alpha^{\wp,f}$-normalization constant $\kappa^{\wp}$.

10.7.2 Tightness of Block-Diagonal and Decoy Methods

Before we present an example of a decoy-state protocol, let us compare the block-diagonal and decoy methods of the two preceding sections to the improved decoy methods introduced in Chapter 6. Specifically, let us discuss the differences between the fixed-length versions, i.e., between h_α^{QKD} and Eq. (6.1.4) with the bounds on the yields from Eq. (6.2.5). We note that previously, in Ref. [2], we also derived a version of the decoy-state key rate expression that asymptotically achieves the Devetak-Winter formula [96]. However, in the finite-size regime, the GEAT key rates are a lower bound on the key rates presented in this section. Therefore, we will instead discuss the methods shown above in more detail.

To reach the asymptotic limit, one can choose $\alpha = 1 + 1/\sqrt{n}$ and then take the limit of n to infinity. Under this assumption, h_α^{QKD} would converge to the following optimization problem for decoy-state protocols:

$$\begin{aligned}
&\inf_{\substack{J_m \in \mathrm{Pos}(A'B),\\ \mathbf{Y}_m}} q^{\mathrm{hon}}(\mathtt{gen}) \sum_{m=0}^{N_{\mathrm{ent}}} p(m|\mathtt{gen}) H(S|\tilde{T}E)_{\nu|\mathtt{gen},m} \qquad (10.7.18)\\
\text{s.t. }\ & \mathrm{Tr}_B[J_m] = \mathbb{I}_{A'},\\
& \mathbf{q}^{\mathrm{hon}} = (\gamma p(\mu_s|\mathtt{test})\mathbf{q}^{\mathrm{hon}}_{|\mathtt{test},\mu_s}, \ldots, 1-\gamma)^T,\\
& \mathbf{q}^{\mathrm{hon}}_{|\mathtt{test},\mu_j} \leq \sum_{m \leq N_{\mathrm{ph}}} p(a,m|\mathtt{test},\mu_j)\mathbf{Y}_m + 1 - \mathbf{p}_{\mathrm{tot}|\mu_j},\\
& \mathbf{q}^{\mathrm{hon}}_{|\mathtt{test},\mu_j} \geq \sum_{m \leq N_{\mathrm{ph}}} p(a,m|\mathtt{test},\mu_j)\mathbf{Y}_m,\\
& p(a|m,\mu_j,\mathtt{test})\mathbf{Y}_m = \mathbf{\Psi}_{m,\mu_j}[J_m]\ \forall m \leq N_{\mathrm{ent}},
\end{aligned}$$

where each ν_m is generated by applying J_m. The distribution $\mathbf{q}^{\mathrm{hon}}$ is the probability distribution which would occur under the honest channel behaviour. In the notation of Chapter 6, it would hold

$$q^{\mathrm{hon}}_{|\mathtt{test},\mu_j}(a,b) = p(a|\mu_j,\mathtt{test})\gamma^{\mu_j}_{b|a} \qquad (10.7.19)$$

for all of Alice's signals a and Bob's measurement outcomes b.

Besides small differences in the conditioning, this optimization problem combines both the key rate calculation, Eq. (6.1.4), and the yields estimation, Eq. (6.2.5), into one single problem. This has several very profound consequences.

The previous numerical decoy methods Section 6.2 or Refs. [68, 71] first solved an optimization problem to bound the yields and then estimated the key rate given these

bounds. Thus, we call them a two-step process. The refined optimization problem of Eq. (6.2.5) already restricts Eve to only physical channels in contrast to the linear program Eq. (6.1.13) originally from [71]. However, Eve could still achieve the worst case for every yield separately.

Our methods here combine the yield estimation and the key rate calculation such that the adversary cannot achieve this worst case for all observations simultaneously. Moreover, this combined problem even considers correlations between each block or photon number. Additionally, as discussed in [2, Remark 3], the optimization problem above even achieves the Devetak-Winter formula [96] in the limit of $N_{\text{ent}}, N_{\text{ph}} \to \infty$. In contrast, any of the previous decoy-state methods Section 6.2 and Refs. [68, 71] did not achieve this limit nor were able to incorporate correlations.

The fact that we actually solve the correct optimization problem (up to introducing the cutoffs $N_{\text{ent}}, N_{\text{ph}}$) and not a weakened version also translates to the finite-size regime. However, here the comparison is more difficult. The methods shown in Corollary 7.2.2 inherit the sub-optimal decoy methods from the asymptotic regime and additionally use worse proof techniques for the finite-size proof. Thus, it is hard to disentangle the effect of the better decoy-state methods from the effects of the proof technique. However, we conjecture that the effect is only stronger in the finite-size regime.

In summary, while the methods in Section 6.2 already significantly tightened the decoy-state analysis and recovered no-decoy results from [10], they could not reach the Devetak-Winter formula. In contrast, the methods we derived in this section indeed reach the Devetak-Winter formula in the limit and thus are tight. The only remaining improvements could be achieved by having a faster convergence to the asymptotic limit, although it is not obvious how to achieve this.

10.7.3 Example: Active Decoy BB84

In this section, we present results for variable-length secret key rates of a decoy version of BB84 protocol with WCP sources and an active detection setup. This setup allows us to use the squashing map from Refs. [62, 63].

In each round, Alice decides with probability γ if it is a test round and with probability $1-\gamma$ if it is a generation round and always selects the signal intensity μ_s and the Z basis if it was a key generation round. Therefore, as in the qubit example of Section 10.5, Alice's basis choices are given by $p_z = 1-\gamma$ and $p_x = \gamma$. In a test round, Alice always selects the X basis and all possible intensities with equal probability, i.e. $p(\mu|\texttt{test}) = 1/3$. For the

intensity choices we assume

$$\mu_2 = 0.02, \quad \mu_3 = 10^{-3}, \tag{10.7.20}$$

and always optimize the signal intensity μ_s. For the one-decoy protocol we show later, we simply omit the third intensity μ_3.

On the receiver side, Bob, actively chooses the Z and X basis, also with probabilities $p_z = 1 - \gamma$ and $p_x = \gamma$, respectively.

We assume the same security parameters $\varepsilon_{\text{PA}} = \varepsilon_{\text{EV}} = \frac{1}{2}10^{-80}$ for a total security of $\varepsilon_{\text{sec}} = 10^{-80}$ and error correction model with efficiency $f_{\text{EC}} = 1.1$ as in Section 10.5. We model channel loss similarly to the qubit protocol and include a misalignment with an angle of $\theta_{\text{hon}}^{\text{misalign}} = 0.03$. The Kraus operators of the maps $\mathcal{G}$ and $\mathcal{Z}$ and the states can be found in Appendix C.3.3.

For comparison, we present the decoy-state key rates using the EUR approach first presented by Lim et al. [22]. However, we use the results of Ref. [24] again incorporating the improvements from Ref. [83] because it removed the detection efficiency mismatch assumption and corrected other issues [97]. Furthermore, this combination yielded, at the time of writing, the highest variable-length secret key rates for decoy-state protocols[1].

In both approaches, we optimize all free parameters, such as the testing probability γ, the signal intensity μ_s and the Rényi parameter α (only for this work's results). The resulting variable-length key rates can be seen in Fig. 10.4.

Additionally, Fig. 10.5 illustrates the optimal values for the signal intensity, μ_s, across a range of transmitted signals, n, from 10^5 to 10^{11}. The optimal intensity decreases linearly with the channel loss in dB until the secret key rate drops to zero, varying between approximately 0.5 and 1, with two outliers at around 0.3. This finite-size scaling diverges from the flat asymptotic behaviour established in Ref. [11]. This discrepancy likely arises from two factors. First, our numerical solver struggles to find the true global optimum; thus, the linear trend could partially be a numerical artifact. Second, the key rates in this framework are calculated in a qualitatively different way, and therefore the scaling behaviour itself could fundamentally differ. A rigorous analytical proof distinguishing these finite-size scalings is left for future work.

As demonstrated in Fig. 10.4, our key rates clearly outperform those using the EUR-based security proof. Compared to the qubit key rates of Section 10.5, the improvements

[1] In Ref. [98] EUR-based key rates were presented with the additional improvement of performing the decoy-state analysis with a linear program. However, the authors' claim of using the Chernoff bound requires the knowledge of the moment generating function of dependent variables which is not derived. Therefore, we left it out of this comparison.

are much more significant. Those improvements in key rate are due to multiple reasons in conjunction. First, the security analysis itself is tighter for our Rényi entropy formulation compared to the EUR approach, especially for small block sizes as already shown by the qubit BB84 example in Section 10.5.

On an intuitive level, for decoy-state protocols, the number of signals sent that are useful for key generation is roughly only those sent with a single photon, effectively reducing the block size. Thus, one would expect that decoy-state protocols magnify the gap in key rate between our work and the EUR approach, which is exactly what we observe. In Fig. 10.2 although key rates were different, one required the same amount of signals sent to generate positive key rates. Now, the EUR approach requires three orders of magnitude more signals to be sent to generate non-zero key rates.

Second, our proof technique allows one to send *only* the signal intensity in key generation rounds. The EUR approach of [24] cannot accommodate such a protocol, yet. For the example we present here this roughly corresponds to a factor of three improvement in key rates.

Third, we employed a decoy-state analysis similar to Ref. [2], which as discussed in Section 10.7.2 achieve better key rates even in the asymptotic limit. To summarize the discussion from there, the main difference of our technique is that we perform the decoy analysis and the key rate optimization in one step in contrast to the "traditional" two-step process [1, 68, 71]. For Rényi entropies, as previously emphasized in Sections 10.6 and 10.7, the key rates presented here can be made arbitrarily close to the "true" decoy-state key rate by simply including more terms.

Finally, we do not distinguish the influence of each individual improvement mentioned in the three points above — disentangling them may not be possible, as certain steps are necessary to be able to calculate variable-length key rates at all under our approach. However, we conclude that the results presented here pose a significant step forward in QKD performance even purely in terms of achievable key rates.

Similarly to the results shown in Refs. [23, 25], we can also employ only a single decoy intensity. To showcase this protocol variant, in Fig. 10.6, we present the same protocol as above with one decoy intensity less, i.e. we use the same decoy intensity $\mu_2 = 0.02$ and drop μ_3. Additionally, we adjust the probability of sending an intensity in a test round to $p(\mu|\texttt{test}) = 1/2$. The signal intensity μ_s is still optimized for each data point.

To compare our results to the literature, we present the fixed-length key rates from Refs. [23, 25]. Ref. [25] improved upon the results from Ref. [23] and corrected some gaps in the analysis. For a survey of the gaps in those methods see Ref. [97]. We will interpret these key rates from Ref. [25] as an approximation of variable-length key rates for the same

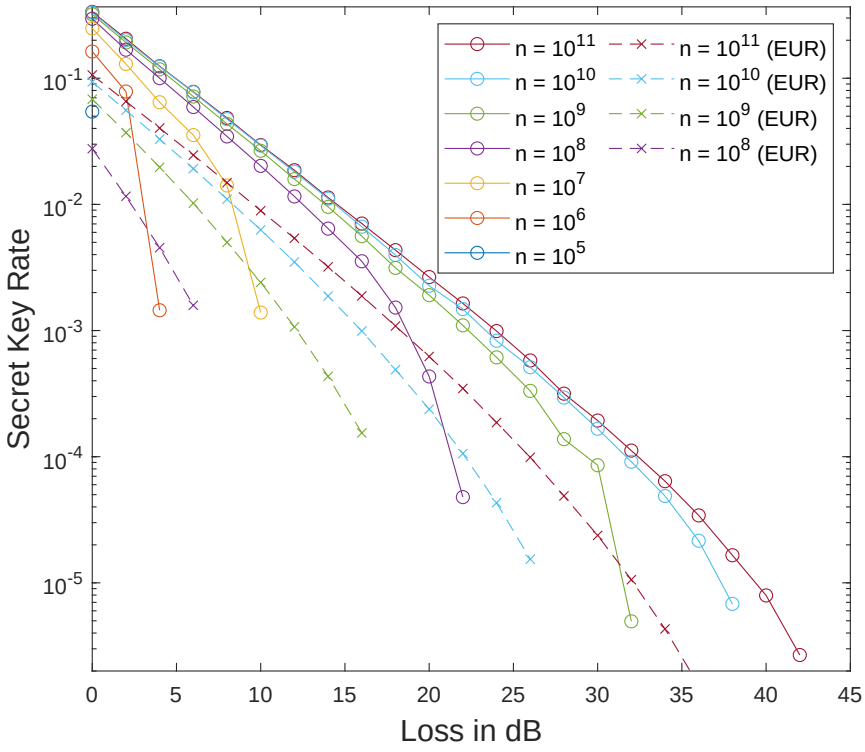


Figure 10.4: Results for variable-length secret key rates of the active decoy BB84 protocol with *two* decoy intensities plotted against the channel loss in dB with a constant misalignment angle of $\theta_{\text{hon}}^{\text{misalign}} = 0.03$ for a varying number of total signals sent $n = 10^5, \ldots, 10^{11}$. The security parameters, and error correction efficiency were chosen as $\varepsilon_{\text{PA}} = \varepsilon_{\text{EV}} = \frac{1}{2}10^{-80}$ and $f_{\text{EC}} = 1.1$. As a comparison, we showcase results using the EUR approach of [24] incorporating the improvements from [83]. For both approaches, we optimize the probability γ of using the X-basis in each round and the signal intensity μ_s. Additionally, for our work we optimize the Rényi parameter α. (Figure 3 from Ref. [4])

protocol, since we expect the key rates to stay roughly the same if the variable-length results from Ref. [24] are incorporated.

Again, we observe that our methods significantly outperform the EUR approach for small block sizes, while the difference reduces as the number of signals grows. Similarly to before, Refs. [23, 25] also require that all intensities are sent in both test and key generation rounds. Hence, we now observe that our key rates at zero loss are roughly a factor of two larger (previously it was a factor of three).

As already observed in [23] and later in [3] (for an IID analysis that is then lifted to

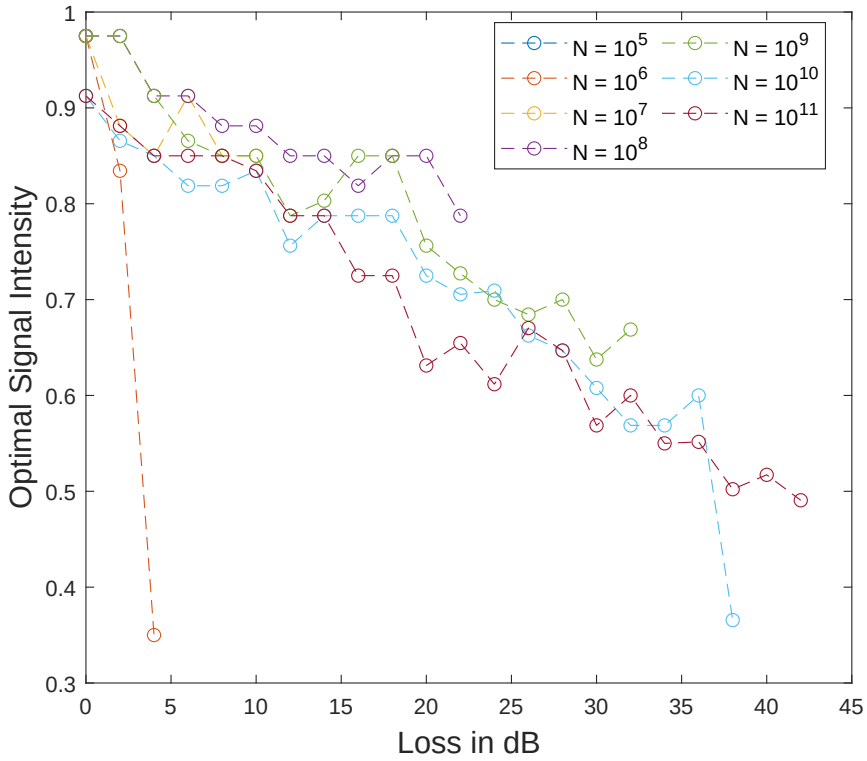


Figure 10.5: Optimal signal intensities for of the variable-length active decoy BB84 protocol with *two* decoy intensities plotted against the channel loss in dB with a constant misalignment angle of $\theta_{\text{hon}}^{\text{misalign}} = 0.03$ for a varying number of total signals sent $n = 10^5, \ldots, 10^{11}$. The security parameters and error correction efficiency were chosen as $\varepsilon_{\text{PA}} = \varepsilon_{\text{EV}} = \frac{1}{2}10^{-80}$ and $f_{\text{EC}} = 1.1$.

coherent attacks via the postselection technique), one single decoy intensity can perform better. Here we observe that especially for small block sizes, the performance of the one-decoy protocol is indeed better, for example see the 10^5 key rates which now extend to 2 dB.

This advantage vanishes for larger block sizes where again two decoy intensities perform better; see for example the last positive key rate for 10^{11} signals. This observation is in line with the expectation that asymptotically three decoy intensities are better than two; see [1], which applied the old two-step process for the decoy analysis.

However, we were unable to determine the point of transition with much precision, as the global optimization routine we used turned out to be very unstable. With better tuning

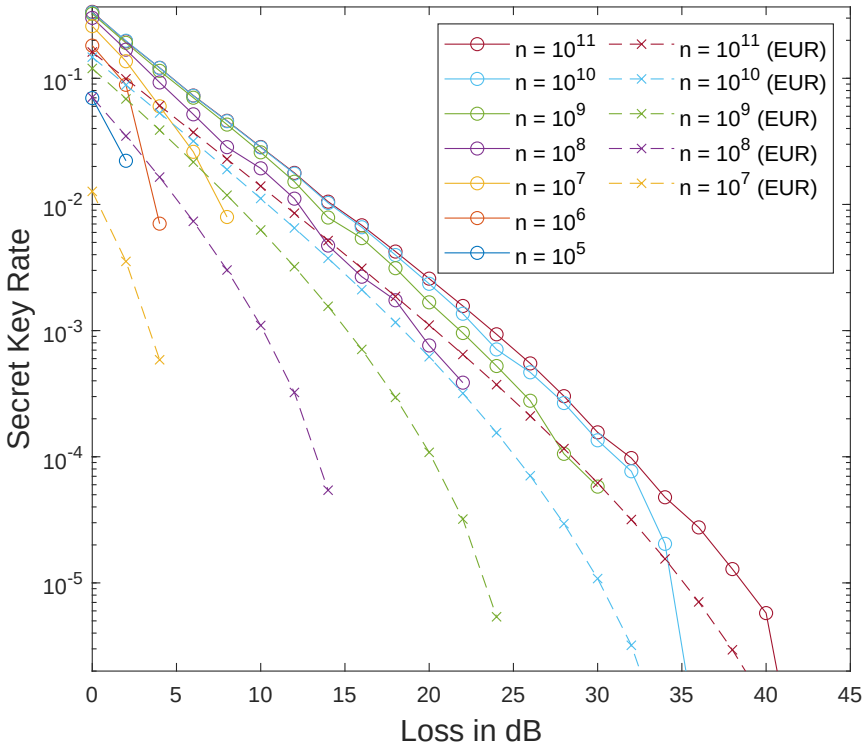


Figure 10.6: Results for variable-length secret key rates of the active decoy BB84 protocol with *one* decoy intensity plotted against the channel loss in dB with a constant misalignment angle of $\theta_{\text{hon}}^{\text{misalign}} = 0.03$ for a varying number of total signals sent $n = 10^5, \ldots, 10^{11}$. The security parameters, and error correction efficiency were chosen as $\varepsilon_{\text{PA}} = \varepsilon_{\text{EV}} = \frac{1}{2}10^{-80}$ and $f_{\text{EC}} = 1.1$. As a comparison, we showcase results using the EUR approach of Refs. [23, 25]. For both approaches, we optimize probability γ of using the X-basis in each round and the signal intensity μ_s. Additionally, for our work we optimize the Rényi parameter α. (Figure 4 from Ref. [4])

of the protocol parameters, one could potentially achieve even higher key rates for both the one- and two-decoy protocol. Crucially, however, we emphasize that the key rates we report here are valid lower bounds.

At this point we also highlight that in contrast to our methods, the global optimization routine does converge well for all EUR key rates we presented here. Thus, those key rates do represent the optimal ones under the respective approaches.

Finally, in Fig. 10.7 we present a comparison between the results of this chapter, the EUR-based key rates, and the postselection technique for the decoy-state BB84 protocol

using *two* decoy intensities. We plot the minimum number of signals sent required to achieve a non-zero key rate, as well as the key rate for $n = 10^{11}$ signals sent. Notably, our results yield non-zero key rates with three orders of magnitude fewer signals than the postselection technique and two orders of magnitude fewer than the EUR-based technique.

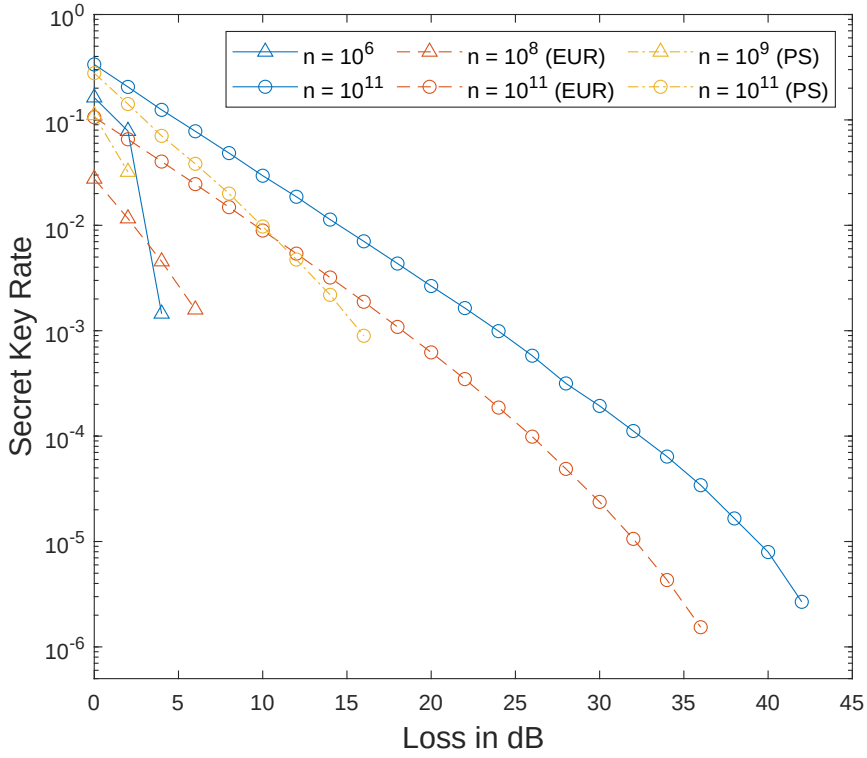


Figure 10.7: Comparison of the variable-length secret key rates for the decoy-state BB84 protocol with *two* decoy intensities. We compare our approach against the EUR [24] and postselection techniques [3], plotted against channel loss in dB for a constant misalignment angle of $\theta_{\text{hon}}^{\text{misalign}} = 0.03$. Lines with circles indicate the key rates for $n = 10^{11}$ signals sent, while lines with triangles show the minimum total signals required to achieve a non-zero key rate. As before, the security parameters and error correction efficiency are set to $\varepsilon_{\text{PA}} = \varepsilon_{\text{EV}} = \frac{1}{2}10^{-80}$ and $f_{\text{EC}} = 1.1$. For all approaches, we optimize the testing probability γ (using the X-basis) and the signal intensity μ_s; for our work, we additionally optimize the Rényi parameter α.

In contrast to the qubit comparison in Fig. 10.3, the postselection technique outperforms the EUR-based proof in the low-loss regime for $n = 10^{11}$. This occurs because the

postselection technique, like the methods in this chapter, supports a more efficient protocol where only the signal intensity is transmitted during generation rounds. Conversely, the EUR-based approach requires all intensities to be sent during key generation. In the high-loss regime, however, the superior finite-size performance of the EUR-based proof becomes dominant, resulting in a crossover.

Overall, our methods substantially outperform both reference techniques, demonstrating a significantly larger advantage than observed in the qubit case.

10.8 Device Imperfections and Implications

The previous section essentially concludes our work on perfect devices, and we will next focus on two examples of imperfections of particular interest. However, before we present those, let us quickly discuss some implications of imperfections.

First, there are different types of imperfections. For example, the source device could prepare an erroneous state as compared to the perfect case, but simply does so in an IID manner in each round (supposing that the signal state choice in each round is IID). Alternatively, the states prepared across rounds could be independent but not identical. The worst case would be that there are correlations between rounds and the devices involved even possess a memory.

Thus far, in the absence of additional assumptions, the MEAT framework [41, 78] can only accommodate imperfections without a memory, i.e. that last scenario cannot currently be handled. However, we note that at the time of writing, no other framework has demonstrated a full security analysis for that scenario either. For source correlations, one could assume a specific correlation length as in Ref. [99]. Then, we conjecture that by analyzing the protocol in terms of several "interleaved virtual protocols" in the manner proposed in that work, our MEAT framework would also address such correlations. However, that work did not present a formal proof of the claim that the "interleaved virtual protocols" analysis indeed yields security of the real protocol, so we leave it at this conjecture for the present work.

Next, let us see how IID or independent imperfections without memories could be included. An example of an independent imperfection (which we will discuss in the upcoming section) is a source sending out a signal with a slightly different intensity in each round without any memory or correlations.

To handle independent imperfections, we consider a scenario where the channel $\mathcal{M}_{\mathrm{QKD}}$ can be different in each round, with the overall channel describing the mapping from $A_1^n B_1^n$

to $S_1^n I_1^n \widehat{C}_1^n$ being a tensor product channel of the form

$$\bigotimes_{i=1}^{n} \mathcal{M}_{\mathrm{QKD}}^{i}. \tag{10.8.1}$$

IID imperfections would simply correspond to having the same channel and marginal state for each round i, so we would be able to immediately apply our previous analysis from Section 8.2. For general non-IID but independent imperfections, we note that [41, Corollary 4.2] (and also [78, Corollary 5.1]) still applies even if the channels and marginal states are different in each round, and hence one can still construct both fixed- and variable-length security proofs by following similar arguments as in the preceding sections, with minor changes.

Specifically, for variable-length proofs, [41, Corollary 4.2] states that Eq. (8.3.15) (the critical step in the proof) still holds if we simply instead define $\hat{f}_{\mathrm{full}}$ in Eq. (8.3.8) with a different $\widetilde{H}_\alpha^{\uparrow,f}$-normalization constant $\kappa_i^{\uparrow}$ for each round i, defined with respect to the channel $\mathcal{M}_{\mathrm{QKD}}^{i}$. If we lower bound each of these constants $\kappa_i^{\uparrow}$ with the worst case over all i, we can then obtain a secure protocol by choosing $\hat{f}_{\mathrm{prot}}$ as in Eq. (8.3.10), with a single lower bound $\underline{\kappa}^{\uparrow}$ for all the rounds. We will make use of this in Sections 10.9 and 10.10 below.

For fixed-length proofs, a technicality arises because when the channels $\mathcal{M}_{\mathrm{QKD}}^{i}$ for each round are different, the formula for h_α^{QKD} according to [41, Corollary 4.2] requires considering a particular convex hull, which may not be simple to describe for some imperfections. Since we focus only on presenting the variable-length key rates, this technicality does not affect the computations we perform later (i.e. the key rates in our subsequent plots are all valid for independent non-IID imperfections), though for the sake of accuracy we restrict our subsequent theorem statements involving h_α^{QKD} to the case of IID imperfections only.[2]

Finally, let us discuss other "simple" imperfections. Detector imperfections under the model of Ref. [100] can be readily included. This model only reduces the weight of the subspace inside the subspace after applying the flag-state squasher [42]. Here in Section 10.3, we explicitly showed that the flag-state can be included in our framework as well.

Furthermore, imperfections that effectively only change Alice's marginal state τ_A or require a ball around it, are also easily implemented in the above framework by adjusting

[2]This does imply that when we modify these h_α^{QKD} optimizations to choose the tradeoff function $\mathbf{f}$ according to the dual solution, we might in principle not be finding the exact optimal $\mathbf{f}$ to use for variable-length key rates against independent imperfections. However, we once again emphasize that our framework is still computing secure lower bounds on the key rates; this suboptimality merely makes the bounds potentially less tight.

the optimization problem. Therefore, in the next two Sections 10.9 and 10.10, we focus on the examples of intensity and phase imperfections, respectively, which are not immediate to include.

10.9 Intensity Imperfections

As described in the previous section, our framework based on Rényi entropies facilitates a straightforward analysis of imperfections, subject to the constraint that they are independent and memoryless. We illustrate this application through two examples, intensity imperfections (discussed here) and phase imperfections in Section 10.10.

Prior works have addressed similar intensity imperfections, though often limited to specific active BB84 protocols [101, 102] or the asymptotic regime [103, 104]. In contrast, our methods are valid for *generic* decoy-state protocols and do not require additional assumptions regarding the receiver, such as an active detection setup.

The analysis in this section is significantly streamlined by the fact that one only needs to find a lower bound on the single-round quantity h_α^{QKD} and the $\widetilde{H}_\alpha^{\downarrow,f}$-normalization constant $\kappa^\downarrow$. The primary requirement is to ensure that the channel $\mathcal{M}_{\mathrm{QKD}}$ with imperfections still satisfies Theorems 8.2.4 and 8.3.3. Once established, we can employ techniques analogous to those used in the asymptotic limit.

We focus here on independent intensity imperfections in decoy-state protocols. We model Alice's source as emitting weak coherent pulses where the exact intensity is unknown. Specifically, in each round, Alice sends a signal with a true intensity μ_j^{true} lying within the interval $[\mu_j, \mu_j + \delta_j]$. Here, μ_j denotes the nominal intensity specified by the protocol, while δ_j represents the maximum deviation.

This model encompasses two distinct physical scenarios. In the first scenario, the source is stable but imperfectly characterized, i.e. the true intensity μ_j^{true} is constant across all rounds, but is only known to lie within a size δ_j interval due to the characterization. In the second scenario, the source fluctuates, i.e. the intensity $\mu_j^{\mathrm{true}} \in [\mu_j, \mu_j + \delta_j]$ is sampled independently in each round from an unknown (and potentially time-varying) probability distribution. Crucially, because the specific distribution of the intensities across rounds is unknown, we cannot exploit its properties. Consequently, both cases are analyzed identically by assuming a worst-case value for each round. Under this model, we demonstrate how to bound the decoy-state version of h_α^{QKD} (and $\kappa^\downarrow$) as presented in Eq. (10.7.11) (and Eq. (10.7.17)).

In principle, minimizing h_α^{QKD} of Eq. (10.7.16) with respect to each intensity in its respective intervals would result in a valid lower bound. However, this is a non-convex problem and general optimization routines do not yield guaranteed lower bounds. Therefore, we bound both the constraints and the objective function separately, such that we still reach a convex optimization problem summarized in Proposition 10.9.1.

Starting with the constraints, for each intensity choice μ_j, we can bound the probability of sending m photons and signal $a = (\alpha, x)$ by

$$p(a, m|\mathtt{test}, \mu_j^{\mathrm{true}}) \geq \min_{\substack{\mu \in \\ [\mu_j, \mu_j + \delta_j]}} p(a, m|\mathtt{test}, \mu) =: p^L(a, m|\mathtt{test}, \mu_j), \tag{10.9.1}$$

$$p(a, m|\mathtt{test}, \mu_j^{\mathrm{true}}) \leq \max_{\substack{\mu \in \\ [\mu_j, \mu_j + \delta_j]}} p(a, m|\mathtt{test}, \mu) =: p^U(a, m|\mathtt{test}, \mu_j). \tag{10.9.2}$$

In Section 10.7.1, we introduced the yields to bound the statistics $\boldsymbol{\nu}_{\widehat{C}|\mathtt{test},\mu_j}$ conditioned on $\mathtt{test}$ and the intensity choice μ_j. These yields were defined as

$$Y_m^{a,b} := \Pr(b|a, \mu_j, m, \mathtt{test}), \tag{10.9.3}$$

and are independent of the intensity choice μ_j for fully phase-randomized weak coherent pulses. Furthermore, for the true intensities μ_j^{true}, the yields satisfy the following relation to $\boldsymbol{\nu}_{\widehat{C}|\mathtt{test},\mu_j}$,

$$\begin{aligned} \boldsymbol{\nu}_{\widehat{C}|\mathtt{test},\mu_j} &\leq \sum_{m \leq N_{\mathrm{ph}}} p(a, m|\mathtt{test}, \mu_j^{\mathrm{true}}) \mathbf{Y}_m + 1 - \mathbf{p}_{\mathrm{tot}|\mu_j^{\mathrm{true}}}, \\ \boldsymbol{\nu}_{\widehat{C}|\mathtt{test},\mu_j} &\geq \sum_{m \leq N_{\mathrm{ph}}} p(a, m|\mathtt{test}, \mu_j^{\mathrm{true}}) \mathbf{Y}_m. \end{aligned} \tag{10.9.4}$$

Then, inserting Eq. (10.9.1), we can immediately bound the constraints on the statistics involving the yields by

$$\begin{aligned} \boldsymbol{\nu}_{\widehat{C}|\mathtt{test},\mu_j} &\leq \sum_{m \leq N_{\mathrm{ph}}} p^U(a, m|\mathtt{test}, \mu_j) \mathbf{Y}_m, + 1 - \mathbf{p}^L_{\mathrm{tot}|\mu_j}, \\ \boldsymbol{\nu}_{\widehat{C}|\mathtt{test},\mu_j} &\geq \sum_{m \leq N_{\mathrm{ph}}} p^L(a, m|\mathtt{test}, \mu_j) \mathbf{Y}_m, \end{aligned} \tag{10.9.5}$$

where

$$\mathbf{p}^L_{\mathrm{tot}|\mu_j} := \min_{\mu \in [\mu_j, \mu_j + \delta_j]} \sum_{\substack{\hat{c} \in \widehat{\mathcal{C}}_{\setminus \mathrm{gen}} \\ a \in \phi^{-1}(\hat{c})}} \sum_{m \leq N_{\mathrm{cons}}} p(a, m|\mathtt{test}, \mu) \hat{e}_{\hat{c}}. \tag{10.9.6}$$

Next, let us see how the objective function depends on the deviations in the intensities. If Alice only sends the signal intensity μ_s in generation rounds and we set $N_{\text{ent}} = 1$, the entropy function $g_\alpha^{N_{\text{ent}}}$ simplifies to

$$g_\alpha^1(J_0, J_1; \mu_s) = \frac{1}{1-\alpha} \log \left\{ 1 - \sum_{m=0,1} p(m|\mu_s) + \sum_{m=0,1} p(m|\mu_s) 2^{(1-\alpha)\widetilde{H}_\alpha^\downarrow(S|\tilde{T}E)_{\nu(J_m)_{|m,\mathtt{gen}}}} \right\}, \tag{10.9.7}$$

where we made the dependence on μ_s explicit. Furthermore, for each m, the state $\nu(J_m)_{|m,\mathtt{gen}}$, is independent of μ_s since

$$\left| \xi_{|m}^g \right\rangle = \sum_a \sqrt{p(a|m,\mathtt{gen})} \left| a \right\rangle_A \left| s_m^{(a)} \right\rangle_{A'}, \tag{10.9.8}$$

is independent of μ_s. Hence, finding μ in $[\mu_s, \mu_s + \delta_s]$ such that the resulting probabilities minimize $g_\alpha^{N_{\text{ent}}}$ will lead to a valid lower bound. We state this result in the following proposition.

Proposition 10.9.1. *Let all intensities μ_j^{true} vary in intervals $[\mu_j, \mu_j + \delta_j]$, but the photon number distribution follows a Poisson distribution for any possible $\mu_j^{\text{true}} \in [\mu_j, \mu_j + \delta_j]$. Let us also assume that, in each round $i = 1, \ldots, n$, the intensity choice of this round is sampled independently from $[\mu_j, \mu_j + \delta_j]$, and that Alice only sends the signal intensity in generation rounds. Additionally, assume it holds*

$$p(a|m, \mu_j, \mathtt{test}) = p(a|m, \mathtt{test}) \; \forall j. \tag{10.9.9}$$

Then, all $\widetilde{H}_\alpha^{\downarrow,f}$-normalization constants $\kappa_i^\uparrow$ can be bounded by

$$\begin{aligned}
\kappa_i^\uparrow \geq \inf_{\substack{J_m\in\mathrm{Pos}(A'B),\\ \boldsymbol{\nu}\in\mathbb{P}(\widehat{\mathcal{C}}_{\setminus\mathrm{gen}}),\\ \mu\in\{\mu_s,\mu_s+\delta_s\}}} & \frac{1}{1-\alpha}\log\left(\gamma\sum_{c\in\widehat{\mathcal{C}}_{\setminus\mathrm{gen}}}\boldsymbol{\nu}(\hat{c})2^{(\alpha-1)f(\hat{c})}+(1-\gamma)\,2^{-(\alpha-1)\left(g_\alpha^1(J_0,J_1;\mu)-f(\mathtt{gen})\right)}\right)\\
s.t.\ \ & \mathrm{Tr}_B[J_m]=\mathbb{I}_{A'},\\
& \boldsymbol{\nu}=(p(\mu_s|\mathtt{test})\boldsymbol{\nu}_{|\mu_s},\dots)^T,\\
& \boldsymbol{\nu}_{|\mu_j}\leq\sum_{m\leq N_{\mathrm{ph}}}p^U(a,m|\mathtt{test},\mu_j)\mathbf{Y}_m+1-\mathbf{p}^L_{\mathrm{tot}|\mu_j},\\
& \boldsymbol{\nu}_{|\mu_j}\geq\sum_{m\leq N_{\mathrm{ph}}}p^L(a,m|\mathtt{test},\mu_j)\mathbf{Y}_m,\\
& p(a|m,\mu_j,\mathtt{test})\mathbf{Y}_m=\boldsymbol{\Psi}_{m,\mu_j}[J_m],\ m=0,1.
\end{aligned}\tag{10.9.10}$$

Furthermore, if the sampling from $[\mu_j,\mu_j+\delta_j]$ is IID, then h_α^{QKD} is bounded from below by

$$\begin{aligned}
h_\alpha^{\mathrm{QKD}}\geq \inf_{\substack{\mathbf{q}\in S_{\mathrm{acc}},\boldsymbol{\nu}_{\widehat{C}}\in\mathbb{P}(\widehat{\mathcal{C}})\\ J_m\in\mathrm{Pos}(A'B),\\ \mu\in\{\mu_s,\mu_s+\delta_s\}}} & \left(\frac{\alpha D\left(\mathbf{q}\middle\|\boldsymbol{\nu}_{\widehat{\mathcal{C}}}\right)}{\alpha-1}+q(\mathtt{gen})g_\alpha^{N_{\mathrm{ent}}}(J_0,J_1;\mu)\right)\\
s.t.\ \ & \mathrm{Tr}_B[J_m]=\mathbb{I}_{A'},\\
& \boldsymbol{\nu}_{\widehat{C}}=(\gamma p(\mu_s|\mathtt{test})\boldsymbol{\nu}_{\widehat{C}|\mathtt{test},\mu_s},\dots,1-\gamma)^T,\\
& \boldsymbol{\nu}_{\widehat{C}|\mathtt{test},\mu_j}\leq\sum_{m\leq N_{\mathrm{ph}}}p^U(a,m|\mathtt{test},\mu_j)\mathbf{Y}_m+1-\mathbf{p}^L_{\mathrm{tot}|\mu_j},\\
& \boldsymbol{\nu}_{\widehat{C}|\mathtt{test},\mu_j}\geq\sum_{m\leq N_{\mathrm{ph}}}p^L(a,m|\mathtt{test},\mu_j)\mathbf{Y}_m,\\
& p(a|m,\mathtt{test})\mathbf{Y}_m=\boldsymbol{\Psi}_{m,\mu_j}[J_m],\ m=0,1.
\end{aligned}\tag{10.9.11}$$

In both equations, we stated the dependence of g_α^1 on μ_s explicitly.

Proof. For simplicity, we only explicitly prove the statement involving h_α^{QKD}. The proof for the $\widetilde{H}_\alpha^{\uparrow,f}$-normalization constants $\kappa_i^\uparrow$ follows similar arguments.

For the statement involving h_α^{QKD}, Alice's source prepares the states in an IID manner. Thus, the channel $\mathcal{M}_{\mathrm{QKD}}$ in Theorem 8.2.4 and Theorem 8.3.3 would in principle be defined in terms of the true intensities μ_j^{true}. Most importantly, although these true intensities are unknown, the channel still satisfies the assumptions of Theorems 8.2.3 and 8.2.4. Then,

one follows the same steps as in Section 10.7.1 to reach the expression of Eq. (10.7.16) with the intensities μ_j^{true} in place of μ_j.

However, since the true intensities μ_j^{true} are unknown, we need to replace the optimization problem for h_α^{QKD} with the "worst possible value" of μ_j^{true} over the entire interval $[\mu_j, \mu_j + \delta_j]$. Our goal is now to show that this is indeed lower bounded by the expression given in the claim.

First, the weakened constraints follow immediately from Eq. (10.9.5) together with the assumption that $p(a|m, \mu, \texttt{test}) = p(a|m, \texttt{test})$. Hence, since $D\big(\mathbf{q}\big\|\boldsymbol{\nu}_{\widehat{C}}\big)$ is independent of the true signal intensity, we are left with showing that

$$\min_{\mu\in[\mu_s,\mu_s+\delta_s]} g_\alpha^1\left(J_0, J_1; \mu\right) = \min_{\mu\in\{\mu_s,\mu_s+\delta_s\}} g_\alpha^1\left(J_0, J_1; \mu\right). \tag{10.9.12}$$

To simplify the notation, let us define

$$\beta_m \coloneqq 1 - 2^{(1-\alpha)\widetilde{H}_\alpha^\downarrow(S|\tilde{T}E)_{\nu(J_m)|m,\texttt{gen}}}, \tag{10.9.13}$$

which is always nonnegative and lets us rewrite g_α^1 as

$$g_\alpha^1\left(J_0, J_1; \mu\right) = \frac{1}{1-\alpha}\log\left\{1 - \sum_{m=0,1} p(m|\mu)\beta_m\right\}. \tag{10.9.14}$$

Next we note that minimizing $-\log(v(x))$ over x, for some function v, is equivalent to maximizing v over x. Therefore, minimizing g_α^1 is equivalent to

$$\max_{\mu\in[\mu_s,\mu_s+\delta_s]} 1 - \sum_{m=0,1} p(m|\mu)\beta_m = 1 - \min_{\mu\in[\mu_s,\mu_s+\delta_s]} \sum_{m=0,1} p(m|\mu)\beta_m \tag{10.9.15}$$

$$= 1 - \min_{x\in[0,\delta_s]} e^{-(\mu_s+x)}\left[\beta_0 + (\mu_s + x)\beta_1\right], \tag{10.9.16}$$

where we introduced the variable $x \in [0, \delta_s]$ to represent the deviation and used that $p(m|\mu + x)$ is distributed according to a Poisson distribution for all $x \in [0, \delta_s]$. By the assumption that Alice sends only the signal intensity in generation rounds, β_m is independent of x, as shown in Eq. (10.9.8).

Next, if we define

$$u(x) \coloneqq e^{-(\mu_s+x)}\left[\beta_0 + (\mu_s + x)\beta_1\right], \tag{10.9.17}$$

one can find the local extrema by taking the first derivative, which leads to an optimal value $x_{\text{opt}} = 1 - \beta_0/\beta_1 - \mu_s$. Taking the second derivative of f yields

$$u''(x_{\text{opt}}) = -\beta_1 \exp\left(\frac{\beta_0}{\beta_1} - 1\right) < 0, \tag{10.9.18}$$

since for $\beta_1 = 0$, $x_{\text{opt}} = -\infty \notin [0, \delta_s]$ and $\beta_0, \beta_1 \geq 0$ Hence, the local extremum is a maximum and by Fermat's theorem the minimum is achieved on the boundary $x \in \{0, \delta_s\}$.

With regards to the $\widetilde{H}_\alpha^{\uparrow,f}$-normalization constants $\kappa_i^\uparrow$ one can follow a very similar line of argument. Here, one needs to consider different channels $\mathcal{M}_{\text{QKD}}^i$ for each $\kappa_i^\uparrow$ as described in Section 10.8. For each $\kappa_i^\uparrow$, we can still apply the simplifications of Section 10.7.1 with the intensities μ_j^{true} in place of μ_j. Then, the loosened constraints again follow from Eq. (10.9.5) for each round i.

Next, one can follow the same argument of minimizing the negative logarithm of a function, and relate the minimum of each $\kappa_i^\uparrow$ to the same function u from above. This yields that for all rounds $i = 1, \ldots, n$, the constants $\kappa_i^\uparrow$ are bounded by the same worst case and the statement follows. □

Remark 10.9.2. The lower bound on the $\widetilde{H}_\alpha^{\uparrow,f}$-normalization constants $\kappa_i^\uparrow$ was based on the $\widetilde{H}_\alpha^{\downarrow,f}$-normalization constant $\kappa^\downarrow$. One can derive a similar expression based on the mixed f-weighting. This will yield essentially the same expression but with different α factors and the proof follows the same line of arguments.

There are several ways in which one could extend the above proposition. For example, one could straightforwardly extend it to allow all intensities to be sent in generation rounds. Another possible extension might be to allow the signal choice $a = (\alpha, x)$ to depend on the intensity, which violates the assumption $p(a|m, \mu, \texttt{test}) = p(a|m, \texttt{test})$, but we will leave these extensions for future work. For now, in the next section we present an example of a biased passive BB84 protocol using WCP sources with intensity imperfections under the model described above.

10.9.1 Example: Biased Passive BB84 with Intensity Imperfections

In this section, we present results for variable-length secret key rates of a decoy-state version of a BB84 protocol with WCP sources and a *biased* passive detection setup. This setup does not allow for simple squashing methods to reduce it back to a qubit space, such as [62, 63]. However, we can apply the flag-state squasher [42] together with the bounds from Section 4.3.

In each round, Alice decides with probability γ if it is a test or a generation round, and if it is a generation round, she always selects the signal intensity μ_s and the Z basis. Therefore, as before, Alice's basis choice probabilities are given by $p_z = 1 - \gamma$ and $p_x = \gamma$.

In a test round, Alice chooses uniformly at random from all the intensities considered, and always uses the X basis. We again present the one- and two-decoy variant of the protocol. In contrast to the active decoy BB84 protocol, we choose a *fixed testing probability* of $\gamma = 0.2$.

Additionally, we assume that Alice's setup has intensity imperfections satisfying the assumptions of the previous section. Hence, we assume that in each round Alice independently sends out a signal with an intensity in the ranges

$$\mu_s^{\text{true}} \in [\mu_s, \mu_s + \delta_s], \quad \mu_2^{\text{true}} \in [\mu_2, \mu_2 + \delta_2], \mu_3^{\text{true}} \in [\mu_3, \mu_3 + \delta_3]. \tag{10.9.19}$$

We set the intensities specified by the protocol to be

$$\mu_s = 0.8, \quad \mu_2 = 0.2, \quad \mu_3 = 0.01, \tag{10.9.20}$$

and omit the second decoy intensity for the one-decoy version of this protocol.

For a better comparison with the literature, we chose δ_j in a similar way as in Refs. [103, 104], such that

$$\left| 1 - \frac{\mu_j}{\mu_j + \delta_j} \right| \leq \epsilon_{\text{int}}, \tag{10.9.21}$$

where for ϵ_{int} we will chose $\epsilon_{\text{int}} = 10\%$, 25% and $\epsilon_{\text{int}} = 0$, which recovers the perfect case. Thus, the intensities can fluctuate and reach a value 10% or 25% higher than the one specified.

On the receiver side (Bob), we assume a passive four-state detection setup that chooses the Z and X basis with splitting ratios $p_z = 1-\gamma$ and $p_x = \gamma$, respectively. To incorporate this into our security proof, we apply the flag-state squasher from [42] with the subspace bound from Theorem 5.3.1, which we already showed in Section 10.3 can be implemented in this framework.

Finally, we assume the same security parameters $\varepsilon_{\text{PA}} = \varepsilon_{\text{EV}} = \frac{1}{2}10^{-80}$ and error correction model with efficiency $f_{\text{EC}} = 1.1$ as in Section 10.5. We again model channel loss similarly to the qubit protocol and include a misalignment with an angle of $\theta_{\text{hon}}^{\text{misalign}} = 0.03$. The Kraus operators of the maps $\mathcal{G}$ and $\mathcal{Z}$ and the states can be found in Appendix C.3.4.

The resulting variable-length key rates for the one- and two-decoy protocols are shown in Fig. 10.9 and Fig. 10.8, respectively, versus channel loss in dB. Both the one- and two-decoy protocol react very similarly to intensity imperfections and, for small imperfections, $\epsilon_{\text{int}} < 10\%$, the difference from a perfect setup is negligible. Moreover, in the perfect case $\epsilon_{\text{int}} = 0$, both variants achieve key rates comparable to the active decoy BB84 protocol as

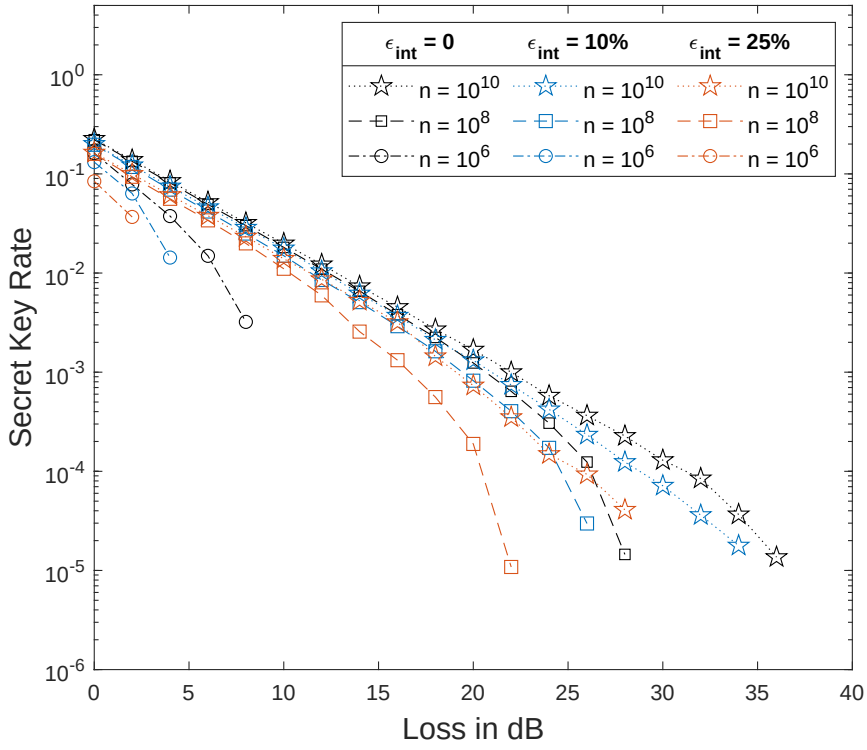


Figure 10.8: Results for variable-length secret key rates of the passive decoy BB84 protocol with *two* decoy intensities and intensity imperfection plotted against the channel loss in dB with a varying number of total signals sent $n = 10^6, 10^8, 10^{10}$. The security parameters, and error correction efficiency were chosen as $\varepsilon_{\mathrm{PA}} = \varepsilon_{\mathrm{EV}} = \frac{1}{2}10^{-80}$ and $f_{\mathrm{EC}} = 1.1$. (Figure 5 from Ref. [4])

shown in Section 10.7.3. Both observations are true across the range of number of signals sent that we present here.

Additionally, the performance of the one-decoy variant of the protocol is better than that of the two-decoy variant. This is especially true if very few signals are sent or if in the high-loss regime. In both scenarios, the largest finite-size penalties would be expected due to small sample sizes. Our observations here are therefore in line with similar results previously shown in [1, 23].

Since the behaviour of the one- and two-decoy protocol with respect to the intensity imperfections is very similar, we discuss their results together. Compared to the perfect case (black curves), for $\epsilon_{\mathrm{int}} = 10\%$ (blue curves), one can see a small but noticeable difference in key rates, which are shifted downward along the y-axis but still roughly follow the same

trend as the perfect setup. The reduction in the key rate is approximately 15%, which is again valid across the full range of the numbers of signals sent presented here.

The influence of intensity imperfections with $\epsilon_{\text{int}} = 25\%$ is significantly more pronounced. Independent of the total number of signals sent, the overall achievable key rate is reduced by approximately 40%. Furthermore, the maximum tolerable loss decreases by 4–6 dB compared to a perfect setup. Therefore, our results demonstrate that we can effectively tolerate experimentally viable levels of intensity imperfections without a drastic reduction in key rates.

Finally, we briefly comment on the capacity of other proof techniques to account for independent intensity imperfections. In principle, the postselection technique could incorporate IID intensity imperfections by loosening the constraints, similar to our approach in Section 10.9. However, general independent imperfections are currently not possible to include, since the postselection technique relies on IID sources (or sources representable as IID sources through source maps).

Conversely, approaches based on phase error correction cannot yet treat these forms of imperfections in the finite-size regime. Within the EUR framework, such imperfections have only been addressed for active one-decoy [102] and two-decoy [101] BB84 protocols. Crucially, those techniques are not applicable to the passive detection setup presented here. Furthermore, it remains unclear how effectively those techniques would translate to other active detection setups, such as the 6-state protocol. In contrast, our methods are tighter, as we solve parts of the optimization problem exactly, and are readily applicable to *generic* decoy-state protocols.

Regarding future improvements, we conclude that the lower bound on the objective function is already optimal within our Rényi framework, as we identified the true minimum of the quantity with respect to intensity imperfections. Future research should therefore focus on tightening the constraints in Eq. (10.9.5), as the looseness introduced there appears to be the only remaining factor limiting the key rates.

10.10 Phase Imperfections

Another common imperfection which occurs especially under high repetition rates of decoy-state protocols are phase imperfections. Typically, decoy-state protocols use a fully phase-randomized weak coherent pulse. However, if the repetition rate becomes too high, the phase-randomization can become imperfect. Then, the sources do not satisfy the assumptions of typical decoy-state protocols anymore. Previously, in Refs. [68, 105] asymptotic

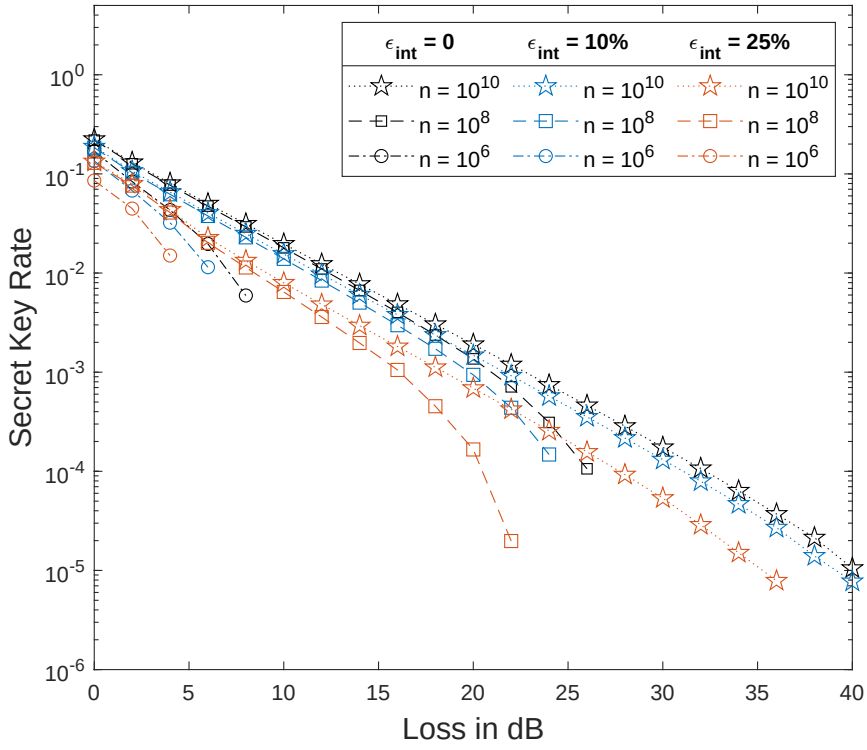


Figure 10.9: Results for variable-length secret key rates of the passive decoy BB84 protocol with *one* decoy intensity and intensity imperfection plotted against the channel loss in dB with a varying number of total signals sent $n = 10^6, 10^8, 10^{10}$. The security parameters, and error correction efficiency were chosen as $\varepsilon_{\mathrm{PA}} = \varepsilon_{\mathrm{EV}} = \frac{1}{2}10^{-80}$ and $f_{\mathrm{EC}} = 1.1$. (Figure 6 from Ref. [4])

key rates of decoy-state protocols with phase imperfections have been derived. Here, we present a finite-size security proof against coherent attacks.

Similarly to intensity imperfections presented in Section 10.9, one can also include phase imperfections as long as there is no memory process involved and states are sent out in an independent manner. Hence, we assume that in each round Alice's source emits a phase-imperfect state that is independent of all other rounds. However, the imperfection might not be identical in all rounds. To simplify notation, as done before, we will abbreviate $a = (\alpha, x)$ in all of this section.

We use the same model of phase imperfections as described in Refs. [68, 105], which connect phase imperfect states to a model state via a source map. In particular, this model assumed that in each round $i = 1, \ldots, n$, instead of a fully phase-randomized weak coherent

pulse, Alice sends out the following mixture

$$\rho_{A'}^{(a,\mu)} = q \sum_{m=0}^{\infty} p(m|\mu) \, |m\rangle\langle m|_a + (1-q) \, |\sqrt{\mu}\rangle\langle\sqrt{\mu}|_a \,, \tag{10.10.1}$$

where $|m\rangle_a$ indicates m photons sent with $a = (\alpha, x)$ in partition α and bit value x. This could be for example horizontal polarized light with m photons. Similarly, $|\sqrt{\mu}\rangle_a$ represents a coherent state sent with the choice $a = (\alpha, x)$ and intensity μ. The quality factor $q \in [0, 1]$ quantifies how close the state is to a fully phase-randomized state. Then the connection to independent phase imperfections is made by a source map Ξ^i_{phase}, which is described in [68, App. A], and can be different in each round $i = 1, \ldots, n$.

In order to prove the security of such a protocol, we follow a similar approach as in Ref. [68] and translate it into our framework. We will mainly focus our derivation on the state $\rho_{A'}^{(a,\mu)}$ because the additional source map will be included in Eve's attack by Section 10.4.

First, we note that each state $\rho_{A'}^{(a,\mu)}$ is block-diagonal in a specific basis, which coincides with the photon-number basis only when $q = 1$. We exploit this block-diagonal structure to derive a formulation compatible with the Rényi framework presented in Section 10.6.

However, a caveat remains: for arbitrary values of q, this block-diagonal structure can only be found approximately using numerical methods [68], which restricts the decomposition to a finite-dimensional subspace. Furthermore, states with phase imperfections, as in Eq. (10.10.1), do not satisfy Assumptions 1–3. Hence, the goal of this section is to derive a formulation of phase imperfections that can be directly incorporated into our framework.

10.10.1 Definition of Block-Tagged States

We start by constructing states that satisfy the assumptions 1–3, incorporating a cutoff required for a numerical block-diagonalization. Later we will show how to rigorously incorporate this cutoff in our key rate optimizations.

For this purpose, let us define the state projected onto the space with $\leq N_A$ photons

$$\sigma_{A'}^{(a,\mu)} \coloneqq \Pi_{\leq N_A} \rho_{A'}^{(a,\mu)} \Pi_{\leq N_A} + \Pi_{>N_A} \rho_{A'}^{(a,\mu)} \Pi_{>N_A} \tag{10.10.2}$$

where the projectors onto the subspaces are

$$\Pi_{\leq N_A} \coloneqq \sum_{n \leq N_A} |n\rangle\langle n| \,, \quad \Pi_{>N_A} \coloneqq \sum_{n > N_A} |n\rangle\langle n| \,. \tag{10.10.3}$$

Notably, this projection onto the $\leq N_A$-photon subspace removes any correlations between the $\leq N_A$-photon and the $> N_A$-photon subspace. Then, $\sigma_{A'}^{(a,\mu)}$ can be diagonalized as

$$\sigma_{A'}^{(a,\mu)} = \sum_{m=0}^{\infty} \omega_m(a,\mu) \left|w_m(a,\mu)\right\rangle\!\left\langle w_m(a,\mu)\right|_{A'} . \tag{10.10.4}$$

Crucially, for different signal and intensity settings, (a,μ), the states $\left|w_m(a,\mu)\right\rangle_{A'}$ are not orthogonal for different values of m. That means that the states $\sigma^{(a,\mu)}$ violate the assumption 1 because the block-diagonal structure over m is not the same for all signal choices (omitting the trivial case of only a single block of the full space). Furthermore, Eve could not apply a QND measurement without disturbing the state. Thus, one cannot without loss of generality assume a block-diagonal channel as in Section 10.6.

To overcome this limitation, we employ block-tagged states to recover the block-diagonal structure essential for the simplifications presented in Section 10.6. We incorporate this tagging through the use of source maps (see Section 10.4). Accordingly, we define the block-tagged counterpart of $\sigma_{A'}$ as follows:

$$\tau_{A'\bar{A}}^{(a,\mu)} = \sum_{m=0}^{\infty} \omega_m(a,\mu) \left|w_m(a,\mu)\right\rangle\!\left\langle w_m(a,\mu)\right|_{A'} \otimes \left|m\right\rangle\!\left\langle m\right|_{\bar{A}} . \tag{10.10.5}$$

The states $\tau_{A'\bar{A}}^{(a,\mu)}$ are now simultaneously block-diagonal in the same basis for all a,μ. Thus, Alice's pure state after the source replacement scheme including a shield system can be written as

$$\left|\psi\right\rangle_{AA_sA'\bar{A}} = \sum_{m=0}^{\infty} \sqrt{\omega_m(a,\mu)} \left|m\right\rangle_{A_s} \left|\psi_{|m}\right\rangle_{AA'\bar{A}} , \tag{10.10.6}$$

where

$$\left|\psi_{|m}\right\rangle_{AA'\bar{A}} = \sum_{a,\mu} \sqrt{p(a,\mu|m)} \left|a,\mu\right\rangle_{A} \left|w_m(a,\mu)\right\rangle_{A'} \left|m\right\rangle_{\bar{A}} . \tag{10.10.7}$$

Most significantly, this state satisfies Assumptions 1–3, implying that the source map technique yields a valid lower bound on the key rate. Building upon $\left|\psi\right\rangle$, we define the conditional states

$$\left|\xi^g\right\rangle_{A_SAA'\bar{A}} = \sum_{m=0}^{\infty} \sqrt{\omega_{m|\mathtt{gen}}} \left|m\right\rangle_{A_S} \left|\xi^g_{|m}\right\rangle_{AA'\bar{A}} , \tag{10.10.8}$$

$$\left|\xi^t\right\rangle_{A_SAA'\bar{A}} = \sum_{m=0}^{\infty} \sqrt{\omega_{m|\mathtt{test}}} \left|m\right\rangle_{A_S} \left|\xi^t_{|m}\right\rangle_{AA'\bar{A}} , \tag{10.10.9}$$

by conditioning on the `test` and `gen` events. Here, the states of each block m are defined as

$$\left|\xi^{g}_{|m}\right\rangle_{AA'\bar{A}} = \sum_{a,\mu} \sqrt{p(a,\mu|m,\texttt{gen})}\,|a,\mu\rangle_{A}\,|w_m(a,\mu)\rangle_{A'}\,|m\rangle_{\bar{A}}\,, \tag{10.10.10}$$

with $\left|\xi^{t}_{|m}\right\rangle_{AA'\bar{A}}$ defined analogously. Thus, we have constructed states that satisfy Item 1. Consistent with Eq. (10.6.1), our notation indicates conditioning on the round type and block m. This concludes the construction of the block-tagged states. In the subsequent subsection, we justify the truncation with a cutoff N_A and prove it yields secure lower bounds on the key rates.

10.10.2 Approximate Diagonalization with Cutoff

Now, we show how to rigorously incorporate the cutoff N_A required for the approximate diagonalization. For both $h^{\mathrm{QKD}}_{\alpha}$ and the $\widetilde{H}^{\downarrow,f}_{\alpha}$-normalization constant $\kappa^{\downarrow}$ we cannot *a priori* assume a cutoff, but intuitively, if we choose N_A large enough, the difference should be insignificant. We will formalize this intuition in this subsection.

Therefore, we obtain the states $|s_m(a,\mu)\rangle_{A'}$ and probabilities $\zeta_m(a,\mu)$ by directly diagonalizing the signal state $\rho^{(a,\mu)}_{A'}$ defined in Eq. (10.10.1). This yields the decomposition

$$\rho^{(a,\mu)}_{A'} = \sum_{m=0}^{\infty} \zeta_m(a,\mu)\,|s_m(a,\mu)\rangle\langle s_m(a,\mu)|_{A'}\,. \tag{10.10.11}$$

Then, similarly to Eq. (10.10.8) let us define the source-replaced states without a cutoff (or equivalently $N_A \to \infty$) as

$$\left|\widetilde{\xi}^{g}\right\rangle_{A_S AA'\bar{A}} = \sum_{m=0}^{\infty} \sqrt{\zeta_{m|\texttt{gen}}}\,|m\rangle_{A_S}\left|\widetilde{\xi}^{g}_{|m}\right\rangle_{AA'\bar{A}}\,, \tag{10.10.12}$$

$$\left|\widetilde{\xi}^{t}\right\rangle_{A_S AA'\bar{A}} = \sum_{m=0}^{\infty} \sqrt{\zeta_{m|\texttt{test}}}\,|m\rangle_{A_S}\left|\widetilde{\xi}^{t}_{|m}\right\rangle_{AA'\bar{A}}\,, \tag{10.10.13}$$

where

$$\left|\widetilde{\xi}^{g}_{|m}\right\rangle_{AA'\bar{A}} = \sum_{a,\mu} \sqrt{p(a,\mu|m,\texttt{gen})}\cdot|a,\mu\rangle_{A}\,|s_m(a,\mu)\rangle_{A'}\,|m\rangle_{\bar{A}}\,, \tag{10.10.14}$$

and similarly for $\left|\widetilde{\xi}^t_{|m}\right\rangle$. Here again the probabilities $\zeta_{m|\mathtt{test}}$ are found by conditioning $\zeta_m(a,\mu)$ appropriately on `test` and `gen` rounds.

Now, as a first step towards rigorously accounting for the cutoff, we aim to relate the conditional entropies with and without a cutoff N_A by

$$\left|\widetilde{H}^{\uparrow}_{\alpha}(S|\tilde{T}E)_{\widetilde{\nu}_{|\mathtt{gen}}} - \widetilde{H}^{\uparrow}_{\alpha}(S|\tilde{T}E)_{\nu_{|\mathtt{gen}}}\right| \leq \Delta, \tag{10.10.15}$$

where $\widetilde{\nu}$ and ν are generated from $\left|\widetilde{\xi}^g\right\rangle$ and $|\xi^g\rangle$, respectively. This form of inequality can be understood as the worst case in which the cutoff could influence the final entropy.

One can derive a bound of this form immediately from the continuity bounds for conditional Rényi entropies found in [106, Corollary 5.1], if a bound on the trace distance,

$$\frac{1}{2}\left\|\nu_{|\mathtt{gen}} - \tilde{\nu}_{|\mathtt{gen}}\right\|, \tag{10.10.16}$$

between the states ν and $\widetilde{\nu}$ is available.

In order to achieve such a bound on the trace distance, for now, let us assume that we have

$$\frac{1}{2}\left\|\rho^{(a,\mu)}_{A'} - \sigma^{(a,\mu)}_{A'}\right\|_1 \leq \epsilon^{a,\mu}_{\mathrm{proj}}, \tag{10.10.17}$$

for all individual (a,μ), as was proven in App. B of Ref. [68]. We will show the exact definition of $\epsilon^{a,\mu}_{\mathrm{proj}}$ in Section 10.10.3, but for now, let us see how we could apply this trace-distance bound to rigorously account for the cutoff.

By applying Lemma A.2.1 from the appendix, together with the data-processing inequality of the trace-distance, see e.g. [48, Theorem 9.2], we conclude

$$\begin{aligned}\frac{1}{2}\left\|\nu_{|\mathtt{gen}} - \tilde{\nu}_{|\mathtt{gen}}\right\| &\leq \frac{1}{2}\left\|\mathrm{Tr}_{A_s}\circ\mathcal{G}\circ(\mathbb{I}_A\otimes\ \mathcal{E})\left(|\xi^g\rangle\langle\xi^g| - \left|\widetilde{\xi}^g\right\rangle\left\langle\widetilde{\xi}^g\right|\right)\right\|_1 &(10.10.18)\\ &\leq \sum_{a,\mu} p(a,\mu|\mathtt{gen})\epsilon^{a,\mu}_{\mathrm{proj}} =: \epsilon_{\mathrm{full}}, &(10.10.19)\end{aligned}$$

where we defined ϵ_{full} as the probabilistic mixture of the bounds $\epsilon^{a,\mu}_{\mathrm{proj}}$ on the signal states.

Using this bound on the trace distance, one could directly apply [106, Eq. (9) in Corollary 5.1] to find a version of Δ. However, we can slightly improve the bound in [106, Corollary 5.1] because the register S is classical, i.e. $\widetilde{H}^{\uparrow}_{\alpha}(S|\tilde{T}E)_{\nu} \geq 0$ for all ν. Hence, appropriately identifying the registers between our notation and the one in [106], one can

replace d_S^2 with d_S in Eq. (9) of Ref. [106, Corollary 5.1] and use this as the definition of Δ. This will reduce the cost due to the continuity bound slightly and results in

$$\Delta(\epsilon) = \min \begin{cases} \log(1+\varepsilon) + \dfrac{1}{\alpha-1}\log\left(1+\varepsilon d_S^{\alpha-1} - \dfrac{\varepsilon^\alpha}{(1+\varepsilon)^{\alpha-1}}\right) \\ \frac{\alpha}{\alpha-1}\log\left(1+\varepsilon d_S^{\frac{\alpha-1}{\alpha}}\right), \\ \log(1+\varepsilon) + \dfrac{\alpha}{\alpha-1}\log\left(1+\varepsilon d_S^{\frac{\alpha-1}{\alpha}} - \dfrac{\varepsilon^{2-\frac{1}{\alpha}}}{(1+\varepsilon)^{\frac{\alpha-1}{\alpha}}}\right). \end{cases} \tag{10.10.20}$$

Thus, we find the following lower bound on the entropy excluding a cutoff in terms of the state including a cutoff N_A

$$\widetilde{H}_\alpha^\uparrow(S|\tilde{T}E)_{\widetilde{\nu}_{|\text{gen}}} \geq \widetilde{H}_\alpha^\uparrow(S|\tilde{T}E)_{\nu_{|\text{gen}}} - \Delta(\epsilon_{\text{full}}) \tag{10.10.21}$$

$$\geq \widetilde{H}_\alpha^\downarrow(S|\tilde{T}E)_{\nu_{|\text{gen}}} - \Delta(\epsilon_{\text{full}}), \tag{10.10.22}$$

where the correction term Δ is given by the tightened version of Eq. (10.10.20) from above.

Next, we need to incorporate the difference in the states due to the cutoff N_A into the constraints, i.e. the map $\mathbf{\Phi}$ (or $\mathbf{\Psi}$). Therefore, let us assume the signal states in system A' and the probabilities satisfy

$$\left\| |s_m(a,\mu)\rangle\langle s_m(a,\mu)| - |w_m(a,\mu)\rangle\langle w_m(a,\mu)| \right\|_1 \leq \epsilon_{a,\mu}, \tag{10.10.23}$$

$$|\zeta_m(a,\mu) - \omega_m(a,\mu)| \leq \epsilon_{\text{prob}}, \tag{10.10.24}$$

which was proven in [68, Theorem 6]. Again in Section 10.10.3 we will present the exact definitions, but for now, let us see how we could apply it to rigorously incorporate the cutoff in the map $\mathbf{\Phi}$. We summarize the bounds in the following lemma.

Lemma 10.10.1 (Bounds on Statistics under Approximate Diagonalization)**.** *Let* $|\xi^t\rangle$ *and* $\left|\widetilde{\xi}^t\right\rangle$ *be defined as in Eq.* (10.10.8) *and Eq.* (10.10.12)*, respectively. Furthermore, let the states and probabilities satisfy*

$$\left\| |s_m(a,\mu)\rangle\langle s_m(a,\mu)| - |w_m(a,\mu)\rangle\langle w_m(a,\mu)| \right\|_1 \leq \epsilon_{a,\mu}, \tag{10.10.25}$$

$$|\zeta_m(a,\mu) - \omega_m(a,\mu)| \leq \epsilon_{\text{prob}}. \tag{10.10.26}$$

Then, it holds

$$\sum_m \zeta_{m|\texttt{test}} \Phi_m \circ \mathcal{E}_m \left[\left|\widetilde{\xi}^t_{|m}\right\rangle\left\langle\widetilde{\xi}^t_{|m}\right|\right] \geq \sum_m \omega^L_{m|\texttt{test}} \left(\Phi_m \circ \mathcal{E}_m \left[|\xi^t_{|m}\rangle\langle\xi^t_{|m}|\right] - \epsilon_m\right), \tag{10.10.27}$$

and

$$\begin{aligned}&\sum_m \zeta_{m|\mathtt{test}} \Phi_m \circ \mathcal{E}_m \left[\left|\widetilde{\xi}^t_{|m}\right\rangle\!\left\langle\widetilde{\xi}^t_{|m}\right|\right] + 1 - \mathbf{p}_{\mathrm{tot}}(\widetilde{\xi}) \\ &\leq \sum_m \omega^U_{m|\mathtt{test}} \left(\Phi_m \circ \mathcal{E}_m \left[\left|\xi^t_{|m}\right\rangle\!\left\langle\xi^t_{|m}\right|\right] + \epsilon_m\right) + 1 - \mathbf{p}^L_{\mathrm{tot}},\end{aligned} \tag{10.10.28}$$

where

$$\omega^U_{m|\mathtt{test}} := \min\{\omega_{m|\mathtt{test}} + \epsilon_{\mathrm{prob}}, 1\}, \tag{10.10.29}$$

$$\omega^L_{m|\mathtt{test}} := \max\{\omega_{m|\mathtt{test}} - \epsilon_{\mathrm{prob}}, 0\}, \tag{10.10.30}$$

$$\mathbf{p}^L_{\mathrm{tot}} := \sum_{\substack{\hat{c}\in\widehat{\mathcal{C}}_{\setminus\mathrm{gen}} \\ a\in\phi^{-1}(\hat{c})}} \sum_{m\leq N_{\mathrm{cons}}} \omega^L_{m|\mathtt{test}} p(a|m,\mathtt{test})\hat{e}_{\hat{c}}, \tag{10.10.31}$$

and

$$\epsilon_m := \sum_{a,\mu} p(a,\mu|\mathtt{test},m)\epsilon_{a,\mu}, \tag{10.10.32}$$

$$\epsilon_{\mathrm{prob}} := \sum_{a,\mu} p(a,\mu|\mathtt{test})\epsilon^{a,\mu}_{\mathrm{proj}}. \tag{10.10.33}$$

Proof. See the proof of the restated Lemma 10.10.1 in Appendix A.2. □

Hence, we have all the tools to state the final expression for h^{QKD}_α and the $\widetilde{H}^{\downarrow,f}_\alpha$-normalization constant $\kappa^\downarrow$ for phase imperfections. We will summarize the result in the next subsection.

10.10.3 Key Rate Construction

In the previous subsections, we found relations that allow us to rigorously incorporate the cutoff N_A, requiring the parameters $\epsilon_m, \epsilon_{\mathrm{prob}}$ and ϵ_{proj}. First, we will show how these can be calculated and then state Theorem 10.10.2 which gives the final expressions required for fixed- and variable-length key rates.

Hence, let us start with the bound ϵ_{full} on the trace distance of the final states in generation rounds. As already mentioned, in App. B of Ref. [68] it was shown that

$$\frac{1}{2}\left\|\rho^{(a,\mu)}_{A'} - \sigma^{(a,\mu)}_{A'}\right\|_1 \leq \epsilon^{a,\mu}_{\mathrm{proj}}, \tag{10.10.34}$$

where

$$\epsilon_{\text{proj}}^{a,\mu} = (1-q)\sqrt{W_{a,\mu}}, \tag{10.10.35}$$

and $W_{a,\mu} \coloneqq \text{Tr}\left[\Pi_{>N_A}\rho_{A'}^{(a,\mu)}\right]$ is the weight contained outside the subspace. Furthermore, note that the weights $W_{a,\mu}$ satisfy $W_{a,\mu} = \text{Tr}\left[\Pi_{>N_A}\sigma_{A'}^{(a,\mu)}\right]$. Thus,

$$\epsilon_{\text{full}} \coloneqq \sum_{a,\mu} p(a,\mu|\texttt{gen})\epsilon_{\text{proj}}^{a,\mu} = (1-q)\sum_{a,\mu} p(a,\mu|\texttt{gen})\sqrt{W_{a,\mu}}. \tag{10.10.36}$$

Next, we continue with ϵ_m and ϵ_{prob}. In Ref. [68, Theorem 6], it was also found that

$$|\zeta_m(a,\mu) - \omega_m(a,\mu)| \leq \epsilon_{\text{proj}}^{a,\mu} \tag{10.10.37}$$

$$\Big\| |s_m(a,\mu)\rangle\langle s_m(a,\mu)| - |w_m(a,\mu)\rangle\langle w_m(a,\mu)| \Big\|_1 \leq \frac{2\epsilon_{\text{proj}}^{a,\mu}}{\delta_m(a,\mu)}, \tag{10.10.38}$$

where for all $m \geq 1$

$$\delta_m(a,\mu) = \min\left\{\omega_m(a,\mu) - \omega_{m-1}(a,\mu) - \epsilon_{\text{proj}}^{a,\mu},\, \omega_{m+1}(a,\mu) - \omega_m(a,\mu) - \epsilon_{\text{proj}}^{a,\mu}\right\}, \tag{10.10.39}$$

and

$$\delta_0(a,\mu) = \omega_1(a,\mu) - \omega_0(a,\mu) - \epsilon_{\text{proj}}^{a,\mu}. \tag{10.10.40}$$

Thus, we find

$$\epsilon_{\text{prob}} = \sum_{a,\mu} p(a,\mu|\texttt{test})\epsilon_{\text{proj}}^{a,\mu} = (1-q)\sum_{a,\mu} p(a,\mu|\texttt{test})\sqrt{W_{a,\mu}}, \tag{10.10.41}$$

$$\epsilon_m = \sum_{a,\mu} p(a,\mu|\texttt{test},m)\epsilon_{a,\mu} = 2(1-q)\sum_{a,\mu} p(a,\mu|\texttt{test},m)\frac{\sqrt{W_{a,\mu}}}{\delta_m(a,\mu)}. \tag{10.10.42}$$

Finally, in the theorem below, we summarize the results for $h_\alpha^{\uparrow,\text{QKD}}$ and the bound on the $\widetilde{H}_\alpha^{\uparrow,f}$-normalization constants $\kappa_i^\uparrow$ assuming the availability of the parameters $\epsilon_m, \epsilon_{\text{prob}}$ and ϵ_{proj}.

Theorem 10.10.2 (Fixed- and Variable-Length key rates with Phase Imperfections)**.** *Let Alice's source prepare states with independent phase imperfections, and let*

$$\Xi_{\text{phase}} = \bigotimes_{i=1}^{n} \Xi_{\text{phase}}^i, \tag{10.10.43}$$

be the source map from [68, App A.]. Furthermore, define the model state

$$\rho_{A'}^{(a,\mu)} = q \sum_{m=0}^{\infty} p(m|\mu) \, |m\rangle\langle m|_a + (1-q) \, |\sqrt{\mu}\rangle\langle\sqrt{\mu}| \, , \tag{10.10.44}$$

with $q \in [0,1]$ *and assume that* $N_{\text{ent}}, N_{\text{cons}} \leq N_A$.

Then, all $\widetilde{H}_\alpha^{\uparrow,f}$*-normalization constants* $\kappa_i^\uparrow$ *can be bounded by*

$$\begin{aligned}
\kappa_i^\uparrow \geq \inf_{\substack{J_m \in \text{Pos}(A'\bar{A}B), \\ \boldsymbol{\nu} \in \mathbb{P}(\widehat{\mathcal{C}}_{\setminus \text{gen}})}} & \left(\frac{1}{1-\alpha} \log \left(\gamma \sum_{c \in \widehat{\mathcal{C}}_{\setminus \text{gen}}} \boldsymbol{\nu}(\hat{c}) 2^{(\alpha-1) f(\hat{c})} \right. \right. \\
& \left. \left. + (1-\gamma) \, 2^{(\alpha-1)(\Delta(\epsilon_{\text{full}}) + f(\mathtt{gen}))} 2^{-(\alpha-1) g_\alpha^{N_{\text{ent}}}(J_0, \ldots J_{N_{\text{ent}}})} \right) \right) \\
s.t. \;\; & \text{Tr}_B[J_m] = \mathbb{I}_{A'\bar{A}}, \\
& \boldsymbol{\nu} \leq \sum_{m \leq N_{\text{cons}}} \omega_{m|\mathtt{test}}^U \left(\boldsymbol{\Psi}_m[J_m] + \epsilon_m \right) + 1 - \mathbf{p}_{\text{tot}}^L, \\
& \boldsymbol{\nu} \geq \sum_{m \leq N_{\text{cons}}} \omega_{m|\mathtt{test}}^L \left(\boldsymbol{\Psi}_m[J_m] - \epsilon_m \right).
\end{aligned} \tag{10.10.45}$$

Furthermore, if the phase imperfections are IID, i.e. all source maps Ξ_{phase}^i *are equal, then* $h_\alpha^{\uparrow,\text{QKD}}$ *is bounded from below by*

$$\begin{aligned}
h_\alpha^{\uparrow,\text{QKD}} \geq \inf_{\substack{\mathbf{q} \in S_{\text{acc}}, \boldsymbol{\nu}_{\widehat{C}} \in \mathbb{P}(\widehat{\mathcal{C}}), \\ J_m \in \text{Pos}(A'\bar{A}B)}} & \left(\frac{\alpha D\big(\mathbf{q} \big\| \boldsymbol{\nu}_{\widehat{C}}\big)}{\alpha - 1} - \Delta(\epsilon_{\text{full}}) + q(\mathtt{gen}) g_\alpha^{N_{\text{ent}}}(J_0, \ldots, J_{N_{\text{ent}}}) \right) \\
s.t. \;\; & \text{Tr}_B[J_m] = \mathbb{I}_{A'\bar{A}}, \\
& \boldsymbol{\nu}_{\widehat{C}} = (\gamma \boldsymbol{\nu}_{\widehat{C}|\mathtt{test}}, 1-\gamma)^T, \\
& \boldsymbol{\nu}_{\widehat{C}|\mathtt{test}} \leq \sum_{m \leq N_{\text{cons}}} \omega_{m|\mathtt{test}}^U \left(\boldsymbol{\Psi}_m[J_m] + \epsilon_m \right) + 1 - \mathbf{p}_{\text{tot}}^L, \\
& \boldsymbol{\nu}_{\widehat{C}|\mathtt{test}} \geq \sum_{m \leq N_{\text{cons}}} \omega_{m|\mathtt{test}}^L \left(\boldsymbol{\Psi}_m[J_m] - \epsilon_m \right).
\end{aligned} \tag{10.10.46}$$

For both quantities, we defined

$$\begin{aligned}
\omega^U_{m|\mathtt{test}} &:= \min\{\omega_{m|\mathtt{test}} + \epsilon_{\text{prob}}, 1\},\\
\omega^L_{m|\mathtt{test}} &:= \max\{\omega_{m|\mathtt{test}} - \epsilon_{\text{prob}}, 0\},\\
\mathbf{p}^L_{\text{tot}} &:= \sum_{\substack{\hat{c}\in\widehat{\mathcal{C}}_{\setminus\text{gen}}\\ a\in\phi^{-1}(\hat{c})}} \sum_{m\leq N_{\text{cons}}} \omega^L_{m|\mathtt{test}} p(a|m,\mathtt{test})\hat{e}_{\hat{c}},
\end{aligned} \tag{10.10.47}$$

and both $g^{N_{\text{ent}}}_{\alpha}(J_0,\dots J_{N_{\text{ent}}})$ *and* $\mathbf{\Psi}_m$ *are defined in terms of the states* $\left|\xi^g_{|m}\right\rangle$ *and* $\left|\xi^t_{|m}\right\rangle$*, respectively.*

Proof. See the proof of the restated Theorem 10.10.2 in Appendix A.2. □

Remark 10.10.3. In the theorem above, the lower bound on the $\widetilde{H}^{\uparrow,f}_{\alpha}$-normalization constants $\kappa^{\uparrow}_i$ was formulated to align with the $\widetilde{H}^{\downarrow,f}_{\alpha}$-normalization constant $\kappa^{\downarrow}$. A similar expression can be derived based on the mixed f-weighting. This yields a structurally similar result but with different α factors, which follows directly from omitting the step in Eq. (A.2.40) of the proof.

10.10.4 Example: 4-6 Protocol with Phase Imperfections

In this section, we present results for variable-length secret key rates for a decoy-state 4-6 protocol with phase imperfections. This specific 4-6 protocol, is directly motivated by the experimental design of the ReFQ project. Notably, prior to our framework, finite-size security for this protocol could only be analyzed using the postselection technique [3, 27], which typically yields poor finite-size key rates, as EUR or phase error correction approaches remain inapplicable here.

Again, we model phase imperfections as in Eq. (10.10.1) derived from earlier works [68, 105]. Under this model, the quality of the phase randomization is characterized by the quality factor $q \in [0,1]$. If $q = 1$ the states are perfectly phase-randomized, i.e. no phase imperfections, and any value of $q < 1$ implies that Alice's source possesses phase imperfections.

As in the previous sections, in each round, Alice decides with probability γ if it is a test or a generation round and always selects the signal intensity μ_s and the Z basis if it was a key generation round. Therefore, as before, Alice's basis choice probabilities are given by $p_z = 1-\gamma$ and $p_x = \gamma$. In a test round, Alice always selects the X basis and all

possible intensities with uniform probability. Now, we present only the one-decoy version of the protocol and choose a *fixed testing probability* of $\gamma = 1/2$. Furthermore, we chose the intensities to be

$$\mu_s = 0.9 \quad \text{and} \quad \mu_2 = 0.2. \tag{10.10.48}$$

On the receiver side (Bob), we now assume a passive six-state detection setup that chooses the Z, X, and Y basis with splitting ratios $p_z = 1 - \gamma$, $p_x = \gamma/2$ and $p_y = \gamma/2$ respectively. To incorporate this into our security proof, we apply the flag-state squasher from [42] with the subspace bound from Theorem 5.3.1. Again, in Section 10.3 we showed that the flag-state squasher can be incorporated.

Finally, we assume the same security parameters $\varepsilon_{\text{PA}} = \varepsilon_{\text{EV}} = \frac{1}{2}10^{-80}$ and error correction model with efficiency $f_{\text{EC}} = 1.1$ as in Section 10.5. We again model channel loss similarly to the qubit protocol and include a misalignment with an angle of $\theta_{\text{hon}}^{\text{misalign}} = 0.03$. For this protocol we calculate the Kraus operators, states and POVM elements numerically and we describe the procedure in Appendix C.3.5.

In Fig. 10.10, we present the variable-length key rates for various signal counts ($n = 10^6, 10^8, 10^{10}$) and quality factors ($q = 1, 0.99$), plotted against channel loss in dB. We compare our results with those of Ref. [3] for the $q = 1$ case using the post-selection technique. Because the key rates in Ref. [3] are significantly lower, we display their results for larger block sizes of $n = 10^{10}, 10^{12}$, and 10^{14}. Focusing first on the perfect case ($q = 1$), the improvements offered by our framework compared to Ref. [3] (which again utilizes the postselection technique from [27]) are dramatic. Even a difference of four orders of magnitude in N is insufficient to bridge the gap; for instance, our rates for $n = 10^6$ signals (blue circles) still outperform the postselection technique with $n = 10^{10}$.

Furthermore, in Fig. 10.10 we present the imperfect case with $q = 0.99$ in orange curves. The reduction in key rates is relatively significant, especially for higher channel losses. For instance, for $n = 10^6$ signals sent, the difference between the perfect case (blue curves) and the imperfect case (orange curves) is rather small, whereas for $n = 10^{10}$ signals sent this difference becomes much bigger.

Here we also mention that similar to all previous examples, at high channel loss, our numerical methods encounter difficulty in converging to a tight enough bound to certify positive key rates. For the imperfect case, this is a particular issue since we cannot use the same simplifications of the optimization problem as for perfect decoy-state protocols. We are required to use N_{cons}-many Choi states as optimization variables. That not only slows down the optimization routine, but also from our experience affects the precision of the underlying solver. Therefore, we again highlight that for future research one should

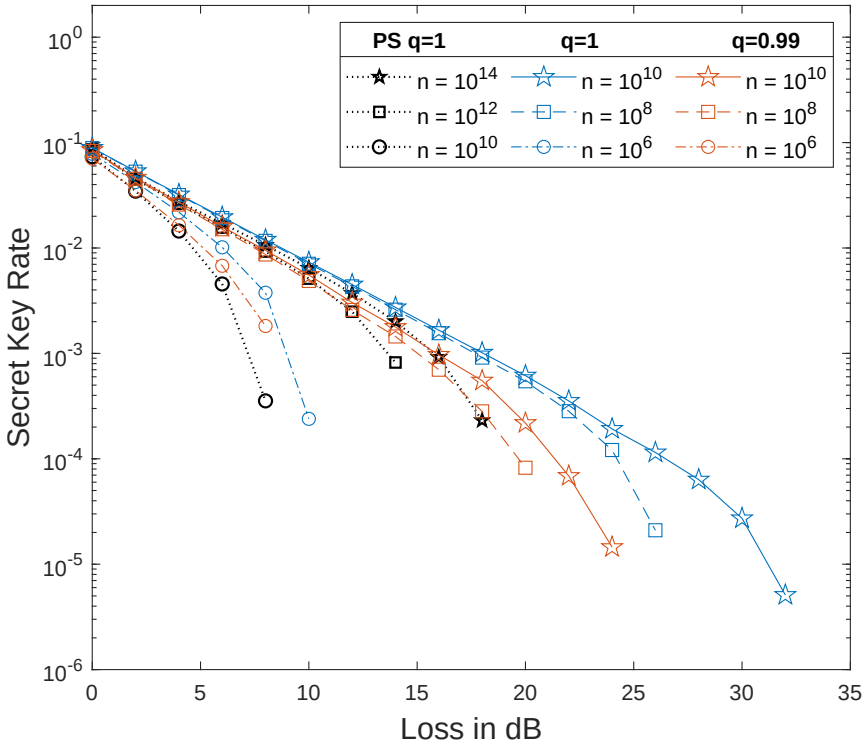


Figure 10.10: Results for variable-length secret key rates of the passive decoy-state 4-6 protocol with *one* decoy intensity and phase imperfections plotted against the channel loss in dB with a varying number of total signals sent $n = 10^6, 10^8, 10^{10}$ and different quality factors $q = 1$ and $q = 0.99$. The security parameters, and error correction efficiency were chosen as $\varepsilon_{\text{PA}} = \varepsilon_{\text{EV}} = \frac{1}{2}10^{-80}$ and $f_{\text{EC}} = 1.1$. As a comparison, we showcase results from [3], which applies the postselection technique [27], for $n = 10^{10}, 10^{12}, 10^{14}$ marked with (PS). For our work we optimize the Rényi parameter α. (Figure 7 from Ref. [4])

focus on improving the performance of the numerical methods. Then, the key rates in the high-loss regime might actually improve even further.

Finally, we discuss the feasibility of integrating phase imperfections into other proof techniques. By applying the methods developed in this section, the postselection technique could, in principle, incorporate independent phase imperfections. However, the key rates would degrade significantly due to the dimensionality scaling of the required correction factor. This would render the performance even worse than the perfect case shown in Fig. 10.10 (black curves). Furthermore, EUR and phase error correction approaches cannot currently account for these imperfections in the finite-size regime. Consequently,

our framework remains the only available option for obtaining practical finite-size key rates in this scenario.

10.11 Conclusion

In this chapter, we developed a flexible framework for security proofs against coherent attacks for both fixed- and variable-length QKD protocols. We introduced new numerical methods for tightly bounding the sandwiched conditional entropy $\widetilde{H}_\alpha^\downarrow$ by relating it to a particular form of Petz divergence and applying a Frank-Wolfe optimization method. In the process, we derived Rényi versions of the dual relations found in [86], which later formed the basis for the numerical framework for asymptotic key rates in Ref. [67].

This complements independent work by Chung et al. [89], which implemented Frank-Wolfe methods to evaluate a tight bound on the Petz conditional entropy $\overline{H}_\alpha^\uparrow$ and calculate IID key rates. Theoretically, the order relation between these two entropies remains an open question. However, any difference between them is expected to be negligible for block sizes exceeding $n = 10^4$.

To demonstrate the versatility of our framework, we applied it to various protocol examples, including a qubit BB84 and an active decoy-state BB84 protocol, comparing it with the current state-of-the-art proof techniques. Additionally, we extended our framework to protocols that possess experimental imperfections, further showcasing its adaptability. Our results demonstrate significant improvements over existing proof techniques, particularly for small block sizes, where our approach consistently outperforms reference methods in achievable secret key rates.

Furthermore, our framework is capable of handling generic block-diagonal protocols, including active and passive decoy-state protocols. Moreover, we can accommodate arbitrary passive detection setups via the flag-state squasher [1, 42], with performance surpassing any proof technique capable of including those setups at the time of writing. In addition, in Section 10.9.1 and Section 10.10.4, we presented the first finite-size security analysis for generic protocols with intensity and phase imperfections. For phase imperfections, we extend beyond prior work that only considered the asymptotic limit, Refs. [68, 105]. For intensity imperfections, Refs. [101, 102] specialized in an active BB84 protocol, and Refs. [103, 104] proved security in the asymptotic limit, whereas our methods provide finite-size security against coherent attacks for generic decoy-state protocols.

From a practical perspective, the ability to simultaneously analyze this combination of independent imperfections represents arguably the most critical result for experimental

QKD systems presented in this thesis. It provides a realistic performance benchmark that can inform the design and optimization of next-generation hardware. Moreover, for experimental implementations, our framework offers significant practical advantages. Specifically, in variable-length protocols, the computational burden of solving optimization problems is a one-time process. Once solved, variable-length rates can be efficiently calculated using a simple affine function, dramatically reducing the overheads and computational demands on QKD devices.

Despite these advancements, some challenges remain. Our current method evaluates the unoptimized version of the sandwiched Rényi entropy, $\widetilde{H}_\alpha^\downarrow$, instead of its optimized counterpart, $\widetilde{H}_\alpha^\uparrow$, leading to minor suboptimalities in high-loss regimes and small block sizes. Additionally, the numerical methods exhibit some instability that requires tuning for reliable results. Future work should focus on improving numerical stability (particularly in high-loss scenarios), better global optimization routines, and exploring alternative optimization techniques (for instance, generalizing the interior-point solvers in [85, 107–109] to handle $\widetilde{H}_\alpha^{\uparrow,f}$ or $\widetilde{H}_\alpha^{\downarrow,f}$) to enhance performance.

Overall, our work establishes a robust and scalable security proof framework for QKD protocols, imposing minimal restrictions while achieving the highest reported secret key rates at the time of writing. Moreover, our framework allows for the inclusion of many protocols and imperfections that are relevant for practical implementations. As such, our work can serve as a foundation for future advancements in security proofs against coherent attacks in practical QKD implementations.

Chapter 11

Conclusion and Outlook

In this thesis, we investigated a spectrum of security proof techniques for QKD protocols, ranging from the asymptotic limit to finite-size security against coherent attacks. We established efficient methods for calculating finite-size key rates for both fixed- and variable-length QKD protocols within the composable security framework of Ref. [54]. Specifically, we focused on practical implementations, such as decoy-state protocols, and incorporated realistic scenarios into the security proofs.

First, we extended the squashing map of Ref. [42] to accommodate arbitrary passive linear optical detection setups. This extension allows for the inclusion of detection inefficiencies and dark counts, ensuring that physically relevant imperfections are rigorously accounted for in security proofs.

Furthermore, we improved standard decoy-state methods in two distinct steps. Initially, we derived asymptotic methods that recover the no-decoy key rates of Ref. [10], offering a significant improvement in terms of tightness. However, these methods could not yet recover the Devetak-Winter formula [96] due to their inability to incorporate correlations between different photon numbers and the necessity of bounding the key rate in a two-step process. Subsequently, we developed a unified framework that performs the key rate optimization and decoy-state analysis in a single step, successfully achieving the Devetak-Winter limit. Notably, this method allows for bounding the relevant entropy with arbitrary precision, even in the finite-size regime.

In addition, we refined the security analysis of QKD protocols against IID collective attacks. Valid for generic QKD protocols, these results yield significantly improved finite-size scaling compared to previous methods of [43]. These gains stem from two main improvements: scaling finite-size corrections by detected signals, not total transmitted ones, and

applying sharper concentration inequalities to tighten the feasible set.

Finally and most notably, we established a flexible numerical framework that delivers finite-size security proofs against coherent attacks for generic QKD protocols. Crucially, by outperforming best-known reference techniques, such as those based on the EUR [22, 24, 37], this framework achieves the highest secret key rates reported to date. This combination of broad applicability to generic protocols within a simple numerical framework and state-of-the-art key rate performance stands as a primary contribution of this thesis. Moreover, we demonstrate efficient variable-length key rates that require complex optimizations only once prior to the protocol start. Building on this, by integrating the aforementioned tight decoy-state methods, we extended these results to model a combination of realistic detection setups with either intensity or phase imperfections, marking the first such key rates derived for generic QKD protocols. Specifically, we established the first finite-size security proof to date for phase imperfections and demonstrated that our framework can tolerate substantial margins of error for intensity fluctuations, directly answering the operational demands of practical implementations such as the ReFQ project.

Despite these advances, a few fundamental gaps remain. While the key rates derived via the MEAT and the Rényi framework in Chapter 10 are among the highest reported to date, the ultimate limit remains unknown. One potential path to close this gap is the use of the optimized sandwiched Rényi entropy instead of the unoptimized version we used in Eq. (10.1.30). However, recent work [110] suggests this yields only diminishing returns. Furthermore, a fundamental gap between achievable results and converse bounds [111] persists. Bridging this gap is essential to determine if the substantial gains provided by the MEAT framework represents the fundamental limit of finite-size performance of QKD.

Additionally, the framework developed in Chapter 10 currently assumes a specific structure for the key rate evaluation. While this does not restrict the choice of QKD protocols, it imposes constraints on how measurement data is processed. Generalizing this framework is conceptually straightforward but requires rigorous formalization; such a generalization could potentially lead to slight increases in key rates. For simplicity, we also employed a tensor product model of the MEAT rather than concatenated channels. While the tensor product model formally precludes so-called "on-the-fly" announcements, extending the proof to concatenated channels should be a simple generalization. Following Ref. [30] we anticipate the numerical results for such a model to remain identical to those presented in this thesis.

While this thesis focused primarily on discrete-variable QKD, continuous-variable (CV) protocols represent an open topic. CV protocols have been analyzed in the finite-size regime against IID collective attacks, see e.g. [73] and coherent attacks [34, 112]. The IID rates

are optimistic, but the rates against coherent attack remain significantly lower. Recently, Ref. [110] applied the MEAT framework to CV protocols; however, their approach, just as Ref. [34], relied on a postulated photon-number cutoff without rigorously reducing the problem to a finite-dimensional one. Therefore, a future direction should be to rigorously incorporate this cutoff to derive a theoretically sound security proof for CV protocols that achieves high key rates.

References

[1] Lars Kamin and Norbert Lütkenhaus. Improved Decoy-State and Flag-State Squashing Methods. *Phys. Rev. Res.* 6.4 (2024), p. 043223.

[2] Lars Kamin, Amir Arqand, Ian George, Norbert Lütkenhaus, and Ernest Y.-Z. Tan. Finite-Size Analysis of Prepare-and-Measure and Decoy-State QKD via Entropy Accumulation. *PRX Quantum* 6.2 (2025), p. 020342. arXiv: 2406.10198 [quant-ph].

[3] Lars Kamin, Devashish Tupkary, and Norbert Lütkenhaus. *Improved Finite-Size Effects in QKD Protocols with Applications to Decoy-State QKD*. 2025. arXiv: 2502.05382 [quant-ph].

[4] Lars Kamin, John Burniston, and Ernest Y.-Z. Tan. *Rényi Security Framework against Coherent Attacks Applied to Decoy-State QKD*. 2025. arXiv: 2504.12248 [quant-ph].

[5] W. K. Wootters and W. H. Zurek. A Single Quantum Cannot Be Cloned. *Nature* 299.5886 (1982), pp. 802–803.

[6] C. H. Bennett and G. Brassard. Quantum Cryptography: Public Key Distribution and Coin Tossing. *Proceedings of International Conference on Computers, Systems and Signal Processing*. Bangalore, India: IEEE Press / IEEE, 1984, p. 175.

[7] A. K. Ekert. Quantum Cryptography Based on Bell's Theorem. *Phys. Rev. Lett.* 67 (1991), p. 661.

[8] Charles H. Bennett, François Bessette, Gilles Brassard, Louis Salvail, and John Smolin. Experimental Quantum Cryptography. *J. Cryptology* 5.1 (1992), pp. 3–28.

[9] Gilles Brassard, Norbert Lütkenhaus, Tal Mor, and Barry C. Sanders. Limitations on Practical Quantum Cryptography. *Physical Review Letters* 85.6 (2000), pp. 1330–1333.

[10] Norbert Lütkenhaus. Security against Individual Attacks for Realistic Quantum Key Distribution. *Phys. Rev. A* 61.5 (2000), p. 052304.

[11] Xiongfeng Ma, Bing Qi, Yi Zhao, and Hoi-Kwong Lo. Practical Decoy State for Quantum Key Distribution. *Phys. Rev. A* 72.1 (2005), p. 012326.

[12] Hoi-Kwong Lo, Xiongfeng Ma, and Kai Chen. Decoy State Quantum Key Distribution. *Phys. Rev. Lett.* 94.23 (2005), p. 230504.

[13] Won-Young Hwang. Quantum Key Distribution with High Loss: Toward Global Secure Communication. *Phys. Rev. Lett.* 91.5 (2003), p. 057901.

[14] Xiang-Bin Wang. Beating the Photon-Number-Splitting Attack in Practical Quantum Cryptography. *Phys. Rev. Lett.* 94.23 (2005), p. 230503.

[15] Masato Koashi. *Simple Security Proof of Quantum Key Distribution via Uncertainty Principle*. 2005. arXiv: quant-ph/0505108.

[16] Masato Koashi. *Efficient Quantum Key Distribution with Practical Sources and Detectors*. 2006. arXiv: quant-ph/0609180.

[17] M Koashi. Simple Security Proof of Quantum Key Distribution Based on Complementarity. *New J. Phys.* 11.4 (2009), p. 045018.

[18] Masahito Hayashi and Toyohiro Tsurumaru. Concise and Tight Security Analysis of the Bennett–Brassard 1984 Protocol with Finite Key Lengths. *New J. Phys.* 14.9 (2012), p. 093014.

[19] Masahito Hayashi and Ryota Nakayama. Security Analysis of the Decoy Method with the Bennett–Brassard 1984 Protocol for Finite Key Lengths. *New J. Phys.* 16.6 (2014), p. 063009.

[20] Marco Tomamichel and Renato Renner. Uncertainty Relation for Smooth Entropies. *Phys. Rev. Lett.* 106.11 (2011), p. 110506.

[21] Marco Tomamichel, Charles Ci Wen Lim, Nicolas Gisin, and Renato Renner. Tight Finite-Key Analysis for Quantum Cryptography. *Nat Commun* 3.1 (2012), p. 634.

[22] Charles Ci Wen Lim, Marcos Curty, Nino Walenta, Feihu Xu, and Hugo Zbinden. Concise Security Bounds for Practical Decoy-State Quantum Key Distribution. *Physical Review A* 89.2 (2014), p. 022307.

[23] Davide Rusca, Alberto Boaron, Fadri Grünenfelder, Anthony Martin, and Hugo Zbinden. Finite-Key Analysis for the 1-Decoy State QKD Protocol. *Applied Physics Letters* 112.17 (2018), p. 171104.

[24] Devashish Tupkary, Shlok Nahar, Pulkit Sinha, and Norbert Lütkenhaus. *Phase Error Rate Estimation in QKD with Imperfect Detectors*. 2024. arXiv: 2408.17349 [quant-ph].

[25] Jerome Wiesemann et al. *A Consolidated and Accessible Security Proof for Finite-Size Decoy-State Quantum Key Distribution*. 2024. arXiv: 2405.16578 [quant-ph].

[26] Matthias Christandl, Robert König, and Renato Renner. Postselection Technique for Quantum Channels with Applications to Quantum Cryptography. *Phys. Rev. Lett.* 102.2 (2009), p. 020504.

[27] Shlok Nahar, Devashish Tupkary, Yuming Zhao, Norbert Lütkenhaus, and Ernest Y.-Z. Tan. Postselection Technique for Optical Quantum Key Distribution with Improved de Finetti Reductions. *PRX Quantum* 5.4 (2024), p. 040315.

[28] Frédéric Dupuis, Omar Fawzi, and Renato Renner. Entropy Accumulation. *Commun. Math. Phys.* 379.3 (2020), pp. 867–913.

[29] Frédéric Dupuis and Omar Fawzi. Entropy Accumulation With Improved Second-Order Term. *IEEE Transactions on Information Theory* 65.11 (2019), pp. 7596–7612.

[30] Tony Metger, Omar Fawzi, David Sutter, and Renato Renner. Generalised Entropy Accumulation. *2022 IEEE 63rd Annual Symposium on Foundations of Computer Science (FOCS)*. 2022, pp. 844–850.

[31] Tony Metger, Omar Fawzi, David Sutter, and Renato Renner. Generalised Entropy Accumulation. *Commun. Math. Phys.* 405.11 (2024), p. 261.

[32] Ian George, Jie Lin, Thomas van Himbeeck, Kun Fang, and Norbert Lütkenhaus. *Finite-Key Analysis of Quantum Key Distribution with Characterized Devices Using Entropy Accumulation*. 2022. arXiv: 2203.06554 [quant-ph].

[33] Tony Metger and Renato Renner. Security of Quantum Key Distribution from Generalised Entropy Accumulation. *Nat Commun* 14.1 (2023), p. 5272.

[34] Stefan Bäuml, Carlos Pascual-García, Victoria Wright, Omar Fawzi, and Antonio Acín. Security of Discrete-Modulated Continuous-Variable Quantum Key Distribution. *Quantum* 8 (2024), p. 1418.

[35] Charles H Bennett and Gilles Brassard. Quantum Cryptography: Public Key Distribution and Coin Tossing. *Theoretical Computer Science* 560 (2014), pp. 7–11.

[36] Renato Renner. *Security of Quantum Key Distribution*. 2006. arXiv: quant-ph/0512258.

[37] Marco Tomamichel and Anthony Leverrier. A Largely Self-Contained and Complete Security Proof for Quantum Key Distribution. *Quantum* 1 (2017), p. 14.

[38] Frédéric Dupuis. Privacy Amplification and Decoupling Without Smoothing. *IEEE Transactions on Information Theory* 69.12 (2023), pp. 7784–7792.

[39] A. Scott et al. The QEYSSAT Mission: On-Orbit Demonstration of Secure Optical Communications Network Technologies. *Environmental Effects on Light Propagation and Adaptive Systems III*. Vol. 11532. SPIE, 2020, pp. 71–76.

[40] Andrew Conrad et al. *Drone- and Vehicle-Based Quantum Key Distribution*. 2025. arXiv: 2505.17587 [quant-ph].

[41] Amir Arqand and Ernest Y.-Z. Tan. *Marginal-Constrained Entropy Accumulation Theorem*. 2025. arXiv: 2502.02563 [quant-ph].

[42] Yanbao Zhang, Patrick J. Coles, Adam Winick, Jie Lin, and Norbert Lütkenhaus. Security Proof of Practical Quantum Key Distribution with Detection-Efficiency Mismatch. *Physical Review Research* 3.1 (2021), p. 013076.

[43] Ian George, Jie Lin, and Norbert Lütkenhaus. Numerical Calculations of Finite Key Rate for General Quantum Key Distribution Protocols. *Physical Review Research* 3.1 (2020), p. 013274. arXiv: 2004.11865.

[44] John Watrous. *The Theory of Quantum Information*. 1st ed. Cambridge University Press, 2018.

[45] Valter Moretti. *Spectral Theory and Quantum Mechanics*. Vol. 110. UNITEXT. Cham: Springer International Publishing, 2017.

[46] Klaas Landsman. *Foundations of Quantum Theory*. Vol. 188. Fundamental Theories of Physics. Cham: Springer International Publishing, 2017.

[47] Marco Tomamichel. *Quantum Information Processing with Finite Resources*. Vol. 5. SpringerBriefs in Mathematical Physics. Cham: Springer International Publishing, 2016.

[48] Michael A. Nielsen and Isaac L. Chuang. *Quantum Computation and Quantum Information*. 10th anniversary edition. Cambridge: Cambridge university press, 2010.

[49] Hisaharu Umegaki. Conditional Expectation in an Operator Algebra. IV. Entropy and Information. *Kodai Mathematical Seminar Reports* 14.2 (1962), pp. 59–85.

[50] Sumeet Khatri and Mark M. Wilde. *Principles of Quantum Communication Theory: A Modern Approach*. 2024. arXiv: 2011.04672 [cond-mat, physics:hep-th, physics:math-ph, physics:quant-ph].

[51] Renato Renner. Security of Quantum Key Distribution. Doctoral Thesis. ETH Zurich, 2005.

[52] Stephen P. Boyd and Lieven Vandenberghe. *Convex Optimization*. Cambridge, UK ; New York: Cambridge University Press, 2004.

[53] Marguerite Frank and Philip Wolfe. An Algorithm for Quadratic Programming. *Naval Research Logistics Quarterly* 3.1-2 (1956), pp. 95–110.

[54] Christopher Portmann and Renato Renner. Security in Quantum Cryptography. *Rev. Mod. Phys.* 94.2 (2022), p. 025008.

[55] Devashish Tupkary, Ernest Y.-Z. Tan, and Norbert Lütkenhaus. Security Proof for Variable-Length Quantum Key Distribution. *Phys. Rev. Res.* 6.2 (2024), p. 023002.

[56] Charles H. Bennett, Gilles Brassard, and N. David Mermin. Quantum Cryptography without Bell's Theorem. *Phys. Rev. Lett.* 68.5 (1992), pp. 557–559.

[57] Agnes Ferenczi and Norbert Lütkenhaus. Symmetries in Quantum Key Distribution and the Connection between Optimal Attacks and Optimal Cloning. *Phys. Rev. A* 85.5 (2012), p. 052310.

[58] Karol Horodecki, Michal Horodecki, Pawel Horodecki, and Jonathan Oppenheim. General Paradigm for Distilling Classical Key From Quantum States. *IEEE Transactions on Information Theory* 55.4 (2009), pp. 1898–1929.

[59] Marco Tomamichel, Christian Schaffner, Adam Smith, and Renato Renner. Leftover Hashing Against Quantum Side Information. *IEEE Transactions on Information Theory* 57.8 (2011), pp. 5524–5535.

[60] Roger A. Horn and Charles R. Johnson. *Matrix Analysis*. Second edition, corrected reprint. New York, NY: Cambridge University Press, 2017.

[61] Rajendra Bhatia and Fuad Kittaneh. Cartesian Decompositions and Schatten Norms. *Linear Algebra and its Applications* 318.1 (2000), pp. 109–116.

[62] Normand J. Beaudry, Tobias Moroder, and Norbert Lütkenhaus. Squashing Models for Optical Measurements in Quantum Communication. *Phys. Rev. Lett.* 101.9 (2008), p. 093601.

[63] O. Gittsovich et al. Squashing Model for Detectors and Applications to Quantum-Key-Distribution Protocols. *Physical Review A* 89.1 (2014), p. 012325.

[64] Nicky Kai Hong Li. Application of the Flag-State Squashing Model to Numerical Quantum Key Distribution Security Analysis. PhD thesis. University of Waterloo, 2020. HDL: 10012/16320.

[65] Ulf Leonhardt. Quantum Physics of Simple Optical Instruments. *Rep. Prog. Phys.* 66.7 (2003), p. 1207.

[66] Nicky Kai Hong Li and Norbert Lütkenhaus. Improving Key Rates of the Unbalanced Phase-Encoded BB84 Protocol Using the Flag-State Squashing Model. *Phys. Rev. Research* 2.4 (2020), p. 043172.

[67] Adam Winick, Norbert Lütkenhaus, and Patrick J. Coles. Reliable Numerical Key Rates for Quantum Key Distribution. *Quantum* 2 (2018), p. 77.

[68] Shlok Nahar, Twesh Upadhyaya, and Norbert Lütkenhaus. Imperfect Phase Randomization and Generalized Decoy-State Quantum Key Distribution. *Phys. Rev. Appl.* 20.6 (2023), p. 064031.

[69] Zhiyao Wang, Devashish Tupkary, and Shlok Nahar. *Phase Error Estimation for Passive Detection Setups with Imperfections and Memory Effects*. 2025. arXiv: 2508.21486 [quant-ph].

[70] Patrick Rice and Jim Harrington. *Numerical Analysis of Decoy State Quantum Key Distribution Protocols*. 2009. arXiv: 0901.0013 [quant-ph].

[71] Wenyuan Wang and Norbert Lütkenhaus. Numerical Security Proof for the Decoy-State BB84 Protocol and Measurement-Device-Independent Quantum Key Distribution Resistant against Large Basis Misalignment. *Phys. Rev. Res.* 4.4 (2022), p. 043097.

[72] *NIST Digital Library of Mathematical Functions*. https://dlmf.nist.gov/, Release 1.2.4 of 2025-03-15. F. W. J. Olver, A. B. Olde Daalhuis, D. W. Lozier, B. I. Schneider, R. F. Boisvert, C. W. Clark, B. R. Miller, B. V. Saunders, H. S. Cohl, and M. A. McClain, eds.

[73] Florian Kanitschar, Ian George, Jie Lin, Twesh Upadhyaya, and Norbert Lütkenhaus. Finite-Size Security for Discrete-Modulated Continuous-Variable Quantum Key Distribution Protocols. *PRX Quantum* 4.4 (2023), p. 040306.

[74] Maurice Fréchet. Généralisation du théorème des probabilités totales. *Fundamenta Mathematicae* 25.1 (1935), pp. 379–387.

[75] A. M. Zubkov and A. A. Serov. A Complete Proof of Universal Inequalities for the Distribution Function of the Binomial Law. *Theory Probab. Appl.* 57.3 (2013), pp. 539–544.

[76] Shun Kawakami, Atsushi Taniguchi, Yoshihide Tonomura, Koichi Takasugi, and Koji Azuma. *Security of the BB84 Protocol with Receiver's Passive Biased Basis Choice*. 2025. arXiv: 2507.04248 [quant-ph].

[77] Amir Arqand, Thomas A. Hahn, and Ernest Y.-Z. Tan. *Generalized Rényi Entropy Accumulation Theorem and Generalized Quantum Probability Estimation*. 2024. arXiv: 2405.05912v4 [quant-ph].

[78] Omar Fawzi, Jan Kochanowski, Cambyse Rouzé, and Thomas Van Himbeeck. *Additivity and Chain Rules for Quantum Entropies via Multi-Index Schatten Norms*. 2025. arXiv: 2502.01611 [quant-ph].

[79] Thomas van Himbeeck and Peter Brown. *Tight and general finite-size security of quantum key distribution*. In preparation. 2023.

[80] Jie Lin, Twesh Upadhyaya, and Norbert Lütkenhaus. Asymptotic Security Analysis of Discrete-Modulated Continuous-Variable Quantum Key Distribution. *Phys. Rev. X* 9.4 (2019), p. 041064.

[81] Wen-Zhao Liu et al. Device-Independent Randomness Expansion against Quantum Side Information. *Nat. Phys.* 17.4 (2021), pp. 448–451.

[82] Marco Tomamichel, Mario Berta, and Masahito Hayashi. Relating Different Quantum Generalizations of the Conditional Rényi Entropy. *Journal of Mathematical Physics* 55.8 (2014), p. 082206.

[83] Vaisakh Mannalath, Víctor Zapatero, and Marcos Curty. Sharp Finite Statistics for Quantum Key Distribution. *Phys. Rev. Lett.* 135.2 (2025), p. 020803.

[84] Hamza Fawzi and James Saunderson. Lieb's Concavity Theorem, Matrix Geometric Means, and Semidefinite Optimization. *Linear Algebra and its Applications* 513 (2017), pp. 240–263.

[85] Kerry He, James Saunderson, and Hamza Fawzi. *Operator Convexity along Lines, Self-Concordance, and Sandwiched Rényi Entropies*. 2025. arXiv: 2502.05627 [math].

[86] Patrick J. Coles. Unification of Different Views of Decoherence and Discord. *Phys. Rev. A* 85.4 (2012), p. 042103.

[87] Gilad Gour and Mark M. Wilde. Entropy of a Quantum Channel. *Phys. Rev. Res.* 3.2 (2021), p. 023096.

[88] Kriss Gutierrez Anco, Tristan Nemoz, and Peter Brown. *How Much Secure Randomness Is in a Quantum State?* 2024. arXiv: 2410.16447 [quant-ph].

[89] Rebecca R. B. Chung, Nelly H. Y. Ng, and Yu Cai. Generalized Numerical Framework for Improved Finite-Sized Key Rates with Rényi Entropy. *Phys. Rev. A* 112.1 (2025), p. 012612.

[90] Marco Tomamichel, Roger Colbeck, and Renato Renner. A Fully Quantum Asymptotic Equipartition Property. *IEEE Transactions on Information Theory* 55.12 (2009), pp. 5840–5847.

[91] Rupert L. Frank and Elliott H. Lieb. Monotonicity of a Relative Rényi Entropy. *Journal of Mathematical Physics* 54.12 (2013), p. 122201.

[92] Martin Müller-Lennert, Frédéric Dupuis, Oleg Szehr, Serge Fehr, and Marco Tomamichel. On Quantum Rényi Entropies: A New Generalization and Some Properties. *Journal of Mathematical Physics* 54.12 (2013), p. 122203.

[93] Salman Beigi. Sandwiched Rényi Divergence Satisfies Data Processing Inequality. *Journal of Mathematical Physics* 54.12 (2013), p. 122202.

[94] Milán Mosonyi and Tomohiro Ogawa. Quantum Hypothesis Testing and the Operational Interpretation of the Quantum Rényi Relative Entropies. *Commun. Math. Phys.* 334.3 (2015), pp. 1617–1648.

[95] Guillermo Currás-Lorenzo, Álvaro Navarrete, Margarida Pereira, and Kiyoshi Tamaki. Finite-Key Analysis of Loss-Tolerant Quantum Key Distribution Based on Random Sampling Theory. *Phys. Rev. A* 104.1 (2021), p. 012406.

[96] Igor Devetak and Andreas Winter. Distillation of Secret Key and Entanglement from Quantum States. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences* 461.2053 (2005), pp. 207–235.

[97] Devashish Tupkary, Ernest Y.-Z. Tan, Shlok Nahar, Lars Kamin, and Norbert Lütkenhaus. *QKD Security Proofs for Decoy-State BB84: Protocol Variations, Proof Techniques, Gaps and Limitations*. 2025. arXiv: 2502.10340 [quant-ph].

[98] Thomas Attema, Joost W. Bosman, and Niels M. P. Neumann. Optimizing the Decoy-State BB84 QKD Protocol Parameters. *Quantum Inf Process* 20.4 (2021), p. 154.

[99] Margarida Pereira, Go Kato, Akihiro Mizutani, Marcos Curty, and Kiyoshi Tamaki. Quantum Key Distribution with Correlated Sources. *Science Advances* 6.37 (2020), eaaz4487.

[100] Shlok Nahar and Norbert Lütkenhaus. *Imperfect Detectors for Adversarial Tasks with Applications to Quantum Key Distribution*. 2025. arXiv: 2503.06328 [quant-ph].

[101] Yang Wang, Wan-Su Bao, Chun Zhou, Mu-Sheng Jiang, and Hong-Wei Li. Tight Finite-Key Analysis of a Practical Decoy-State Quantum Key Distribution with Unstable Sources. *Phys. Rev. A* 94.3 (2016), p. 032335.

[102] Chun Zhou et al. Finite-Key Analysis of 1-Decoy Method Quantum Key Distribution with Intensity Fluctuation. *Applied Sciences* 12.9 (2022), p. 4709.

[103] Víctor Zapatero, Álvaro Navarrete, Kiyoshi Tamaki, and Marcos Curty. Security of Quantum Key Distribution with Intensity Correlations. *Quantum* 5 (2021), p. 602.

[104] Xoel Sixto, Víctor Zapatero, and Marcos Curty. Security of Decoy-State Quantum Key Distribution with Correlated Intensity Fluctuations. *Phys. Rev. Appl.* 18.4 (2022), p. 044069.

[105] Guillermo Currás-Lorenzo, Shlok Nahar, Norbert Lütkenhaus, Kiyoshi Tamaki, and Marcos Curty. Security of Quantum Key Distribution with Imperfect Phase Randomisation. *Quantum Sci. Technol.* 9.1 (2023), p. 015025.

[106] Andreas Bluhm, Ángela Capel, Paul Gondolf, and Tim Möbus. Unified Framework for Continuity of Sandwiched Rényi Divergences. *Ann. Henri Poincaré* (2024).

[107] Hao Hu, Jiyoung Im, Jie Lin, Norbert Lütkenhaus, and Henry Wolkowicz. Robust Interior Point Method for Quantum Key Distribution Rate Computation. *Quantum* 6 (2022), p. 792.

[108] Andrés González Lorente, Pablo V. Parellada, Miguel Castillo-Celeita, and Mateus Araújo. Quantum Key Distribution Rates from Non-Symmetric Conic Optimization. *Quantum* 9 (2025), p. 1657.

[109] Kerry He, James Saunderson, and Hamza Fawzi. *Exploiting Structure in Quantum Relative Entropy Programs*. 2024. arXiv: 2407.00241 [quant-ph].

[110] Mariana Navarro, Andrés González Lorente, Pablo V. Parellada, Carlos Pascual-García, and Mateus Araújo. *Finite-Size Quantum Key Distribution Rates from Rényi Entropies Using Conic Optimization*. 2025. arXiv: 2511.10584 [quant-ph].

[111] Yu-Chen Shen, Li Gao, and Hao-Chung Cheng. Strong Converse for Privacy Amplification against Quantum Side Information. *2022 IEEE International Symposium on Information Theory (ISIT)*. 2022, pp. 1880–1885.

[112] Ignatius William Primaatmaja, Wen Yu Kon, and Charles Lim. *Discrete-Modulated Continuous-Variable Quantum Key Distribution Secure against General Attacks*. 2024. arXiv: 2409.02630 [quant-ph].

[113] H. Epstein. Remarks on Two Theorems of E. Lieb. *Commun.Math. Phys.* 31.4 (1973), pp. 317–325.

[114] Lars Kamin and Norbert Lütkenhaus. *Improved Decoy-state and Flag-state Squashing Methods - GitHub Repository*. Available at https://github.com/Optical-Quantum-Communication-Theory/ImprovedDecoyStateAndFlagStateSquashingMethods. 2024.

[115] Lars Kamin, Devashish Tupkary, and Norbert Lütkenhaus. *Improved Finite-Size Effects in QKD Protocols with Applications to Decoy-State QKD - GitHub Repository*. Available at https://github.com/Optical-Quantum-Communication-Theory/Improved-finite-size-effects-in-QKD-protocols-with-applications-to-decoy-state-QKD. 2025.

[116] Lars Kamin, Amir Arqand, Ian George, Norbert Lütkenhaus, and Ernest Y.-Z. Tan. *Finite-Size Analysis of Prepare-and-Measure and Decoy-State QKD via Entropy Accumulation - GitHub Repository*. Available at https://github.com/Optical-Quantum-Communication-Theory/FiniteSizeAnalysisOfPrepareAndMeasureAndDecoyStateQKDViaEntropyAccumulation. 2024.

[117] Lars Kamin, John Burniston, and Ernest Y.-Z. Tan. *Rényi Security Framework against Coherent Attacks Applied to Decoy-State QKD - GitHub Repository*. Available at https://github.com/Optical-Quantum-Communication-Theory/Renyi-security-framework. 2025.

Appendix A

Additional Proofs for Rényi Framework

A.1 Convexity Proofs

In this section, we will present the numerical algorithm and prove the joint convexity of the optimization problems for both the single-round quantity h_α^{QKD} and the normalization constants $\kappa^{\downarrow}$ and $\kappa^{\supset}$.

The convexity statements required for our numerical algorithm do not follow immediately from [41, 77] for several reasons. First, we will employ a Frank-Wolfe algorithm, see Algorithm 1, which requires the knowledge of the derivative. Even more, the derivative needs to exist anywhere in the domain. However, just like in the case of the von-Neumann entropy [67] for a pure state the derivative does not exist. Thus, we will apply perturbations to the output of the $\mathcal{G}$ to perturb any state that does not posses a derivative, such that the overall derivative indeed exists.

Therefore, we will first construct a perturbed protocol map $\mathcal{G}_\delta$ and show the joint convexity for the perturbed optimization problems.The exact details of the implementation of the Frank-Wolfe algorithm and the exact perturbation can be found in [4, Appendix G]

For brevity of notation, in all of this section, wherever necessary, we use the following identification of registers to the main text,

$$E \mapsto \tilde{T}E \quad \text{and} \quad Q \mapsto S_Q I A B X Y. \tag{A.1.1}$$

Additionally, for completeness, similar to Theorem 10.1.1, we define a generic pinching map $\mathcal{Z} \in \mathrm{CPTP}(Q)$ and its Stinespring dilation.

Definition A.1.1. Let $\{Z_i\}_{i=1\ldots n} \subset \mathrm{Pos}(Q)$ be a set of projection operators on Q satisfying $\sum_{i=1}^{n} Z_i = \mathbb{I}_Q$ and $Z_i Z_j = \delta_{i,j} Z_i$. We define the isometry $V_{\mathcal{Z}} \in \mathrm{U}(Q, SQ)$ and the Z-map, $\mathcal{Z} \in \mathrm{CPTP}(Q)$, as

$$V_{\mathcal{Z}} = \sum_z |i\rangle_S \otimes Z_i, \tag{A.1.2}$$

$$\mathcal{V}_{\mathcal{Z}}(\rho) = V_{\mathcal{Z}} \rho V_{\mathcal{Z}}^\dagger, \quad \mathcal{Z}(\rho) = \mathrm{Tr}_S \circ \mathcal{V}_{\mathcal{Z}}(\rho). \tag{A.1.3}$$

A.1.1 Convexity of Quantities related to Petz Conditional Rényi entropy

As an initial step, we prove the convexity and concavity of several quantities related to the Petz conditional Rényi entropy.

In Theorem 10.1.1, we used a dual relation to relate $\widetilde{H}_\alpha^\downarrow$ to $\overline{H}_{\frac{1}{\alpha}}^\uparrow$ and the following expression to simplify $\overline{H}_\alpha^\uparrow(S|Q)$,

$$\overline{H}_\alpha^\uparrow(S|Q)_\sigma = \frac{\alpha}{1-\alpha} \log \mathrm{Tr}\left[\mathcal{Z}\left(\rho_Q^\alpha\right)^{\frac{1}{\alpha}}\right], \tag{A.1.4}$$

where the channel $\mathcal{Z}$ is again a pinching channel, as defined above, and σ_{SQE} is created from ρ_{QE} as in Theorem 10.1.1 by

$$\sigma_{SQE} \coloneqq V_{\mathcal{Z}} \rho_{QE} V_{\mathcal{Z}}^\dagger. \tag{A.1.5}$$

Thus, for simplicity, we define the quantity $\widehat{Q}_\alpha$ as

$$\widehat{Q}_\alpha(\rho_Q) \coloneqq \mathrm{Tr}\left[\mathcal{Z}\left(\rho_Q^\alpha\right)^{\frac{1}{\alpha}}\right], \tag{A.1.6}$$

for all $\rho \in \mathrm{S}_=(Q)$ and $\alpha \in [0, 2]$. With this definition, we can write

$$\overline{H}_\alpha^\uparrow(S|Q)_\sigma = \frac{\alpha}{1-\alpha} \log\left(\widehat{Q}_\alpha(\rho_Q)\right). \tag{A.1.7}$$

Theorem A.1.2. *The function $\rho \mapsto \widehat{Q}_\alpha(\rho)$ is concave for all $\rho \in \mathrm{S}_\leq(Q)$ and $\alpha \in [0, 1]$.*

Proof. From [91, Lemma 5] (originally shown by Epstein [113] for $t \in [0, 1]$), it follows that the function

$$\mathrm{Tr}\left[\left(K X^t K^\dagger\right)^{\frac{1}{t}}\right], \tag{A.1.8}$$

is concave in $X \in \mathrm{Pos}(Q)$ for $t \in [-1,1], t \neq 0$ given any fixed operator K. Thus, we will define an operator K such that we can rewrite $\widehat{Q}_\alpha$ in the form above. First, note that we can write

$$\mathcal{Z}(\rho_Q) = \mathrm{Tr}_S[V_{\mathcal{Z}} \rho_Q V_{\mathcal{Z}}^\dagger], \tag{A.1.9}$$

where $V_{\mathcal{Z}}$ is given as in Definition A.1.1. Next, let S' be a Hilbert space of the same dimension as S and let $|\Phi^+\rangle_{SS'}$ be the unnormalized maximally entangled state on SS'. This allows us to rewrite the partial trace of any trace-class operator $M_{SQ} \in \mathrm{Pos}(SQ)$ by

$$\mathrm{Tr}_S[M_{SQ}] = \left\langle \Phi^+ \right|_{SS'} (M_{SQ} \otimes \mathbb{I}_{S'}) \left| \Phi^+ \right\rangle_{SS'}. \tag{A.1.10}$$

Specifically we find

$$\mathcal{Z}(\rho_Q) = \left\langle \Phi^+ \right|_{SS'} \left(V_{\mathcal{Z}} \rho_Q V_{\mathcal{Z}}^\dagger \otimes \mathbb{I}_{S'} \right) \left| \Phi^+ \right\rangle_{SS'}. \tag{A.1.11}$$

Hence, let us define K as

$$K \coloneqq \left(\left\langle \Phi^+ \right|_{SS'} \otimes \mathbb{I}_Q \right) (V_{\mathcal{Z}} \otimes \mathbb{I}_{S'}), \tag{A.1.12}$$

which allows us to rewrite $\widehat{Q}_\alpha(\rho_Q)$ as

$$\widehat{Q}_\alpha(\rho_Q) = \mathrm{Tr}\left[\left(K \left(\rho_Q \otimes \mathbb{I}_{S'} \right)^\alpha K^\dagger \right)^{\frac{1}{\alpha}} \right]. \tag{A.1.13}$$

Using this construction of K the statement follows immediately from [91, Lemma 5]. □

A.1.2 Perturbed Objective Function

Next, similar to Eq. (10.1.29) we define the objective function g_α in terms of the state ρ instead of the channel, which we will indicate by $\bar{g}_\alpha$.

Definition A.1.3. Let $\mathcal{Z}$ be the pinching channel of Corollary 10.1.2 and $\mathcal{G}$ be the CPTP protocol map as in Definition 8.2.1. We define $\bar{g}_\alpha : \mathrm{S}_=(AB) \to \mathbb{R}$ as

$$\begin{aligned} \bar{g}_\alpha(\rho_{AB}) &= \widetilde{H}_\alpha^\downarrow(S|\tilde{T}E)_{\mathcal{M}_{\mathrm{QKD}}(\rho_{ABE})}, \\ &= \frac{-1}{\alpha - 1} \log \mathrm{Tr}\left[\left(\mathcal{Z}\left(\mathcal{G}(\rho_{AB})^{\frac{1}{\alpha}} \right) \right)^\alpha \right], \end{aligned} \tag{A.1.14}$$

where we made the negativity of the prefactor in front of the logarithm explicit.

For simplicity of the later convexity proofs, we did not use the final simplification of the expression to CPTNI maps as in Corollary 10.1.2. Importantly, this definition allows us to rewrite g_α from Eq. (10.1.29) as

$$g_\alpha(J) = \bar{g}_\alpha \circ \chi_g(J). \tag{A.1.15}$$

However, for the existence of gradients everywhere on the domain, we require a perturbed version of $\bar{g}_\alpha$, which adds an affine perturbation to $\mathcal{G}$. Thus, as the next step, define the perturbed protocol map $\mathcal{G}_\delta$.

Definition A.1.4. Let $\mathcal{G}$ and $\hat{\mathcal{G}}$ be the CPTP and CPTNI protocol maps as in Definition 8.2.1. We define the perturbed map $\mathcal{G}_\delta$ for $\varepsilon_1, \varepsilon_2 \in [0,1)$ with $\delta = \varepsilon_1 + \varepsilon_2 - \varepsilon_1\varepsilon_2$ ($\delta \in [0,1)$) as

$$\mathcal{G}_\delta(\rho) \coloneqq (1-\delta)\,\mathcal{G}(\rho) + \delta\tau(\rho), \tag{A.1.16}$$

where

$$\tau(\rho) = \left(\frac{p\mathbb{I}_{\hat{S}_Q}}{\delta \dim(\hat{S}_Q)} + \left(\mathrm{Tr}[\rho] - \frac{p}{\delta}\right)|\perp\rangle\langle\perp|\right) \otimes \frac{\mathbb{I}_{ABXYI}}{\dim(ABXYI)}, \tag{A.1.17}$$

and

$$p = \varepsilon_1 \,\mathrm{Tr}\left[\hat{\mathcal{G}}(\rho)\right](1-\varepsilon_2) + \varepsilon_2 \,\mathrm{Tr}[\rho]. \tag{A.1.18}$$

Then, the perturbed objective $\bar{g}_\alpha[\delta]$ in terms of the state is given by

$$\bar{g}_\alpha[\delta](\rho_{AB}) \coloneqq \frac{-1}{\alpha - 1} \log\left[\widehat{Q}_{\frac{1}{\alpha}} \circ \mathcal{G}_\delta(\rho_{AB})\right], \tag{A.1.19}$$

and the perturbed objective in terms of Eve's channel is

$$g_\alpha[\delta](J) \coloneqq \frac{-1}{\alpha - 1} \log\left[\widehat{Q}_{\frac{1}{\alpha}} \circ \mathcal{G}_\delta \circ \chi_g(J)\right]. \tag{A.1.20}$$

We note that for $\delta = 0$ in both cases we recover the unperturbed objective functions and most importantly the perturbed map $\mathcal{G}_\delta$ is still a linear function on $\mathrm{S}_{\leq}(AB)$.

Next, in preparation of the joint convexity proof of the optimization problem of h_α^{QKD} in Eq. (10.1.30) (and its perturbed version), we show that the term inside the logarithm, i.e. $\widehat{Q}_{\frac{1}{\alpha}} \circ \mathcal{G}_\delta$ is still concave for $\alpha > 1$.

Corollary A.1.5. *The function $\rho_{AB} \mapsto \widehat{Q}_{\frac{1}{\alpha}} \circ \mathcal{G}_\delta(\rho_{AB})$ is concave and $\bar{g}_\alpha[\delta]$ is convex for all $\rho \in \mathrm{S}_{\leq}(AB)$ and $\alpha \in (1, \infty)$.*

Proof. The perturbed protocol map $\mathcal{G}_\delta$ applies an affine transformation to ρ. Furthermore, by Theorem A.1.2, $\overline{Q}_{\frac{1}{\alpha}}$ is concave for $\alpha \in [1, \infty)$. Thus, since $\mathcal{G}_\delta$ is an affine map by Lemma 2.4.4, $\rho_{AB} \mapsto \widehat{Q}_{\frac{1}{\alpha}} \circ \mathcal{G}_\delta(\rho_{AB})$ is concave.

Additionally, since the negative logarithm, $-\log(x)$, is convex and non-increasing, by [52, Section 3.2.4] the function $\bar{g}_\alpha[\delta]$ is convex. □

The perturbed objective is a lower bound on the unperturbed version and satisfies the following property.

Lemma A.1.6. *For all $\rho_{AB} \in \mathrm{S}_=(AB)$, $\alpha \in (1, \infty)$ and $\delta \in [0, 1)$, $\bar{g}_\alpha[\delta]$ satisfies*

$$\frac{\bar{g}_\alpha[\delta](\rho_{AB})}{1-\delta} \leq \bar{g}_\alpha(\rho_{AB}) \tag{A.1.21}$$

Proof. First, a simple calculation shows that $\log[\widehat{Q}_{\frac{1}{\alpha}}(\tau(\rho))] = 0$ for all $\rho \in \mathrm{S}_=(AB)$. Next, using this and the fact that $\bar{g}_\alpha$ is a convex function we find

$$\bar{g}_\alpha[\delta](\rho) \leq (1-\delta)\bar{g}_\alpha(\rho) - \frac{\delta}{\alpha - 1} \log[\widehat{Q}_{\frac{1}{\alpha}}(\tau(\rho))] = (1-\delta)\bar{g}_\alpha(\rho). \tag{A.1.22}$$

Then, rearranging this equation yields the bound. □

Next, we show another property which is required for the joint convexity of the $\widetilde{H}_\alpha^{\downarrow,f}$-normalization constant $\kappa^\downarrow$.

Corollary A.1.7. *The functions $\rho \mapsto 2^{-(\alpha-1)\frac{\bar{g}_\alpha[\delta](\rho)}{1-\delta}}$ and $\rho \mapsto 2^{-\frac{\alpha-1}{\alpha}\frac{\bar{g}_\alpha[\delta](\rho)}{1-\delta}}$ are concave for all $\rho \in \mathrm{S}_=(AB)$, $\alpha > 1$ and $\delta \in [0, 1)$.*

Proof. We start by rewriting $2^{-(\alpha-1)\frac{\bar{g}_\alpha[\delta](\rho)}{1-\delta}}$ and a simple calculation yields,

$$2^{-(\alpha-1)\frac{\bar{g}_\alpha[\delta](\rho)}{1-\delta}} = \frac{\mathrm{Tr}\left[\left(\mathcal{Z}(\mathcal{G}_\delta(\rho)^{\frac{1}{\alpha}})\right)^\alpha\right]}{1-\delta} \tag{A.1.23}$$

$$= \frac{\widehat{Q}_{\frac{1}{\alpha}} \circ \mathcal{G}_\delta(\rho_{AB})}{1-\delta}. \tag{A.1.24}$$

Thus, using Corollary A.1.5 we can conclude that $\rho \mapsto 2^{-(\alpha-1)\frac{\bar{g}_\alpha[\delta](\rho)}{1-\delta}}$ is concave for all $\rho \in \mathrm{S}_=(AB)$, $\alpha > 1$ and $\delta \in [0, 1)$.

Similarly for $\rho \mapsto 2^{-\frac{\alpha-1}{\alpha}\frac{\bar{g}_\alpha[\delta](\rho)}{1-\delta}}$, it holds

$$2^{-\frac{\alpha-1}{\alpha}\frac{\bar{g}_\alpha[\delta](\rho)}{1-\delta}} = \left(\frac{\widehat{Q}_{\frac{1}{\alpha}} \circ \mathcal{G}_\delta(\rho_{AB})}{1-\delta}\right)^{\frac{1}{\alpha}}. \tag{A.1.25}$$

The function $x^{\frac{1}{\alpha}}$ is non-decreasing and concave for $\alpha \in (1, \infty)$ and due to Corollary A.1.5 $\widehat{Q}_{\frac{1}{\alpha}} \circ \mathcal{G}_\delta(\rho_{AB})$ is concave. Thus, by Lemma 2.4.5 the composition of $x^{\frac{1}{\alpha}}$ with $\widehat{Q}_{\frac{1}{\alpha}} \circ \mathcal{G}_\delta(\rho_{AB})$ is concave and the function $\rho \mapsto 2^{-\frac{\alpha-1}{\alpha}\frac{\bar{g}_\alpha[\delta](\rho)}{1-\delta}}$ is concave as well. □

Now, we have all tools to prove the joint convexity of both the optimization problem of h_α^{QKD}, the $\widetilde{H}_\alpha^{\downarrow,f}$-normalization constant $\kappa^{\downarrow}$ and the $\widetilde{H}_\alpha^{\supset,f}$-normalization constant $\kappa^{\supset}$.

A.1.3 Joint Convexity of h_α^{QKD}

Before we state the theorem, let us introduce some more notation regarding the relative entropy for convenience. For two scalars $x, y \in \mathbb{R}_{\geq} 0$, we denote $x \log(x/y)$ by

$$D(x\|y) \coloneqq x \log(x/y), \tag{A.1.26}$$

such that for two vectors $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n_{\geq 0}$ we can write

$$D(\mathbf{u}\|\mathbf{v}) = \sum_{i=1}^{n} D(u_i\|v_i)\,. \tag{A.1.27}$$

Theorem A.1.8 (Joint Convexity of perturbed h_α^{QKD} (State))**.** *Let the feasible set of h_α^{QKD} of Eq. (10.1.12) be defined as*

$$\begin{aligned}\mathcal{D} \coloneqq \{\mathbf{q} \in S_{\mathrm{acc}}, \rho \in \mathrm{S}_=(AB), \boldsymbol{\nu}_{\widehat{C}} \in \mathbb{P}(\widehat{\mathcal{C}})| \\ \mathrm{Tr}_B[\rho_{AB}] = \tau_A, \boldsymbol{\nu}_{\widehat{C}} = (\gamma \boldsymbol{\nu}_{\widehat{C}|\mathtt{test}}, 1-\gamma)^T, \\ \boldsymbol{\nu}_{\widehat{C}|\mathtt{test}} = \boldsymbol{\Phi}[\rho_{|\mathtt{test}}]\},\end{aligned} \tag{A.1.28}$$

and for all $\alpha > 1$ and $\delta \in [0, 1)$ define the function

$$\bar{h}(\mathbf{q}, \rho, \boldsymbol{\nu}_{\widehat{C}}) \coloneqq \frac{\alpha}{\alpha-1} D\big(\mathbf{q}\big\|\boldsymbol{\nu}_{\widehat{C}}\big) + q(\mathtt{gen}) \frac{\bar{g}_\alpha[\delta](\rho_{AB|\mathtt{gen}})}{1-\delta}. \tag{A.1.29}$$

Then, $\bar{h}$ is jointly convex in $(\mathbf{q}, \rho, \boldsymbol{\nu}_{\widehat{C}})$ over $\mathcal{D}$ for all $\alpha > \frac{1}{1-\delta} > 1$ and $\delta \in [0, 1)$. Furthermore, $\bar{h}$ satisfies

$$h_\alpha^{\mathrm{QKD}} \geq \inf_{(\mathbf{q},\rho,\boldsymbol{\nu}_{\widehat{C}})\in\mathcal{D}} \bar{h}(\mathbf{q}, \rho, \boldsymbol{\nu}_{\widehat{C}}), \tag{A.1.30}$$

for all $\alpha > 1$ and $\delta \in [0, 1)$, with equality for $\delta = 0$.

Proof. The fact that $\bar{h}$ is a lower bound on h_α^{QKD} for all $\alpha > 1$ and $\delta \in [0, 1)$ follows from Lemma A.1.6.

The proof for convexity will proceed by rewriting $\bar{h}$ into a sum of functions that are known to be jointly convex. We start by rephrasing $\rho_{|\mathtt{gen}}$ and following Definition 8.2.1, we find

$$\rho_{AB|\mathtt{gen}} = \frac{\rho_{AB\wedge\mathtt{gen}}}{\mathrm{Tr}[\rho_{AB\wedge\mathtt{gen}}]}, \quad \rho_{AB\wedge\mathtt{gen}} = \sqrt{\Pi^{\mathtt{gen}}}\rho_{AB}\sqrt{\Pi^{\mathtt{gen}}} \tag{A.1.31}$$

Furthermore, $\boldsymbol{\nu}_{\widehat{C}}$ satisfies in the generation component

$$\nu_{\widehat{C}}(\mathtt{gen}) = 1 - \gamma = \mathrm{Tr}[\Pi^{\mathtt{gen}}\rho_{AB}]. \tag{A.1.32}$$

Inserting this into $\bar{h}$ yields

$$\bar{h}(\mathbf{q}, \rho, \boldsymbol{\nu}_{\widehat{C}}) = \frac{\alpha}{\alpha - 1} D\big(\mathbf{q}\big\|\boldsymbol{\nu}_{\widehat{C}}\big) + q(\mathtt{gen})\frac{\bar{g}_\alpha[\delta]\left(\frac{\rho_{AB\wedge\mathtt{gen}}}{1-\gamma}\right)}{1-\delta} \tag{A.1.33}$$

The main step of this proof will be to rearrange the terms of the relative entropy. Thus, we expand the definition of $\bar{g}_\alpha[\delta]$ to find

$$\bar{h}(\mathbf{q}, \rho, \boldsymbol{\nu}_{\widehat{C}}) = \frac{\alpha}{\alpha - 1} D\big(\mathbf{q}\big\|\boldsymbol{\nu}_{\widehat{C}}\big) - \frac{q(\mathtt{gen})}{(\alpha-1)(1-\delta)} \log\left[\widehat{Q}_{\frac{1}{\alpha}} \circ \mathcal{G}_\delta\left(\frac{\rho_{AB\wedge\mathtt{gen}}}{1-\gamma}\right)\right]. \tag{A.1.34}$$

Next, by using our definition of the relative entropy for scalars shown in Eq. (A.1.26), we separate the test and generation components of the relative entropy between $\mathbf{q}$ and $\boldsymbol{\nu}_{\widehat{C}}$,

$$D\big(\mathbf{q}\big\|\boldsymbol{\nu}_{\widehat{C}}\big) = D\big(q(\mathtt{gen})\big\|\nu_{\widehat{C}}(\mathtt{gen})\big) + D\big(\mathbf{q}_{\wedge\mathtt{test}}\big\|\boldsymbol{\nu}_{\widehat{C}\wedge\mathtt{test}}\big). \tag{A.1.35}$$

Here $\mathbf{q}_{\wedge\mathtt{test}}$ denotes the subnormalized distribution containing only the test components of $\mathbf{q}$ (and similarly for $\boldsymbol{\nu}_{\widehat{C}}$), so we have e.g. $\sum_{\hat{c}} q_{\wedge\mathtt{test}}(\hat{c}) + q(\mathtt{gen}) = 1$.

Finally, we note the following identities involving several combinations of α and δ

$$\frac{\alpha}{\alpha-1} = \frac{1}{\alpha-1} + 1, \tag{A.1.36}$$

$$\frac{1}{\alpha-1} = \frac{1}{(\alpha-1)(1-\delta)} - \frac{\delta}{(\alpha-1)(1-\delta)}, \tag{A.1.37}$$

$$\frac{\alpha(1-\delta)-1}{(\alpha-1)(1-\delta)} = 1 - \frac{\delta}{(\alpha-1)(1-\delta)}, \tag{A.1.38}$$

$$-\frac{\alpha(1-\delta)-1}{(\alpha-1)(1-\delta)} = \frac{1}{(\alpha-1)(1-\delta)} - \frac{\alpha}{\alpha-1}. \tag{A.1.39}$$

A simple calculation exploiting the properties of the relative entropy while inserting the identities above together with Eq. (A.1.35), lets us write

$$\begin{aligned}\bar{h}(\mathbf{q},\rho,\boldsymbol{\nu}_{\widehat{C}}) =& \frac{D\Big(q(\mathtt{gen})\Big\|\widehat{Q}_{\frac{1}{\alpha}}\circ\mathcal{G}_\delta\left(\rho_{AB\wedge\mathtt{gen}}\right)\Big)}{(\alpha-1)(1-\delta)} \\ &+ \frac{\alpha}{\alpha-1}D\big(\mathbf{q}_{\wedge\mathtt{test}}\big\|\boldsymbol{\nu}_{\widehat{C}\wedge\mathtt{test}}\big) \\ &+ \frac{\alpha(1-\delta)-1}{(\alpha-1)(1-\delta)}D\big(q(\mathtt{gen})\big\|\nu_{\widehat{C}}(\mathtt{gen})\big)\,.\end{aligned} \tag{A.1.40}$$

This reformulation of $\bar{h}$ allows us to derive the claimed convexity.

First, we note that the relative entropy is jointly convex in its arguments, see e.g. [52, Example 3.19], and that a sum of jointly convex functions remains jointly convex. Thus, we will argue that each term is jointly convex. The second term is obviously jointly convex as it is a relative entropy between $\mathbf{q}$ and $\boldsymbol{\nu}_{\widehat{C}}$ omitting their $\mathtt{gen}$ component.

By the same reasoning the third term, the relative entropy between the two scalars $q(\mathtt{gen}), \nu_{\widehat{C}}(\mathtt{gen})$ (in the sense of Eq. (A.1.26)) is jointly convex as long as

$$0 \leq \alpha(1-\delta)-1 \Leftrightarrow \frac{1}{1-\delta} \leq \alpha. \tag{A.1.41}$$

To see that the first term is jointly convex, consider

$$\widehat{Q}_{\frac{1}{\alpha}}\circ\mathcal{G}_\delta\left(\rho_{AB\wedge\mathtt{gen}}\right) = \widehat{Q}_{\frac{1}{\alpha}}\circ\mathcal{G}_\delta\left(\sqrt{\Pi^{\mathtt{gen}}}\rho_{AB}\sqrt{\Pi^{\mathtt{gen}}}\right). \tag{A.1.42}$$

This is concave due to Corollary A.1.5 and the composition with the affine transformation $\sqrt{\Pi^{\mathtt{gen}}}(\cdot)\sqrt{\Pi^{\mathtt{gen}}}$. Additionally, the relative entropy between two scalars (again in the sense

of Eq. (A.1.26)) is nonincreasing in the second argument. Thus, for any $\mathbf{q} \in S_{\text{acc}}$, $\lambda \in [0,1]$ and $\rho_1, \rho_2 \in \mathrm{S}_=(AB)$, we find

$$D\Big(q(\mathtt{gen})\Big\|\widehat{Q}_{\frac{1}{\alpha}} \circ \mathcal{G}_\delta\left(\lambda\rho_{1\wedge\mathtt{gen}} + (1-\lambda)\rho_{2\wedge\mathtt{gen}}\right)\Big) \tag{A.1.43}$$

$$\leq D\Big(q(\mathtt{gen})\Big\|\lambda\widehat{Q}_{\frac{1}{\alpha}} \circ \mathcal{G}_\delta\left(\rho_{1\wedge\mathtt{gen}}\right) + (1-\lambda)\widehat{Q}_{\frac{1}{\alpha}} \circ \mathcal{G}_\delta\left(\rho_{2\wedge\mathtt{gen}}\right)\Big). \tag{A.1.44}$$

Hence, using the joint convexity of the relative entropy, we find for all $\mathbf{q}_1, \mathbf{q}_2 \in S_{\text{acc}}$, $\lambda \in [0,1]$ and $\rho_1, \rho_2 \in \mathrm{S}_=(AB)$

$$D\Big(\lambda q_1(\mathtt{gen}) + (1-\lambda)q_2(\mathtt{gen})\Big\|\widehat{Q}_{\frac{1}{\alpha}} \circ \mathcal{G}_\delta\left(\lambda\rho_{1\wedge\mathtt{gen}} + (1-\lambda)\rho_{2\wedge\mathtt{gen}}\right)\Big) \tag{A.1.45}$$

$$\leq D\Big(\lambda q_1(\mathtt{gen}) + (1-\lambda)q_2(\mathtt{gen})\Big\|\lambda\widehat{Q}_{\frac{1}{\alpha}} \circ \mathcal{G}_\delta\left(\rho_{1\wedge\mathtt{gen}}\right) \tag{A.1.46}$$

$$+ (1-\lambda)\widehat{Q}_{\frac{1}{\alpha}} \circ \mathcal{G}_\delta\left(\rho_{2\wedge\mathtt{gen}}\right)\Big)$$

$$\leq \lambda D\Big(q_1(\mathtt{gen})\Big\|\widehat{Q}_{\frac{1}{\alpha}} \circ \mathcal{G}_\delta\left(\rho_{1\wedge\mathtt{gen}}\right)\Big) + (1-\lambda)D\Big(q_2(\mathtt{gen})\Big\|\widehat{Q}_{\frac{1}{\alpha}} \circ \mathcal{G}_\delta\left(\rho_{2\wedge\mathtt{gen}}\right)\Big), \tag{A.1.47}$$

and thus the first term in $\bar{h}$ is also jointly convex.

Finally, the set $\mathcal{D}$ is a convex set since it only contains linear trace constraints and we can write

$$\gamma\boldsymbol{\nu}_{\widehat{C}|\mathtt{test}} = \gamma\boldsymbol{\Phi}[\rho_{|\mathtt{test}}] = \boldsymbol{\Phi}[\sqrt{\Pi^{\mathtt{test}}}\rho\sqrt{\Pi^{\mathtt{test}}}], \tag{A.1.48}$$

which concludes the proof. □

From this theorem we can immediately conclude that the optimization problem for h_α^{QKD} in Eq. (10.1.30) in terms of Eve's channel is jointly convex.

Corollary A.1.9 (Joint Convexity of perturbed h_α^{QKD} (Channel))**.** *Let the feasible set of h_α^{QKD} of Eq. (10.1.30) be defined as*

$$\begin{aligned}\mathcal{D}' \coloneqq \{&\mathbf{q} \in S_{\text{acc}}, J \in \mathrm{Pos}(A'B), \boldsymbol{\nu}_{\widehat{C}} \in \mathbb{P}(\widehat{\mathcal{C}})| \\ &\mathrm{Tr}_B[J] = \mathbb{I}_{A'}, \boldsymbol{\nu}_{\widehat{C}} = (\gamma\boldsymbol{\nu}_{\widehat{C}|\mathtt{test}}, 1-\gamma)^T, \\ &\boldsymbol{\nu}_{\widehat{C}|\mathtt{test}} = \boldsymbol{\Psi}[J]\},\end{aligned} \tag{A.1.49}$$

and for all $\alpha > 1$ and $\delta \in [0,1)$ define the function

$$h(\mathbf{q}, J, \boldsymbol{\nu}_{\widehat{C}}) \coloneqq \frac{\alpha D\big(\mathbf{q}\big\|\boldsymbol{\nu}_{\widehat{C}}\big)}{\alpha - 1} + q(\mathtt{gen})\frac{g_\alpha[\delta](J)}{1-\delta}. \tag{A.1.50}$$

Then, h is jointly convex in $(\mathbf{q}, J, \nu_{\widehat{C}})$ over $\mathcal{D}'$ for all $\alpha > \frac{1}{1-\delta} > 1$ and $\delta \in [0,1)$. Furthermore, h satisfies

$$h_\alpha^{\mathrm{QKD}} \geq \inf_{(\mathbf{q},J,\boldsymbol{\nu}_{\widehat{C}})\in\mathcal{D}'} h(\mathbf{q}, J, \boldsymbol{\nu}_{\widehat{C}}), \tag{A.1.51}$$

for all $\alpha > 1$ and $\delta \in [0,1)$, with equality for $\delta = 0$.

Proof. The map χ_g is an affine map and its composition with $\bar{g}_\alpha[\delta]$ remains convex by Lemma 2.4.4. □

A.1.4 Joint Convexity of Normalization Constants $\kappa^\downarrow$ and $\kappa^\supset$

Finally, we turn our attention to the normalization constants $\kappa^\downarrow$ and $\kappa^\supset$ for our f-weighted entropies. With our previous lemmas at hand, we can immediately prove the convexity of the optimization problems to find either the $\widetilde{H}_\alpha^{\downarrow,f}$-normalization or the $\widetilde{H}_\alpha^{\supset,f}$-normalization constant. For brevity, in this section we shall refer to these optimizations as the "κ computations".

We first prove the joint convexity of the optimization in the κ-computations in terms of the state ρ and deduce the joint convexity in terms of Eve's channel from it.

Theorem A.1.10 (Joint Convexity $\kappa^\downarrow$ and $\kappa^\supset$ (State))**.** *The optimization problems to find the $\widetilde{H}_\alpha^{\downarrow,f}$-normalization constant $\kappa^\downarrow$,*

$$\begin{aligned}\kappa^\downarrow = \inf_{\substack{\rho\in\mathrm{S}_=(AB),\\ \boldsymbol{\nu}\in\mathbb{P}(\widehat{\mathcal{C}}_{\setminus\mathrm{gen}})}} \Bigg(\frac{1}{1-\alpha} \log\Big(\gamma \sum_{c\in\widehat{\mathcal{C}}_{\setminus\mathrm{gen}}} \nu(c) 2^{(\alpha-1)f(c)} \\ + (1-\gamma)\, 2^{-(\alpha-1)\left(\frac{\bar{g}_\alpha[\delta](\rho)}{1-\delta} - f(\mathtt{gen})\right)}\Big)\Bigg) \\ \text{s.t.}\quad \mathrm{Tr}_B[\rho] = \tau_A, \\ \boldsymbol{\nu} = \boldsymbol{\Phi}[\rho_{|\mathtt{test}}].\end{aligned} \tag{A.1.52}$$

and the $\widetilde{H}_\alpha^{\wp,f}$-normalization constant $\kappa^{\wp}$,

$$\begin{aligned}\kappa^{\downarrow} = \inf_{\substack{\rho\in \mathrm{S}_=(AB),\\ \boldsymbol{\nu}\in\mathbb{P}(\widehat{\mathcal{C}}_{\setminus\mathtt{gen}})}} \Bigg(\frac{1}{1-\alpha}\log\Big(\gamma \sum_{c\in\widehat{\mathcal{C}}_{\setminus\mathtt{gen}}} \nu(c) 2^{(\alpha-1)f(c)} \\ + (1-\gamma)\, 2^{-(\alpha-1)\left(\frac{\bar{g}_\alpha[\delta](\rho)}{1-\delta} - f(\mathtt{gen})\right)}\Big)\Bigg) \\ s.t.\ \ \mathrm{Tr}_B[\rho] = \tau_A, \\ \boldsymbol{\nu} = \boldsymbol{\Phi}[\rho_{|\mathtt{test}}].\end{aligned} \tag{A.1.53}$$

are jointly convex in $(\rho, \boldsymbol{\nu})$ for all $\alpha > 1$ and $\delta \in [0, 1)$.

Proof. Since $-\log(x)$ is convex, by [52, Section 3.2.4] the $\widetilde{H}_\alpha^{\downarrow,f}$-normalization constant $\kappa^{\downarrow}$ will be convex in $(\rho, \boldsymbol{\nu})$ if the sum within the logarithm is concave. We will show that this is indeed the case for all $\alpha > 1$ and $\delta \in [0, 1)$, yielding the claimed result (the set over which the optimization is performed is clearly convex).

First, note that the function

$$\boldsymbol{\nu} \mapsto \gamma \sum_{c\in\widehat{\mathcal{C}}_{\setminus\mathtt{gen}}} \boldsymbol{\nu}(c) 2^{(\alpha-1)f(c)} \tag{A.1.54}$$

is trivially concave because it is a linear transformation of $\boldsymbol{\nu}$. Next, the concavity with respect to ρ follows immediately from Corollary A.1.7.

Thus, it remains to show that the sum of both functions is jointly concave in $(\rho, \boldsymbol{\nu})$. However, this also follows trivially, since any function f defined from concave functions f_1 and f_2 by

$$f(x, y) = f_1(x) + f_2(y), \tag{A.1.55}$$

is concave. Therefore, the sum inside the logarithm is jointly concave, as claimed.

Finally, the statement regarding the the $\widetilde{H}_\alpha^{\wp,f}$-normalization constant $\kappa^{\wp}$ follows by the same steps. □

Similarly, we prove the joint convexity of the optimizations in the $\kappa^{\downarrow}$- and $\kappa^{\wp}$-computation characterized in terms of Eve's channel.

Theorem A.1.11 (Joint Convexity $\kappa^{\downarrow}$ and $\kappa^{\wr}$ (Channel))**.** *The optimization problem to find the $\widetilde{H}_\alpha^{\downarrow,f}$-normalization constant $\kappa^{\downarrow}$,*

$$\begin{aligned}\kappa^{\downarrow} = \inf_{\substack{J\in\mathrm{Pos}(A'B),\\ \boldsymbol{\nu}\in\mathbb{P}(\widehat{\mathcal{C}}_{\setminus\mathrm{gen}})}} \Bigg(\frac{1}{1-\alpha}\log\Big(\gamma \sum_{c\in\widehat{\mathcal{C}}_{\setminus\mathrm{gen}}} \nu(c) 2^{(\alpha-1)f(c)} \\ + (1-\gamma)\, 2^{-(\alpha-1)\left(\frac{g_\alpha[\delta](J)}{1-\delta} - f(\mathtt{gen})\right)}\Big)\Bigg) \\ s.t.\ \ \mathrm{Tr}_B[J] = \mathbb{I}_{A'}, \\ \boldsymbol{\nu} = \boldsymbol{\Psi}[J],\end{aligned} \tag{A.1.56}$$

and the $\widetilde{H}_\alpha^{\wr,f}$-normalization constant $\kappa^{\wr}$

$$\begin{aligned}\kappa^{\wr} = \inf_{\substack{J\in\mathrm{Pos}(A'B),\\ \boldsymbol{\nu}\in\mathbb{P}(\widehat{\mathcal{C}}_{\setminus\mathrm{gen}})}} \Bigg(\frac{\alpha}{1-\alpha}\log\Big(\gamma \sum_{\hat{c}\in\widehat{\mathcal{C}}_{\setminus\mathrm{gen}}} \boldsymbol{\nu}(\hat{c}) 2^{\frac{\alpha-1}{\alpha}f(\hat{c})} \\ + (1-\gamma)\, 2^{-\frac{\alpha-1}{\alpha}(g_\alpha(J) - f(\mathtt{gen}))}\Big)\Bigg) \\ s.t.\ \ \mathrm{Tr}_B[J] = \mathbb{I}_{A'}, \\ \boldsymbol{\nu} = \boldsymbol{\Psi}[J],\end{aligned} \tag{A.1.57}$$

are jointly convex in $(J, \boldsymbol{\nu})$ for all $\alpha > 1$ and $\delta \in [0, 1)$ and equal to Eq. (10.2.34) and Eq. (10.2.33) for $\delta = 0$, respectively.

Proof. The proof has the same structure as Theorem A.1.10 with one small change.

The concavity of

$$2^{-(\alpha-1)\left(\frac{g_\alpha[\delta](J)}{1-\delta} - f(\mathtt{gen})\right)} \tag{A.1.58}$$

and

$$2^{-\frac{\alpha-1}{\alpha}\left(\frac{g_\alpha[\delta](J)}{1-\delta} - f(\mathtt{gen})\right)} \tag{A.1.59}$$

follows from Corollary A.1.7 and Lemma 2.4.4, because $g_\alpha[\delta]$ is a composition of $\bar{g}_\alpha[\delta]$ with the affine map χ_g. The remainder of the proof remains the same.

Equality to Eq. (10.2.34) and Eq. (10.2.33), respectively, for $\delta = 0$ is also straightforward since $g_\alpha[0] = g_\alpha$. □

A.2 Proofs for Phase Imperfections

To include phase imperfections in our framework, we require a statement allowing us to relate the approximately diagonalized states to exactly diagonalized states, as laid out in Section 10.10.2. The following lemma allows for exactly this application.

Lemma A.2.1 (Trace distance bounds under source replacement scheme)**.** *Let* $\rho_{AA'A_s}$ *and* $\sigma_{AA'A_s} \in \mathrm{S}_=(AA'A_s)$ *be pure states of the form*

$$\rho_{AA'A_s} = \sum_{i,j} \sqrt{p_i p_j}\, |i\rangle\langle j|_A \, |\psi_i\rangle\langle\psi_j|_{A'A_s}\,, \tag{A.2.1}$$

$$\sigma_{AA'A_s} = \sum_{i,j} \sqrt{p_i p_j}\, |i\rangle\langle j|_A \, |\phi_i\rangle\langle\phi_j|_{A'A_s}\,, \tag{A.2.2}$$

defined by the source replacement scheme. Furthermore, let

$$\mathrm{Tr}_{A_s}[|\psi_i\rangle\langle\psi_i|] \coloneqq \rho^i_{A'} \quad \textit{and} \quad \mathrm{Tr}_{A_s}[|\phi_i\rangle\langle\phi_i|] \coloneqq \sigma^i_{A'} \tag{A.2.3}$$

satisfy

$$\left\|\rho^i_{A'} - \sigma^i_{A'}\right\|_1 \leq \epsilon_i, \tag{A.2.4}$$

for all i*. Then, it holds*

$$\left\|\mathrm{Tr}_{A_s} \circ \mathcal{G} \circ (\mathbb{I}_A \otimes\ \mathcal{E})\left(\rho_{AA'A_s} - \sigma_{AA'A_s}\right)\right\|_1 \leq \sum_i p_i \epsilon_i, \tag{A.2.5}$$

where $\mathcal{E} : A' \to B$ *is Eve's channel.*

Proof. First, we make use of the special properties of Alice's measurements under the source replacement scheme. Therefore, let Ξ_{meas} be the channel measuring Alice's system in the computational basis defined by

$$\Xi_{\mathrm{meas}}(\rho_A) = \sum_i |i\rangle\langle i|\, \rho_A\, |i\rangle\langle i|\,, \tag{A.2.6}$$

for all ρ_A. Under the source replacement scheme [56, 57], we can always write Ξ_{meas} this way and it satisfies

$$\Xi_{\mathrm{meas}} \circ \Xi_{\mathrm{meas}}(\rho_A) = \Xi_{\mathrm{meas}}(\rho_A), \tag{A.2.7}$$

for all $\rho_A \in \mathrm{S}_=(A)$. Therefore, based on the construction of the map $\mathcal{G}$ in [67], it follows that $\mathcal{G}$ satisfies

$$\mathcal{G} = \mathcal{G} \circ \Xi_{\mathrm{meas}}. \tag{A.2.8}$$

Next, we will exploit this property of the map $\mathcal{G}$ to bound the trace distance. We find

$$\|\mathrm{Tr}_{A_s} \circ\ \mathcal{G} \circ\ (\mathbb{I}_A \otimes \mathcal{E})\,(\rho_{AA'A_s} - \sigma_{AA'A_s})\|_1 \tag{A.2.9}$$
$$= \|\,\mathcal{G} \circ\ (\mathbb{I}_A \otimes \mathcal{E}) \circ \mathrm{Tr}_{A_s} \circ (\Xi_{\mathrm{meas}} \otimes \mathbb{I}_{A'})\,(\rho_{AA'A_s} - \sigma_{AA'A_s})\|_1 \tag{A.2.10}$$

Applying the data-processing inequality for the trace distance, see e.g. [48, Theorem 9.2], we find

$$\|\mathrm{Tr}_{A_s} \circ\ \mathcal{G} \circ\ (\mathbb{I}_A \otimes \mathcal{E})\,(\rho_{AA'A_s} - \sigma_{AA'A_s})\|_1 \tag{A.2.11}$$
$$\leq \|\mathrm{Tr}_{A_s} \circ (\Xi_{\mathrm{meas}} \otimes \mathbb{I}_{A'})\,(\rho_{AA'A_s} - \sigma_{AA'A_s})\|_1. \tag{A.2.12}$$

We can simplify $(\Xi_{\mathrm{meas}} \otimes \mathbb{I}_{A'})\,(\rho_{AA'A_s} - \sigma_{AA'A_s})$ further. For this, observe that

$$\Xi_{\mathrm{meas}} \otimes \mathbb{I}_{A'}\,(\rho_{AA'A_s}) = \sum_i p_i\,|i\rangle\langle i| \otimes |\psi_i\rangle\langle\psi_i|\,, \tag{A.2.13}$$

and similarly for $\sigma_{AA'A_s}$. A simple calculation applying strong convexity [48, Theorem 9.3] and sub-additivity with respect to tensor products of the trace distance, yields

$$\|\mathrm{Tr}_{A_s} \circ (\Xi_{\mathrm{meas}} \otimes \mathbb{I}_{A'})\,(\rho_{AA'A_s} - \sigma_{AA'A_s})\|_1 \tag{A.2.14}$$
$$\leq \sum_i p_i \|\mathrm{Tr}_{A_s}[|\psi_i\rangle\langle\psi_i| - |\phi_i\rangle\langle\phi_i|]\|_1 \tag{A.2.15}$$
$$= \sum_i p_i \|\rho^i_{A'} - \sigma^i_{A'}\|_1. \tag{A.2.16}$$

After inserting the assumption $\|\rho^i_{A'} - \sigma^i_{A'}\|_1 \leq \epsilon_i$ the claim follows. □

In order to incorporate the cutoff N_A into the map $\mathbf{\Phi}$ or $\mathbf{\Psi}$, we require the following lemma.

Lemma 10.10.1 (Bounds on Statistics under Approximate Diagonalization)**.** *Let* $|\xi^t\rangle$ *and* $\left|\widetilde{\xi}^t\right\rangle$ *be defined as in Eq.* (10.10.8) *and Eq.* (10.10.12)*, respectively. Furthermore, let the states and probabilities satisfy*

$$\left\| |s_m(a,\mu)\rangle\langle s_m(a,\mu)| - |w_m(a,\mu)\rangle\langle w_m(a,\mu)| \right\|_1 \leq \epsilon_{a,\mu}, \tag{10.10.25}$$
$$|\zeta_m(a,\mu) - \omega_m(a,\mu)| \leq \epsilon_{\mathrm{prob}}. \tag{10.10.26}$$

Then, it holds

$$\sum_m \zeta_{m|\mathtt{test}} \Phi_m \circ \mathcal{E}_m \left[\left|\widetilde{\xi}^t_{|m}\right\rangle\!\left\langle\widetilde{\xi}^t_{|m}\right|\right] \geq \sum_m \omega^L_{m|\mathtt{test}} \left(\Phi_m \circ \mathcal{E}_m \left[|\xi^t_{|m}\rangle\langle\xi^t_{|m}|\right] - \epsilon_m\right), \tag{10.10.27}$$

and

$$\begin{aligned}&\sum_m \zeta_{m|\mathtt{test}} \Phi_m \circ \mathcal{E}_m \left[\left| \widetilde{\xi}^t_{|m} \right\rangle\!\left\langle \widetilde{\xi}^t_{|m} \right| \right] + 1 - \mathbf{p}_{\mathrm{tot}}(\widetilde{\xi}) \\ &\leq \sum_m \omega^U_{m|\mathtt{test}} \left(\Phi_m \circ \mathcal{E}_m \left[\left| \xi^t_{|m} \right\rangle\!\left\langle \xi^t_{|m} \right| \right] + \epsilon_m \right) + 1 - \mathbf{p}^L_{\mathrm{tot}},\end{aligned} \tag{10.10.28}$$

where

$$\omega^U_{m|\mathtt{test}} \coloneqq \min\{\omega_{m|\mathtt{test}} + \epsilon_{\mathrm{prob}}, 1\}, \tag{10.10.29}$$

$$\omega^L_{m|\mathtt{test}} \coloneqq \max\{\omega_{m|\mathtt{test}} - \epsilon_{\mathrm{prob}}, 0\}, \tag{10.10.30}$$

$$\mathbf{p}^L_{\mathrm{tot}} \coloneqq \sum_{\substack{\hat{c} \in \widehat{\mathcal{C}}_{\setminus \mathrm{gen}} \\ a \in \phi^{-1}(\hat{c})}} \sum_{m \leq N_{\mathrm{cons}}} \omega^L_{m|\mathtt{test}} p(a|m, \mathtt{test}) \hat{e}_{\hat{c}}, \tag{10.10.31}$$

and

$$\epsilon_m \coloneqq \sum_{a,\mu} p(a, \mu|\mathtt{test}, m) \epsilon_{a,\mu}, \tag{10.10.32}$$

$$\epsilon_{\mathrm{prob}} \coloneqq \sum_{a,\mu} p(a, \mu|\mathtt{test}) \epsilon^{a,\mu}_{\mathrm{proj}}. \tag{10.10.33}$$

Proof. First, note that under the source replacement scheme we can define $M^{A|\mathtt{test},m}_{\alpha,x}$ such that $(M^{A|\mathtt{test},m}_{\alpha,x})^2 = M^{A|\mathtt{test},m}_{\alpha,x}$. Then, it holds for every component $\hat{c} = \phi(\alpha, x, \beta, y)$ of $\mathbf{\Phi}$

$$\left| \left(\Phi_m \circ (\mathbb{I}_A \otimes \mathcal{E}_m) \left[\left| \xi^t_{|m} \right\rangle\!\left\langle \xi^t_{|m} \right| - \left| \widetilde{\xi}^t_{|m} \right\rangle\!\left\langle \widetilde{\xi}^t_{|m} \right| \right] \right)_{\hat{c}} \right| \tag{A.2.17}$$

$$= \left| \mathrm{Tr} \left[M^B_{\beta,y} (\mathbb{I}_A \otimes \mathcal{E}_m) [Q] \right] \right|, \tag{A.2.18}$$

where we defined

$$Q \coloneqq M^{A|\mathtt{test},m}_{\alpha,x} \left(\left| \xi^t_{|m} \right\rangle\!\left\langle \xi^t_{|m} \right| - \left| \widetilde{\xi}^t_{|m} \right\rangle\!\left\langle \widetilde{\xi}^t_{|m} \right| \right) M^{A|\mathtt{test},m}_{\alpha,x}. \tag{A.2.19}$$

Next, it follows by Hölder's inequality

$$\left| \left(\Phi_m \circ (\mathbb{I}_A \otimes \mathcal{E}_m) \left[\left| \xi^t_{|m} \right\rangle\!\left\langle \xi^t_{|m} \right| - \left| \widetilde{\xi}^t_{|m} \right\rangle\!\left\langle \widetilde{\xi}^t_{|m} \right| \right] \right)_{\hat{c}} \right| \tag{A.2.20}$$

$$\leq \mathrm{Tr} \left[\left| M^B_{\beta,y} (\mathbb{I}_A \otimes \mathcal{E}_m) [Q] \right| \right] \tag{A.2.21}$$

$$\leq \left\| \mathbb{I}_A \otimes M^B_{\beta,y} \right\|_\infty \cdot \left\| (\mathbb{I}_A \otimes \mathcal{E}_m) [Q] \right\|_1. \tag{A.2.22}$$

Since $M^B_{\beta,y}$ are POVM elements, they satisfy $\left\|M^B_{\beta,y}\right\|_\infty \leq 1$, and we can bound the first term by one. Additionally, we apply the data-processing inequality for the trace distance on the second term. Combining both, we find in summary

$$\left|\left(\Phi_m \circ (\mathbb{I}_A \otimes \mathcal{E}_m)\left[\left|\xi^t_{|m}\right\rangle\!\left\langle\xi^t_{|m}\right| - \left|\widetilde{\xi}^t_{|m}\right\rangle\!\left\langle\widetilde{\xi}^t_{|m}\right|\right]\right)_{\hat{c}}\right| \tag{A.2.23}$$

$$\leq \left\|M^{A|\mathtt{test},m}_{\alpha,x}\left(\left|\xi^t_{|m}\right\rangle\!\left\langle\xi^t_{|m}\right| - \left|\widetilde{\xi}^t_{|m}\right\rangle\!\left\langle\widetilde{\xi}^t_{|m}\right|\right)M^{A|\mathtt{test},m}_{\alpha,x}\right\|_1. \tag{A.2.24}$$

Hence, we continue by bounding the right-hand-side of this equation. Here, a simple calculation applying strong convexity [48, Theorem 9.3] and sub-additivity with respect to tensor products of the trace distance, yields

$$\left\|M^{A|\mathtt{test},m}_{\alpha,x}\left(\left|\xi^t_{|m}\right\rangle\!\left\langle\xi^t_{|m}\right| - \left|\widetilde{\xi}^t_{|m}\right\rangle\!\left\langle\widetilde{\xi}^t_{|m}\right|\right)M^{A|\mathtt{test},m}_{\alpha,x}\right\|_1 \tag{A.2.25}$$

$$\leq \sum_{a,\mu} p(a,\mu|\mathtt{test},m)\left\| \left|s_m(a,\mu)\right\rangle\!\left\langle s_m(a,\mu)\right| - \left|w_m(a,\mu)\right\rangle\!\left\langle w_m(a,\mu)\right| \right\|_1. \tag{A.2.26}$$

Now, by assumption, it holds

$$\left\| \left|s_m(a,\mu)\right\rangle\!\left\langle s_m(a,\mu)\right| - \left|w_m(a,\mu)\right\rangle\!\left\langle w_m(a,\mu)\right| \right\|_1 \leq \epsilon_{a,\mu}. \tag{A.2.27}$$

Therefore, we find

$$\left|\left(\Phi_m\left[\left|\xi^t_{|m}\right\rangle\!\left\langle\xi^t_{|m}\right|\right]\right)_{\hat{c}} - \left(\Phi_m\left[\left|\widetilde{\xi}^t_{|m}\right\rangle\!\left\langle\widetilde{\xi}^t_{|m}\right|\right]\right)_{\hat{c}}\right| \leq \epsilon_m, \tag{A.2.28}$$

for all components $\hat{c} = \phi(\alpha, x, \beta, y)$ of $\mathbf{\Phi}$.

Turning our attention to the probabilities $\zeta_{m|\mathtt{test}}$ and $\omega_{m|\mathtt{test}}$. We can bound the difference between $\zeta_{m|\mathtt{test}}$ and $\omega_{m|\mathtt{test}}$ by

$$\left|\zeta_{m|\mathtt{test}} - \omega_{m|\mathtt{test}}\right| \leq \sum_{a,\mu} p(a,\mu|\mathtt{test})\left|\zeta_m(a,\mu) - \omega_(a,\mu)\right| \tag{A.2.29}$$

$$\leq \sum_{a,\mu} p(a,\mu|\mathtt{test})\epsilon^{a,\mu}_{\mathrm{proj}} = \epsilon_{\mathrm{prob}}, \tag{A.2.30}$$

where the last inequality follows by assumption. Furthermore, since both $\zeta_{m|\mathtt{test}}$ and $\omega_{m|\mathtt{test}}$ are probabilities, we can rewrite their difference as upper and lower bounds on $\zeta_{m|\mathtt{test}}$ such as

$$\zeta_{m|\mathtt{test}} \geq \max\{\omega_{m|\mathtt{test}} - \epsilon_{\mathrm{prob}}, 0\} = \omega^L_{m|\mathtt{test}}, \tag{A.2.31}$$

$$\zeta_{m|\mathtt{test}} \leq \min\{\omega_{m|\mathtt{test}} + \epsilon_{\mathrm{prob}}, 1\} = \omega^U_{m|\mathtt{test}}. \tag{A.2.32}$$

Next, by definition, we see that $1 - \mathbf{p}_{\text{tot}}(\widetilde{\xi})$ is bounded by

$$1 - \mathbf{p}_{\text{tot}}(\widetilde{\xi}) = 1 - \sum_{\substack{\hat{c}\in\widehat{\mathcal{C}}_{\setminus\text{gen}} \\ a\in\phi^{-1}(\hat{c})}} \sum_{m\leq N_{\text{cons}}} \zeta_{m|\texttt{test}} p(a|m,\texttt{test})\hat{e}_{\hat{c}} \tag{A.2.33}$$

$$\leq 1 - \sum_{\substack{\hat{c}\in\widehat{\mathcal{C}}_{\setminus\text{gen}} \\ a\in\phi^{-1}(\hat{c})}} \sum_{m\leq N_{\text{cons}}} \omega^{L}_{m|\texttt{test}} p(a|m,\texttt{test})\hat{e}_{\hat{c}} \tag{A.2.34}$$

$$= 1 - \mathbf{p}^{L}_{\text{tot}}. \tag{A.2.35}$$

Finally, combining these results for the probabilities with Eq. (A.2.28) yields the statement of the lemma. □

Theorem 10.10.2 (Fixed- and Variable-Length key rates with Phase Imperfections)**.** *Let Alice's source prepare states with independent phase imperfections, and let*

$$\Xi_{\text{phase}} = \bigotimes_{i=1}^{n} \Xi^{i}_{\text{phase}}, \tag{10.10.43}$$

be the source map from [68, App A.]. Furthermore, define the model state

$$\rho^{(a,\mu)}_{A'} = q \sum_{m=0}^{\infty} p(m|\mu)\, |m\rangle\langle m|_{a} + (1-q)\, |\sqrt{\mu}\rangle\langle\sqrt{\mu}|\,, \tag{10.10.44}$$

with $q \in [0,1]$ and assume that $N_{\text{ent}}, N_{\text{cons}} \leq N_A$.

Then, all $\widetilde{H}^{\uparrow,f}_{\alpha}$-normalization constants $\kappa^{\uparrow}_{i}$ can be bounded by

$$\begin{aligned}
\kappa^{\uparrow}_{i} \geq \inf_{\substack{J_m\in\text{Pos}(A'\bar{A}B), \\ \boldsymbol{\nu}\in\mathbb{P}(\widehat{\mathcal{C}}_{\setminus\text{gen}})}} \Bigg(& \frac{1}{1-\alpha} \log\Big(\gamma \sum_{c\in\widehat{\mathcal{C}}_{\setminus\text{gen}}} \boldsymbol{\nu}(\hat{c}) 2^{(\alpha-1)f(\hat{c})} \\
& + (1-\gamma)\, 2^{(\alpha-1)(\Delta(\epsilon_{\text{full}})+f(\texttt{gen}))} 2^{-(\alpha-1)g^{N_{\text{ent}}}_{\alpha}(J_0,\ldots J_{N_{\text{ent}}})}\Big)\Bigg) \\
\text{s.t.}\ \ & \text{Tr}_B[J_m] = \mathbb{I}_{A'\bar{A}}, \\
& \boldsymbol{\nu} \leq \sum_{m\leq N_{\text{cons}}} \omega^{U}_{m|\texttt{test}} \left(\boldsymbol{\Psi}_m[J_m] + \epsilon_m\right) + 1 - \mathbf{p}^{L}_{\text{tot}}, \\
& \boldsymbol{\nu} \geq \sum_{m\leq N_{\text{cons}}} \omega^{L}_{m|\texttt{test}} \left(\boldsymbol{\Psi}_m[J_m] - \epsilon_m\right).
\end{aligned} \tag{10.10.45}$$

Furthermore, if the phase imperfections are IID, i.e. all source maps Ξ^i_{phase} are equal, then $h^{\uparrow,\text{QKD}}_\alpha$ is bounded from below by

$$h^{\uparrow,\text{QKD}}_\alpha \geq \inf_{\substack{\mathbf{q}\in S_{\text{acc}},\boldsymbol{\nu}_{\widehat{C}}\in\mathbb{P}(\widehat{\mathcal{C}}),\\ J_m\in\text{Pos}(A'\bar{A}B)}} \left(\frac{\alpha D(\mathbf{q}\|\boldsymbol{\nu}_{\widehat{C}})}{\alpha-1} - \Delta(\epsilon_{\text{full}}) + q(\texttt{gen}) g^{N_{\text{ent}}}_\alpha \left(J_0,\ldots,J_{N_{\text{ent}}}\right) \right) \tag{10.10.46}$$

$$\begin{aligned} s.t.\ \ & \text{Tr}_B[J_m] = \mathbb{I}_{A'\bar{A}}, \\ & \boldsymbol{\nu}_{\widehat{C}} = (\gamma \boldsymbol{\nu}_{\widehat{C}|\texttt{test}}, 1-\gamma)^T, \\ & \boldsymbol{\nu}_{\widehat{C}|\texttt{test}} \leq \sum_{m\leq N_{\text{cons}}} \omega^U_{m|\texttt{test}} \left(\boldsymbol{\Psi}_m[J_m] + \epsilon_m\right) + 1 - \mathbf{p}^L_{\text{tot}}, \\ & \boldsymbol{\nu}_{\widehat{C}|\texttt{test}} \geq \sum_{m\leq N_{\text{cons}}} \omega^L_{m|\texttt{test}} \left(\boldsymbol{\Psi}_m[J_m] - \epsilon_m\right). \end{aligned}$$

For both quantities, we defined

$$\begin{aligned} \omega^U_{m|\texttt{test}} &:= \min\{\omega_{m|\texttt{test}} + \epsilon_{\text{prob}}, 1\}, \\ \omega^L_{m|\texttt{test}} &:= \max\{\omega_{m|\texttt{test}} - \epsilon_{\text{prob}}, 0\}, \\ \mathbf{p}^L_{\text{tot}} &:= \sum_{\substack{\hat{c}\in\widehat{\mathcal{C}}_{\setminus\text{gen}}\\ a\in\phi^{-1}(\hat{c})}} \sum_{m\leq N_{\text{cons}}} \omega^L_{m|\texttt{test}} p(a|m,\texttt{test}) \hat{e}_{\hat{c}}, \end{aligned} \tag{10.10.47}$$

and both $g^{N_{\text{ent}}}_\alpha(J_0,\ldots J_{N_{\text{ent}}})$ and $\boldsymbol{\Psi}_m$ are defined in terms of the states $\left|\xi^g_{|m}\right\rangle$ and $\left|\xi^t_{|m}\right\rangle$, respectively.

Proof. For simplicity we start the proof with $h^{\uparrow,\text{QKD}}_\alpha$, where we assume the special case of IID imperfections, i.e. $\Xi^i_{\text{phase}} \equiv \Xi_{\text{phase}}$ for all i.

Furthermore, we note that since we only had continuity bounds available for $\widetilde{H}^\uparrow_\alpha$, we are required to start from $h^{\uparrow,\text{QKD}}_\alpha$, which is defined in Theorem 8.2.4.

First, we observe that the untagged states are recovered from the block-tagged states $\left|\widetilde{\xi}\right\rangle$ via the partial trace $\text{Tr}_{\bar{A}}$. Additionally, we must incorporate the source map Ξ_{phase} from [68, App. A], which connects the model state to the phase-imperfect laser pulses. Consequently, we can apply Theorem 10.4.3 to bound the entropy by absorbing the combined map $\Xi_{\text{phase}}\circ$

$\mathrm{Tr}_{\bar{A}}$ into Eve's attack. This yields

$$h_\alpha^{\uparrow,\mathrm{QKD}} \geq \inf_{\substack{\mathbf{q}\in S_{\mathrm{acc}},\\ J\in\mathrm{Pos}(A'\bar{A}B)}} \left(\frac{\alpha D\big(\mathbf{q}\big\|\boldsymbol{\nu}_{\widehat{C}}\big)}{\alpha-1} + q(\mathtt{gen})\widetilde{H}_\alpha^{\uparrow}(S|\tilde{T}E)_{\tilde{\nu}_{|\mathtt{gen}}(J)} \right) \tag{A.2.36}$$
$$\begin{aligned} \text{s.t.}\quad & \mathrm{Tr}_B[J] = \mathbb{I}_{A'\bar{A}}, \\ & \boldsymbol{\nu}_{\widehat{C}} = (\gamma\boldsymbol{\nu}_{\widehat{C}|\mathtt{test}}, 1-\gamma)^T, \\ & \boldsymbol{\nu}_{\widehat{C}|\mathtt{test}} = \boldsymbol{\Psi}_{\tilde{\xi}}[J], \end{aligned}$$

where we made the dependence of $\boldsymbol{\Psi}$ on the state $\left|\widetilde{\xi}^t\right\rangle$ explicit. Next, we can insert the continuity bound of Eq. (10.10.21) to bound the conditional Rényi entropy and find

$$h_\alpha^{\uparrow,\mathrm{QKD}} \geq \inf_{\substack{\mathbf{q}\in S_{\mathrm{acc}},\\ J\in\mathrm{Pos}(A'\bar{A}B)}} \left(\frac{\alpha D\big(\mathbf{q}\big\|\boldsymbol{\nu}_{\widehat{C}}\big)}{\alpha-1} - \Delta(\epsilon_{\mathrm{full}}) + q(\mathtt{gen})\widetilde{H}_\alpha^{\downarrow}(S|\tilde{T}E)_{\nu_{|\mathtt{gen}}(J)} \right) \tag{A.2.37}$$
$$\begin{aligned} \text{s.t.}\quad & \mathrm{Tr}_B[J] = \mathbb{I}_{A'\bar{A}}, \\ & \boldsymbol{\nu}_{\widehat{C}} = (\gamma\boldsymbol{\nu}_{\widehat{C}|\mathtt{test}}, 1-\gamma)^T, \\ & \boldsymbol{\nu}_{\widehat{C}|\mathtt{test}} = \boldsymbol{\Psi}_{\tilde{\xi}}[J]. \end{aligned}$$

As noted in Section 10.10.1, the block-tagged states satisfy assumptions 1–3. Hence, we can include the simplifications of Section 10.6 to reach

$$h_\alpha^{\uparrow,\mathrm{QKD}} \geq \inf_{\substack{\mathbf{q}\in S_{\mathrm{acc}},\boldsymbol{\nu}_{\widehat{C}}\in\mathbb{P}(\widehat{\mathcal{C}})\\ J_m\in\mathrm{Pos}(A'\bar{A}B)}} \left(\frac{\alpha D\big(\mathbf{q}\big\|\boldsymbol{\nu}_{\widehat{C}}\big)}{\alpha-1} - \Delta(\epsilon_{\mathrm{full}}) + q(\mathtt{gen})g_\alpha^{N_{\mathrm{ent}}}\left(J_0,\ldots,J_{N_{\mathrm{ent}}}\right) \right) \tag{A.2.38}$$
$$\begin{aligned} \text{s.t.}\quad & \mathrm{Tr}_B[J_m] = \mathbb{I}_{A'\bar{A}}, \\ & \boldsymbol{\nu}_{\widehat{C}} = (\gamma\boldsymbol{\nu}_{\widehat{C}|\mathtt{test}}, 1-\gamma)^T, \\ & \boldsymbol{\nu}_{\widehat{C}|\mathtt{test}} \leq \sum_{m\leq N_{\mathrm{cons}}} \zeta_{m|\mathtt{test}}\boldsymbol{\Psi}_{\tilde{\xi}_m}[J_m] + 1 - \mathbf{p}_{\mathrm{tot}}(\widetilde{\xi}), \\ & \boldsymbol{\nu}_{\widehat{C}|\mathtt{test}} \geq \sum_{m\leq N_{\mathrm{cons}}} \zeta_{m|\mathtt{test}}\boldsymbol{\Psi}_{\tilde{\xi}_m}[J_m], \end{aligned}$$

where we now explicitly highlighted that each $\boldsymbol{\Psi}_{\tilde{\xi}_m}$ depends on the state $\left|\widetilde{\xi}_{|m}^t\right\rangle$. However, note that $g_\alpha^{N_{\mathrm{ent}}}$ is already defined in terms of the states $\left|\xi_{|m}^g\right\rangle$ with a cutoff.

Lastly, the substitution of the unbounded states $\left|\widetilde{\xi}^t_{|m}\right\rangle$ with the truncated states $\left|\xi^t_{|m}\right\rangle$ within the constraints is justified by Lemma 10.10.1. This completes the derivation for $h_\alpha^{\uparrow,\mathrm{QKD}}$.

Next, we show the claim for the $\widetilde{H}_\alpha^{\uparrow,f}$-normalization constants. Since here we assume independent imperfections the source maps Ξ^i_{phase} are different for each j. Again since we only had continuity bounds available for $\widetilde{H}_\alpha^\uparrow$, we start from $\kappa_i^\uparrow$ which includes the source map Ξ^i_{phase} for each round i.

As before, we first relate the block-tagged states $\left|\widetilde{\xi}\right\rangle$ to the states without tagging via the source map $\mathrm{Tr}_{\bar{A}}$. Furthermore, the model state is related to the phase imperfect state via the source map Ξ^i_{phase}. Applying Theorem 10.4.1 we bound the f-weighted entropy by including the combined source map $\Xi^i_{\mathrm{phase}} \circ \mathrm{Tr}_{\bar{A}}$ in Eve's attack, which lets us write

$$\begin{aligned}\kappa_i^\uparrow \geq \inf_{\substack{J\in\mathrm{Pos}(A'\bar{A}B),\\ \boldsymbol{\nu}\in\mathbb{P}(\widehat{\mathcal{C}}_{\setminus\mathrm{gen}})}} \Bigg(\frac{\alpha}{1-\alpha}\log\Big(\gamma \sum_{c\in\widehat{\mathcal{C}}_{\setminus\mathrm{gen}}} \boldsymbol{\nu}(\hat{c}) 2^{\frac{\alpha-1}{\alpha}f(\hat{c})} \\ + (1-\gamma)\, 2^{-\frac{\alpha-1}{\alpha}\left(\widetilde{H}_\alpha^\uparrow(S|\tilde{T}E)_{\widetilde{\nu}_{|\mathtt{gen}}} - f(\mathtt{gen})\right)}\Big)\Bigg) \\ \text{s.t. } \mathrm{Tr}_B[J] = \mathbb{I}_{A'\bar{A}}, \\ \boldsymbol{\nu} = \boldsymbol{\Psi}_{\tilde{\xi}}[J],\end{aligned} \tag{A.2.39}$$

Next, we make use of the monotonicity of the "log-mean-exponential", see [4, Lemma 21], to obtain a lower bound by replacing the factors of $\frac{\alpha}{1-\alpha}$ with $\frac{1}{1-\alpha}$ (noting that since $\alpha > 1$, we have $\frac{1-\alpha}{\alpha} \geq 1-\alpha$). This leaves us with

$$\begin{aligned}\kappa_i^\uparrow \geq \inf_{\substack{J\in\mathrm{Pos}(A'\bar{A}B),\\ \boldsymbol{\nu}\in\mathbb{P}(\widehat{\mathcal{C}}_{\setminus\mathrm{gen}})}} \Bigg(\frac{1}{1-\alpha}\log\Big(\gamma \sum_{c\in\widehat{\mathcal{C}}_{\setminus\mathrm{gen}}} \boldsymbol{\nu}(\hat{c}) 2^{(\alpha-1)f(\hat{c})} \\ + (1-\gamma)\, 2^{-(\alpha-1)\left(\widetilde{H}_\alpha^\uparrow(S|\tilde{T}E)_{\widetilde{\nu}_{|\mathtt{gen}}} - f(\mathtt{gen})\right)}\Big)\Bigg) \\ \text{s.t. } \mathrm{Tr}_B[J] = \mathbb{I}_{A'\bar{A}}, \\ \boldsymbol{\nu} = \boldsymbol{\Psi}_{\tilde{\xi}}[J],\end{aligned} \tag{A.2.40}$$

Furthermore, we apply the same continuity bound as in Eq. (10.10.21) to bound $\widetilde{H}_\alpha^\uparrow(S|\tilde{T}E)_{\widetilde{\nu}}$,

which gives

$$\kappa_i^{\uparrow} \geq \inf_{\substack{J\in \mathrm{Pos}(A'\bar{A}B),\\ \boldsymbol{\nu}\in\mathbb{P}(\widehat{\mathcal{C}}_{\setminus \mathrm{gen}})}} \Bigg(\frac{1}{1-\alpha} \log \Big(\gamma \sum_{c\in\widehat{\mathcal{C}}_{\setminus \mathrm{gen}}} \boldsymbol{\nu}(\hat{c}) 2^{(\alpha-1)f(\hat{c})} \\ + (1-\gamma)\, 2^{(\alpha-1)(\Delta(\epsilon_{\mathrm{full}})+f(\mathtt{gen}))} 2^{-(\alpha-1)\widetilde{H}_\alpha^{\downarrow}(S|\tilde{T}E)_{\nu_{|\mathtt{gen}}}} \Big) \Bigg) \tag{A.2.41}$$

$$\text{s.t. } \mathrm{Tr}_B[J] = \mathbb{I}_{A'\bar{A}},$$
$$\boldsymbol{\nu} = \boldsymbol{\Psi}_{\tilde{\xi}}[J].$$

Again, the block-tagged states satisfy the assumptions 1–3. Hence, as for $h_\alpha^{\uparrow,\mathrm{QKD}}$ one can proceed by applying the simplifications of Section 10.6 and with Lemma 10.10.1 the claim follows. □

Appendix B

Code Availability

The numerical simulations and data analysis presented in this thesis were performed using MATLAB. The source code corresponding to the results in each chapter is available in the following public repositories.

Please note that these repositories contain the full codebase for the associated publications, including simulations for results that may have been omitted from this thesis.

- **Chapters 5 and 6 (based on Ref. [1]):** Source code [114] available at https://github.com/Optical-Quantum-Communication-Theory/ImprovedDecoyStateAndFlagStateSquashingMethods.
- **Chapter 7 (based on Ref. [3]):** Source code [115] available at https://github.com/Optical-Quantum-Communication-Theory/Improved-finite-size-effects-in-QKD-protocols-with-applications-to-decoy-state-QKD.
- **Chapter 9 (based on Ref. [2]):** Source code [116] available at https://github.com/Optical-Quantum-Communication-Theory/FiniteSizeAnalysisOfPrepareAndMeasureAndDecoyStateQKDViaEntropyAccumulation.
- **Chapter 10 (based on Ref. [4]):** Source code [117] available at https://github.com/Optical-Quantum-Communication-Theory/Renyi-security-framework.

All code is open-source and documented within the respective repositories.

Appendix C

Kraus Operators for Various Example Protocols

C.1 Kraus Operators for Example Protocols in Asymptotic Regime

C.1.1 Kraus Operators BB84 Protocol

Again, we are using the framework of [67] together with the simplifications of [80, App. A]. Furthermore, as mentioned in the main text, we use the squashing model from [62, 63] and recover Bob's qubit POVM elements, which are

$$
\begin{aligned}
F^B_{(Z,0)} &= p_z \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \; F^B_{(Z,1)} = p_z \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \\
F^B_{(X,0)} &= \frac{p_x}{2} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix}, \; F^B_{(X,1)} = \frac{p_x}{2} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & -1 & 1 \end{pmatrix}, \\
F^B_{\perp} &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.
\end{aligned}
\tag{C.1.1}
$$

For Alice's side, the POVM elements are determined by the source replacement scheme [56, 57] and given by

$$\begin{aligned} F^A_{(Z,0)} &= |0\rangle\langle 0|\,, F^A_{(Z,1)} = |1\rangle\langle 1|\,, \\ F^A_{(X,0)} &= |2\rangle\langle 2|\,, F^A_{(X,1)} = |3\rangle\langle 3|\,. \end{aligned} \tag{C.1.2}$$

The resulting Kraus operators are

$$\begin{aligned} K_Z = &\left[\begin{pmatrix} 1 \\ 0 \end{pmatrix}_R \otimes \begin{pmatrix} 1 & & \\ & 0 & \\ & & 0 \end{pmatrix}_A + \begin{pmatrix} 0 \\ 1 \end{pmatrix}_R \otimes \begin{pmatrix} 0 & & \\ & 1 & \\ & & 0 \end{pmatrix}_A \right] \\ &\otimes \sqrt{p_z} \begin{pmatrix} 0 & \\ & 1 \end{pmatrix}_B \otimes \begin{pmatrix} 1 \\ 0 \end{pmatrix}_C, \\ K_X = &\left[\begin{pmatrix} 1 \\ 0 \end{pmatrix}_R \otimes \begin{pmatrix} 0 & & \\ & 0 & \\ & & 1 \end{pmatrix}_A + \begin{pmatrix} 0 \\ 1 \end{pmatrix}_R \otimes \begin{pmatrix} 0 & & \\ & 0 & \\ & & 1 \end{pmatrix}_A \right] \\ &\otimes \sqrt{p_x} \begin{pmatrix} 0 & \\ & 1 \end{pmatrix}_B \otimes \begin{pmatrix} 0 \\ 1 \end{pmatrix}_C, \end{aligned} \tag{C.1.3}$$

and

$$\begin{aligned} Z_1 &= \begin{pmatrix} 1 & \\ & 0 \end{pmatrix} \otimes \mathbb{1}_{\dim_A \times \dim_B \times 2}, \\ Z_2 &= \begin{pmatrix} 0 & \\ & 1 \end{pmatrix} \otimes \mathbb{1}_{\dim_A \times \dim_B \times 2}. \end{aligned} \tag{C.1.4}$$

C.1.2 POVM Elements and Kraus Operators Six-State Protocol

For simplicity, we only present the Kraus operators for the case $N_B = 1$. In all cases, the original POVM elements (before squashing) can be found in [63].

After applying the flag-state squasher, the POVM elements on Bob's $\leq N_B$-photon

subspace are

$$\begin{aligned}
\tilde{F}^B_{(Z,0)} &= p_z \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \; \tilde{F}^B_{(Z,1)} = p_z \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \\
\tilde{F}^B_{(X,0)} &= \frac{p_x}{2} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix}, \; \tilde{F}^B_{(X,1)} = \frac{p_x}{2} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & -1 & 1 \end{pmatrix}, \\
\tilde{F}^B_{(Y,0)} &= \frac{p_y}{2} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & -i \\ 0 & i & 1 \end{pmatrix}, \; \tilde{F}^B_{(Y,1)} = \frac{p_y}{2} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & i \\ 0 & -i & 1 \end{pmatrix}, \\
\tilde{F}^B_{\perp} &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.
\end{aligned} \tag{C.1.5}$$

These need to be padded with flags. Therefore, let $E_i := \mathrm{diag}(0, \ldots, 0, 1, 0, \ldots, 0) \in \mathbb{R}^{10\times 10}$, be a diagonal matrix with 1 at the i-th entry. Furthermore, one needs to include double (dc) and cross-clicks (cc). For $N_B = 1$, these are only part of the flags. Hence, Bob's full squashed POVM elements are

$$\begin{aligned}
F^B_{(Z,0)} &= \tilde{F}^B_{(Z,0)} \oplus E_1, \quad F^B_{(Z,1)} = \tilde{F}^B_{(Z,1)} \oplus E_2, \\
F^B_{(X,0)} &= \tilde{F}^B_{(X,0)} \oplus E_3, \quad F^B_{(X,1)} = \tilde{F}^B_{(X,1)} \oplus E_4, \\
F^B_{(Y,0)} &= \tilde{F}^B_{(Y,0)} \oplus E_5, \quad F^B_{(Y,1)} = \tilde{F}^B_{(Y,1)} \oplus E_6, \\
F^B_{\mathrm{dc},Z} &= \bar{0}_3 \oplus E_7, \quad F^B_{\mathrm{dc},X} = \bar{0}_3 \oplus E_8, \\
F^B_{\mathrm{dc},Y} &= \bar{0}_3 \oplus E_9, \\
F^B_{\mathrm{cc}} &= \bar{0}_3 \oplus E_{10}, \\
F^B_{\perp} &= \tilde{F}^B_{\perp} \oplus \bar{0}_{10},
\end{aligned} \tag{C.1.6}$$

where $\bar{0}_m$ indicates a matrix with only zeros of dimension m. Regarding Alice's POVM elements, for numerical stability we apply a Schmidt decomposition after the source replacement scheme to reduce the dimensions and arrive at

$$\begin{aligned}
F^A_{(Z,0)} &= p_z \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, F^A_{(Z,1)} = p_z \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \\
F^A_{(X,0)} &= \frac{p_x}{2} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}, F^A_{(X,1)} = \frac{p_x}{2} \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}, \\
F^A_{(Y,0)} &= \frac{p_y}{2} \begin{pmatrix} 1 & i \\ -i & 1 \end{pmatrix}, F^A_{(Y,1)} = \frac{p_y}{2} \begin{pmatrix} 1 & -i \\ i & 1 \end{pmatrix}.
\end{aligned} \tag{C.1.7}$$

The resulting Kraus operators in the framework of [67] are

$$\begin{aligned}
K_Z &= \sqrt{p_z}\left[\begin{pmatrix}1\\0\end{pmatrix}_R \otimes \begin{pmatrix}1 & 0\\0 & 0\end{pmatrix} + \begin{pmatrix}0\\1\end{pmatrix}_R \otimes \begin{pmatrix}0 & 0\\0 & 1\end{pmatrix}\right] \\
&\quad \otimes \left[0 \oplus \sqrt{p_z}\mathbb{1}_2 \oplus (E_1 + E_2)\right]_B \otimes \begin{pmatrix}1\\0\\0\end{pmatrix}_C , \\
K_X &= \frac{\sqrt{p_x}}{2}\left[\begin{pmatrix}1\\0\end{pmatrix}_R \otimes \begin{pmatrix}1 & 1\\1 & 1\end{pmatrix} + \begin{pmatrix}0\\1\end{pmatrix}_R \otimes \begin{pmatrix}1 & -1\\-1 & 1\end{pmatrix}\right] \\
&\quad \otimes \left[0 \oplus \sqrt{p_x}\mathbb{1}_2 \oplus (E_3 + E_4)\right]_B \otimes \begin{pmatrix}0\\1\\0\end{pmatrix}_C , \\
K_Y &= \frac{\sqrt{p_y}}{2}\left[\begin{pmatrix}1\\0\end{pmatrix}_R \otimes \begin{pmatrix}1 & i\\-i & 1\end{pmatrix} + \begin{pmatrix}0\\1\end{pmatrix}_R \otimes \begin{pmatrix}1 & -i\\i & 1\end{pmatrix}\right] \\
&\quad \otimes \left[0 \oplus \sqrt{p_y}\mathbb{1}_2 \oplus (E_5 + E_6)\right]_B \otimes \begin{pmatrix}0\\0\\1\end{pmatrix}_C ,
\end{aligned} \tag{C.1.8}$$

and

$$\begin{aligned}
Z_1 &= \begin{pmatrix}1 & \\ & 0\end{pmatrix} \otimes \mathbb{1}_{\dim_A \times \dim_B \times 3}, \\
Z_2 &= \begin{pmatrix}0 & \\ & 1\end{pmatrix} \otimes \mathbb{1}_{\dim_A \times \dim_B \times 3}.
\end{aligned} \tag{C.1.9}$$

C.2 Kraus Operators for Example Protocols in IID Regime

Again, we are using the framework of [67] together with the simplifications of [80, App. A].

C.2.1 Qubit BB84

In the main text, we set $p_z^B = p_x^B = 1/2$, but for generality, we provide the Kraus operators for arbitrary basis choices. Given a perfect qubit protocol, Bob's qubit POVM elements

are

$$
\begin{aligned}
M^B_{(Z,0)} &= p^B_z \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \; M^B_{(Z,1)} = p^B_z \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \\
M^B_{(X,0)} &= \frac{p^B_x}{2} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix}, \; M^B_{(X,1)} = \frac{p^B_x}{2} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & -1 & 1 \end{pmatrix}, \\
M^B_{\perp} &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.
\end{aligned} \tag{C.2.1}
$$

For Alice's side, the POVM elements are determined by the source replacement scheme [56, 57] and given by

$$
\begin{aligned}
M^A_{(Z,0)} &= |0\rangle\langle 0|, M^A_{(Z,1)} = |1\rangle\langle 1|, \\
M^A_{(X,0)} &= |2\rangle\langle 2|, M^A_{(X,1)} = |3\rangle\langle 3|.
\end{aligned} \tag{C.2.2}
$$

We only generate secret key from the key generation rounds. This can be viewed as a post-processing where a generation round happens for each of Alice's choices with probability $p(\text{gen}|a)$. However, the generation round decision is independent of Alice's choices in this particular protocol (see main text for details Section 7.5.1), hence $p(\text{gen}|a) = p(\text{gen})$. The resulting Kraus operators are

$$
\begin{aligned}
K_Z = &\left[\begin{pmatrix} 1 \\ 0 \end{pmatrix}_R \otimes \begin{pmatrix} 1 & & \\ & 0 & \\ & & 0 \end{pmatrix}_A + \begin{pmatrix} 0 \\ 1 \end{pmatrix}_R \otimes \begin{pmatrix} 0 & & \\ & 1 & \\ & & 0 \end{pmatrix}_A \right] \\
&\otimes \sqrt{p^B_z} \begin{pmatrix} 0 & & \\ & 1 & \\ & & 1 \end{pmatrix}_B \otimes \begin{pmatrix} 1 \\ 0 \end{pmatrix}_C \sqrt{p(\text{gen})}, \\
K_X = &\left[\begin{pmatrix} 1 \\ 0 \end{pmatrix}_R \otimes \begin{pmatrix} 0 & & \\ & 0 & \\ & & 0 \end{pmatrix}_A + \begin{pmatrix} 0 \\ 1 \end{pmatrix}_R \otimes \begin{pmatrix} 0 & & \\ & 0 & \\ & & 1 \end{pmatrix}_A \right] \\
&\otimes \sqrt{p^B_x} \begin{pmatrix} 0 & & \\ & 1 & \\ & & 1 \end{pmatrix}_B \otimes \begin{pmatrix} 0 \\ 1 \end{pmatrix}_C \sqrt{p(\text{gen})},
\end{aligned} \tag{C.2.3}
$$

and

$$Z_1 = \begin{pmatrix} 1 & \\ & 0 \end{pmatrix} \otimes \mathbb{1}_{\dim_A \times \dim_B \times 2},$$
$$Z_2 = \begin{pmatrix} 0 & \\ & 1 \end{pmatrix} \otimes \mathbb{1}_{\dim_A \times \dim_B \times 2}. \tag{C.2.4}$$

C.2.2 Decoy BB84

After applying the squashing map from [62, 63], Bob's measurements act on a qubit again. In the main text, we set $p_z^B = p_x^B = 1/2$, but for generality, we provide the Kraus operators for arbitrary basis choices. The resulting POVM elements coincide with the ones for a perfect qubit protocol, which again are

$$M^B_{(Z,0)} = p_z^B \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \; M^B_{(Z,1)} = p_z^B \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$
$$M^B_{(X,0)} = \frac{p_x^B}{2} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix}, \; M^B_{(X,1)} = \frac{p_x^B}{2} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & -1 & 1 \end{pmatrix}, \tag{C.2.5}$$
$$M^B_{\perp} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

For Alice's side, the POVM elements are determined by the source replacement scheme [56, 57] and given by

$$M^A_{(Z,0)} = |0\rangle\langle 0|, M^A_{(Z,1)} = |1\rangle\langle 1|,$$
$$M^A_{(X,0)} = |2\rangle\langle 2|, M^A_{(X,1)} = |3\rangle\langle 3|. \tag{C.2.6}$$

Still, as in the qubit case we only generate secret key from the key generation rounds. This can again be viewed as a post-processing, but now we additionally need to condition on single photons. Thus, a generation round happens for each of Alice's choices with probability $p(\text{gen}|a, 1)$. The generation round decision remains independent of Alice's choices in this particular protocol (see main text for details Section 7.5.2), hence $p(\text{gen}|a, 1) = p(\text{gen}|1)$.

The resulting Kraus operators are

$$\begin{aligned} K_Z = &\left[\begin{pmatrix} 1 \\ 0 \end{pmatrix}_R \otimes \begin{pmatrix} 1 & & \\ & 0 & \\ & & 0 \end{pmatrix}_A + \begin{pmatrix} 0 \\ 1 \end{pmatrix}_R \otimes \begin{pmatrix} 0 & & \\ & 1 & \\ & & 0 \end{pmatrix}_A \right] \\ &\otimes \sqrt{p_z^B} \begin{pmatrix} 0 & & \\ & 1 & \\ & & 1 \end{pmatrix}_B \otimes \begin{pmatrix} 1 \\ 0 \end{pmatrix}_C \cdot \sqrt{p(\text{gen}|1)}, \\ K_X = &\left[\begin{pmatrix} 1 \\ 0 \end{pmatrix}_R \otimes \begin{pmatrix} 0 & & \\ & 0 & \\ & & 1 \end{pmatrix}_A + \begin{pmatrix} 0 \\ 1 \end{pmatrix}_R \otimes \begin{pmatrix} 0 & & \\ & 0 & \\ & & 1 \end{pmatrix}_A \right] \\ &\otimes \sqrt{p_x^B} \begin{pmatrix} 0 & & \\ & 1 & \\ & & 1 \end{pmatrix}_B \otimes \begin{pmatrix} 0 \\ 1 \end{pmatrix}_C \cdot \sqrt{p(\text{gen}|1)}, \end{aligned} \tag{C.2.7}$$

and

$$\begin{aligned} Z_1 &= \begin{pmatrix} 1 & \\ & 0 \end{pmatrix} \otimes \mathbb{1}_{\dim_A \times \dim_B \times 2}, \\ Z_2 &= \begin{pmatrix} 0 & \\ & 1 \end{pmatrix} \otimes \mathbb{1}_{\dim_A \times \dim_B \times 2}. \end{aligned} \tag{C.2.8}$$

C.3 Kraus Operators and States for Example Protocols in Rényi Security Framework

In this section, we summarize the states and Kraus operators required to calculate the key rates for each protocol. It is important to note that all Kraus operators are generated from Alice's POVM elements *conditioned* on a generation round. Therefore, we will first state how those conditional POVM elements can be constructed in a generic way.

C.3.1 Conditional POVM Construction

Here we especially focus on the case where a Schmidt decomposition was performed to reduce dimensions and show how the conditional POVM elements can be constructed in a generic way.

First, as in the main text, let $\{M_k^A\}_{k=1\ldots d_A}$ be Alice's POVM elements, however, here we do not set the "test" POVM elements via the relation $M_k^{A|\mathtt{test}} = M_k^A$, and similarly for the "generation" POVM elements. This is because that approach is not entirely suitable for handling the POVM elements arising after a Schmidt decomposition.

Next, as in the main text, we partition the POVM elements into test and generation subsets. Furthermore, under a Schmidt decomposition the POVM elements M_k^A contain a factor proportional to the testing or generation fraction, depending on which subset they fall into. Therefore, for convenience of notation we indicate these POVM elements with a "$\wedge$" similar to probabilities to indicate this factor, e.g. $M_1^{A\wedge\mathtt{test}}$. With this notation, the split into test and generation subsets can be written as

$$\{M_{\alpha,x}^{A\wedge\mathtt{test}}\}_{\alpha\in\mathcal{C}^{A,\mathtt{test}},x\in\mathcal{X}_\alpha^{\mathtt{test}}} = \{M_k^{A\wedge\mathtt{test}}\}, \tag{C.3.1}$$

$$\{M_{\alpha,x}^{A\wedge\mathtt{gen}}\}_{\alpha\in\mathcal{C}^{A,\mathtt{gen}},x\in\mathcal{X}_\alpha^{\mathtt{gen}}} = \{M_k^{A\wedge\mathtt{gen}}\}. \tag{C.3.2}$$

Again, a "$\wedge$" is supposed to indicate that

$$\mathrm{Tr}\big[\rho M_k^{A\wedge\mathtt{test}}\big] = p(k\wedge\mathtt{test}). \tag{C.3.3}$$

Then, similar to the main text, we define the partition operators for test (generation) rounds as

$$\Lambda^{\mathtt{test}} = \sum_{\substack{\alpha\in\mathcal{C}^{A,\mathtt{test}},\\ x\in\mathcal{X}_\alpha^{\mathtt{test}}}} M_{\alpha,x}^{A\wedge\mathtt{test}}, \quad \Lambda^{\mathtt{gen}} = \sum_{\substack{\alpha\in\mathcal{C}^{A,\mathtt{gen}},\\ x\in\mathcal{X}_\alpha^{\mathtt{gen}}}} M_{\alpha,x}^{A\wedge\mathtt{gen}}, \tag{C.3.4}$$

though now we allow for the possibility that these might not be projectors. Importantly however, note that summing positive semidefinite operators cannot result in an operator with smaller support, which implies we still have

$$\forall\alpha\in\mathcal{C}^{A,\mathtt{test}}, x\in\mathcal{X}_\alpha^{\mathtt{test}}, \quad \mathrm{supp}(M_{\alpha,x}^{A\wedge\mathtt{test}})\subseteq\mathrm{supp}(\Lambda^{\mathtt{test}}), \tag{C.3.5}$$

$$\forall\alpha\in\mathcal{C}^{A,\mathtt{gen}}, x\in\mathcal{X}_\alpha^{\mathtt{gen}}, \quad \mathrm{supp}(M_{\alpha,x}^{A\wedge\mathtt{gen}})\subseteq\mathrm{supp}(\Lambda^{\mathtt{gen}}), \tag{C.3.6}$$

a property we will use later. Similarly we define states conditioned on test and generation as

$$\big|\xi^t\big\rangle_{AA'} \coloneqq \sqrt{\Lambda^{\mathtt{test}}}\frac{|\psi\rangle_{AA'}}{\sqrt{\gamma}}, \tag{C.3.7}$$

$$|\xi^g\rangle_{AA'} \coloneqq \sqrt{\Lambda^{\mathtt{gen}}}\frac{|\psi\rangle_{AA'}}{\sqrt{1-\gamma}}. \tag{C.3.8}$$

We now aim to find the POVM elements conditioned on test and generation rounds, which we indicate as in the main text by

$$\{M^{A|\mathtt{test}}_{\alpha,x}\}_{\alpha\in\mathcal{C}^{A,\mathtt{test}},x\in\mathcal{X}^{\mathtt{test}}_{\alpha}} = \{M^{A|\mathtt{test}}_{k}\}, \tag{C.3.9}$$

$$\{M^{A|\mathtt{gen}}_{\alpha,x}\}_{\alpha\in\mathcal{C}^{A,\mathtt{gen}},x\in\mathcal{X}^{\mathtt{gen}}_{\alpha}} = \{M^{A|\mathtt{gen}}_{k}\}. \tag{C.3.10}$$

We define the conditional POVM elements by

$$M^{A|\mathtt{test}}_{\alpha,x} \coloneqq \sqrt{(\Lambda^{\mathtt{test}})^{-1}}^{\dagger} M^{A\wedge\mathtt{test}}_{\alpha,x} \sqrt{(\Lambda^{\mathtt{test}})^{-1}}, \tag{C.3.11}$$

$$M^{A|\mathtt{gen}}_{\alpha,x} \coloneqq \sqrt{(\Lambda^{\mathtt{gen}})^{-1}}^{\dagger} M^{A\wedge\mathtt{gen}}_{\alpha,x} \sqrt{(\Lambda^{\mathtt{gen}})^{-1}}, \tag{C.3.12}$$

for all $\alpha \in \mathcal{C}^{A,\mathtt{test}}, x \in \mathcal{X}^{\mathtt{test}}_{\alpha}$ and $\alpha \in \mathcal{C}^{A,\mathtt{gen}}, x \in \mathcal{X}^{\mathtt{gen}}_{\alpha}$, respectively. Crucially, the $^{-1}$ notation in the above expression is to be understood as the Moore-Penrose pseudo-inverse rather than the usual inverse (which would be undefined if the operators are not full-support). By construction these operators are positive semidefinite and hermitian. Moreover, due to the properties of the pseudo-inverse, it holds

$$\sqrt{(\Lambda^{\mathtt{test}})^{-1}}\sqrt{(\Lambda^{\mathtt{test}})} = \Pi_{\mathrm{supp}(\Lambda^{\mathtt{test}})}, \tag{C.3.13}$$

$$\sqrt{(\Lambda^{\mathtt{gen}})^{-1}}\sqrt{(\Lambda^{\mathtt{gen}})} = \Pi_{\mathrm{supp}(\Lambda^{\mathtt{gen}})}, \tag{C.3.14}$$

where with $\Pi_{\mathrm{supp}(A)}$ we indicate the projector onto the support of an operator A. In addition, for each $M^{A\wedge\mathtt{test}}_{\alpha,x}$ and $M^{A\wedge\mathtt{gen}}_{\alpha,x}$ we have

$$\Pi_{\mathrm{supp}(\Lambda^{\mathtt{test}})} M^{A\wedge\mathtt{test}}_{\alpha,x} \Pi_{\mathrm{supp}(\Lambda^{\mathtt{test}})} = M^{A\wedge\mathtt{test}}_{\alpha,x}, \tag{C.3.15}$$

$$\Pi_{\mathrm{supp}(\Lambda^{\mathtt{gen}})} M^{A\wedge\mathtt{gen}}_{\alpha,x} \Pi_{\mathrm{supp}(\Lambda^{\mathtt{gen}})} = M^{A\wedge\mathtt{gen}}_{\alpha,x}. \tag{C.3.16}$$

Furthermore, it holds for the POVM elements conditioned on test rounds

$$\sum_{\substack{\alpha\in\mathcal{C}^{A,\mathtt{test}},\\ x\in\mathcal{X}^{\mathtt{test}}_{\alpha}}} M^{A|\mathtt{test}}_{\alpha,x} \tag{C.3.17}$$

$$= \sum_{\substack{\alpha\in\mathcal{C}^{A,\mathtt{test}},\\ x\in\mathcal{X}^{\mathtt{test}}_{\alpha}}} \sqrt{(\Lambda^{\mathtt{test}})^{-1}}^{\dagger} M^{A\wedge\mathtt{test}}_{\alpha,x} \sqrt{(\Lambda^{\mathtt{test}})^{-1}} \tag{C.3.18}$$

$$= \Pi_{\mathrm{supp}(\Lambda^{\mathtt{test}})}, \tag{C.3.19}$$

and similarly for generation rounds,

$$\sum_{\substack{\alpha\in\mathcal{C}^{A,\mathtt{gen}},\\ x\in\mathcal{X}^{\mathtt{gen}}_{\alpha}}} M^{A|\mathtt{gen}}_{\alpha,x} = \Pi_{\mathrm{supp}(\Lambda^{\mathtt{gen}})}. \tag{C.3.20}$$

Therefore, the sets $\{M_{\alpha,x}^{A|\mathtt{test}}\}$ and $\{M_{\alpha,x}^{A|\mathtt{gen}}\}$ indeed form valid POVMs, but on the subspaces $\mathrm{supp}(\Lambda^{\mathtt{test}})$ and $\mathrm{supp}(\Lambda^{\mathtt{gen}})$ respectively rather than the entire Hilbert space of A.[1]

Next, inserting the above properties, we find (omitting identities on A') that these operators indeed yield the correct conditional probabilities,

$$\mathrm{Tr}\big[\big|\xi^t\big\rangle\!\big\langle\xi^t\big|_{AA'} M_{\alpha,x}^{A|\mathtt{test}}\big] \tag{C.3.21}$$

$$= \frac{1}{\gamma}\,\mathrm{Tr}\Big[\sqrt{\Lambda^{\mathtt{test}}}\,|\psi\rangle\!\langle\psi|_{AA'}\,\sqrt{\Lambda^{\mathtt{test}}}^{\dagger}\sqrt{(\Lambda^{\mathtt{test}})^{-1}}^{\dagger} M_{\alpha,x}^{A\wedge\mathtt{test}}\sqrt{(\Lambda^{\mathtt{test}})^{-1}}\Big] \tag{C.3.22}$$

$$= \frac{1}{\gamma}\,\mathrm{Tr}\Big[|\psi\rangle\!\langle\psi|_{AA'}\,\sqrt{\Lambda^{\mathtt{test}}}^{\dagger}\sqrt{(\Lambda^{\mathtt{test}})^{-1}}^{\dagger} M_{\alpha,x}^{A\wedge\mathtt{test}}\sqrt{(\Lambda^{\mathtt{test}})^{-1}}\sqrt{\Lambda^{\mathtt{test}}}\Big] \tag{C.3.23}$$

$$= \frac{1}{\gamma}\,\mathrm{Tr}\big[|\psi\rangle\!\langle\psi|_{AA'}\,\Pi_{\mathrm{supp}(\Lambda^{\mathtt{test}})} M_{\alpha,x}^{A\wedge\mathtt{test}}\Pi_{\mathrm{supp}(\Lambda^{\mathtt{test}})}\big] \tag{C.3.24}$$

$$= \frac{1}{\gamma}\,\mathrm{Tr}\big[|\psi\rangle\!\langle\psi|_{AA'}\,M_{\alpha,x}^{A\wedge\mathtt{test}}\big] = p(\alpha,x|\mathtt{test}), \tag{C.3.25}$$

and similarly for generation rounds.

In addition, they result in the correct state after the measurement. To see this, let us see what happens if we apply $\sqrt{M_{\alpha,x}^{A|\mathtt{test}}}$ onto $|\xi^t\rangle_{AA'}$. Since $\mathrm{supp}(M_{\alpha,x}^{A\wedge\mathtt{test}}) \subseteq \mathrm{supp}(\Lambda^{\mathtt{test}})$, it follows

$$\sqrt{M_{\alpha,x}^{A|\mathtt{test}}}\,\big|\xi^t\big\rangle_{AA'} \tag{C.3.26}$$

$$= \sqrt{M_{\alpha,x}^{A\wedge\mathtt{test}}}\sqrt{(\Lambda^{\mathtt{test}})^{-1}}\sqrt{\Lambda^{\mathtt{test}}}\frac{|\psi\rangle_{AA'}}{\sqrt{\gamma}} \tag{C.3.27}$$

$$= \sqrt{M_{\alpha,x}^{A\wedge\mathtt{test}}}\Pi_{\mathrm{supp}(\Lambda^{\mathtt{test}})}\frac{|\psi\rangle_{AA'}}{\sqrt{\gamma}} \tag{C.3.28}$$

$$= \sqrt{M_{\alpha,x}^{A\wedge\mathtt{test}}}\frac{|\psi\rangle_{AA'}}{\sqrt{\gamma}}. \tag{C.3.29}$$

A similar calculation yields the equivalent result for the POVM elements conditioned on generation rounds.

Finally, we note that any operator and state unitarily equivalent to the ones above satisfies the same properties and could equally well be used. Now, in the following subsections, we will present the POVM elements constructed for each example shown in the main text.

[1] In some cases though, such as the qubit BB84 protocol we consider below with the Schmidt-decomposition dimension reduction, these subspaces can turn out to be equal to the full Hilbert space. However, this fact does not affect any claims in this section.

C.3.2 Qubit BB84

For this protocol, we reduced the dimensions using a Schmidt-decomposition for better numerical performance. The final dimensions of the systems involved, are

$$\dim_A = 2, \quad \dim_{A'} = 2, \quad \dim_B = 3. \tag{C.3.30}$$

Then, Bob's POVM elements are given by

$$\begin{aligned} M^B_{(Z,0)} &= p^B_z \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad M^B_{(Z,1)} = p^B_z \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \\ M^B_{(X,0)} &= \frac{p^B_x}{2} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix}, \quad M^B_{(X,1)} = \frac{p^B_x}{2} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & -1 & 1 \end{pmatrix}, \\ M^B_{\perp} &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \end{aligned} \tag{C.3.31}$$

where $p^B_z = \gamma$ and $p^B_x = 1 - \gamma$. For Alice's side, the POVM elements are determined by the source replacement scheme [56, 57] and after reducing the dimensions through the Schmidt-decomposition, they are given by

$$\begin{aligned} M^A_{(Z,0)} &= p^A_z \begin{pmatrix} 1 & \\ & 0 \end{pmatrix}, \quad M^A_{(Z,1)} = p^A_z \begin{pmatrix} 0 & \\ & 1 \end{pmatrix}, \\ M^A_{(X,0)} &= \frac{p^A_x}{2} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}, \quad M^A_{(X,1)} = \frac{p^A_x}{2} \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}, \end{aligned} \tag{C.3.32}$$

where $p^A_z = \gamma$ and $p^A_x = 1 - \gamma$. The conditioning on `gen` and `test` follows immediately by dividing by p^A_z and p^A_x, respectively, since the Z-basis is used exclusively for generation rounds and the X-basis is used exclusively for test rounds.

To define the signal states $|\xi^t\rangle$ and $|\xi^g\rangle$, for notational convenience, let $|\phi^+\rangle_{AA'} \coloneqq \frac{1}{\sqrt{2}}(|00\rangle_{AA'} + |11\rangle_{AA'})$, be a maximally entangled Bell-state on systems AA'. Then, the signal states are given by

$$|\xi^t\rangle = |\phi^+\rangle_{AA'}, \qquad |\xi^g\rangle = |\phi^+\rangle_{AA'}. \tag{C.3.33}$$

Next, we define the Kraus operators of the map $\mathcal{G}$. Since in key generation rounds we only send the Z-basis, only one Kraus operator for the map $\mathcal{G}$ is required, which is

$$K_Z = \sqrt{p^B_z} \left[\begin{pmatrix} 1 & \\ & 0 \end{pmatrix}_{A \to S_Q} + \begin{pmatrix} 0 & \\ & 1 \end{pmatrix}_{A \to S_Q} \right] \otimes \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}_B \otimes 1_I, \tag{C.3.34}$$

where $p_z^B = \gamma$. Furthermore, 1_I is just a scalar since there is only one announcement surviving sifting in generation rounds.

Finally, the Kraus operators of the map $\mathcal{Z}$ are

$$Z_1 = \begin{pmatrix} 1 & \\ & 0 \end{pmatrix} \otimes \mathbb{I}_{\dim_B - 1}, \quad Z_2 = \begin{pmatrix} 0 & \\ & 1 \end{pmatrix} \otimes \mathbb{I}_{\dim_B - 1}. \tag{C.3.35}$$

C.3.3 Active Decoy BB84

This protocol uses the qubit squasher [62, 63] for Bob's POVM elements, hence we recover the qubit POVM elements stated in Eq. (C.3.31). Furthermore, we set $N_{\text{ent}} = 1$, i.e. we only use $m = 0$ and $m = 1$ photons for key generation.

We chose the dimension of A consistent over all photon numbers to be $\dim_A = 2$ as it simplifies the expressions of the states, POVM elements and Kraus operators. For $m = 0$ photons, this could be reduced to $\dim_A = 1$ via a another Schmidt-decomposition.

Thus, given this setup, Alice's POVM elements are those of the qubit protocol Eq. (C.3.32) for both photon numbers. Defining the signal states with $m = 0$ and $m = 1$ photons, consistent with our choices above leads to

$$\left|\xi^t_{|m=0}\right\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}_{AA'}, \qquad \left|\xi^g_{|m=0}\right\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}_{AA'} \tag{C.3.36}$$

$$\left|\xi^t_{|m=1}\right\rangle = \left|\phi^+\right\rangle_{AA'}, \qquad \left|\xi^g_{|m=1}\right\rangle = \left|\phi^+\right\rangle_{AA'}. \tag{C.3.37}$$

Next, we require the Kraus operators of the $\mathcal{G}$ for $m = 0$ and $m = 1$ photons. Since in a key generation round we only send the Z-basis, there are only two Kraus operators for the $\mathcal{G}$; one for $m = 0$ and one for $m = 1$ photons. However, with our current representation, both Kraus operators are equal, and both are given by the qubit Kraus operator shown in Eq. (C.3.34). The same holds for the pinching map $\mathcal{Z}$.

C.3.4 Passive Decoy BB84

This protocol uses the flag-state squasher [42]. Thus, we need to construct Bob's POVM elements including flags. First, Bob's POVM elements on the $\leq N_B$-photon subspace with

$N_B = 1$ are

$$\begin{aligned}
\tilde{M}^B_{(Z,0)} &= p_z^B \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \; \tilde{M}^B_{(Z,1)} = p_z^B \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \\
\tilde{M}^B_{(X,0)} &= \frac{p_x^B}{2} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix}, \; \tilde{M}^B_{(X,1)} = \frac{p_x^B}{2} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & -1 & 1 \end{pmatrix}, \\
\tilde{M}^B_{\perp} &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},
\end{aligned} \tag{C.3.38}$$

where $p_z^B = \gamma$ and $p_x^B = 1 - \gamma$. Next, these POVM elements need to be padded with flags. Therefore, let $E_i := \mathrm{diag}(0, \ldots, 0, 1, 0, \ldots, 0) \in \mathbb{R}^{5\times 5}$, be a diagonal matrix with 1 at the i-th entry. Hence, Bob's full squashed POVM elements are

$$\begin{aligned}
M^B_{(Z,0)} &= \tilde{M}^B_{(Z,0)} \oplus E_1, \quad M^B_{(Z,1)} = \tilde{M}^B_{(Z,1)} \oplus E_2, \\
M^B_{(X,0)} &= \tilde{M}^B_{(X,0)} \oplus E_3, \quad M^B_{(X,1)} = \tilde{M}^B_{(X,1)} \oplus E_4, \\
M^B_{\mathrm{mult}} &= \bar{0}_3 \oplus E_5, \quad M^B_{\perp} = \tilde{M}^B_{\perp} \oplus \bar{0}_5,
\end{aligned} \tag{C.3.39}$$

where $\bar{0}_k$ indicates a matrix with only zeros of dimension k.

Next, we construct the states and Alice's POVM elements. Again, we chose the dimension of A consistent over all photon numbers to be $\dim_A = 4$ and $N_{\mathrm{ent}} = 1$, i.e. we only use $m = 0$ and $m = 1$ photons for key generation. Then, the signal states are given by

$$\left|\xi^t_{|m=0}\right\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix}_{AA'}, \qquad \left|\xi^g_{|m=0}\right\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}_{AA'} \tag{C.3.40}$$

$$\left|\xi^t_{|m=1}\right\rangle = \frac{1}{\sqrt{2}} \left(\begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}_{A} \otimes \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}_{A'} + \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}_{A} \otimes \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}_{A'} \right), \tag{C.3.41}$$

$$\left|\xi^g_{|m=1}\right\rangle = \frac{1}{\sqrt{2}} \left(\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}_{A} \otimes \begin{pmatrix} 1 \\ 0 \end{pmatrix}_{A'} + \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}_{A} \otimes \begin{pmatrix} 0 \\ 1 \end{pmatrix}_{A'} \right). \tag{C.3.42}$$

For all photon numbers Alice's POVM elements are,

$$M^A_{(Z,0)} = \begin{pmatrix} 1 & & & \\ & 0 & & \\ & & 0 & \\ & & & 0 \end{pmatrix}, \quad M^A_{(Z,1)} = \begin{pmatrix} 0 & & & \\ & 1 & & \\ & & 0 & \\ & & & 0 \end{pmatrix},$$
$$M^A_{(X,0)} = \begin{pmatrix} 0 & & & \\ & 0 & & \\ & & 1 & \\ & & & 0 \end{pmatrix}, \quad M^A_{(X,1)} = \begin{pmatrix} 0 & & & \\ & 0 & & \\ & & 0 & \\ & & & 1 \end{pmatrix}. \tag{C.3.43}$$

Here, no explicit conditioning on `test` or `gen` required, because these POVM elements are defined based on the generic source replacement scheme, i.e. no Schmidt decomposition was applied.

Next, we define the Kraus operators of the $\mathcal{G}$ for $m = 0$ and $m = 1$ photons. Since in key generation rounds we only send the Z-basis, there are only two Kraus operators, for the $\mathcal{G}$. However, with our current representation, both Kraus operators are equal and given by

$$K_Z = \left[\begin{pmatrix} 1 \\ 0 \end{pmatrix}_S \begin{pmatrix} 1 & 0 & 0 & 0 \end{pmatrix}_A + \begin{pmatrix} 0 \\ 1 \end{pmatrix}_S \begin{pmatrix} 0 & 1 & 0 & 0 \end{pmatrix}_A \right] \otimes \left[0 \oplus \sqrt{p_z^B}\mathbb{I}_2 \oplus (E_1 + E_2) \right]_B \otimes 1_I \quad , \tag{C.3.44}$$

where 1_I is just a scalar. Finally, the Kraus operators for the pinching map $\mathcal{Z}$ are

$$Z_1 = \begin{pmatrix} 1 & \\ & 0 \end{pmatrix} \otimes \mathbb{I}_{\dim_B}, \quad Z_2 \ = \begin{pmatrix} 0 & \\ & 1 \end{pmatrix} \otimes \mathbb{I}_{\dim_B}. \tag{C.3.45}$$

C.3.5 Passive Decoy 4-6 Protocol

We again are required to use the flag-state squasher [42] for this protocol. Therefore, let us first define POVM elements on Bob's $\leq N_B$-photon subspace for $N_B = 1$. Those are

$$\begin{aligned}
\tilde{M}^B_{(Z,0)} &= p^B_z \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \; \tilde{M}^B_{(Z,1)} = p^B_z \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \\
\tilde{M}^B_{(X,0)} &= \frac{p^B_x}{2} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix}, \; \tilde{M}^B_{(X,1)} = \frac{p^B_x}{2} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & -1 & 1 \end{pmatrix}, \\
\tilde{M}^B_{(Y,0)} &= \frac{p^B_y}{2} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & -i \\ 0 & i & 1 \end{pmatrix}, \; \tilde{M}^B_{(Y,1)} = \frac{p^B_y}{2} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & i \\ 0 & -i & 1 \end{pmatrix}, \\
\tilde{M}^B_{\perp} &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.
\end{aligned} \tag{C.3.46}$$

As for the passive BB84 protocol, these POVM elements need to be padded with flags. Therefore, let $E_i := \mathrm{diag}(0, \ldots, 0, 1, 0, \ldots, 0) \in \mathbb{R}^{7\times 7}$, be a diagonal matrix with 1 at the i-th entry. Furthermore, one needs to include multi-clicks (mult) where multiple of Bob's detectors click, which for $N_B = 1$, are only part of the flags. Hence, Bob's full squashed POVM elements are

$$\begin{aligned}
M^B_{(Z,0)} &= \tilde{M}^B_{(Z,0)} \oplus E_1, \quad M^B_{(Z,1)} = \tilde{M}^B_{(Z,1)} \oplus E_2, \\
M^B_{(X,0)} &= \tilde{M}^B_{(X,0)} \oplus E_3, \quad M^B_{(X,1)} = \tilde{M}^B_{(X,1)} \oplus E_4, \\
M^B_{(Y,0)} &= \tilde{M}^B_{(Y,0)} \oplus E_5, \quad M^B_{(Y,1)} = \tilde{M}^B_{(Y,1)} \oplus E_6, \\
M^B_{\mathrm{mult}} &= \bar{0}_3 \oplus E_7, \quad M^B_{\perp} = \tilde{M}^B_{\perp} \oplus \bar{0}_7,
\end{aligned} \tag{C.3.47}$$

where $\bar{0}_k$ indicates a matrix with only zeros of dimension k.

For this protocol we construct Alice's signal states and POVM elements numerically since we do the diagonalization of Eq. (10.10.2) numerically.

In the example presented in the main text Alice sends four states with two intensities. Therefore, let $\{|i\rangle\}$ for $i = 1 \ldots 8$ label Alice's state and intensity choices.

Next, we calculate the states $|w_m(i)\rangle$ in Eq. (10.10.2), where we relabeled $i = (a, \mu)$.

For each $i = (a, \mu)$, we construct

$$\rho_{A'}^{(a,\mu)} = q \sum_{m=0}^{\infty} p(m|\mu) \, |m\rangle\langle m|_a + (1-q) \, |\sqrt{\mu}\rangle\langle\sqrt{\mu}|_a \,, \tag{C.3.48}$$

and diagonalize

$$\Pi_{\leq N_A} \rho_{A'}^{(a,\mu)} \Pi_{\leq N_A}. \tag{C.3.49}$$

This gives us the following quantities

$$\{\omega_m(i)\}_{i=1\ldots8}, \quad \{|w_m(i)\rangle\}_{i=1\ldots8}, \tag{C.3.50}$$

from which, as described in the main text, we can construct the state

$$|\psi\rangle_{AA_sA'\bar{A}} = \sum_{m=0}^{\infty} \sqrt{\omega_m} \, |m\rangle_{A_s} \left|\psi_{|m}\right\rangle_{AA'\bar{A}} , \tag{C.3.51}$$

where

$$\left|\psi_{|m}\right\rangle_{AA'\bar{A}} = \sum_i \sqrt{p(i|m)} \, |i\rangle_A \, |w_m(i)\rangle_{A'} \, |m\rangle_{\bar{A}} \,. \tag{C.3.52}$$

Next, we perform a Schmidt decomposition to reduce the dimensions of each block m. For each m we can therefore write

$$\left|\psi_{|m}\right\rangle_{AA'\bar{A}} = (U_m \otimes V_m) \, (\mathrm{diag}(\boldsymbol{\lambda}) \otimes \mathbb{I}_{\tilde{A}'}) \left|\bar{\phi}^+\right\rangle_{\tilde{A}\tilde{A}'} \otimes |m\rangle_{\bar{A}} \,, \tag{C.3.53}$$

where $\boldsymbol{\lambda}$ is the vector containing the Schmidt coefficients, U and V isometries and $\left|\bar{\phi}\right\rangle_{\tilde{A}\tilde{A}'}$ the unnormalized maximally entangled state.

Then, we define the states after the Schmidt-decomposition as

$$\left|\tilde{\psi}_{|m}\right\rangle_{\tilde{A}\tilde{A}'\bar{A}} := (\mathrm{diag}(\boldsymbol{\lambda}) \otimes \mathbb{I}_{\tilde{A}'}) \left|\bar{\phi}^+\right\rangle_{\tilde{A}\tilde{A}'} \otimes |m\rangle_{\bar{A}} \,, \tag{C.3.54}$$

and the POVM elements as

$$M_i^{\tilde{A}} := U_m^\dagger \, |i\rangle\langle i|_A \, U_m. \tag{C.3.55}$$

Afterwards, to find the POVM elements conditioned on test and generation rounds for each block m, we apply the procedure described in Appendix C.3.1. Having calculated these conditional POVM elements, the Kraus operators of the map $\mathcal{G}$ are also found numerically

from the resulting POVM elements conditioned on generation rounds. For each block m, they are given by

$$K_{Z|m} = \left[\begin{pmatrix} 1 \\ 0 \end{pmatrix}_S \otimes \sqrt{M_{Z,0}^{\tilde{A}|\mathbf{gen},m}} + \begin{pmatrix} 0 \\ 1 \end{pmatrix}_S \otimes \sqrt{M_{Z,1}^{\tilde{A}|\mathbf{gen},m}} \right] \otimes \left[0 \oplus \sqrt{p_z^B}\mathbb{I}_2 \oplus (E_1 + E_2) \right]_B \otimes 1_I \quad , \tag{C.3.56}$$

where 1_I is again a scalar. Finally, the Kraus operators of the pinching channel $\mathcal{Z}$ are for each block m given by

$$Z_{1|m} = \begin{pmatrix} 1 & \\ & 0 \end{pmatrix} \otimes \mathbb{I}_{\dim_{\tilde{A}|m} \cdot \dim_B}, \quad Z_{2|m} = \begin{pmatrix} 0 & \\ & 1 \end{pmatrix} \otimes \mathbb{I}_{\dim_{\tilde{A}|m} \cdot \dim_B}. \tag{C.3.57}$$

In addition to the Schmidt decomposition, one can reduce the dimensions of the optimization variables, i.e. the Choi states of Eve's channels acting on each block m, even further by noting the following.

Let J_m be the Choi state of Eve's channel $\mathcal{E}_m : \tilde{A}'\bar{A} \to B$ acting on the m-th block. Then, for any state ρ in system $\tilde{A}'$ it holds

$$\mathcal{E}_m[\rho_{\tilde{A}'} \otimes |m\rangle\langle m|_{\bar{A}}] = \mathrm{Tr}_{\tilde{A}'\bar{A}} \left[\left(\rho^T \otimes |m\rangle\langle m|^T \otimes \mathbb{I}_B \right) J_m \right] = \mathrm{Tr}_{\tilde{A}'} \left[\left(\rho^T \otimes \mathbb{I}_B \right) \langle m| J_m |m\rangle \right]. \tag{C.3.58}$$

If we define the operators $\tilde{J}_m$ as

$$\tilde{J}_m \coloneqq \langle m| J_m |m\rangle , \tag{C.3.59}$$

it is easy to verify that these are valid Choi states for a channel from $\tilde{A}'$ to B. Hence, we can reduce the dimensions of the Choi states and the signal states by optimizing over $\tilde{J}_m$ instead.