

Finite-Size Quantum Key Distribution Rates from Rényi Entropies Using Conic Optimization

Mariana Navarro^{1,2,*}, Andrés González Lorente^{3,4,†}, Pablo V. Parellada^{3,4},
 Carlos Pascual-García^{1,2} and Mateus Araújo^{3,4}

¹*Luxquantum Technologies S.L., Av. Joan Carles I, 30,
 1º1ª. 08908 L'Hospitalet de Llobregat, Barcelona, Spain*

²*ICFO—Institut de Ciències Fotoniques, The Barcelona Institute of Science and Technology,
 Castelldefels, Spain*

³*Departamento de Física Teórica, Atómica y Óptica, Universidad de Valladolid, 47011, Valladolid, Spain*

⁴*Laboratory for Disruptive Interdisciplinary Science (LaDIS),
 Universidad de Valladolid, 47011, Valladolid, Spain*



(Received 22 December 2025; accepted 24 June 2026; published 10 September 2026)

Finite-size general security proofs for quantum key distribution based on Rényi entropies have recently been introduced. These approaches are more flexible and provide tighter bounds on the secret key rate than traditional formulations based on the von Neumann entropy. However, deploying them requires minimizing the conditional Rényi entropy, a difficult optimization problem that has hitherto been tackled using *ad hoc* techniques based on the Frank-Wolfe algorithm, which are unstable and can only handle particular cases. In this work, we introduce a method based on non-symmetric conic optimization for solving this problem. Our technique is fast, reliable, and completely general. We illustrate its performance on several protocols and demonstrate a dramatic runtime improvement over the state of the art.

DOI: 10.1103/bf9s-m4jb

I. INTRODUCTION

Quantum key distribution (QKD) [1,2] allows two honest parties, traditionally named Alice and Bob, to create a shared key that remains secret against any quantum third parties thanks to the principles of quantum information theory. In particular, QKD is secure against adversaries with unlimited resources and the ability to perform cross-correlations between the different signals that compose the protocol—a scenario known as general security [3,4].

Recent works have successfully achieved general security for different protocols (for instance, see Refs. [5–7]) thanks to different versions of the so-called entropy accumulation theorem (EAT) [8,9]. The EAT framework reduces any multi-round QKD protocol into a single-round scenario, together with a finite-size analysis, which allows security to be proven for real QKD protocols that run for a finite number of rounds. Among its variants, the marginal-constrained EAT (MEAT) [10] constitutes the most up-to-date version. It relies on a formulation based on Rényi

entropies, removing impractical limitations from previous approaches [11] while providing tighter finite-size corrections compared to former results [12].

However, Rényi-based secret key rate computations require optimizing the conditional Rényi entropy, which is challenging to perform and is not covered by standard optimization software. Existing works [7,13] handle it with *ad hoc* methods based on the Frank-Wolfe algorithm [14,15]. Unfortunately, this approach suffers from low convergence rates and instability. Furthermore, the existing implementations skipped facial reduction [16,17], a critical step necessary to make the gradient well-defined, opting instead to add artificial depolarization noise, which reduces both precision and computational performance. Another approach would be to use existing non-symmetric conic methods for optimizing the sandwiched Rényi relative entropy [18]. However, this would also require adding artificial depolarization noise, because performing facial reduction makes the method no longer applicable. This is analogous to the case of von Neumann relative entropy, where facial reduction necessitated introducing a new cone dedicated to the QKD problem [19,20].

In this work we perform facial reduction on the sandwiched Rényi relative entropy cone and introduce a dedicated cone for optimizing over the resulting problem, named RényiQKD. We do the same for a lower bound introduced in Ref. [13], resulting in another cone, named FastRényiQKD. The latter cone gives very close bounds and is much faster, and is therefore the most appropriate

*Contact author: mariana.navarro@icfo.eu

†Contact author: andres.gonzalez.lorente@uva.es

‡These authors contributed equally to this work.

Published by the American Physical Society under the terms of the [Creative Commons Attribution 4.0 International license](#). Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

one for numerical computations. These cones can then be used for computing secret key rates using Rényi entropies within the MEAT framework combined with the Rényi leftover hashing lemma [21]. We provide numerical examples for several protocols: qubit BB84 [1,22], mutually unbiased bases (MUB) [23,24], and discrete-modulated continuous variable (DMCV) QKD [25,26]. Our implementation of the cones is developed in the programming language Julia [27] using the conic solver Hypatia [28,29], which uses an improved version of the non-symmetric optimization algorithm by Skajaa and Ye [30,31]. It is interfaced by the modeller JuMP [32], which makes it easy to use and mix with extra conic constraints that may come from specific QKD protocols, such as the partial state characterization constraints in Ref. [33].

The rest of this document goes as follows: in Sec. II we introduce diverse mathematical concepts that will be required along the manuscript, in Sec. III we provide a generic QKD protocol amenable for an analysis under the MEAT, as well as the main results to derive the secret key rate. Section IV shows the application of facial reduction and the construction of our cones. Section V presents our numerical results, and Sec. VI benchmarks our method against previous approaches. We conclude in Sec. VII.

II. NOTATION AND DEFINITIONS

Given a Hilbert space $\mathcal{H}_X$, we define $\mathcal{D}(X)$ as the set of all trace one, positive semidefinite matrices acting on this space, whereas $\mathbb{H}_{\geq}^d$ (respec. $\mathbb{H}_{>}^d$) will be used to denote Hermitian positive semidefinite (respec. definite) matrices of dimension d . For a register X , we denote n identical copies as $X^{(n)}$, and define the notation $[XY]_j := X_j Y_j$. Let us denote a state ρ_{XY} as classical-quantum (cq-state) when it can be expressed as

$$\rho_{XY} = \sum_{x \in \mathcal{X}} p(x) |x\rangle\langle x|_X \otimes \rho_Y^x, \quad (1)$$

for some countable alphabet $\mathcal{X}$, which may induce a space of probabilities denoted as $\mathbb{P}_{\mathcal{X}}$. Any subset $\Omega \subseteq \mathcal{X}$ can be used to define an event, which allows a restriction for cq-states according to a conditioning, as

$$\rho_{XY|\Omega} = \frac{1}{\Pr_{\rho}[\Omega]} \sum_{x \in \Omega} p(x) |x\rangle\langle x|_X \otimes \rho_Y^x, \quad (2)$$

with $\Pr_{\rho}[\Omega] = \sum_{x \in \Omega} p(x)$ representing the probability of observing the event. When only one case is contained in the event, i.e. $\Omega = \{x\}$ for some $x \in \mathcal{X}$, we will directly refer to the conditioned state as $\rho_{|x}$. We define the generalized trace distance as

$$T(\rho, \sigma) = \frac{1}{2} \|\rho - \sigma\|_1 + \frac{1}{2} |\text{Tr}[\rho - \sigma]|, \quad (3)$$

and $\|p\|_1$ to denote the trace norm of a vector p . We also make use of several entropic quantities; particularly the Rényi entropy, expressed as $H_{\alpha}(A)_{\rho}$ for $\alpha \in (0, \infty) \setminus \{1\}$, and its sandwiched conditional version, defined as

$$H_{\alpha}^{\uparrow}(A|B)_{\rho} = \sup_{\sigma_B \in \mathcal{D}(\mathcal{H}_B)} -D_{\alpha}(\rho_{AB} \| \mathbb{1} \otimes \sigma_B), \quad (4)$$

where we use the (unnormalized) sandwiched Rényi relative entropy

$$D_{\alpha}(\rho \| \sigma) = \frac{1}{\alpha - 1} \log [\Psi_{\alpha}(\rho, \sigma)], \quad (5)$$

where

$$\Psi_{\alpha}(\rho, \sigma) = \text{Tr} \left[\left(\sigma^{\frac{1-\alpha}{2\alpha}} \rho \sigma^{\frac{1-\alpha}{2\alpha}} \right)^{\alpha} \right], \quad (6)$$

provided that $\alpha \in (0, \infty) \setminus \{1\}$ and $\text{supp}(\rho) \subseteq \text{supp}(\sigma)$. The case $\alpha \rightarrow 1$ is well-defined and leads to the usual quantum relative entropy

$$D(\rho \| \sigma) = \text{Tr}[\rho \log(\rho) - \rho \log(\sigma)], \quad (7)$$

which we will denote as the Kullback-Leibler divergence $D_{\text{KL}}(\cdot \| \cdot)$ when both states are classical. To conclude, we caution the reader that the numerical optimization methods used in this manuscript typically employ natural logarithms, while the calculations for the key rates make use of the base 2 logarithm.

III. PROTOCOL OUTLINE

Let us follow [13] and provide a brief, generic description of a prepare-and-measure QKD protocol that is valid under the MEAT [10].

Preparation and measurement. For each round $j \in \{1, \dots, n\}$, Alice prepares a state $|\psi^x\rangle_{A'}$ according to the value x of a random variable X_j , and sends it to Bob. Bob receives the state, applies one (of possibly many) POVM, and records the outcome in a register Y_j . Then, Alice (possibly also Bob) designates the round as either a key generation or parameter estimation round, with probabilities p^K and $1 - p^K$, respectively. Additionally, they build a reference key register R_j for the key generation.

Public announcement. Alice and Bob publicly disclose information about the registers $[XY]^{(n)}$, with all such communication stored in a register $I^{(n)}$. They use this register to construct $C^{(n)}$, which contains all the public classical information associated to parameter estimation rounds.

Sifting. Using $[RI]^{(n)}$, Alice builds her key register $S^{(n)}$ [34]. Optionally, Alice may also create a register $C'^{(n)}$ with private information for statistical checks [35].

Parameter estimation. Alice and Bob use $C^{(n)}$ (optionally, Alice uses $C'^{(n)}$) to perform parameter estimation in order to bound Eve's information. If the process reveals that Eve holds excessive information, they abort.

Information reconciliation. Alice and Bob use an error correction subroutine, which might follow a reverse reconciliation, to remove the discrepancies between their private keys. Alice and Bob then certify that their keys are the same using a universal₂ hash function. If their hashed keys disagree, they abort.

Privacy amplification. Alice and Bob distill a shorter, fully private shared key recorded as K_A and K_B respectively, by using another universal₂ hash function.

For future reference, we define $\mathcal{C}$ as the alphabet for the registers $C^{(n)}$. To specify the values observed by Alice and Bob during parameter estimation rounds, let us also define $\tilde{\mathcal{C}} := \mathcal{C} \setminus \{\perp\}$. On similar grounds, we denote Alice's key registers $S^{(n)}$ as defined on an alphabet $\mathcal{S}$, where $S_j = \perp$ indicates a non-key round, and $\tilde{\mathcal{S}} := \mathcal{S} \setminus \{\perp\}$ for non-trivial key values. Finally, we assume that, within a given session of the protocol, no public announcements are needed before all measurements are performed, which allows an simplified analysis via Ref. [10], Corollary 4.2]. Furthermore, this approach allows an interpretation of the protocol in terms of an entanglement-based picture [36,37], where Alice prepares an entangled state $|\psi\rangle_{AA'}$ in every round.

A. Security and completeness

The security of QKD protocols is customarily defined via the ε -security framework by Renner [3], which establishes that Alice's and Bob's final key will be ε_s -secure if their shared key is ε_s -close to being completely decoupled from Eve's total information. Quantitatively, this is expressed according to the formula

$$\Pr_\rho[\Omega]T(\rho_{K_A K_B E|\Omega}, \bar{\tau}_{K_A K_B E|\Omega}) \leq \varepsilon_s, \quad (8)$$

for a given $\varepsilon_s > 0$. Here, E represents Eve's information, and $\rho_{K_A K_B E}$ is the final state shared by the three parties after the execution of the protocol. The set Ω represents the event of not aborting at any step of the protocol, and $\bar{\tau}_{K_A K_B E}$ is the ideal output state, where the keys share ideal local randomness and no correlations with respect to any other registers. Formally, we define

$$\bar{\tau}_{K_A K_B E|\Omega} = \tau_{K_A K_B} \otimes \rho_{E|\Omega} = \frac{1}{d} \sum_{x=0}^{d-1} |x, x\rangle\langle x, x|_{K_A K_B} \otimes \rho_{E|\Omega}. \quad (9)$$

Condition (8) can be simplified according to a triangle inequality into the statements of correctness and secrecy, such that we may define an ε_{EC} -correct implementation according to

$$\Pr[(K_A \neq K_B) \wedge \Omega] \leq \varepsilon_{\text{EC}}. \quad (10)$$

In particular, correctness can be enforced by Alice sending her key to Bob after applying a universal₂ hash function once error correction has finished, such that the total length of the hashed key does not exceed $\lceil \log(1/\varepsilon_{\text{EC}}) \rceil$ bits (see [38] for a complete derivation).

Secrecy is induced by the fact that Eve does not hold excessive information about one of the keys. Namely, a protocol is ε_{PA} -secret whenever

$$\Pr_\rho[\Omega]T(\rho_{K_A E|\Omega}, \bar{\tau}_{K_A E|\Omega}) \leq \varepsilon_{\text{PA}}. \quad (11)$$

We defer the discussion about the security of our approach to Appendix B, which will be based on a proof against general attacks via the MEAT [10].

Although the above definitions indeed certify that the final key remains unknown against any quantum adversary, it is still necessary to ensure that an execution will not frequently abort. This is given by the notion of completeness, which certifies that any honest implementation of the protocol (i.e., when Eve performs no attack) will be successful with very high probability. Formally, a protocol is ε_{cm} -complete whenever

$$1 - \Pr^h[\Omega] \leq \varepsilon_{\text{cm}}, \quad (12)$$

where the superscript $(\cdot)^h$ denotes that the implementation is honest. In general, completeness must be enforced for both parameter estimation and error correction. For the former, this is done via $C^{(n)}$ and a reference probability distribution $p \in \mathbb{P}_{\mathcal{C}}$ on the alphabet $\mathcal{C}$, which constitutes the vector of probabilities expected by Alice and Bob in such an honest implementation. We define $\hat{\Sigma}_\Omega$ as the set of all accepted frequency estimators according to $\{p(c)\}_{c \in \mathcal{C}}$, and a scalar $\delta > 0$. Explicitly,

$$\hat{\Sigma}_\Omega = \{\text{freq}_{c^{(n)}} \in \mathbb{P}_{\mathcal{C}} : \|\text{freq}_{c^{(n)}} - p\|_1 \leq \delta\}. \quad (13)$$

During parameter estimation, Alice and Bob will verify that their observed finite distribution $\text{freq}_{c^{(n)}}$ is actually contained in $\hat{\Sigma}_\Omega$, or abort otherwise. This analysis is carried out according to the margin of confidence for the statistical estimators; given an honest implementation (such that we may assume that the rounds are iid), there exists a probability $\bar{\varepsilon}_{\text{PE}} > 0$ such that

$$\Pr^h[\|\text{freq}_{c^{(n)}} - p\|_1 \leq \delta] \geq 1 - \bar{\varepsilon}_{\text{PE}}. \quad (14)$$

In order to bound this probability quantitatively, we make use of the following concentration bound.

Proposition III.1. (Bretagnolle-Huber-Carol inequality). [39,40] Let p be a multinomial distribution defined on an alphabet $\mathcal{X}$. After $n \in \mathbb{N}$ iid trials, the empirical

frequency distribution $\hat{p}$ verifies

$$\Pr[\|\hat{p} - p\|_1 \leq \delta] \geq 1 - 2^{|\mathcal{X}|} e^{-n\delta^2/2}. \quad (15)$$

Provided that this proposition holds in an implementation, and identifying the right-hand side with the one of (14), the step of parameter estimation will be successful. Hence, by splitting the nonabortion event as $\Omega = \Omega_{\text{PE}} \cap \Omega_{\text{EC}}$ (i.e. success at both parameter estimation and error correction), we can certify via the union bound that the protocol is $(\bar{\epsilon}_{\text{PE}} + \bar{\epsilon}_{\text{EC}})$ -complete. This conclusion assumes the observation of an $\bar{\epsilon}_{\text{EC}}$ -complete information reconciliation [41], which we further develop in Sec. III B.

B. Finite-size secret key

Using the definitions of Secs. II and III, we characterize the final secret key achievable via the following result.

Theorem III.2. Let $n \in \mathbb{N}$, $\sigma_A \in \mathcal{D}(A)$ and $\tilde{\mathcal{M}}: [AB]_j \rightarrow [SCI]_j$ a quantum channel concatenated n times to form the protocol. Let further C a classical register defined on an alphabet $\mathcal{C}$ and Ω the nonabortion event for the protocol. Then, for a choice of coefficients $\epsilon_{\text{EC}}, \epsilon_{\text{PA}}, \bar{\epsilon}_{\text{PE}}, \bar{\epsilon}_{\text{EC}} > 0$, the protocol is ϵ_{EC} -correct, ϵ_{PA} -secret and $(\bar{\epsilon}_{\text{PE}} + \bar{\epsilon}_{\text{EC}})$ -complete providing a binary key whose length verifies

$$\ell \leq nh_\alpha^\uparrow - \frac{\alpha}{\alpha-1} \log\left(\frac{1}{\epsilon_{\text{PA}}}\right) - \text{leak}_{\text{EC}} + 2, \quad (16)$$

where $\alpha \in (1, 2)$; leak_{EC} is the information lost to Eve during an ϵ_{EC} -correct and $\bar{\epsilon}_{\text{EC}}$ -complete error correction, and

$$h_\alpha^\uparrow \geq \inf_{q \in \tilde{\Sigma}_\Omega} \inf_{\omega \in \mathcal{D}(AB)} \inf_{\psi \in \mathcal{D}(RAB)} \frac{\alpha}{\alpha-1} D_{\text{KL}}(q \| \tilde{\mathcal{M}}(\omega)_C) + q(\perp) D_\gamma(\mathcal{G}(\omega) \| \mathcal{Z}(\psi)) \quad (17)$$

where $\gamma = \alpha/(2\alpha-1) \in (2/3, 1)$, $\text{Tr}_B[\omega_{AB}] = \sigma_A$, and $\tilde{\Sigma}_\Omega$ is the set defined in (13) given a reference distribution $p \in \mathbb{P}_\mathcal{C}$ and $\delta \geq 0$ verifying $\bar{\epsilon}_{\text{PE}} \geq 2^{|\mathcal{C}|} e^{-n\delta^2/2}$. The maps $\mathcal{G}: AB \rightarrow RAB$ and $\mathcal{Z}: RAB \rightarrow RAB$ denote the CPTP coherent and pinching maps involved in a single-round key distillation.

Additionally, we provide the following lower bound, which greatly simplifies the computation of the secret key rate at the cost of a slightly suboptimal yield.

Corollary III.3. Given the conditions stated in Theorem III.2, the bound $h_\alpha^\uparrow \geq h_\alpha^\downarrow$ holds at (16), where

$$h_\alpha^\downarrow = \inf_{q \in \tilde{\Sigma}_\Omega} \inf_{\omega \in \mathcal{D}(AB)} \frac{\alpha}{\alpha-1} D_{\text{KL}}(q \| \tilde{\mathcal{M}}(\omega)_C) + q(\perp) D_\beta(\mathcal{G}(\omega) \| \mathcal{Z} \circ \mathcal{G}(\omega)), \quad (18)$$

with $\beta = 1/\alpha \in (1/2, 1)$.

We defer the security proof of Theorem III.2 to Appendix B.

C. Error correction cost

To conclude the theoretical framework, we define the information leaked during error correction. Albeit this step is carried through an authenticated classical channel, there is no universal, information-theoretic bound on the length of leak_{EC} . Some implementations can indeed characterize said length [3,42], although they require algorithms that are challenging to apply.

As an efficient alternative, Alice and Bob can impose a limit to the total information exchanged during error correction before executing the protocol and verify the success of the implementation by using a universal₂ hash function. Taking this approach, leak_{EC} can be reasonably estimated via the Slepian-Wolf bound [43], defining an ϵ_{EC} -correct and ϵ_{EC} -complete implementation as

$$\text{leak}_{\text{EC}} \leq p^s n f H(S|Y) + \left\lceil \log\left(\frac{1}{\epsilon_{\text{EC}}}\right) \right\rceil, \quad (19)$$

where p^s denotes the sifting probability, i.e., the fraction of key rounds that are not lost due to no-clicks and/or postselection, and thus require to be error-corrected. The conditional Shannon entropy $H(S|Y)$ is provided by Alice's key register conditioned on Bob's (after removing the $\{\perp\}$ symbol from said registers), which can be analytically calculated from their measurements. Furthermore, $\lceil \log(1/\epsilon_{\text{EC}}) \rceil$ denotes the cost in bits of using the verification hash, and $f \geq 1$ represents the error correction efficiency, where $f = 1$ constitutes the ideal case (known as the Shannon limit).

D. Non-trace-preserving maps

Although in Theorem III.2 the coherent map $\mathcal{G}$ is assumed to be trace-preserving, we often need to model non-detections or postselection, which leads to trace-non-increasing maps. To address this, we extend such maps into trace-preserving ones by adding a symbol $\{\perp\}$ corresponding to non-detection or discarded events, similar to Ref. [7], Corollary 9].

Let $\mathcal{G}': AB \rightarrow \tilde{RAB}$ be the original map with register $\tilde{R}$ defined on an alphabet $\tilde{\mathcal{R}}$ as embedded on an extended register R on $\mathcal{R} = \{\perp\} \cup \tilde{\mathcal{R}}$. We then consider its trivial extension $\mathcal{G}_\top: AB \rightarrow RAB$, given explicitly by $\mathcal{G}_\top(\omega) = (V \otimes \mathbb{1}_{AB}) \mathcal{G}'(\omega) (V^\dagger \otimes \mathbb{1}_{AB})$ where $V = \sum_{r \in \tilde{\mathcal{R}}} |r\rangle_R \langle r|_{\tilde{R}}$ is an isometry. Analogously, for the map $\mathcal{Z}': \tilde{RAB} \rightarrow \tilde{RAB}$ we define its extension $\mathcal{Z}_\top: RAB \rightarrow RAB$. Our goal now is to construct additional maps $\mathcal{G}_\perp$ and $\mathcal{Z}_\perp$ such that $\mathcal{G} := \mathcal{G}_\top + \mathcal{G}_\perp$ and $\mathcal{Z} := \mathcal{Z}_\top + \mathcal{Z}_\perp$ are trace-preserving and satisfy the orthogonality relations

$$\begin{aligned}
\mathcal{G}_\top(\omega)\mathcal{G}_\perp(\omega') &= \mathcal{Z}_\top(\psi)\mathcal{Z}_\perp(\psi') \\
&= \mathcal{G}_\top(\omega)\mathcal{Z}_\perp(\psi') \\
&= \mathcal{Z}_\top(\psi)\mathcal{G}_\perp(\omega') = 0.
\end{aligned} \tag{20}$$

This construction is not unique, and a generic choice is

$$\begin{aligned}
\mathcal{G}_\perp(\omega) &:= [\text{Tr}(\omega) - \text{Tr}(\mathcal{G}_\top(\omega))]|_\perp\rangle\langle_\perp|_R \otimes \frac{1}{d_{AB}}\mathbb{1}_{AB}, \\
\mathcal{Z}_\perp(\psi) &:= (|\perp\rangle\langle_\perp|_R \otimes \mathbb{1}_{AB})\psi(|\perp\rangle\langle_\perp|_R \otimes \mathbb{1}_{AB}).
\end{aligned} \tag{21}$$

Using the orthogonality relations in (20) on these maps, the expression $\Psi_\gamma(\mathcal{G}(\omega), \mathcal{Z}(\psi))$ used in (17) decomposes as

$$\begin{aligned}
\Psi_\gamma(\mathcal{G}(\omega), \mathcal{Z}(\psi)) &= \Psi_\gamma(\mathcal{G}_\top(\omega), \mathcal{Z}_\top(\psi)) \\
&\quad + \Psi_\gamma(\mathcal{G}_\perp(\omega), \mathcal{Z}_\perp(\psi)),
\end{aligned} \tag{22}$$

such that we can split the optimization into two smaller cones. In addition, the fact that $\mathcal{Z}_\perp \circ \mathcal{G}_\perp(\omega) = \mathcal{G}_\perp(\omega)$ allows a further simplification for the expression $\Psi_\beta(\mathcal{G}(\omega), \mathcal{Z} \circ \mathcal{G}(\omega))$ used in (18):

$$\begin{aligned}
\Psi_\beta(\mathcal{G}(\omega), \mathcal{Z} \circ \mathcal{G}(\omega)) &= \Psi_\beta(\mathcal{G}_\top(\omega), \mathcal{Z}_\top \circ \mathcal{G}_\top(\omega)) \\
&\quad + \text{Tr}(\omega) - \text{Tr}(\mathcal{G}_\top(\omega)),
\end{aligned} \tag{23}$$

which implies that for the lower bound $h_\alpha^\downarrow$, we never need to construct $\mathcal{G}_\perp$ and $\mathcal{Z}_\perp$ explicitly.

IV. COMPUTING THE KEY RATE

The problem of computing the secret key rate thus reduces to solving the optimization problem in (17) (or (18)). It consists of a Kullback-Leibler divergence and a Rényi relative entropy. The Kullback-Leibler divergence is straightforward to optimize: it is a jointly convex function that can be tackled by standard convex optimization techniques. Moreover, it can be handled by conic solvers that support non-symmetric cones, such as Hypatia [28] and QICS [44], which makes the optimization particularly effective and convenient.

The challenging term is the sandwiched Rényi relative entropy. It is also jointly convex, and thus in principle it can also be tackled by standard convex optimization techniques, such as the Frank-Wolfe algorithm [14]. However, its arguments $\mathcal{G}(\omega), \mathcal{Z}(\psi)$ often have eigenvalues that are constrained to be equal to zero. This makes the derivatives undefined, which causes the optimization to become unstable and fail. Indeed, previous attempts suffered from numerical instability and could only handle particular cases [7,13]. Technically, we say that the problem is not strictly feasible, and therefore Slater's condition is not satisfied. There is a standard technique to handle this, consisting of reformulating the problem in terms of the support of the matrices, known as facial

reduction [16]. It has been successfully applied in the case of von Neumann entropies [17,19].

A. Facial reduction

Before going further, note that the Rényi relative entropy is a composition of the function Ψ_α with the logarithm. Since the logarithm is a monotonic function, it is enough to optimize over Ψ_α , and we will do so for simplicity.

What we need is to reformulate the function $\Psi_\alpha(\mathcal{G}(\omega), \mathcal{Z}(\psi))$ in terms of the supports of $\omega \in \mathbb{H}_{\geq}^p$, $\psi \in \mathbb{H}_{\geq}^m$, $\mathcal{G}(\omega) \in \mathbb{H}_{\geq}^m$, and $\mathcal{Z}(\psi) \in \mathbb{H}_{\geq}^m$. First of all, let us consider the support of ω . In general, there are several constraints acting on it, which might constrain some of its eigenvalues to be zero. We write then $\omega = V\omega'V^\dagger$ for some isometry $V \in \mathbb{C}^{p \times q}$ such that ω' is not constrained to be singular. On the other hand, the only constraint acting on ψ is that it has trace one, so it can already be full rank; $\psi \in \mathbb{H}_{\geq}^m$, and no reduction is needed. Let us define then the strictly positive maps $\hat{\mathcal{G}}, \hat{\mathcal{Z}}$ such that

$$\mathcal{G}(V\omega'V^\dagger) = W_{\hat{\mathcal{G}}} \hat{\mathcal{G}}(\omega') W_{\hat{\mathcal{G}}}^\dagger, \tag{24a}$$

$$\mathcal{Z}(\psi) = W_{\hat{\mathcal{Z}}} \hat{\mathcal{Z}}(\psi) W_{\hat{\mathcal{Z}}}^\dagger, \tag{24b}$$

for some isometries $W_{\hat{\mathcal{G}}} \in \mathbb{C}^{m \times k}$, $W_{\hat{\mathcal{Z}}} \in \mathbb{C}^{m \times n}$. Consequently, we have that

$$\Psi_\alpha(\mathcal{G}(\omega), \mathcal{Z}(\psi)) = \Psi_\alpha(W_{\hat{\mathcal{G}}} \hat{\mathcal{G}}(\omega') W_{\hat{\mathcal{G}}}^\dagger, W_{\hat{\mathcal{Z}}} \hat{\mathcal{Z}}(\psi) W_{\hat{\mathcal{Z}}}^\dagger). \tag{25}$$

In order to put this in a convenient form, we will use the identities

$$\begin{aligned}
\Psi_\alpha(\rho, \sigma) &= \text{Tr} \left[(\sigma^{\frac{1-\alpha}{2\alpha}} \rho \sigma^{\frac{1-\alpha}{2\alpha}})^\alpha \right] \\
&= \|\sigma^{\frac{1-\alpha}{2\alpha}} \rho^{\frac{1}{2}}\|_{2\alpha}^{2\alpha} \\
&= \text{Tr} \left[(\rho^{\frac{1}{2}} \sigma^{\frac{1-\alpha}{\alpha}} \rho^{\frac{1}{2}})^\alpha \right].
\end{aligned} \tag{26}$$

which then imply that

$$\begin{aligned}
\Psi_\alpha(W_{\hat{\mathcal{G}}} \hat{\mathcal{G}}(\omega') W_{\hat{\mathcal{G}}}^\dagger, W_{\hat{\mathcal{Z}}} \hat{\mathcal{Z}}(\psi) W_{\hat{\mathcal{Z}}}^\dagger) \\
= \|\hat{\mathcal{Z}}(\psi)^{\frac{1-\alpha}{2\alpha}} W_{\hat{\mathcal{Z}}}^\dagger W_{\hat{\mathcal{G}}} \hat{\mathcal{G}}(\omega')^{\frac{1}{2}}\|_{2\alpha}^{2\alpha}
\end{aligned} \tag{27}$$

Letting then $S = W_{\hat{\mathcal{Z}}}^\dagger W_{\hat{\mathcal{G}}}$, which as we prove in Appendix D is an isometry, we define

$$\hat{\Psi}_\alpha(\omega', \psi) := \|\hat{\mathcal{Z}}(\psi)^{\frac{1-\alpha}{2\alpha}} S \hat{\mathcal{G}}(\omega')^{\frac{1}{2}}\|_{2\alpha}^{2\alpha}, \tag{28}$$

and the facial reduction of the sandwiched Rényi relative entropy is done.

It is also necessary to do the facial reduction of $\Psi_\alpha(\mathcal{G}(\omega), \mathcal{Z} \circ \mathcal{G}(\omega))$. It proceeds along the same lines, with the difference that we define the strictly positive map $\widehat{\mathcal{ZG}}$ such that

$$\mathcal{Z} \circ \mathcal{G}(V\omega'V^\dagger) = W_{\mathcal{ZG}} \widehat{\mathcal{ZG}}(\omega') W_{\mathcal{ZG}}^\dagger, \quad (29)$$

for some isometry $W_{\mathcal{ZG}} \in \mathbb{C}^{m \times r}$, we define $S = W_{\mathcal{ZG}}^\dagger W_{\mathcal{G}}$, and end up with

$$\hat{\Psi}_\alpha(\omega') := \|\widehat{\mathcal{ZG}}(\omega')^{\frac{1-\alpha}{2\alpha}} S \hat{\mathcal{G}}(\omega')^{\frac{1}{2}}\|_{2\alpha}^{2\alpha}. \quad (30)$$

One can always find adequate isometries $V, W_{\mathcal{G}}, W_{\mathcal{Z}}$, and $W_{\mathcal{ZG}}$ numerically via the spectral decomposition, as shown in Ref. [17]. The numerical solution inevitably has numerical noise, however, which might negatively affect the performance of the solver. It is frequently the case, though, that the structure of the maps is simple enough that one can find suitable isometries analytically. Indeed we do so in all the examples shown in Sec. V.

B. The Rényi QKD cones

The expressions (28) and (30) are suitable for convex optimization. They can, for example, be used to implement a Frank-Wolfe algorithm with facial reduction. However, our goal here is to implement a conic method, and for that we need to represent these expressions via convex cones:

$$\mathcal{K}_{\text{RényiQKD}}^{\alpha, \hat{\mathcal{G}}, \hat{\mathcal{Z}}, S} = \text{cl}\{(u, \rho, \sigma) \in \mathbb{R} \times \mathbb{H}_{\geq}^q \times \mathbb{H}_{\geq}^m; u \geq s_\alpha \hat{\Psi}_\alpha(\rho, \sigma)\}, \quad (31)$$

suitable for optimizing expression (17), and

$$\mathcal{K}_{\text{FastRényiQKD}}^{\alpha, \hat{\mathcal{G}}, \widehat{\mathcal{ZG}}, S} = \text{cl}\{(u, \rho) \in \mathbb{R} \times \mathbb{H}_{\geq}^q; u \geq s_\alpha \hat{\Psi}_\alpha(\rho)\}, \quad (32)$$

suitable for optimizing the lower bound (18). The latter we denominate fast cone as it has lower dimension and is thus considerably faster to optimize over. In both expressions $s_\alpha = \text{sign}(\alpha - 1)$ is a convenience factor to make the cones convex for all $\alpha \geq 1/2$ [45]. Note, however, that for computing key rates the relevant ranges are $(2/3, 1)$ and $(1/2, 1)$, respectively, as given by Theorem III.2 and Corollary III.3.

In order to implement optimization over these cones, the most crucial part is to find logarithmically homogeneous self-concordant barriers for them (see Definition E.1). We use the natural barriers

$$f(u, \rho, \sigma) = -\log(u - s_\alpha \hat{\Psi}_\alpha(\rho, \sigma)) - \log \det(\rho) - \log \det(\sigma), \quad (33)$$

$$f(u, \rho) = -\log(u - s_\alpha \hat{\Psi}_\alpha(\rho)) - \log \det(\rho), \quad (34)$$

which can be proven to be self-concordant for $\alpha \in [1/2, 1)$ following the argument of Ref. [18], Theorem 1.2], where it is shown that the natural barrier of the quasi-Rényi cone

$$\mathcal{K}_{\text{quasi-Rényi}}^\alpha = \text{cl}\{(u, \rho, \sigma) \in \mathbb{R} \times \mathbb{H}_{\geq}^n \times \mathbb{H}_{\geq}^n; u \geq s_\alpha \Psi_\alpha(\rho, \sigma)\} \quad (35)$$

is self-concordant [46]. The key difference in our proof is that we need to do a composition with the strictly positive map $X \mapsto S^\dagger X S$ when using Ref. [47], Theorem 2.1]. Note that this map is strictly positive because S is an isometry. The complete proof is shown in Appendix E.

It is also necessary to compute the gradient, Hessian, and third-order derivatives of the barriers to run the conic solver. This computation also follows along the lines of Ref. [18], except that we need to find an analogue of the identities (26) where all the operators involved are full rank. It is given by

$$\begin{aligned} \hat{\Psi}_\alpha(\rho, \sigma) &= \text{Tr}[(S^\dagger \hat{\mathcal{Z}}(\sigma)^{\frac{1-\alpha}{2\alpha}} S)^{\frac{1}{2}} \hat{\mathcal{G}}(\rho) (S^\dagger \hat{\mathcal{Z}}(\sigma)^{\frac{1-\alpha}{2\alpha}} S)^{\frac{1}{2}}]^\alpha \\ &= \|\hat{\mathcal{Z}}(\sigma)^{\frac{1-\alpha}{2\alpha}} S \hat{\mathcal{G}}(\rho)^{\frac{1}{2}}\|_{2\alpha}^{2\alpha} \\ &= \text{Tr}[(\hat{\mathcal{G}}(\rho)^{\frac{1}{2}} S^\dagger \hat{\mathcal{Z}}(\sigma)^{\frac{1-\alpha}{2\alpha}} S \hat{\mathcal{G}}(\rho)^{\frac{1}{2}})^\alpha]. \end{aligned} \quad (36)$$

Note that, in the first line, the right-hand side simplifies to $\text{Tr}[(\hat{\mathcal{Z}}(\sigma)^{\frac{1-\alpha}{2\alpha}} S \hat{\mathcal{G}}(\rho) S^\dagger \hat{\mathcal{Z}}(\sigma)^{\frac{1-\alpha}{2\alpha}})^\alpha]$, but the simpler expression is not usable as the matrix inside the trace has null eigenvalues that make the derivatives ill-defined. The complete derivation is shown in Appendix F. Finally, we need to compute a starting point for the optimization of each cone, which is shown in Appendix G.

We implemented these cones in the programming language Julia [27], as an extension of the solver Hypatia [28]. It is interfaced with the modeller JuMP [32] and uses a generic type system, allowing the computation of key rates with precision greater than 64 bits when desired. The code is part of the ConicQKD.jl package [48].

C. Explicit conic formulation

We are now in a position to explicitly reformulate the optimization problems (17) and (18) as conic programs, suitable for our conic solver. First of all, let us rewrite (17) with all its constraints explicitly

$$\begin{aligned} h_\alpha^\dagger &\geq \min_{q, \omega, \psi} \frac{\alpha}{\alpha - 1} D_{\text{KL}}(q \| \tilde{\mathcal{M}}(\omega)_c) + q(\perp) D_\gamma(\mathcal{G}(\omega) \| \mathcal{Z}(\psi)) \\ \text{s.t. } &\omega_{AB}, \psi_{RAB} \geq 0, \\ &\text{Tr}_B[\omega_{AB}] = \sigma_A, \quad \text{Tr}[\psi_{RAB}] = 1, \\ &\sum_{c \in \mathcal{C}} q(c) = 1, \quad q \geq 0, \\ &\|p - q\|_1 \leq \delta. \end{aligned} \quad (37)$$

Here, we point out that the Kullback-Leibler divergence and the 1-norm can be expressed via the well-known cones [28]

$$\begin{aligned}\mathcal{K}_{\text{KL}} &= \text{cl}\{(u, q, p) \in \mathbb{R} \times \mathbb{R}_{\geq}^d \times \mathbb{R}_{\geq}^d; u \geq D_{\text{KL}}(q \| p)\}, \\ \mathcal{K}_1 &= \{(u, w) \in \mathbb{R} \times \mathbb{R}^d; u \geq \|w\|_1\}.\end{aligned}\quad (38)$$

In conjunction with the RényiQKD cone (31), this allows us to rewrite (37) as

$$\begin{aligned}h_a^\uparrow &\geq \min_{h_{\text{KL}}, u, q, \omega, \psi} \frac{\alpha}{\alpha-1} h_{\text{KL}} + \frac{q(\perp)}{\gamma-1} \log(s_\gamma u) \\ \text{s.t. } \text{Tr}_B[\omega_{AB}] &= \sigma_A, \quad \text{Tr}[\psi_{RAB}] = 1, \quad \sum_{c \in \mathcal{C}} q(c) = 1 \\ (u, \omega_{AB}, \psi_{RAB}) &\in \mathcal{K}_{\text{RényiQKD}}^{\gamma, \hat{G}, \hat{Z}, S} \\ (h_{\text{KL}}, q, \tilde{\mathcal{M}}(\omega_{AB})_C) &\in \mathcal{K}_{\text{KL}} \\ (\delta, p - q) &\in \mathcal{K}^1.\end{aligned}\quad (39)$$

Note that the factor $s_\gamma = -1$ is necessary to cancel out the factor s_γ inside the definition of the RényiQKD cone, and positivity constraints are implicitly enforced by the cones.

We are still not done because the objective function must be affine. The logarithm term can be handled as before by using the logarithm cone [49]

$$\mathcal{K}_{\log} = \text{cl}\left\{(u, v, w) \in \mathbb{R} \times \mathbb{R}_{\geq} \times \mathbb{R}_{\geq}; u \leq v \log\left(\frac{w}{v}\right)\right\}, \quad (40)$$

but the term $q(\perp)$ multiplying the logarithm requires more care. One possibility is to do a nested optimization, where the inner is a conic optimization with fixed $q(\perp)$, and the outer optimizes over the single parameter $q(\perp)$ via a non-convex method. Another, more practical option is to use the bound on the probabilities at the acceptance set

$$p(\perp) - q(\perp) \leq \|p - q\|_1 \leq \delta, \quad (41)$$

to lower bound $q(\perp)$ at the objective function, resulting in the relaxed problem

$$\begin{aligned}h_a^\uparrow &\geq \min_{h_{\text{KL}}, h_{\text{QKD}}, u, q, \omega, \psi} \frac{\alpha}{\alpha-1} h_{\text{KL}} + \frac{p(\perp) - \delta}{\gamma-1} h_{\text{QKD}} \\ \text{s.t. } \text{Tr}_B[\omega_{AB}] &= \sigma_A, \quad \text{Tr}[\psi_{RAB}] = 1, \quad \sum_{c \in \mathcal{C}} q(c) = 1 \\ (u, \omega_{AB}, \psi_{RAB}) &\in \mathcal{K}_{\text{RényiQKD}}^{\gamma, \hat{G}, \hat{Z}, S} \\ (h_{\text{KL}}, q, \tilde{\mathcal{M}}(\omega_{AB})_C) &\in \mathcal{K}_{\text{KL}} \\ (h_{\text{QKD}}, 1, s_\gamma u) &\in \mathcal{K}_{\log} \\ (\delta, p - q) &\in \mathcal{K}^1.\end{aligned}\quad (42)$$

Numerical experiments show that the effect of this relaxation on the key rate is small but noticeable for small block sizes. A more elegant solution can be applied to the FastRényiQKD cone (32), following the argument from [7], Appendix F]. Let $q = (q_{\tilde{\mathcal{C}}}, q(\perp))$ and $\tilde{\mathcal{M}}(\omega)_C = (\tilde{\mathcal{M}}(\omega)_{\tilde{\mathcal{C}}}, p(\perp))$ where $\tilde{\mathcal{C}}$ is defined on the alphabet $\tilde{\mathcal{C}}$, so that

$$\begin{aligned}D_{\text{KL}}(q \| \tilde{\mathcal{M}}(\omega)_C) &= D_{\text{KL}}(q_{\tilde{\mathcal{C}}} \| \tilde{\mathcal{M}}(\omega)_{\tilde{\mathcal{C}}}) \\ &\quad + q(\perp) \log\left(\frac{q(\perp)}{p(\perp)}\right).\end{aligned}\quad (43)$$

Then write $\hat{\Psi}_\beta = \frac{p(\perp)\hat{\Psi}_\beta}{q(\perp)p(\perp)}$, so that

$$\begin{aligned}q(\perp) \log(\hat{\Psi}_\beta) &= q(\perp) \log\left(\frac{p(\perp)\hat{\Psi}_\beta}{q(\perp)}\right) \\ &\quad + q(\perp) \log_2\left(\frac{q(\perp)}{p(\perp)}\right)\end{aligned}\quad (44)$$

Then the expression in (18) can be rewritten as

$$\begin{aligned}&\frac{\alpha}{\alpha-1} D_{\text{KL}}(q \| \tilde{\mathcal{M}}(\omega)_C) + \frac{q(\perp)}{\beta-1} \log_2 \hat{\Psi}_\beta(\omega) \\ &= \frac{\alpha}{\alpha-1} \left(D_{\text{KL}}(q_{\tilde{\mathcal{C}}} \| \tilde{\mathcal{M}}(\omega)_{\tilde{\mathcal{C}}}) - q(\perp) \log\left(\frac{p(\perp)\hat{\Psi}_\beta}{q(\perp)}\right) \right)\end{aligned}\quad (45)$$

and using the same cones as before gives us the exact conic formulation:

$$\begin{aligned}h_a^\downarrow &\geq \min_{h_{\text{QKD}}, h_{\text{KL}}, u, q, \omega} \frac{\alpha}{\alpha-1} (h_{\text{KL}} - h_{\text{QKD}}) \\ \text{s.t. } \text{Tr}_B[\omega_{AB}] &= \sigma_A, \quad \sum_{c \in \mathcal{C}} q(c) = 1 \\ (u, \omega_{AB}) &\in \mathcal{K}_{\text{FastRényiQKD}}^{\beta, \hat{G}, \hat{Z}, S} \\ (h_{\text{KL}}, q_{\tilde{\mathcal{C}}}, \tilde{\mathcal{M}}(\omega)_{\tilde{\mathcal{C}}}) &\in \mathcal{K}_{\text{KL}} \\ (h_{\text{QKD}}, q(\perp), p(\perp)s_\beta u) &\in \mathcal{K}_{\log} \\ (\delta, p - q) &\in \mathcal{K}^1.\end{aligned}\quad (46)$$

V. EXAMPLES

In this section, we define diverse QKD protocols, the necessary tools to adapt them into the formalism of the MEAT, and compute key rates for them using our technique. We exclusively use the FastRényiQKD cone (32), because it is significantly faster and more stable than the RényiQKD cone (31), and as illustrated in Fig. 4, the key rates they produce are essentially indistinguishable.

Unless explicitly stated otherwise, we assume the following values for the tolerance coefficients

$$\varepsilon_{\text{EC}} = 10^{-11}, \quad (47a)$$

$$\varepsilon_{\text{PA}} = 9 \times 10^{-11}, \quad (47b)$$

$$\bar{\varepsilon}_{\text{PE}} = 9 \times 10^{-11}, \quad (47c)$$

and set $\bar{\varepsilon}_{\text{PE}} = 2^{|C|} e^{-n\delta^2/2}$, which defines the value of δ . On the other hand, we define for each protocol the corresponding entanglement-based picture, the state $|\sigma\rangle_{AA'}$ prepared by Alice in every round, the channel representing Alice and Bob's measurements, and the maps $\mathcal{G}$ and $\mathcal{Z}$ representing the key distillation. Additionally, the entanglement-based approach enforces a constraint on Alice's marginal σ_A in the minimization (since said register is not accessible to Eve), which may require a further facial reduction on the state.

A. Qubit BB84

Following the protocol outlined in [[7], Sec. VII], we consider an implementation of the qubit BB84 protocol [1] that accounts for no-click events [50]. Under the source-replacement scheme, qubit BB84 can be reformulated as an entanglement-based protocol [2,22], in which Alice prepares the maximally entangled state

$$|\sigma\rangle_{AA'} = \frac{1}{\sqrt{2}}(|0\rangle \otimes |0\rangle + |1\rangle \otimes |1\rangle)_{AA'}, \quad (48)$$

so that her marginal state is maximally mixed, $\sigma_A = \mathbb{1}/2$. To model imperfections, we consider that the state sent by Alice in the subsystem A' is subjected to a depolarizing channel

$$\mathcal{E}(\rho) = v\rho + (1-v)\text{Tr}[\rho]\frac{\mathbb{1}}{d}, \quad (49)$$

where $v \in [0, 1]$ describes the visibility. Subsequently, Alice sends her state through a lossy channel with transmittance $\eta = 10^{-(\chi/10)}$, where χ represents the total channel transmittance in dB.

In each round, Alice chooses with probability p^K to measure her subsystem A in the Z -basis, which will be used exclusively for key generation rounds. With probability $1 - p^K$, she instead measures in the X -basis, which is reserved for parameter estimation. The measurements performed by Alice to generate the state are given by the operators

$$P^{X,0} = \frac{1-p^K}{2} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}, \quad P^{X,1} = \frac{1-p^K}{2} \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix},$$

$$P^{Z,0} = p^K \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad P^{Z,1} = p^K \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}. \quad (50)$$

Bob performs the same measurements with the same probabilities as Alice, but also includes the possibility of observing a no-click event denoted by the symbol $\perp$. Thus, Bob's measurement operators are given by

$$Q^{X,0} = \frac{1-p^K}{2} \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad Q^{X,1} = \frac{1-p^K}{2} \begin{pmatrix} 1 & -1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

$$Q^{Z,0} = p^K \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad Q^{Z,1} = p^K \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

$$Q^{Z,\perp} = p^K \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad Q^{X,\perp} = (1-p^K) \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \quad (51)$$

Once all quantum states have been sent and measured, Alice and Bob proceed with the public announcements. She announces her basis choice Z for key rounds, while for parameter estimation rounds she reveals her basis choice X together with her measurement outcome (either 0 or 1). Bob's announcement depends on Alice's choice of measurement; if she indicates a key generation round and Bob measures in the Z -basis, he discloses only whether a detection occurred, using the symbol $I = \top$ for a successful detection and $I = \perp$ for a non-detection event. If their bases do not coincide, the round is discarded and Bob sets $I = \perp$ as well. For parameter estimation rounds, he announces both his basis, which could be either Z or X , and the corresponding outcome, which may be 0, 1, or $\perp$.

From these announcements, Alice and Bob construct the register C with alphabet $\mathcal{C} = \{\perp\} \cup (\{0, 1\} \times \{(Z, 0), (Z, 1), (Z, \perp), (X, 0), (X, 1), (X, \perp)\})$ where $C = \perp$ corresponds to a key generation round, independent of Bob's basis choice. We also define the restricted alphabet $\tilde{\mathcal{C}} = \mathcal{C} \setminus \{\perp\}$. On the other hand, Alice stores her key in the secret register S . Whenever she announces a generation round and Bob measures in the Z basis with a detection event, she sets $S \in \{0, 1\}$ according to her measurement outcome; for no-clicks, she assigns $S = \perp$. Given this description, the channel is as follows

$$\begin{aligned}
\tilde{\mathcal{M}}_{AB \rightarrow SCI}(\cdot) = & \sum_{s=0,1} \text{Tr}_{AB} [P_A^{Z,s} \otimes Q_B^{Z,\top}(\cdot)] \otimes |s\rangle\langle s|_S \otimes |\perp\rangle\langle\perp|_C \otimes |\top\rangle\langle\top|_I \\
& + \sum_{s=0,1} \text{Tr}_{AB} [P_A^{Z,s} \otimes Q_B^{Z,\perp}(\cdot)] \otimes |\perp\rangle\langle\perp|_S \otimes |\perp\rangle\langle\perp|_C \otimes |\perp\rangle\langle\perp|_I \\
& + \sum_{s=0,1} \sum_{b=0,1,\perp} \text{Tr}_{AB} [P_A^{Z,s} \otimes Q_B^{X,b}(\cdot)] \otimes |\perp\rangle\langle\perp|_S \otimes |\perp\rangle\langle\perp|_C \otimes |\perp\rangle\langle\perp|_I \\
& + \sum_{a,(y,b) \in \tilde{\mathcal{C}}} \text{Tr} [P_A^{X,a} \otimes Q_B^{y,b}(\cdot)] |\perp\rangle\langle\perp|_S \otimes |a, (y,b)\rangle\langle a, (y,b)|_C \otimes |\perp\rangle\langle\perp|_I.
\end{aligned} \tag{52}$$

From this structure, we can find the trace-preserving key map

$$\text{Tr}_I[\mathcal{G}(\omega)_{ABRI}] = \mathcal{G}(\omega) = \mathcal{G}_\top(\omega) + \mathcal{G}_\perp(\omega), \tag{53}$$

where $\mathcal{G}_\top$ and $\mathcal{G}_\perp$ have a single Kraus operator each, namely,

$$\begin{aligned}
G_\top &= \sqrt{p^K} \sum_{r=0,1} |r\rangle_R \otimes |r\rangle\langle r|_A \otimes (|0\rangle\langle 0| + |1\rangle\langle 1|)_B, \\
G_\perp &= |\perp\rangle_R \otimes \mathbb{1}_A \otimes [\sqrt{1-p^K}(|0\rangle\langle 0| + |1\rangle\langle 1|)_B + |\perp\rangle\langle\perp|_B].
\end{aligned} \tag{54}$$

In turn, the pinching map is defined as $\mathcal{Z}(\psi) = \mathcal{Z}_\top(\psi) + \mathcal{Z}_\perp(\psi)$, where $\mathcal{Z}_\top$ has Kraus operators $|r\rangle\langle r|_R \otimes \mathbb{1}_{AB}$, with $r \in \{0, 1\}$, and $\mathcal{Z}_\perp$ has a single Kraus operator $|\perp\rangle\langle\perp|_R \otimes \mathbb{1}_{AB}$. Note that these maps obey the structure introduced in Sec. III C, and therefore, we can use the simplifications discussed there.

For the RényiQKD cone (31) we must perform the facial reduction of the maps $\mathcal{G}_\top$, $\mathcal{Z}_\top$ and $\mathcal{G}_\perp$, $\mathcal{Z}_\perp$. In the case of $\mathcal{G}_\top$, we require an isometry $W_{\mathcal{G}_\top}$ such that $W_{\mathcal{G}_\top}^\dagger W_{\mathcal{G}_\top}$ is a projector onto the support of $G_\top G_\top^\dagger$. A suitable choice is

$$W_{\mathcal{G}_\top} = \sum_{r=0,1} |r\rangle_R \otimes |r\rangle\langle r|_A \otimes (|0\rangle_B\langle 0|_{B'} + |1\rangle_B\langle 1|_{B'}), \tag{55}$$

where $\{|0\rangle_{B'}, |1\rangle_{B'}\}$ is an orthonormal basis. Then, $\hat{\mathcal{G}}_\top$ has a single Kraus operator given by

$$W_{\mathcal{G}_\top}^\dagger G_\top = \sqrt{p^K} \mathbb{1}_A \otimes (|0\rangle_{B'}\langle 0|_B + |1\rangle_{B'}\langle 1|_B), \tag{56}$$

which satisfies (24a) when $V = \mathbb{1}$. In its turn, $\mathcal{Z}_\top$ is strictly positive when restricted to the subspace $\{|0\rangle, |1\rangle\}$ of the R subsystem. The isometry that does the job is therefore

$$W_{\mathcal{Z}_\top} = (|0\rangle_R\langle 0|_{R'} + |1\rangle_R\langle 1|_{R'}) \otimes \mathbb{1}_{AB}, \tag{57}$$

where $\{|0\rangle_{R'}, |1\rangle_{R'}\}$ is an orthonormal basis. Thus, $\hat{\mathcal{Z}}_\top$ has Kraus operators $|r\rangle_{R'}\langle r|_R \otimes \mathbb{1}_{AB}$ for $r \in \{0, 1\}$, which can

be verified to satisfy Eq. (24b). The facial reduction of $\mathcal{G}_\perp$ and $\mathcal{Z}_\perp$ is trivial: we eliminate the R subsystem employing $W_{\mathcal{G}_\perp} = W_{\mathcal{Z}_\perp} = |\perp\rangle_R \otimes \mathbb{1}_{AB}$. Hence, $\hat{\mathcal{G}}_\perp$ has a single Kraus operator $\mathbb{1}_A \otimes [\sqrt{1-p^K}(|0\rangle\langle 0| + |1\rangle\langle 1|)_B + |\perp\rangle\langle\perp|_B]$, and $\hat{\mathcal{Z}}_\perp$ has a single Kraus operator $|\perp\rangle_R \otimes \mathbb{1}_{AB}$. All together, we use the facial reduction of the maps $\mathcal{G}_\top$, $\mathcal{Z}_\top$ and $\mathcal{G}_\perp$, $\mathcal{Z}_\perp$ with parameters $S = W_{\mathcal{Z}_\top}^\dagger W_{\mathcal{G}_\top}$ and $S = W_{\mathcal{Z}_\perp}^\dagger W_{\mathcal{G}_\perp} = \mathbb{1}_{AB}$, respectively.

For the FastRényiQKD cone (32), we only need the facial reduction of $\mathcal{G}_\top$ and $\mathcal{Z}_\top \circ \mathcal{G}_\top$ (see Sec. III C). While the former is already done above, the latter is given by the map $\hat{\mathcal{Z}}\mathcal{G}_\top$ with Kraus operators $\sqrt{p^K}|r\rangle\langle r|_A \otimes (|0\rangle_{B'}\langle 0|_B + |1\rangle_{B'}\langle 1|_B)$ for $r \in \{0, 1\}$. One can check that Eq. (29) is respected with $W_{\mathcal{Z}_\top \mathcal{G}_\top}^\dagger W_{\mathcal{G}_\top} = \mathbb{1}_{AB'}$, and thus the parameter S is given by $W_{\mathcal{Z}_\top \mathcal{G}_\top}^\dagger W_{\mathcal{G}_\top} = \mathbb{1}_{AB'}$.

Our results are shown in Fig. 1, where we plot the secret key rates against the total transmittance χ for the qubit BB84 protocol via the FastRényiQKD cone. Here we

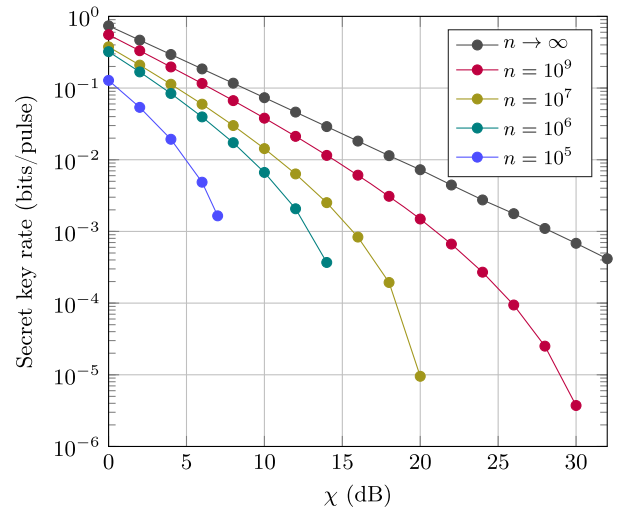


FIG. 1. Finite secret key rate for qubit BB84 protocol using the FastRényiQKD cone for a different number of rounds n . All curves consider $v = 0.97$, and $f = 1.16$. The probability p^K was optimized for each point according to a coarse-grained tuning, while the Rényi parameter α was optimized numerically. The asymptotic key rate is included as a reference.

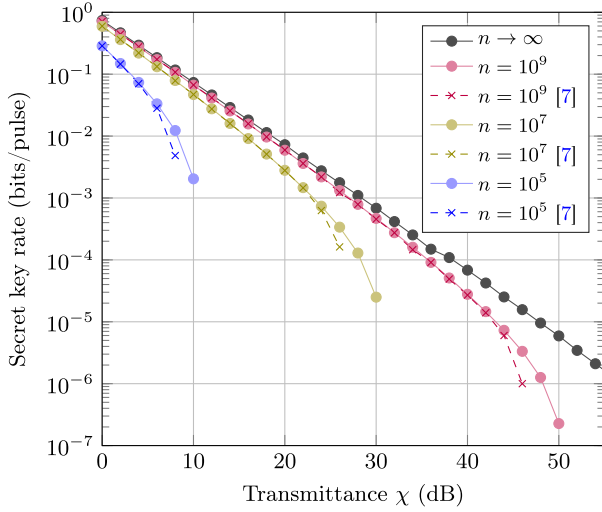


FIG. 2. Variable-length finite secret key rate for qubit BB84 protocol using the FastRényiQKD cone (solid lines), compared with the results showcased in [7] (dashed lines) for different numbers of rounds n . All curves consider $v = 0.97$, and $f = 1.16$. The probability p^K was optimized for each point according to a coarse-grained tuning, while the Rényi parameter α was optimized numerically. For a fair comparison, we set the security parameters to $\varepsilon_{\text{PA}} = \varepsilon_{\text{EC}} = \frac{1}{2} 10^{-80}$. The asymptotic key rate is included as a reference.

consider $f = 1.16$, different numbers of rounds n , visibility $v = 0.97$, and optimal values for α and p^K in every point. We observe that our implementation reaches 14 dB of total loss for a number of rounds $n \leq 10^6$, compatible with current industrial capacity [51]. To further validate our approach, we implement the variable-length QKD framework described in Ref. [7, Sec. VII]. In contrast to the fixed-length scenario discussed above, this formulation requires solving two distinct optimization problems to compute the final key rate (see Appendix C). As shown in Fig. 2, this improved numerical convergence leads to slightly better key rates and an increase in maximal transmittance. Moreover, the use of our conic methods provides a neat speedup as indicated in Sec. VI.

B. MUBs

The protocol described in [52] (originally outlined in [23,24]) consists of Alice and Bob using a set of m MUBs to measure a state whose dimension is d , such that $2 \leq m \leq d + 1$. This protocol is directly provided in an entanglement-based picture, where the initial state prepared by an external source is the generalized, maximally entangled state

$$|\sigma\rangle_{AA'} = \frac{1}{\sqrt{d}} \sum_{j=0}^{d-1} |j\rangle_A \otimes |j\rangle_{A'}. \quad (58)$$

This state when distributed to Alice and Bob evolves into an isotropic state $\sigma_{AA'}^{\text{iso}}(v) = v\sigma_{AA'} + (1-v)\mathbb{1}/d^2$, where Alice's marginal is always $\sigma_A = \mathbb{1}/d$. Following [52], the key is distilled according to Alice's projective measurements on the computational basis $\{|s\rangle\langle s|_A\}_{s=0}^{d-1}$. On the other hand, since no-click events are not considered in this formulation (or in a broader sense, postselection events), register I becomes trivial and thus can be eliminated for this protocol.

Parameter estimation is performed by Alice using the POVMs $\{\{\Pi^{j,l}\}_{j=0}^{d-1}\}_{l=0}^{m-1}$ where $\Pi^{j,l}$ denotes the j -th vector of the l -th MUB (in particular, $l = 0$ represents the computational basis). Bob's POVMs are the transposed version of Alice's. The choice of basis is defined via the probability vector

$$\{p(l)\}_{l=0}^{m-1} = (p^K, (1-p^K)/(m-1), \dots, (1-p^K)/(m-1)). \quad (59)$$

Hence, only rounds where both Alice and Bob use the zeroth basis will be spent for the key distillation, implying that the sifting probability is given by $p^s = (p^K)^2$. Note further that the original implementation requires only the statistics generated by Alice and Bob when their basis choices coincide, so we can merge the statistics from all other cases into a single probability outcome for parameter estimation. All in all, we define the alphabet $\mathcal{C} = \{\perp, 1, \dots, m\}$, such that $\tilde{\mathcal{C}} = \mathcal{C} \setminus \{\perp\}$. Furthermore, $\mathcal{S} = \{\perp, 0, \dots, d-1\}$ with $\tilde{\mathcal{S}} = \mathcal{S} \setminus \{\perp\}$ for the secret key alphabet. According to this notation and reducing the POVMs that add up to the identity, one round of the MUB protocol is described by the following channel

$$\begin{aligned} \tilde{\mathcal{M}}_{AB \rightarrow SC}(\cdot) = & p(0)^2 \sum_{s \in \tilde{\mathcal{S}}} \text{Tr}[|s\rangle\langle s|_A \otimes \mathbb{1}_B(\cdot)] |s\rangle\langle s|_S \otimes |\perp\rangle\langle \perp|_C \\ & + \sum_{l=1}^{m-1} \sum_{j=0}^{d-1} p(l)^2 \text{Tr}[\Pi^{j,l} \otimes (\Pi^{j,l})^T(\cdot)] |\perp\rangle\langle \perp|_S \otimes |l\rangle\langle l|_C \\ & + \left[1 - \sum_{l=0}^{m-1} p(l)^2\right] |\perp\rangle\langle \perp|_S \otimes |m\rangle\langle m|_C. \end{aligned} \quad (60)$$

The key map $\mathcal{G}$ is given by $\mathcal{G}(\rho) = G\rho G^\dagger$, where

$$G = \sum_{r=0}^{d-1} |r\rangle_R \otimes |r\rangle_{A'} \otimes \mathbb{1}_B. \quad (61)$$

As this is an isometry, the facial reduction is trivial: $\hat{\mathcal{G}}$ is the identity map, and thus Eq. (24a) is satisfied with $V = \mathbb{1}$ and $W_{\mathcal{G}} = G$. In its turn, $\mathcal{Z}$ is provided by the usual dephasing superoperators $Z^r = |r\rangle\langle r|_R \otimes \mathbb{1}_{AB}$ for

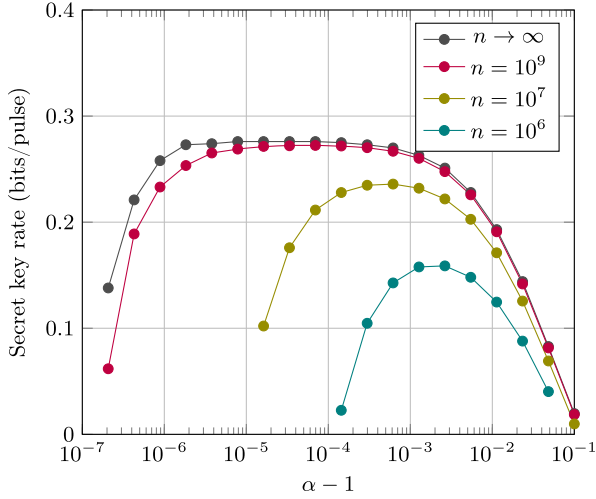


FIG. 3. Finite secret key generation rate for the MUB protocol using the FastRényiQKD (32) cone with $d = 5$, $m = 6$, $v = 0.9$, $f = 1.16$, $p^K = 0.5$, and different values for the number of rounds n and the Rényi parameter α . The asymptotic key rate is included as a reference.

$r \in \{0, \dots, d-1\}$ and is already a strictly positive map, so Eq. (24b) is satisfied with $\hat{\mathcal{Z}} = \mathcal{Z}$ and $W_{\mathcal{Z}} = 1$. The last parameter we need for the RényiQKD cone (31) is $S = W_{\mathcal{Z}}^\dagger W_{\mathcal{G}} = G$. For the FastRényiQKD cone (32) we need in addition the facial reduction of $\mathcal{Z} \circ \mathcal{G}$. It is the map $\widehat{\mathcal{ZG}}$ with Kraus operators $|r\rangle\langle r|_A \otimes \mathbb{1}_B$ for $r \in \{0, \dots, d-1\}$, and Eq. (29) is satisfied with $W_{\mathcal{ZG}} = G$, and thus the parameter S is simply $W_{\mathcal{ZG}}^\dagger W_{\mathcal{G}} = G^\dagger G = 1$.

In Fig. 3 we illustrate the secret key generation rates against the Rényi parameter α for diverse values of the block size n with parameters $d = 5$, $m = 6$, and $v = 0.9$ for the FastRényiQKD cone (32). Overall, we observe (as is the case for the other protocols) that the optimal value for α does not require fine-tuning beyond the order of magnitude to yield values close to the actual maximum at the curves. Figure 4 compares the FastRényiQKD cone (32) with the RényiQKD cone (31) for α varying from 1.1 to 2. We see that the RényiQKD cone delivers superior performance, as expected, but the advantage is diminutive and only visible for values of α where the bounds on the secret key rate drop below zero. We emphasize that negative key rates have no operational meaning in QKD, and we consider this regime strictly to compare the two cones. In order to deliver a fair comparison, we did not use the relaxed formulation (42) for the RényiQKD cone, but the nested optimization discussed just above it.

C. DMCV

For the QKD protocol based on a discrete-modulated continuous variable (DMCV) approach, we implement a variant [6] of the well-known quadrature phase shift

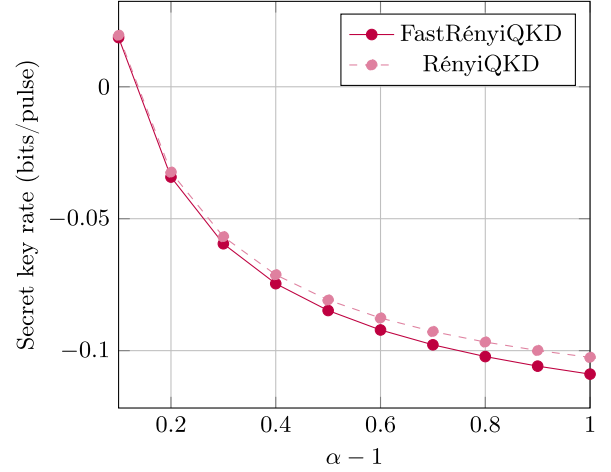


FIG. 4. Lower bound on the finite secret key rate for the MUB protocol using the RényiQKD (31) and FastRényiQKD (32) cones with $d = 5$, $m = 6$, $v = 0.9$, $f = 1.16$, $p^K = 0.5$, $n = 10^9$, and Rényi parameter α varying from 1.1 to 2. We emphasize that this range of α is not relevant for secret key generation, and negative key rates have no operational meaning. The purpose here is only to illustrate the difference between the two cones.

keying (QPSK) [53–55] based on Alice sampling states with equal probabilities from a distribution of coherent states $\{|i^x \gamma\rangle\}_{x=0}^3$, and Bob implementing a full discretization for the measurement outcomes. We caution the reader to understand S as the key register generated by Bob, as it is customary in CVQKD due to its reverse nature for information reconciliation [56].

Under the entanglement-based picture, Alice will send for all rounds the state

$$|\sigma\rangle_{AA'} = \frac{1}{2} \sum_{x=0}^3 |x\rangle_A \otimes |i^x \gamma\rangle_{A'} \quad (62)$$

from which her marginal follows directly. While Alice generates her key registers via measurements on her computational basis, Bob extracts said register using a discretization according to a coarse-grained heterodyne detection

$$\bar{R} = \frac{1}{\pi} \int_{\mathcal{R}} |\gamma\rangle\langle\gamma| d\gamma^2. \quad (63)$$

Here, $\bar{R}$ is known as a region operator, and $\mathcal{R} \subseteq \mathbb{C}$ an element of a partition in $\mathbb{C}$, such that the collection of region operators completely covers the phase space, constituting a POVM. In particular, we consider the sets of region operators $\{\hat{R}_B^z\}_{z=0}^3$ and $\{\tilde{R}_B^z\}_{z=0}^5$ for key and parameter estimation rounds, respectively, defined in [26] according to a modulation (Δ, Δ_s) . We refer the reader to [26] for further implementation specifics.

Provided these considerations we define the alphabets $\mathcal{C} = \{\perp, \perp\} \cup (\{0, \dots, 3\} \times \{0, \dots, 5\})$ and $\mathcal{S} = \{\perp, 0, \dots, 3\}$. We can now formulate the map constituting the DMCV protocol via

$$\begin{aligned} \mathcal{M}_{AB \rightarrow SC}(\cdot) &= p^K \sum_{s \in \mathcal{S}} \text{Tr}[\mathbb{1}_A \otimes \hat{R}_B^s(\cdot)] |s\rangle\langle s|_S \otimes |\perp, \perp\rangle\langle \perp, \perp|_C + (1 - p^K) \\ &\times \sum_{(x,z) \in \tilde{\mathcal{C}}} \text{Tr}[|x\rangle\langle x|_A \otimes \tilde{R}_B^z(\cdot)] |\perp\rangle\langle \perp|_S \otimes |x, z\rangle\langle x, z|_C. \end{aligned} \quad (64)$$

As in the MUB protocol, we have that no postselection is applied on the key distillation, therefore I can be removed. Furthermore, Alice always implements the same measurement such that the sifting probability coincides with the key probability $p^s = p^K$. From the key generation rounds, we can build the key map $\mathcal{G}(\rho) = G\rho G^\dagger$ where [55]

$$G = \sum_{r=0}^3 |r\rangle_R \otimes \mathbb{1}_A \otimes \sqrt{\hat{R}_B^r}. \quad (65)$$

The $\mathcal{Z}$ map is again provided by a set of pinching operators $\mathcal{Z}^r = |r\rangle\langle r|_R \otimes \mathbb{1}_{AB}$ for $r \in \{0, \dots, 3\}$.

In order to do the facial reduction, we first note that $\mathcal{G}$ is an isometry [58]. Therefore $\hat{\mathcal{G}}$ is the identity map, and Eq. (24a) is satisfied with $V = \mathbb{1}$ and $W_G = G$. In turn, $\mathcal{Z}$ is already a strictly positive map, so Eq. (24b) is satisfied with $\hat{\mathcal{Z}} = \mathcal{Z}$ and $W_{\mathcal{Z}} = \mathbb{1}$. Then, the isometry S is given

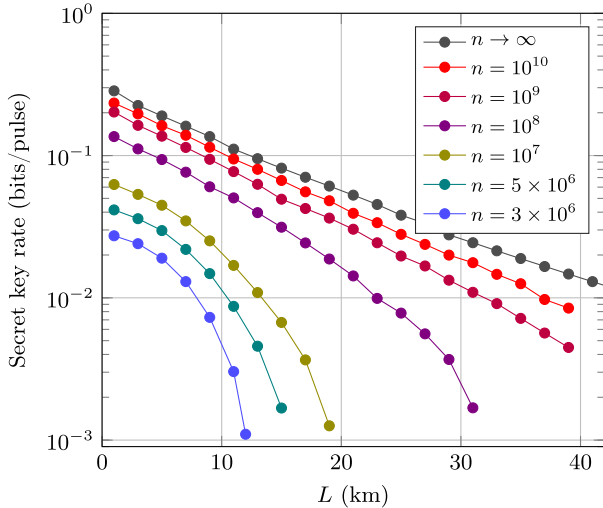


FIG. 5. Secret key generation rates for the DMCV protocol using the FastRényiQKD cone (32) for different values of the number of rounds n , using $N_c = 10$, $\Delta = 4.0$ and $\Delta_s = 1.5$ and error correction at the Shannon limit. α , p^K and the coherent state amplitude were optimized for each point according to a coarse-grained tuning. The asymptotic key rate is included as a reference.

by $S = W_{\mathcal{Z}}^\dagger W_G = G$, and so the facial reduction for the RényiQKD cone (31) is done. For the FastRényiQKD cone (32) we need in addition the facial reduction of $\mathcal{Z} \circ \mathcal{G}$, which is trivial, since $\mathcal{Z} \circ \mathcal{G}$ is already strictly positive [19]. Thus, $\hat{\mathcal{Z}}\mathcal{G} = \mathcal{Z} \circ \mathcal{G}$ and Eq. (29) is satisfied with $W_{\mathcal{Z}} = \mathbb{1}$, implying $S = G$ once again.

Figure 5 provides several finite secret key rates, using the modulation presented in [26] while adjusting all other parameters to their optimal value. In particular, we observe that key rates can be generated with $n = 5 \times 10^6$ at 15 km, which improves in orders of magnitude over [26], where $n = 5 \times 10^8$ was reported as the minimum block size. On the other hand, our method provides an enhanced numerical scalability with respect to other approaches [19,52] and does not require a min-tradeoff linearization, which should allow a removal of the cutoff assumption without incurring in unfortunate penalties [59].

VI. BENCHMARKS

We conclude our numerical analysis by showcasing the performance of our algorithm in contrast with the repository Renyi-security-framework [60] and the package QICS [44] (used in Refs. [7] and [18] respectively), which are also capable of optimizing the sandwiched Rényi entropy.

In Table I, we compare the time and iterations taken to optimize the cones corresponding to the DMCV protocol between our implementation and QICS [44]. Since the Rényi cone in QICS does not include facial reduction, a small depolarization noise ($\epsilon = 10^{-8}$) was introduced in $\mathcal{G}$ to avoid null eigenvalues, which would make the gradient undefined and the optimization unstable. Moreover, given that QICS exploits the sparsity in the optimization of the Rényi entropy cone, we also exploited the block structure of the $\hat{\mathcal{Z}}$ in the FastRényiQKD and RényiQKD cones to make a fair comparison.

TABLE I. Average iterations and time to optimize the cones corresponding to the DMCV protocol. For each value of N_c , the iterations and time are averaged over 5 runs with the same parameters: $L = 13$ km, $n = 10^7$, $f = 1.0$, $\Delta_s = 1.5$, and $\Delta = 4.0$. In the more relevant case $N_c = 10$ [17,55,57], QICS results in the computer running out of RAM due to the higher dimension of the cones, which forces it to use SWAP memory, making the optimization very slow. The benchmarks were performed using ConicQKD.jl v0.3.0 and QICS v1.1.3 [44] on a workstation equipped with an Apple M4 Max and 128 GB DDR5 RAM.

N_c	FastRényiQKD		RényiQKD		QICS [44]	
	Iterations	Time [s]	Iterations	Time [s]	Iterations	Time [s]
5	67	16.87	70	594.7	53	1731
7	54	49.64	80	3374	60	8701
10	60	347.1	73	18430	Out of RAM	

TABLE II. Average iterations and computational time over 20 runs of the qubit BB84 protocol at transmittance $\chi = 8$ dB and different block sizes. FastRényiQKD benchmarks were executed with ConicQKD.jl v0.3.0, whereas the Rényi-security-framework [60] was evaluated using MATLAB R2024a. All computations were performed on a workstation equipped with an 11th Gen Intel Core i5-11500 processor and 16 GB of RAM.

n	FastRényiQKD		Rényi-security-framework [60]	
	Iterations	Time [s]	Iterations	Time [s]
10^9	62	0.133	60	34.427
10^7	42	0.098	60	32.945
10^5	36	0.049	60	19.682

As we can see, our RényiQKD cone (31) is considerably faster than QICS. This is due to a factor of 4 reduction in the output dimension of $\mathcal{G}$ caused by the facial reduction. Additionally, switching to the FastRényiQKD (32), which is not available in QICS, provides a dramatic speedup with respect to the RényiQKD cone. Regarding stability, in some runs (for example, for $N_c = 7$, $n = 10^7$, and $L = 9$ km) the optimization in QICS failed to converge, unless we increased the depolarization noise.

In Table II, we provide a numerical comparison between our FastRényiQKD cone (32) and the repository Rényi-security-framework [60] used by Ref. [7], which implements the Frank-Wolfe algorithm [14,15]. Unlike our approach, this method needs to add an artificial depolarization noise ($\epsilon = 10^{-10}$ in this specific realization) to prevent undefined gradients. The values reported represent the average iterations and computation time required to solve the two optimization problems leading to the final key rate (see Appendix C), averaged over 20 runs of the variable-length qubit BB84 protocol at transmittance $\chi = 8$ dB. The results indicate that our conic method is orders of magnitude faster than the Frank-Wolfe approach.

VII. CONCLUSION

This work introduces the most advanced framework for QKD, combining the improved privacy amplification via Rényi hashing [21] and the security proof provided by the MEAT [10], which admits the calculation of secret key rates under finite-size effects and general security by means of sandwiched Rényi entropies. The core technical contribution is the introduction of two convex cones that allow the fast and reliable optimization of the conditional Rényi entropy via non-symmetric conic optimization.

We have tested this framework using instances of different protocols, including the full-discrete QPSK protocol from the DMCV family [6,26], which is well-known for its resource-intensive numerical implementation [59]. In said case, we have observed secret key rates at smaller block sizes thanks to the tightness of the finite-size analysis of the MEAT. Further improvements can be

achieved by using more sophisticated approaches, such as a fine-tuned postselection on Bob's measurements [57,61]. Furthermore, we have applied our method to the qubit BB84 protocol using the variable-length framework described in Ref. [7]. We benchmarked our conic approach in these two protocols against QICS [44] and the Frank-Wolfe algorithm [14,15] used in previous implementations, respectively. We observed orders of magnitude faster computation time, as well as improved stability, and slightly better secret key rates and critical transmittances.

Regarding future work, it would be interesting to apply our conic methods to compute key rates in more complex and realistic protocols, such as DMCV without a cutoff assumption [61] or BB84 using decoy states. From a practical perspective, we note that our framework opens the possibility of performing QKD implementations using reduced block sizes and ensuring long-distance general security without the repetition-rate limitations of previous entropy accumulation techniques [11,12], which paves the path for a practical, reliable, and large-scale implementation for QKD.

ACKNOWLEDGMENTS

We thank Lars Kamin and Ernest Tan for insightful discussions about entropy accumulation, variable-length approaches, and theoretical aspects of QKD. M. N. and C. P. G. were supported by the Government of Spain (Severo Ochoa CEX2019-000910-S, FUNQIP and NextGeneration EU PRTR-C17.I1) and European Union (QSNP, 101114043). M. N. acknowledges funding from the European Union's Horizon Europe research and innovation programme under the MSCA Grant Agreement No. 101081441. C. P. G. has received funding from the European Union's Digital Europe Programme under the project QUARTER (101091588) and Horizon 2020 Research and Innovation Programme under the project QSNP (101114043), and from the European Innovation Council's Horizon Europe EIC Accelerator Programme under the project MIQRO (101161539). The research of A. G. L., P. V. P., and M. A. was supported by the Q-CAYLE project, funded by the European Union-Next Generation UE/MICIU/Plan de Recuperación, Transformación y Resiliencia/Junta de Castilla y León (PRTRC17.11), and also by the Department of Education of the Junta de Castilla y León and FEDER Funds (Reference: CLU-2023-1-05). P. V. P. has also been funded under the UVa 2024 predoctoral contract, co-financed by Banco Santander. M. A. was also supported by the Spanish Agencia Estatal de Investigación, Grants No. RYC2023-044074-I and PID2024-161725OA-I00.

DATA AVAILABILITY

The implementation of the cones introduced in this work is available in the package ConicQKD.jl [48]. The results

presented in Sec. V can be reproduced using the repository in Ref. [62]. The comparative results and benchmarks presented in Fig. 2 and Table II were obtained using the repository in Ref. [60], which is heavily based on the Open QKD Security package [63].

APPENDIX A: MARGINAL-CONSTRAINED ENTROPY ACCUMULATION THEOREM

This section provides a theoretical security analysis based on the marginal-constrained entropy accumulation theorem (MEAT) [10] and the required notation changes to adapt it to our approach. As a first step, we define the corresponding results of the MEAT itself, considering that no public announcements are made before all quantum signals have been measured. In addition, we apply the special implementation provided by [10], Theorem 4.2b], in which only the classical register C is considered for test rounds.

Proposition A.1. [10], Corollary 4.2, Theorem 4.2a] For each $j \in \{1, \dots, n\}$, take a state $\sigma^j \in \mathcal{D}(A_j)$ and a CPTP map $\tilde{\mathcal{M}}_j: [AB]_j \rightarrow [SCI]_j$, such that C_j and I_j are classical. Let $\rho_{[SCI]^{(n)}\hat{E}} = \tilde{\mathcal{M}}_n \circ \dots \circ \tilde{\mathcal{M}}_1(\bar{\omega})$ for $\bar{\omega} \in \mathcal{D}([AB]^{(n)}\hat{E})$, verifying $\bar{\omega}_{A_1^n} = \sigma_{A_1}^1 \otimes \dots \otimes \sigma_{A_n}^n$.

Suppose furthermore that $\rho = p_\Omega \rho_{|\Omega} + (1 - p_\Omega) \rho_{|\neg\Omega}$ for $p_\Omega \in (0, 1]$ and all C_j are isomorphic to a single register C with alphabet $\mathcal{C}$. Let S_Ω be the convex set of probability distributions defined on the alphabet $\mathcal{C}$ such that any $c^{(n)}$ with nonzero probability in $\rho_{|\Omega}$ verifies that $\text{freq}_{c^{(n)}} \in S_\Omega$. Then, for any $\alpha \in (1, \infty)$,

$$H_\alpha^\uparrow(S^{(n)}|[CI]^{(n)}\hat{E})_{\rho_{|\Omega}} \geq n h_\alpha^\uparrow - \frac{\alpha}{\alpha - 1} \log \frac{1}{p_\Omega}, \quad (\text{A1})$$

where

$$h_\alpha^\uparrow = \inf_{q \in S_\Omega} \inf_{\nu \in \Sigma_j} \frac{\alpha}{\alpha - 1} D_{\text{KL}}(q \| \nu_C) + \sum_{c \in \text{supp}(\nu_C)} q(c) H_\alpha^\uparrow(S|IE)_{\nu_{|c}} \quad (\text{A2})$$

provided that E is a large-enough purifying register for any $[AB]_j$, and Σ_j the set of all states $\tilde{\mathcal{M}}_j(\omega_{[AB]_j E})$ for some initial $\omega \in \mathcal{D}([AB]_j E)$ such that $\omega_{A_j} = \sigma_{A_j}^j$.

Here, the register I_j stores the public announcements, with alphabet $\mathcal{I}$ satisfying $\mathcal{C} \subseteq \mathcal{I}$. In particular, the equality holds whenever there are no public announcements in key generation rounds. For further reference, we introduce a change of notation by writing the probability p_Ω as $\text{Pr}[\Omega]$, and the set Σ as Σ_Ω .

For simplicity, in the following discussion we assume that all maps are identical, i.e. $\tilde{\mathcal{M}}_1 = \dots = \tilde{\mathcal{M}}_n = \tilde{\mathcal{M}}$, and consequently, so is the marginal σ_A . We then divide map $\tilde{\mathcal{M}}$ in two channels as

$$\tilde{\mathcal{M}}(\cdot)_{SIC} = p(\perp) \mathcal{M}^K(\cdot)_{SI} \otimes |\perp\rangle\langle\perp|_C + [1 - p(\perp)] |\perp\rangle\langle\perp|_S \otimes \mathcal{M}^T(\cdot)_{IC} \quad (\text{A3})$$

where $\tilde{C} := C \setminus \{\perp\}$, and $\mathcal{M}^K(\cdot)$ and $\mathcal{M}^T(\cdot)$ are quantum channels for key generation and parameter estimation, respectively. Given this redefined channel and provided that $\nu_C = \text{Tr}_{SI}[\tilde{\mathcal{M}}(\omega)_{SCI}]$, for some $\omega \in \mathcal{D}(AB)$, the Kullback-Leibler divergence in Eq. (A2) can be written as

$$D_{\text{KL}}(q \| \nu_C) = q(\perp) \log \left(\frac{q(\perp)}{p(\perp) \mathcal{M}^K(\omega_{AB})_{|\perp}} \right) + \sum_{c \in \tilde{C}} q(c) \log \left(\frac{q(c)}{[1 - p(\perp)] \mathcal{M}^T(\omega_{AB})_{|c}} \right). \quad (\text{A4})$$

For the right-hand side in (A2), the sandwiched Rényi entropy can be simplified as

$$\sum_{c \in \text{supp}(\nu_C)} q(c) H_\alpha^\uparrow(S|IE)_{\nu_{|c}} \geq q(\perp) H_\alpha^\uparrow(S|IE)_{\nu_{|\perp}}. \quad (\text{A5})$$

Within our discussion, this relation holds as an equality because $S = \perp$ when $C \neq \perp$. Moreover, note that Eve also holds a copy $\tilde{I}$ of the classical register I , which implies

$$H_\alpha^\uparrow(S|IE)_{\nu_{|\perp}} = H_\alpha^\uparrow(S|\tilde{E})_{\nu_{|\perp}} = H_\alpha^\uparrow(S|\tilde{I}E)_{\nu_{|\perp}}, \quad (\text{A6})$$

where E represents Eve's quantum side information, and $\tilde{E} := \tilde{I}E$ all her quantum and classical information. The conditional entropy in (A6) can be further lower-bounded by following [13], Supplementary material I]. First, note that our final state for a key round ρ_{RABSI} comes as the combination of two actions: a coherent key map, acting on AB to produce a key reference R and a public discussion I ; and a pinching operation that tells Alice and Bob how to distill the raw key at register S according to the key reference R . Hence, we have

- (i) The quantum state shared by Alice and Bob ω_{AB} .
- (ii) The state after the key isometry σ_{RABI} , and in particular its marginal $\mathcal{G}(\omega_{AB}) = \text{Tr}_I[\sigma_{RABI}]$.
- (iii) The final state after the key extraction, $\rho_{SRABI} = V(\sigma_{RABI})V^\dagger$ where V acts on register R , as well as its marginal $\rho_{SRAB} = V(\sigma_{RAB})V^\dagger$.

In particular, $V = \sum_s |s\rangle_S \otimes Z_R^s$ where Z_R is a projective operator, and R is an intermediate key register between Alice and Bob, which may include postselection and/or removal of no-clicks.

Let us recall, for $\alpha \geq 0$, the up-arrow sandwiched Rényi entropy

$$H_\alpha^\uparrow(A|B)_\rho := - \inf_{\psi \in \mathcal{D}(B)} D_\alpha(\rho_{AB} \| \mathbb{1}_A \otimes \psi_B), \quad (\text{A7})$$

and define the down-arrow sandwiched Rényi entropy

$$H_\alpha^\downarrow(A|B)_\rho := -D_\alpha(\rho_{AB} \| \mathbb{1}_A \otimes \rho_B), \quad (\text{A8})$$

as well as the up-arrow Petz-Rényi entropy

$$\tilde{H}_\alpha^\uparrow(A|B)_\rho := - \inf_{\psi \in \mathcal{D}(B)} \bar{D}_\alpha(\rho_{AB} \| \mathbb{1}_A \otimes \psi_B) \quad (\text{A9})$$

where $\bar{D}_\alpha(\rho \| \psi)$ is the Petz-Rényi divergence. Using these definitions and considering $\rho_{SRAB|\tilde{E}}$ as the purification of the state ρ_{SRAB} , we find for $\gamma = \alpha/(2\alpha - 1)$ when $\alpha \in (1/2, \infty]$ that [13], Supplementary material I]

$$\begin{aligned} H_\alpha^\uparrow(S|\tilde{E})_{\nu_{|\perp}} &= -H_\gamma^\uparrow(S|RAB)_{\nu_{|\perp}} \\ &= \inf_{\psi \in \mathcal{D}(RAB)} D_\gamma(\rho_{SRAB} \| \mathbb{1}_S \otimes \psi_{RAB}) \\ &= \inf_{\psi \in \mathcal{D}(RAB)} D_\gamma(V\sigma_{RAB}V^\dagger \| \mathbb{1}_S \otimes \psi_{RAB}) \\ &\geq \inf_{\psi \in \mathcal{D}(RAB)} D_\gamma(\sigma_{RAB} \| V^\dagger[\mathbb{1}_S \otimes \psi_{RAB}]V) \\ &= \inf_{\psi \in \mathcal{D}(RAB)} D_\gamma\left(\sigma_{RAB} \| \sum_s Z_R^s \psi_{RAB} Z_R^s\right) \\ &= \inf_{\psi \in \mathcal{D}(RAB)} D_\gamma(\mathcal{G}(\omega_{AB}) \| \mathcal{Z}(\psi_{RAB})). \end{aligned} \quad (\text{A10})$$

Here, we used the fact that $\tilde{I} = I$ and the duality relation shown in [64], Theorem 3], while the inequality arises due to the data processing inequality [65]. We also note that [13] proves that the inequality can be set as an equality, which requires deriving an upper bound for the up-arrow Rényi entropy.

Alternatively, we can lower bound the up-arrow Rényi entropy in (A6) using the down-arrow entropy (A8). This eventually defines a new bound $h_\alpha^\downarrow$ at the finite-size analysis, which yields faster numerical results at the cost of a suboptimal secret key rate. This bound goes as follows

$$\begin{aligned} H_\alpha^\uparrow(S|\tilde{E})_{\nu_{|\perp}} &\geq \tilde{H}_\alpha^\uparrow(S|\tilde{E})_{\nu_{|\perp}} \\ &= -H_\beta^\downarrow(S|RAB)_{\nu_{|\perp}} \\ &= D_\beta(\rho_{SRAB} \| \mathbb{1}_S \otimes \rho_{RAB}) \\ &= D_\beta(V\sigma_{RAB}V^\dagger \| \mathbb{1}_S \otimes \rho_{RAB}) \\ &\geq D_\beta(\sigma_{RAB} \| V^\dagger[\mathbb{1}_S \otimes \rho_{RAB}]V) \\ &= D_\beta\left(\sigma_{RAB} \| \sum_s Z_R^s \rho_{RAB} Z_R^s\right) \\ &= D_\beta(\mathcal{G}(\omega_{AB}) \| \mathcal{Z} \circ \mathcal{G}(\omega_{AB})). \end{aligned} \quad (\text{A11})$$

For the first line, we used [64], Corollary 4], while in the second we again applied duality relations shown in [64], Theorem 3] for $\beta = 1/\alpha$ whenever $\alpha \in (0, \infty)$. The second

inequality arises due to the data processing inequality, and the last line comes from calculating $\rho_{RAB} = \text{Tr}_{SI}[\rho_{RABSI}]$.

With the new notation and bounds, we redefine Proposition A.1 as

Proposition A.2. Let $n \in \mathbb{N}$, $\rho_{[SCI]^{(n)}\tilde{E}} = \tilde{\mathcal{M}}^\otimes(\tilde{\omega})$ for some $\tilde{\omega} \in \mathcal{D}([AB]^{(n)}\hat{E})$ and $\tilde{\mathcal{M}}: AB \rightarrow SCI$ a quantum channel, and $\sigma \in \mathcal{D}(A)$. Given a nonabortion event Ω , we have

$$H_\alpha^\uparrow(S^{(n)}|[CI]^{(n)}\hat{E})_{\rho|\Omega} \geq nh_\alpha^\uparrow - \frac{\alpha}{\alpha-1} \log \frac{1}{\text{Pr}[\Omega]} \quad (\text{A12})$$

where $\alpha \in (1, \infty)$ and

$$\begin{aligned} h_\alpha^\uparrow &\geq \inf_{q \in \hat{\Sigma}_\Omega} \inf_{\omega \in \mathcal{D}(AB)} \inf_{\psi \in \mathcal{D}(RAB)} \frac{\alpha}{\alpha-1} D_{\text{KL}}(q \| \tilde{\mathcal{M}}(\omega)_C) \\ &\quad + q(\perp) D_\gamma(\mathcal{G}(\omega) \| \mathcal{Z}(\psi)) \end{aligned} \quad (\text{A13})$$

for $\hat{\Sigma}_\Omega$ the set in (13), all ω such that $\text{Tr}_B[\omega] = \sigma_A$ and $\gamma = \alpha/(2\alpha - 1)$. Furthermore, the bound $h_\alpha^\uparrow \geq h_\alpha^\downarrow$ holds at (A12) where

$$\begin{aligned} h_\alpha^\downarrow &\geq \inf_{q \in \hat{\Sigma}_\Omega} \inf_{\omega \in \mathcal{D}(AB)} \frac{\alpha}{\alpha-1} D_{\text{KL}}(q \| \tilde{\mathcal{M}}(\omega)_C) \\ &\quad + q(\perp) D_\beta(\mathcal{G}(\omega) \| \mathcal{Z} \circ \mathcal{G}(\omega)), \end{aligned} \quad (\text{A14})$$

with $\beta = 1/\alpha$.

Remark 1. At (A6), it is possible to apply a classical conditioning on $\tilde{I}$ (e.g. to discard a key round due to a no-click). Using Ref. [66], Proposition 9], this induces

$$H_\alpha^\uparrow(S|\tilde{I}E)_\rho = \frac{\alpha}{1-\alpha} \log \left(\sum_{i \in \tilde{I}} p(\tilde{I} = i) 2^{\frac{1-\alpha}{\alpha} H_\alpha^\uparrow(S|E)_{\rho_i}} \right). \quad (\text{A15})$$

This formulation constitutes an essential step in decoy-state BB84 and variable-length formulations [7]. Within our context, and provided the aforementioned considerations, we have that the state can be decomposed as

$$\rho_{SRAB} = \sum_{i \in \tilde{I}} \mathcal{G}^i(\omega)_{SRAB} \otimes |i\rangle\langle i|_I \quad (\text{A16})$$

where $\mathcal{G}^i$ is the unnormalized CP map applied according to the value of register I . Let us now expand this expression by using (A10). In the following, we take shorthand notation for the normalized, sandwiched Rényi divergence (5)

$$\begin{aligned} H_\alpha^\uparrow(S|E)_{\rho_i} &\geq \inf_{\psi \in \mathcal{D}(RAB)} D_\gamma(\mathcal{G}^i(\omega_{AB}) \| \mathcal{Z}(\psi_{RAB})). \\ &=: \frac{1}{\gamma-1} \log \left[\frac{\tilde{\Psi}_\gamma^i(\omega)}{\text{Tr}[\mathcal{G}^i(\omega)]} \right], \end{aligned} \quad (\text{A17})$$

where $\tilde{\Psi}_\gamma^i$ already incorporates the optimization with respect to ψ_{RAB} , and the action of the maps $\mathcal{G}^i$ and $\mathcal{Z}$. Then,

$$H_\alpha^\dagger(S|\tilde{I}E)_\rho \geq \frac{\alpha}{1-\alpha} \log \left(\sum_{i \in \tilde{\mathcal{I}}} \text{Tr}[\mathcal{G}^i(\omega)] \left[\frac{\tilde{\Psi}_\gamma^i(\omega)}{\text{Tr}[\mathcal{G}^i(\omega)]} \right]^{\frac{1}{\gamma}} \right). \quad (\text{A18})$$

We note that each term in the sum constitutes a perspective of the function $\tilde{\Psi}_\gamma^i(\omega)^{(1/\gamma)}$, which for $\gamma^{-1} > 1$ can be modelled using the well-known power cone

$$\mathcal{K}_p = \{(x, y, z) \in \mathbb{R} \times \mathbb{R} \times \mathbb{R} : x^p y^{1-p} \geq |z|\}. \quad (\text{A19})$$

This formulation provides a tighter bound than (A10). Nevertheless, it does not achieve a noticeable improvement in the secret key rate, while it greatly increases the cost of the numerical implementation.

In the case of using the bound (A11) on the exponent, we may explicitly decompose the final expression using the normalized divergence (5) to find

$$\begin{aligned} H_\alpha^\dagger(S|\tilde{I}E)_\rho &\geq \frac{1}{\beta-1} \log \left(\sum_{i \in \tilde{\mathcal{I}}} \text{Tr}[\mathcal{G}^i(\omega)] 2^{\log[\Psi_\beta^i(\omega)/\text{Tr}[\mathcal{G}^i(\omega)]]} \right) \\ &= \frac{1}{\beta-1} \log \left(\sum_{i \in \tilde{\mathcal{I}}} \Psi_\beta^i(\omega) \right). \end{aligned} \quad (\text{A20})$$

APPENDIX B: PROOF OF THEOREM III.2

The security for finite-size QKD protocols is based on the leftover hash lemma [67,68], which bounds the length of an extractable secret key according to the raw key given all the information accessible to Eve. Within our discussion, we use a variant of this lemma based on Rényi entropies.

Proposition B.1. [21], Theorem 9] Let $\sigma_{AE} \in \mathcal{D}(AE)$ be a cq-state and $\{\mathcal{R}_{A \rightarrow C}^h(\cdot), h \in \mathcal{H}\}$ a family of λ -randomizing hash functions on register A. Then,

$$T(\rho_{CFE}^{\mathcal{R}}, \frac{1}{|C|} \mathbb{1}_C \otimes \rho_{FE}) \leq 2^{\frac{2(1-\alpha)}{\alpha}} 2^{\frac{\alpha-1}{\alpha} (\log |C| - H_\alpha^\dagger(A|E)_\sigma + 2 \log \lambda)} \quad (\text{B1})$$

where $\rho_{CFE}^{\mathcal{R}} = (\sum_h p(h) \mathcal{R}^h \otimes |h\rangle\langle h|_F \otimes \mathbb{1}_E)(\sigma_{AE})$, $\alpha \in (1, 2)$ and $\{|h\rangle_F\}_{h \in \mathcal{H}}$ constitutes a basis in the register F of hash functions.

This result can be adapted to our analysis by performing the following changes.

$$\begin{aligned} C &\rightarrow K_B, \\ A &\rightarrow S^{(n)}, \\ E &\rightarrow E'L, \\ \rho_{CFE}^{\mathcal{R}} &\rightarrow \rho_{K_B F E' L | \Omega}, \\ \sigma_{AE} &\rightarrow \rho_{S^{(n)} E' L | \Omega}, \\ \frac{1}{|C|} \mathbb{1}_C \otimes \rho_{FE} &\rightarrow \bar{\tau}_{K_B F E' L | \Omega}, \end{aligned}$$

where $E' := [CI]^{(n)}E$, $\bar{\tau}$ is defined as in (9), ρ is the state after a concatenation of n channels constituting our protocol, and more explicitly, F denotes the family of universal₂ hash functions used for error correction and privacy amplification. Furthermore, $|K_B| = 2^\ell$ since the secret key is actually a binary register of length ℓ ; and L is the register of all the public information exchanged by Alice and Bob during error correction. Since we are interested in universal₂ families of hash functions, we set $\lambda = 1$ [21], Lemma 7]. Therefore, (B1) can be rewritten as

$$T(\rho_{K_B F E' L | \Omega}, \bar{\tau}_{K_B F E' L | \Omega}) \leq 2^{\frac{1-\alpha}{\alpha} (H_\alpha^\dagger(S^{(n)}|E'L)_\rho - \ell + 2)} \quad (\text{B2})$$

Let now define a scalar coefficient $\varepsilon_{PA} > 0$ verifying that

$$T(\rho_{K_B F E' L | \Omega}, \bar{\tau}_{K_B F E' L | \Omega}) \leq \frac{\varepsilon_{PA}}{\text{Pr}[\Omega]_\rho} \quad (\text{B3})$$

To ensure the above expression follows from the bound in Eq. (B2), it is sufficient to require its right-hand side to be no larger than $\varepsilon_{PA} \backslash \text{Pr}[\Omega]_\rho$. Solving this condition for the key length and considering that $(1-\alpha)/\alpha < 0$, we have

$$\begin{aligned} \ell &\leq H_\alpha^\dagger(S^{(n)}|E'L)_\rho - \frac{\alpha}{\alpha-1} \log \left(\frac{1}{\varepsilon_{PA}} \right) \\ &\quad + \frac{\alpha}{\alpha-1} \log \left(\frac{1}{\text{Pr}[\Omega]_\rho} \right) + 2. \end{aligned} \quad (\text{B4})$$

Next, we split off register L using [67], Lemma 5.14, Lemma 5.15] and the notation $\text{leak}_{EC} := \log(|L|)$, such that

$$H_\alpha^\dagger(S^{(n)}|E'L)_\rho \geq H_\alpha^\dagger(S^{(n)}|E')_\rho - \text{leak}_{EC}. \quad (\text{B5})$$

Thus, the inequality in (B4) can be written as

$$\begin{aligned} \ell &\leq H_\alpha^\dagger(S^{(n)}|E')_\rho - \frac{\alpha}{\alpha-1} \log \left(\frac{1}{\varepsilon_{PA}} \right) \\ &\quad + \frac{\alpha}{\alpha-1} \log \left(\frac{1}{\text{Pr}[\Omega]_\rho} \right) - \text{leak}_{EC} + 2. \end{aligned} \quad (\text{B6})$$

We can now apply the MEAT following Proposition A.2, such that the final expression of the key length is

$$\ell \leq nh_\alpha - \frac{\alpha}{\alpha-1} \log\left(\frac{1}{\varepsilon_{\text{PA}}}\right) - \text{leak}_{\text{EC}} + 2, \quad (\text{B7})$$

where h_α could be equal to $h_\alpha^\uparrow$ or $h_\alpha^\downarrow$. Thus, we can prove the security of Theorem III.2.

Proof. Let us consider the state after the execution of the protocol $\rho_{K_A K_B F E' L}$ and an ideal state $\bar{\tau}_{K_A K_B F E' L}$. We calculate its trace distance, noting that both states are the same when the protocol aborts (albeit no key rate is produced then),

$$\begin{aligned} T(\rho_{K_A K_B F E' L}, \bar{\tau}_{K_A K_B F E' L}) \\ = \Pr[\Omega]_\rho T(\rho_{K_A K_B F E' L|\Omega}, \bar{\tau}_{K_A K_B F E' L|\Omega}). \end{aligned} \quad (\text{B8})$$

Then, we bound the resulting trace distance as follows

$$\begin{aligned} T(\rho_{K_A K_B F E' L|\Omega}, \bar{\tau}_{K_A K_B F E' L|\Omega}) \\ \leq T(\rho_{K_A K_B F E' L|\Omega}, \rho_{K_A K_B F E' L|\Omega \wedge K_A = K_B}) \\ + T(\rho_{K_A K_B F E' L|\Omega \wedge K_A = K_B}, \bar{\tau}_{K_A K_B F E' L|\Omega}) \\ \leq \frac{\varepsilon_{\text{EC}}}{\Pr[\Omega]_\rho} + T(\rho_{K_B F E' L|\Omega}, \bar{\tau}_{K_B F E' L|\Omega}). \end{aligned} \quad (\text{B9})$$

We applied the triangle inequality in the first line and our hypothesis of ε_{EC} -correctness in the second (which is eventually reflected in the value of leak_{EC}), such that we can trace out register K_A in the second line. For the final trace distance, we may assume the hypothetical value for $\varepsilon_{\text{PA}} \in (0, 1)$ in (B3). Thus,

$$\Pr[\Omega]_\rho T(\rho_{K_B F E' L|\Omega}, \bar{\tau}_{K_B F E' L|\Omega}) \leq \varepsilon_{\text{PA}}. \quad (\text{B10})$$

Using Proposition B.1, we certify that the protocol is $(\varepsilon_{\text{EC}} + \varepsilon_{\text{PA}})$ -secret, as long as the resulting secret key has a length bounded by (B7). For any choice of parameters, the protocol verifies

$$\Pr[\Omega]_\rho T(\rho_{K_A K_B F E' L|\Omega}, \bar{\tau}_{K_A K_B F E' L|\Omega}) \leq \varepsilon_{\text{EC}} + \varepsilon_{\text{PA}}, \quad (\text{B11})$$

which constitutes our security statement, in accordance with (8). ■

APPENDIX C: VARIABLE-LENGTH SECRET KEY RATES VIA CONIC OPTIMIZATION

In this section, we develop a conic approach for the variable-length secret key rate for qubit BB84 using the framework described in Ref. [7]. To do so, we first introduce a tradeoff function $f: \mathcal{C} \rightarrow \mathbb{R}$, which allows us to express the f -weighted Rényi entropy [10]

$$\begin{aligned} H_\alpha^{\uparrow, f}(S|CIE)_{\mathcal{M}(\omega)} \\ = \frac{\alpha}{1-\alpha} \log \left(\sum_{c \in \mathcal{C}} p(c) 2^{\frac{1-\alpha}{\alpha} (-f(c) + H_\alpha^{\uparrow}(S|IE))_{\mathcal{M}(\omega)|c}} \right). \end{aligned} \quad (\text{C1})$$

This function is necessary to define κ as an $H_\alpha^{\uparrow, f}$ -normalization constant corresponding to the set of all states that can be produced via a single-round map $\mathcal{M}$, such that

$$\begin{aligned} \kappa &:= \inf_{\omega_{AB} \in \mathcal{D}(AB)} H_\alpha^{\uparrow, f}(S|CIE)_{\mathcal{M}(\omega)}, \\ \text{s.t. } \text{Tr}_B[\omega_{AB}] &= \sigma_A. \end{aligned} \quad (\text{C2})$$

Based on this, the security proof given by Ref. [7, Theorem 12] establishes that a variable-length protocol is secure when the final key length ℓ_{var} is bounded by a function $\hat{f}_{\text{prot}}(c_1^n) = \sum_i (f(c_i) + \kappa)$. As further discussed in Ref. [7, Sec. VI], rewriting this function in terms of the observed frequency vector and matching its expectation to a *fixed* honest probability distribution q^h , eventually provides an average secret key rate

$$\mathbb{E}[\ell_{\text{var}}] \geq n(f \cdot q^h + \kappa) - \text{leak}_{\text{EC}} - \frac{\alpha}{\alpha-1} \log \frac{1}{\varepsilon_{\text{PA}}} + 2, \quad (\text{C3})$$

where the leakage is provided by inequality (19), and is evaluated according to q^h . To find the optimal tradeoff function f , we introduce an auxiliary variable $\lambda \in \mathbb{P}_{\mathcal{C}}$. Then, similarly to Ref. [7], we formulate the following fixed-length optimization program

$$\begin{aligned} \inf_{\omega, \lambda} \frac{\alpha}{\alpha-1} D_{\text{KL}}(\lambda | \mathcal{M}(\omega)_C) + q^h(\perp) D_\beta(\mathcal{G}(\omega) \| \mathcal{Z} \circ \mathcal{G}(\omega)) \\ \text{s.t. } \omega_{AB} \geq 0, \quad \text{Tr}_B[\omega_{AB}] = \sigma_A \\ \sum_{c \in \mathcal{C}} \lambda(c) = 1, \quad \lambda \geq 0, \\ \lambda - q^h = 0. \end{aligned} \quad (\text{C4})$$

Using the techniques described in Sec. IV, we can reformulate this optimization as the conic program

$$\begin{aligned} \inf_{h_{\text{KL}}, h_{\text{QKD}}, u, \lambda, \omega} \frac{\alpha}{\alpha-1} [h_{\text{KL}} - p^K h_{\text{QKD}}] \\ \text{s.t. } \text{Tr}_B[\omega_{AB}] = \sigma_A, \quad \sum_{c \in \mathcal{C}} \lambda(c) = 1, \\ (u, \omega_{AB}) \in \mathcal{K}_{\text{FastRényiQKD}}^{\beta, \hat{\mathcal{G}}, \hat{\mathcal{Z}}\hat{\mathcal{G}}, S} \\ (h_{\text{KL}}, \lambda, \mathcal{M}(\omega)_C) \in \mathcal{K}_{\text{KL}} \\ (h_{\text{QKD}}, 1, s_\beta u + 1 - \text{Tr}[\mathcal{G}^\top(\omega)]) \in \mathcal{K}_{\log} \\ \lambda - q^h = 0. \end{aligned} \quad (\text{C5})$$

As stated in Ref. [[10], Lemma 4.12], the dual value of the last constraint provides an optimal choice for f , which can be readily introduced in the optimization for κ . Before continuing, note that f -weighted Rényi entropy can be lower-bounded as

$$\begin{aligned}
 H_{\alpha}^{\uparrow f}(S|CIE)_{\rho} &= \frac{\alpha}{1-\alpha} \log \left(\sum_{c \in \mathcal{C}} p(c) 2^{\frac{\alpha-1}{\alpha} f(c)} \right. \\
 &\quad \left. + p^K 2^{\frac{1-\alpha}{\alpha} (-f(\perp) + H_{\alpha}^{\uparrow}(S|IE)_{\mathcal{M}(\omega)_{\perp}})} \right) \\
 &\geq \frac{\alpha}{1-\alpha} \log \left(\sum_{c \in \mathcal{C}} p(c) 2^{\frac{\alpha-1}{\alpha} f(c)} \right. \\
 &\quad \left. + p^K 2^{\frac{\alpha-1}{\alpha} f(\perp)} \left[\sum_{i \in \mathcal{I}} \Psi_{\rho}^i(\omega) \right] \right), \tag{C6}
 \end{aligned}$$

where in the last line we used Eq. (A20) from Remark 1. In the particular case of qubit BB84, we have $\tilde{\mathcal{I}} = \{\top, \perp\}$. Using then the description in Sec. III C, we can explicitly write the optimization of κ in its conic form via the FastRényiQKD cone

$$\begin{aligned}
 \kappa &\geq \inf_{u, \omega} \frac{\alpha}{1-\alpha} \log \left((1-p^K) \sum_{c \in \mathcal{C}} \text{Tr}[\Gamma^c \omega_{AB}] 2^{\frac{\alpha-1}{\alpha} f(c)} \right. \\
 &\quad \left. + p^K 2^{\frac{\alpha-1}{\alpha} f(\perp)} (s_{\beta} u + 1 - \text{Tr}[\mathcal{G}^{\top}(\omega)]) \right) \\
 &\quad \text{s.t. } \text{Tr}_B[\omega_{AB}] = \sigma_A \\
 &\quad (u, \omega_{AB}) \in \mathcal{K}_{\text{FastRényiQKD}}^{\beta, \hat{\mathcal{G}}, \hat{\mathcal{Z}}, S}
 \end{aligned} \tag{C7}$$

where $\{\Gamma^c\}_{c \in \mathcal{C}}$ succinctly denotes all the POVM elements used by Alice and Bob in parameter estimation rounds, weighted according to the probabilities $(p^K, 1-p^K)$ of Bob's choices.

APPENDIX D: PROPERTIES OF S

As noted in Sec. II, we require $\text{supp}(\rho) \subseteq \text{supp}(\sigma)$ in our definition of the sandwiched Rényi relative entropy. This is not necessary for the entropy itself to be well-defined in our range of interest, $\alpha \in [1/2, 1)$, but we require it in order to prove that S is an isometry, which is needed for the derivatives to be well-defined and the proof of self-concordance.

In terms of our maps $\mathcal{G}, \mathcal{Z}$, this means that $\text{supp}(\mathcal{G}) \subseteq \text{supp}(\mathcal{Z})$, where

$$\begin{aligned}
 \text{supp}(\mathcal{G}) &= \text{cl}\{|\phi\rangle; \langle\phi|\mathcal{G}(V\omega'V^{\dagger})|\phi\rangle > 0 \ \forall \ \omega' \succ 0\} \\
 &= \{W_{\mathcal{G}}\mathbf{x}; \mathbf{x} \in \mathbb{C}^k\}, \tag{D1a}
 \end{aligned}$$

$$\begin{aligned}
 \text{supp}(\mathcal{Z}) &= \text{cl}\{|\phi\rangle; \langle\phi|\mathcal{Z}(\psi)|\phi\rangle > 0 \ \forall \ \psi \succ 0\} \\
 &= \{W_{\mathcal{Z}}\mathbf{y}; \mathbf{y} \in \mathbb{C}^n\}, \tag{D1b}
 \end{aligned}$$

Proposition D.1. $S = W_{\mathcal{Z}}^{\dagger} W_{\mathcal{G}}$ is an isometry.

Proof. Let $\mathbf{x} \in \mathbb{C}^k$. $W_{\mathcal{G}}\mathbf{x} \in \text{supp}(\mathcal{G})$; since $\text{supp}(\mathcal{G}) \subseteq \text{supp}(\mathcal{Z})$, there exists $\mathbf{y} \in \mathbb{C}^n$ such that

$$W_{\mathcal{G}}\mathbf{x} = W_{\mathcal{Z}}\mathbf{y}. \tag{D2}$$

On the one hand, multiplying both sides of (D2) by $W_{\mathcal{G}}^{\dagger}$,

$$\mathbf{x} = W_{\mathcal{G}}^{\dagger} W_{\mathcal{Z}}\mathbf{y}. \tag{D3}$$

On the other hand, multiplying both sides of (D2) by $W_{\mathcal{Z}}^{\dagger}$,

$$W_{\mathcal{Z}}^{\dagger} W_{\mathcal{G}}\mathbf{x} = \mathbf{y}. \tag{D4}$$

Then, replacing (D4) in (D3),

$$\mathbf{x} = W_{\mathcal{G}}^{\dagger} W_{\mathcal{Z}} W_{\mathcal{Z}}^{\dagger} W_{\mathcal{G}}\mathbf{x} = S^{\dagger} S\mathbf{x}. \tag{D5}$$

As it is true for all $\mathbf{x} \in \mathbb{C}^k$, it must be $S^{\dagger} S = \mathbb{1}_k$. ■

The proof for the $\mathcal{G}, \mathcal{Z} \circ \mathcal{G}$ case follows by replacing $W_{\mathcal{Z}}$ with $W_{\mathcal{Z}\mathcal{G}}$. We will also check that the map $X \mapsto S^{\dagger} X S$ is strictly positive, as we will need it to prove that the barrier function is self-concordant:

Proposition D.2. The map $X \mapsto S^{\dagger} X S$ is strictly positive.

Proof. First note that this map is positive, as $Y \geq 0$ iff $Y = K^{\dagger} K$, and therefore $S^{\dagger} Y S = S^{\dagger} K^{\dagger} K S \geq 0$. Now note that $Z \succ 0$ iff there exists $\varepsilon > 0$ such that $Z - \varepsilon \mathbb{1} \geq 0$. Then

$$0 \leq S^{\dagger} (Z - \varepsilon \mathbb{1}) S = S^{\dagger} Z S - \varepsilon \mathbb{1}, \tag{D6}$$

which implies that $S^{\dagger} Z S \succ 0$. ■

APPENDIX E: SELF-CONCORDANCE

In this section, we adapt the proof given in [18] to prove that the barrier functions (33) and (34) are *logarithmically homogeneous self-concordant barriers* (see the definition below) for their respective cones, (31) and (32). In particular, we are interested in $\alpha \in [1/2, 1)$, and therefore, $s_{\alpha} = \text{sign}(\alpha - 1) = -1$.

Definition E.1. [69], Definitions 1.3.2, 5.1.1], [70], Definition 2.3.2] Let $\mathcal{V}$ be a finite-dimensional real vector space and let $\mathcal{K} \subset \mathcal{V}$ a closed convex cone. A ν -logarithmically homogeneous self-concordant barrier for $\mathcal{K}$ is a $\mathcal{C}^3$ function $f: \text{int } \mathcal{K} \rightarrow \mathbb{R}$ such that

(i) For all $x_0 \in \partial \mathcal{K}$.

$$\lim_{x \rightarrow x_0} f(x) = +\infty. \tag{E1}$$

(ii) f is ν -logarithmically homogeneous:

$$f(\lambda x) = f(x) - \nu \log(\lambda) \tag{E2}$$

for all $x \in \text{int } \mathcal{K}$, for all $\lambda > 0$.

(iii) f is self-concordant:

$$|D^3 f(x)[h, h, h]| \leq 2(D^2 f(x)[h, h])^{3/2} \quad (\text{E3})$$

for all $x \in \text{int } \mathcal{K}$, for all $h \in \mathcal{V}$.

In the proofs of self-concordance of our barriers, we use the following concepts:

1. Conic monotony and conic convexity. Compatibility

Remember that a closed convex cone $\mathcal{K} \subset \mathcal{V}$ in a real vector space defines a partial ordering in $\mathcal{V}$ given by:

$$x \preceq_{\mathcal{K}} y \Leftrightarrow y - x \in \mathcal{K}.$$

Definition E.2. [[18], Sect. 2.1] Let $\mathcal{V}, \mathcal{V}'$ be finite-dimensional real vector spaces and let $\mathcal{K} \subset \mathcal{V}, \mathcal{K}' \subset \mathcal{V}'$ closed convex cones. A map $f: \mathcal{V} \rightarrow \mathcal{V}'$ is $(\mathcal{K}, \mathcal{K}')$ -monotone if for all $x, y \in \mathcal{V}$ such that

$$x \preceq_{\mathcal{K}} y$$

then,

$$f(x) \preceq_{\mathcal{K}'} f(y).$$

Definition E.3. [[18], Sec. 2.1] Let $\mathcal{U}, \mathcal{V}$ be finite-dimensional real vector spaces and let $\mathcal{K} \subset \mathcal{V}$ a closed convex cone. A map $f: \mathcal{U} \rightarrow \mathcal{V}$ is $\mathcal{K}$ -convex if for all $x, y \in \mathcal{U}, \lambda \in [0, 1]$,

$$f(\lambda x + (1 - \lambda)y) \preceq_{\mathcal{K}} \lambda f(x) + (1 - \lambda)f(y). \quad (\text{E4})$$

A map $f: \mathcal{U} \rightarrow \mathcal{V}$ is $\mathcal{K}$ -concave if $-f$ is $\mathcal{K}$ -convex.

To define operator monotony and operator convexity, we establish the notation

$$\mathbb{H}_{\sigma \subset I}^d := \{X \in \mathbb{H}^d; \sigma(X) \subset I\} \quad (\text{E5})$$

for an interval $I \subset \mathbb{R}$, where $\sigma(X)$ denotes the set of eigenvalues of X .

Definition E.4. A real function $f: (a, b) \subset \mathbb{R} \rightarrow \mathbb{R}$ is operator-monotone if for all $d \in \mathbb{N}$, for all $X, Y \in \mathbb{H}_{\sigma \subset (a, b)}^d$ with $X \preceq Y$,

$$f(X) \preceq f(Y).$$

Lemma E.5. Let $I \subset \mathbb{R}$ be an interval. The set $\mathbb{H}_{\sigma \subset I}^d$ is convex.

Proof. Let $X, Y \in \mathbb{H}_{\sigma \subset I}^d$ and $\lambda \in [0, 1]$. Let $Z = \lambda X + (1 - \lambda)Y$. As a consequence of Weyl's theorem ([71], Theorem 4.3.1)],

$$\left. \begin{aligned} \min \sigma(Z) &\geq \lambda \min \sigma(X) + (1 - \lambda) \min \sigma(Y) \\ \max \sigma(Z) &\leq \lambda \max \sigma(X) + (1 - \lambda) \max \sigma(Y) \end{aligned} \right\}, \quad (\text{E6})$$

where

$$\min \sigma(X), \min \sigma(Y), \max \sigma(X), \max \sigma(Y) \in I. \quad (\text{E7})$$

We conclude that $[\min \sigma(Z), \max \sigma(Z)] \subset I$, so $\sigma(Z) \subset I$. ■

Definition E.6. A real function $f: (a, b) \subset \mathbb{R}$ is operator-convex if for all $d \in \mathbb{N}$, for all $X, Y \in \mathbb{H}_{\sigma \subset (a, b)}^d, \lambda \in [0, 1]$,

$$f(\lambda X + (1 - \lambda)Y) \preceq \lambda f(X) + (1 - \lambda)f(Y) \quad (\text{E8})$$

A function $f: (a, b) \subset \mathbb{R}$ is operator concave if $-f$ is operator-convex.

Definition E.7. [[70], Definition 5.1.1], [18], Definition 2.3]. Let $\mathcal{U}, \mathcal{V}$ be finite-dimensional real vector spaces and let $\mathcal{K} \subset \mathcal{V}$ a closed convex cone. Let $C \subset \mathcal{U}$ be a closed convex subset and $f: \text{int } C \rightarrow \mathcal{V}$ a $\mathcal{C}^3$ function. Given a real number $\beta > 0$, f is $(\mathcal{K}, \beta)$ -compatible with C if

- (i) f is $\mathcal{K}$ -concave.
- (ii) For all $x \in \text{int } C$, for all $h \in \mathcal{U}$ such that $x \pm h \in C$,

$$D^3 f(x)[h, h, h] \preceq_{\mathcal{K}} -3\beta D^2 f(x)[h, h]. \quad (\text{E9})$$

Definition E.8. [[70], Definition 5.1.2]. Let $f: \text{int } C \rightarrow \mathcal{V}$ verifying hypothesis of Definition E.7 and let $G: \text{int } C \rightarrow \mathbb{R}$ a self-concordant barrier for C . Given a real number $\beta > 0$, f is $(\mathcal{K}, \beta)$ -compatible with the barrier G if

- (i) f is $\mathcal{K}$ -concave.
- (ii) For all $x \in \text{int } C$, for all $h \in \mathcal{U}$ such that $x \pm h \in C$,

$$D^3 f(x)[h, h, h] \preceq_{\mathcal{K}} -3\beta D^2 f(x)[h, h](D^2 G(x)[h, h])^{\frac{1}{2}}. \quad (\text{E10})$$

Concepts in Definitions E.7 and E.8 are closely related, as shown in [[70], Remark 5.1.2]:

Proposition E.9. Let $f: \text{int } C \rightarrow \mathcal{V}$ verifying hypothesis of Definitions E.7 and E.8, and let $\beta > 0$.

- (i) If f is $(\mathcal{K}, \beta)$ -compatible with C , it is $(\mathcal{K}, \beta)$ -compatible with any barrier of C .
- (ii) Conversely, if f is $(\mathcal{K}, \beta)$ -compatible with a θ -self-concordant barrier $G: \text{int } C \rightarrow \mathbb{R}$ for C , it is $(\mathcal{K}, (3\theta + 1)\beta)$ -compatible with C .

What we want now is to prove that the function

$$\Psi_{a, S}(G, Z) := \|Z^{\frac{1-a}{2a}} S G^{\frac{1}{2}}\|_{2a}^{2a} = \text{Tr} \left[(G^{\frac{1}{2}} S^{\dagger} Z^{\frac{1-a}{2a}} S G^{\frac{1}{2}})^a \right] \quad (\text{E11})$$

is $(\mathbb{R}_{\geq}, 1)$ -compatible so that we can prove that the functions (28) and (30) also are. We will do that by doing the composition with the appropriate linear map applying the following lemma:

Lemma E.10. [18], lemma 2.5], [70], lemma 5.1.3 (iii)]. Let $\mathcal{U}, \mathcal{V}$ be finite-dimensional real vector spaces and let $C \subset \mathcal{V}$ be a closed convex subset. Let $f: \text{int } C \rightarrow \mathbb{H}^n$ be a

$\mathcal{C}^3\mathbb{H}_{\geq}^n$ -concave function. Suppose f is $(\mathbb{H}_{\geq}^n, \beta)$ -compatible with C , $\beta > 0$.

- (i) Let $\mathcal{L}: \mathbb{H}^n \rightarrow \mathbb{H}^m$ a positive linear map. Then, $\mathcal{L} \circ f: \text{int } C \rightarrow \mathbb{H}^n \rightarrow \mathbb{H}^m$ is $(\mathbb{H}_{\geq}^m, \beta)$ -compatible with C .
- (ii) Let $\varphi: \mathcal{U} \rightarrow \mathcal{V}$ be an affine map such that $\text{int } C \cap \text{Im } \varphi \neq \emptyset$. Then, $f \circ \varphi: \varphi^{-1}(\text{int } C) \rightarrow \text{int } C \rightarrow \mathbb{H}^n$ is $(\mathbb{H}_{\geq}^n, \beta)$ -compatible with $\varphi^{-1}(C)$.

2. $(\mathbb{R}_{\geq}, 1)$ -compatibility

We are proving the $(\mathbb{R}_{\geq}, 1)$ -compatibility of (E11) this way:

Theorem E.11. [[18], Theorem 3.1] Let $\mathcal{U}, \mathcal{V}$ be finite-dimensional real vector spaces. Let $C \subset \mathcal{U}$ be a closed convex set and $\mathcal{K} \subset \mathcal{V}$ a closed convex cone. Let $\Phi: \text{int } C \rightarrow \mathcal{V}$ be a $\mathcal{C}^3$ function. Let $\mathcal{K}^*$ be the dual cone, and suppose that for all $z \in \mathcal{K}^*$ and for all $x \in \text{int } C$, $h \in \mathcal{U}$ such that $x \pm h \in \text{int } C$, the function

$$\begin{aligned} \phi: (-1, 1) &\rightarrow \mathbb{R} \\ t &\mapsto \langle z, \Phi(x + th) \rangle \end{aligned} \quad (\text{E12})$$

is operator concave in $(-1, 1)$. Then $\Phi: \text{int } C \rightarrow \mathcal{V}$ is $(\mathcal{K}, 1)$ -compatible with C .

Let

$$\begin{aligned} \psi: (-1, 1) &\rightarrow \mathbb{R} \\ t &\mapsto \Psi_{\alpha, s}(G + tH, Z + tK) \end{aligned} \quad (\text{E13})$$

where $(G, Z) \in \mathbb{H}_{\geq}^k \times \mathbb{H}_{\geq}^n$, $(H, K) \in \mathbb{H}^k \times \mathbb{H}^n$ are such that $(G \pm H, Z \pm K) \in \mathbb{H}_{\geq}^k \times \mathbb{H}_{\geq}^n$. The idea is to prove that ψ is operator concave in $(-1, 1)$ so that we can apply Theorem E.11.

For the following lemma, we define the the upper half-plane $\mathbb{C}^+ := \{z \in \mathbb{C}; \text{Im}(z) > 0\}$.

Lemma E.12. Let

$$\begin{aligned} \hat{\psi}: (1, +\infty) &\rightarrow \mathbb{R} \\ t &\mapsto t\psi\left(\frac{1}{t}\right) \end{aligned} \quad (\text{E14})$$

$\hat{\psi}$ has an analytic continuation to $\mathbb{C}^+$ such that $\hat{\psi}(\mathbb{C}^+) \subset \mathbb{C}^+$.

Proof. Note that this is a particular case of [[47], Theorem 2.1 (1)] with $p = \frac{1-\alpha}{\alpha}$, $q = 1$, $s = \alpha$, $\Psi(X) = S^\dagger X S$, and Φ the identity map. First of all, let us write $\hat{\psi}$:

$$\hat{\psi}(t) = \text{Tr}[(tG + H)^{\frac{1}{2}} S^\dagger (tZ + K)^{\frac{1-\alpha}{\alpha}} S (tG + H)^{\frac{1}{2}}]^\alpha. \quad (\text{E15})$$

Let us remember the Cartesian decomposition ([72], Sec. 4.4]) of a matrix $C \in \mathbb{C}^{d \times d}$: the Hermitian matrices

$$\begin{aligned} \text{Re}(C) &= \frac{C + C^\dagger}{2}, \\ \text{Im}(C) &= \frac{C - C^\dagger}{2i}, \end{aligned} \quad (\text{E16})$$

are called real part and imaginary part of C , respectively. In this context, $C = \text{Re}(C) + i\text{Im}(C)$ and they are the unique decomposition of $C = A + iB$ with A, B Hermitian matrices.

We define the set ([73], Sec. 1])

$$\mathcal{I}_d^+ := \{A + iB; A, B \in \mathbb{H}^d, B \succ 0\} \subset \mathbb{C}^{d \times d}. \quad (\text{E17})$$

If $c \in \mathbb{C}^+$, $cG + H \in \mathcal{I}_k^+$ and $cZ + K \in \mathcal{I}_n^+$. By [[73], lemma 1.2], then $(cG + H)^{\frac{1}{2}} \in \mathcal{I}_k^+$ and $(cZ + K)^{\frac{1-\alpha}{\alpha}} \in \mathcal{I}_n^+$, because $0 < \frac{1}{2}, \frac{1-\alpha}{\alpha} \leq 1$. As the map $X \mapsto S^\dagger X S$ is strictly positive (see proposition D.2), $S^\dagger (tZ + K)^{\frac{1-\alpha}{\alpha}} S \in \mathcal{I}_n^+$. Let us call

$$F(t) = (tG + H)^{\frac{1}{2}} S^\dagger (tZ + K)^{\frac{1-\alpha}{\alpha}} S (tG + H)^{\frac{1}{2}} \quad (\text{E18})$$

and $\gamma = \frac{1-\alpha}{\alpha} + 1 = (1/\alpha) \in (1, 2] \subset (0, 2]$. By [[47], proof of Theorem 2.1, Eq. (2.5)], if $\sigma(F(c))$ is the set of the eigenvalues of $F(c)$, $\sigma(F(c)) \subset \Gamma_\gamma$ for $c \in \mathbb{C}^+$, where

$$\Gamma_\Theta = \{re^{i\theta}; r > 0, 0 < \theta < \Theta\}, \quad \Theta \in (0, 2\pi). \quad (\text{E19})$$

Therefore, $\sigma(F(c)^\alpha) \subset \Gamma_\pi = \mathcal{I}_d^+$ (remember that $\gamma = (1/\alpha)$). Then, $\text{Tr}[(F(c)^\alpha)] \in \mathcal{I}_d^+$. ■

Once we have lemma E.12, by Lowner's theorem ([18], lemma 2.1]), we conclude that $\hat{\psi}$ is operator-monotone in $(1, +\infty)$. Now we can get that $\hat{\psi}$ is operator concave in $(0, 1)$ by [18], lemma 4.3], since

$$\hat{f}(t) = t\hat{f}\left(\frac{1}{t}\right) = t\frac{1}{t}f\left(\frac{1}{t}\right) = f(t). \quad (\text{E20})$$

Lemma E.13. Let ϕ the function (E12) in the conditions of Theorem E.11. If ϕ is operator concave in $(0, 1)$, it is also operator concave in $(-1, 0)$.

Proof. Let us recall ϕ_+ the function (E12) ϕ . Because of the conditions of Theorem E.11, the function (E12) is well defined in $(-1, 1)$ swapping h for $-h$. We call ϕ_- this new function.

$$\begin{aligned} \phi_-: (-1, 0) &\rightarrow (0, 1) \xrightarrow{\phi_+} \mathbb{R} \\ t &\mapsto -t \mapsto \phi_+(-t) \end{aligned}$$

Let $X, Y \in \mathbb{H}^d$ whose eigenvalues are in $(-1, 0)$. By hypothesis, ϕ_+ is operator concave in $(0, 1)$. Then, given $\lambda \in [0, 1]$

$$\begin{aligned} \phi_-(\lambda X + (1 - \lambda)Y) &= \phi_+(-(\lambda X + (1 - \lambda)Y)) \\ &= \phi_+(\lambda(-X) + (1 - \lambda)(-Y)) \\ &\geq \lambda\phi_+(-X) + (1 - \lambda)\phi_+(-Y) \\ &= \lambda\phi_-(X) + (1 - \lambda)\phi_-(Y). \end{aligned}$$

■

By lemma E.13, the function (E13) ψ is also operator concave in $(-1, 0)$. Using [18], lemma 4.4] for all $d \in \mathbb{N}$,

we get that ψ is operator concave in $(-1, 1)$. Finally, applying Theorem E.11, we conclude that the function (E11) is $(\mathbb{R}_{\geq}, 1)$ -compatible with $\mathbb{H}_{\geq}^k \times \mathbb{H}_{\geq}^m$ for $\alpha \in [\frac{1}{2}, 1)$.

3. Proof of self-concordance for the RényiQKD cone

Let us rewrite the function (28) in terms of (E11):

$$\hat{\Psi}_\alpha(\rho, \sigma) = \Psi_{\alpha, S} \circ \mathcal{L}_{\hat{G}, \hat{Z}}(\rho, \sigma), \quad (\text{E21})$$

where $\mathcal{L}_{\hat{G}, \hat{Z}}(\rho, \sigma) = (\hat{G}(\rho), \hat{Z}(\sigma))$. Applying lemma E.10 (ii), $\hat{\Psi}_\alpha(\rho, \sigma)$ is $(\mathbb{R}_{\geq}, 1)$ -compatible with $\mathbb{H}_{\geq}^q \times \mathbb{H}_{\geq}^m$. Once we know that, we can prove that the function

$$f(u, \rho, \sigma) = -\log(\hat{\Psi}_\alpha(\rho, \sigma) - u) - \log \det(\rho) - \log \det(\sigma) \quad (\text{E22})$$

is a self-concordant barrier for $\text{cl hypo}(\hat{\Psi}_\alpha(\rho, \sigma))$ by [[18], lemma 2.4]:

Theorem E.14. Let $\mathcal{U}, \mathcal{V}$ be finite-dimensional real vector spaces. Let $C \subset \mathcal{U}$ be a closed convex set and $\mathcal{K} \subset \mathcal{V}$ a closed convex cone. Let $G: \text{int } C \subset \mathcal{U} \rightarrow \mathbb{R}$ be a $\mathcal{C}^2 \nu$ -SCB for C and $F: \text{int } C \subset \mathcal{U} \rightarrow \mathcal{V}(\mathcal{K}, \beta)$ -compatible with C , $\beta \geq 0$. Let $H: \text{int } \mathcal{K} \rightarrow \mathbb{R}$ be a η -SCB for $\mathcal{K}$. Then, the function

$$f: \text{int } C \times \mathcal{V} \rightarrow \mathbb{R} \cup \{+\infty\} \\ (x, y) \mapsto H(F(x) - y) + \beta^3 G(x)$$

is a $(\eta + \beta^3 \nu)$ -SCB for $\text{cl hypo } F$, where

$$\text{hypo } F = \{(y, x) \in \mathcal{W} \times \text{int } C; y \preceq_{\mathcal{K}} F(x)\}.$$

This is a particular case of [[69], Theorem 5.4.4], noticing proposition E.9 (i).

Corollary E.15. The function (E22) is a $(1 + q + m)$ -SCB for $\text{cl hypo}(\hat{\Psi}_\alpha(\rho, \sigma))$.

Proof. We apply Theorem E.14 with

- (i) $\mathcal{U} = \mathbb{H}^q \times \mathbb{H}^m$, $C = \mathbb{H}_{\geq}^q \times \mathbb{H}_{\geq}^m$, $G(\rho, \sigma) = -\log \det(\rho) - \log \det(\sigma)$, which is a $(q + m)$ -SCB for $\mathbb{H}_{\geq}^q \times \mathbb{H}_{\geq}^m$ ($\nu = q + m$).
- (ii) $\mathcal{V} = \mathbb{R}$, $\mathcal{K} = \mathbb{R}_{\geq}$.
- (iii) $F = \hat{\Psi}_\alpha: \mathbb{H}_{\geq}^q \times \mathbb{H}_{\geq}^m \rightarrow \mathbb{R}$, which is $(\mathbb{R}_{\geq}, 1)$ -compatible with $\mathbb{H}_{\geq}^q \times \mathbb{H}_{\geq}^m$ ($\beta = 1$).
- (iv) $H = -\log$, which is a 1-SCB for $\mathbb{R}_{\geq}$ ($\eta = 1$).

4. Proof of self-concordance for the FastRényiQKD cone

Following an analogous argument, we can rewrite the function (30) in terms of (E11):

$$\hat{\Psi}_\alpha(\rho) = \Psi_{\alpha, S} \circ \mathcal{L}_{\hat{G}, \hat{Z}\hat{G}}(\rho), \quad (\text{E23})$$

where $\mathcal{L}_{\hat{G}, \hat{Z}\hat{G}}(\rho) = (\hat{G}(\rho), \hat{Z}\hat{G}(\rho))$. Applying again lemma E.10 (ii), $\hat{\Psi}_\alpha(\rho)$ is $(\mathbb{R}_{\geq}, 1)$ -compatible with $\mathbb{H}_{\geq}^q$. We can prove then that the function

$$f(u, \rho) = -\log(\hat{\Psi}_\alpha(\rho) - u) - \log \det(\rho) - \log \det(\sigma) \quad (\text{E24})$$

is a self-concordant barrier for $\text{cl hypo}(\hat{\Psi}_\alpha(\rho))$:

Corollary E.16. The function (E24) is a $(1 + q)$ -SCB for $\text{cl hypo}(\hat{\Psi}_\alpha(\rho))$.

Proof. We apply Theorem E.14 with

- (i) $\mathcal{U} = \mathbb{H}^q$, $C = \mathbb{H}_{\geq}^q$, $G(\rho) = -\log \det(\rho)$, which is a q -SCB for $\mathbb{H}_{\geq}^q$ ($\nu = q$).
- (ii) $\mathcal{V} = \mathbb{R}$, $\mathcal{K} = \mathbb{R}_{\geq}$.
- (iii) $F = \hat{\Psi}_\alpha: \mathbb{H}_{\geq}^q \rightarrow \mathbb{R}$, which is $(\mathbb{R}_{\geq}, 1)$ -compatible with $\mathbb{H}_{\geq}^q$ ($\beta = 1$).
- (iv) $H = -\log$, which is a 1-SCB for $\mathbb{R}_{\geq}$ ($\eta = 1$).

APPENDIX F: DERIVATIVES

1. Background

Real functions can be extended to act on Hermitian matrices by diagonalizing the matrix and applying the function to each (real) eigenvalue. More precisely, let $f: \mathbb{R} \mapsto \mathbb{R}$ and $X \in \mathbb{H}^n$ with spectral decomposition $X = U \text{diag}(\vec{\lambda}) U^\dagger$, where $\vec{\lambda} \in \mathbb{R}^n$ is the vector of eigenvalues. Then we define $f(X)$ as

$$f(X) = U \text{diag}(f(\vec{\lambda})) U^\dagger, \quad (\text{F1})$$

where f acts on $\vec{\lambda}$ elementwise. We will call $f: \mathbb{H}^n \rightarrow \mathbb{H}^n$ a *spectral function*.

The barriers of our cones, Eqs. (33) and (34), are defined in terms of spectral functions. In order to differentiate spectral functions, we need to extend the usual rules of multivariable calculus to these functions. In this section, we collect some of these rules for convenience, but they can also be found in matrix analysis books like [74,75]. We will refer to these properties in the computation of the gradient, Hessian and third derivatives of the barrier.

First, let's recall some rules valid for general functions. The Fréchet derivative of the composition of two functions applied to a vector direction H obeys the chain rule:

$$D(f \circ g)(X)[H] = Df(g(X))[Dg(X)[H]], \quad (\text{F2})$$

and the derivative of a product obeys the product rule:

$$D(fg)(X)[H] = Df(X)[H]g(X) + f(X)Dg(X)[H]. \quad (\text{F3})$$

In particular, we will use it when the product is composed with the trace. Since the trace is a linear operator, its

differential is itself and we obtain

$$D(\text{Tr} \circ fg)(X)[H] = \langle Df(X)[H], g(X) \rangle + \langle f(X), Dg(X)[H] \rangle. \quad (\text{F4})$$

Turning our attention to spectral functions, the derivative is given by

$$Df(X)[H] = U(f^{[1]}(\Lambda) \odot (U^\dagger HU))U^\dagger, \quad (\text{F5})$$

where $X = U\Lambda U^\dagger$ is the diagonalization of X , $f^{[1]}$ are the first divided differences of f and $\odot$ is the Hadamard or elementwise matrix product.

Another important property is that the derivative of a spectral function composed with the trace is equal to:

$$D(\text{Tr} \circ f)(X)[H] = \langle H, f'(X) \rangle, \quad (\text{F6})$$

where $\langle A, B \rangle = \text{Tr}(AB^\dagger)$ is the Hilbert-Schmidt inner product. We will call this equation the *trace derivative*. We can combine the chain rule and the trace derivative to obtain:

$$D(\text{Tr} \circ f \circ g)(X)[H] = D(\text{Tr} \circ f)(g(X))[Dg(X)[H]] = \langle Dg(X)[H], f'(g(X)) \rangle. \quad (\text{F7})$$

These equations will be used extensively in the upcoming derivations. To compute the Hessian and third derivatives of the barriers, we will need higher-order Fréchet derivatives of spectral functions. The second-order derivative applied to two vector directions is given by

$$D^2f(X)[H, K] = U \left[\sum_{k=1}^n f_k^{[2]}(\Lambda) \odot (U^\dagger HU|k\rangle\langle k|U^\dagger KU + U^\dagger KU|k\rangle\langle k|U^\dagger HU) \right] U^\dagger, \quad (\text{F8})$$

where $X = U\Lambda U^\dagger$ and $f_k^{[2]}(\Lambda) = [f^{[2]}(\lambda_i, \lambda_j, \lambda_k)]_{i,j=1}^n$ (second divided differences of f). The third derivative applied twice to the same vector direction is

$$D^3f(X)[H, H, K] = U \left[2 \sum_{k,\ell=1}^n f_{k,\ell}^{[3]}(\Lambda) \odot [U^\dagger HU|k\rangle\langle k|U^\dagger HU|\ell\rangle\langle \ell|U^\dagger KU + U^\dagger HU|k\rangle\langle k|U^\dagger KU|\ell\rangle\langle \ell|U^\dagger HU + U^\dagger KU|k\rangle\langle k|U^\dagger HU|\ell\rangle\langle \ell|U^\dagger HU] \right] U^\dagger, \quad (\text{F9})$$

where $X = U\Lambda U^\dagger$ and $f_{k,\ell}^{[3]}(\Lambda) = [f^{[3]}(\lambda_i, \lambda_k, \lambda_\ell, \lambda_j)]_{i,j=1}^n$ (third divided differences of f).

Finally, to isolate some of the vector directions later on, we will use the property that Fréchet derivatives of spectral functions are self-adjoint:

$$\begin{aligned} \langle A, Df(X)[H] \rangle &= \langle A, U(f^{[1]}(\Lambda) \odot (U^\dagger HU))U^\dagger \rangle \\ &= \langle U^\dagger AU, f^{[1]}(\Lambda) \odot (U^\dagger HU) \rangle \\ &= \langle f^{[1]}(\Lambda) \odot (U^\dagger AU), U^\dagger HU \rangle \\ &= \langle U(f^{[1]}(\Lambda) \odot (U^\dagger AU))U^\dagger, H \rangle \end{aligned} \quad (\text{F10})$$

$$= \langle Df(X)[A], H \rangle. \quad (\text{F11})$$

By applying this property successive times, we can prove that higher-order derivatives are also self-adjoint:

$$\langle A, D^2f(X)[H, K] \rangle = \langle D^2f(X)[H, A], K \rangle, \quad (\text{F12})$$

$$\langle A, D^3f(X)[H, K, L] \rangle = \langle D^3f(X)[H, K, A], L \rangle. \quad (\text{F13})$$

2. RényiQKD cone

a. Gradient

The gradient of the barrier (33) is

$$\nabla_h f(h, \rho, \sigma) = -\frac{1}{z}, \quad (\text{F14})$$

$$\nabla_\rho f(h, \rho, \sigma) = \frac{s_\alpha}{z} \nabla_\rho \hat{\Psi}_\alpha(\rho, \sigma) - \rho^{-1}, \quad (\text{F15})$$

$$\nabla_\sigma f(h, \rho, \sigma) = \frac{s_\alpha}{z} \nabla_\sigma \hat{\Psi}_\alpha(\rho, \sigma) - \sigma^{-1}, \quad (\text{F16})$$

where $z = h - s_\alpha \hat{\Psi}_\alpha(\rho, \sigma)$ and $\hat{\Psi}_\alpha$ is given by Eq. (28).

For convenience let us define the intermediate variables $G := \hat{G}(\rho)$ and $Z := \hat{Z}(\sigma)$, and the functions $g(x) = x^\alpha$ and $h(x) = x^{\frac{1-\alpha}{\alpha}}$, so that we can write

$$\begin{aligned} \hat{\Psi}_\alpha(G, Z) &:= \hat{\Psi}_\alpha(\rho, \sigma) \\ &= \text{Tr}[g(G^{\frac{1}{2}} S^\dagger h(Z) S G^{\frac{1}{2}})] \\ &= \text{Tr}[g(Z_S^{\frac{1}{2}} G Z_S^{\frac{1}{2}})], \end{aligned} \quad (\text{F17})$$

where we defined $Z_S := S^\dagger h(Z)S$ for brevity. The two expressions we give for $\hat{\Psi}_\alpha(G, Z)$ come from Eq. (36), and as mentioned before, the operators appearing there are positive definite, which is necessary for the derivatives to be well-defined.

The chain rule (F2) gives us the derivatives of $\hat{\Psi}_\alpha$ in terms of the derivatives of $\hat{\Psi}_\alpha$:

$$D_\rho \hat{\Psi}_\alpha(\rho, \sigma)[H] = D_G \hat{\Psi}_\alpha(G, Z)[D_\rho G(\rho)[H]], \quad (\text{F18})$$

$$D_\sigma \hat{\Psi}_\alpha(\rho, \sigma)[H] = D_Z \hat{\Psi}_\alpha(G, Z)[D_\sigma Z(\sigma)[H]]. \quad (\text{F19})$$

Applying the chain rule and the trace derivative (F7) to the expression $\text{Tr}[g(Z_S^\dagger G Z_S^\dagger)]$ in Eq. (F17) gives us

$$\begin{aligned} D_G \hat{\Psi}_\alpha(G, Z)[H_G] &= \left\langle Z_S^\dagger H_G Z_S^\dagger, g'(Z_S^\dagger G Z_S^\dagger) \right\rangle \\ &= \left\langle H_G, Z_S^\dagger g'(Z_S^\dagger G Z_S^\dagger) Z_S^\dagger \right\rangle. \end{aligned} \quad (\text{F20})$$

Now, using the expression $\text{Tr}[g(G^\dagger S^\dagger h(Z) S G^\dagger)]$ from Eq. (F17) to compute the derivative with respect to Z , we obtain

$$\begin{aligned} D_Z \hat{\Psi}_\alpha(G, Z)[H_Z] &= \langle G^\dagger S^\dagger D h(Z)[H_Z] S G^\dagger, g'(G^\dagger Z_S G^\dagger) \rangle \\ &= \langle D h(Z)[H_Z], S G^\dagger g'(G^\dagger Z_S G^\dagger) G^\dagger S^\dagger \rangle \\ &= \langle H_Z, D h(Z)[S G^\dagger g'(G^\dagger Z_S G^\dagger) G^\dagger S^\dagger] \rangle, \end{aligned} \quad (\text{F21})$$

where in the last step we used the self-adjointness of the spectral derivative (F10). Now, substituting $G = \hat{G}(\rho)$, $Z = \hat{Z}(\sigma)$, $H_G = D_\rho G(\rho)[H_\rho] = \hat{G}(H_\rho)$ and $H_Z = D_\sigma Z(\sigma)[H_\sigma] = \hat{Z}(H_\sigma)$, we obtain

$$D_\rho \hat{\Psi}_\alpha(\rho, \sigma)[H_\rho] = \left\langle H_\rho, \hat{G}^\dagger (Z_S^\dagger g'(Z_S^\dagger G Z_S^\dagger) Z_S^\dagger) \right\rangle, \quad (\text{F22})$$

$$D_\sigma \hat{\Psi}_\alpha(\rho, \sigma)[H_\sigma] = \left\langle H_\sigma, \hat{Z}^\dagger (D h(Z)[S G^\dagger g'(G^\dagger Z_S G^\dagger) G^\dagger S^\dagger]) \right\rangle, \quad (\text{F23})$$

and therefore the gradient is

$$\nabla_\rho \hat{\Psi}_\alpha(\rho, \sigma) = \hat{G}^\dagger (Z_S^\dagger g'(Z_S^\dagger G Z_S^\dagger) Z_S^\dagger), \quad (\text{F24})$$

$$\nabla_\sigma \hat{\Psi}_\alpha(\rho, \sigma) = \hat{Z}^\dagger (D h(Z)[S G^\dagger g'(G^\dagger Z_S G^\dagger) G^\dagger S^\dagger]). \quad (\text{F25})$$

The gradient can be significantly simplified if we write it in terms of the singular value decomposition of $Z_S^\dagger G^\dagger = U \Lambda V^\dagger$:

$$\nabla_\rho \hat{\Psi}_\alpha(\rho, \sigma) = \hat{G}^\dagger (Z_S^\dagger U g'(\Lambda^2) U^\dagger Z_S^\dagger), \quad (\text{F26})$$

$$\nabla_\sigma \hat{\Psi}_\alpha(\rho, \sigma) = \hat{Z}^\dagger (D h(Z)[S G^\dagger V g'(\Lambda^2) V^\dagger G^\dagger S^\dagger]). \quad (\text{F27})$$

b. Hessian

The Hessian of the barrier (33) is

$$\nabla_{hh}^2 f(h, \rho, \sigma) = \frac{1}{z^2}, \quad (\text{F28})$$

$$\nabla_{h\rho}^2 f(h, \rho, \sigma) = -\frac{s_\alpha}{z^2} \nabla_\rho \hat{\Psi}_\alpha(\rho, \sigma), \quad (\text{F29})$$

$$\nabla_{h\sigma}^2 f(h, \rho, \sigma) = -\frac{s_\alpha}{z^2} \nabla_\sigma \hat{\Psi}_\alpha(\rho, \sigma), \quad (\text{F30})$$

$$\nabla_{\rho\rho}^2 f(h, \rho, \sigma) = H_\rho \mapsto \frac{1}{z^2} \langle \nabla_\rho \hat{\Psi}_\alpha(\rho, \sigma), H_\rho \rangle \nabla_\rho \hat{\Psi}_\alpha(\rho, \sigma) + \frac{s_\alpha}{z} \nabla_{\rho\rho}^2 \hat{\Psi}_\alpha(\rho, \sigma) H_\rho + \rho^{-1} H_\rho \rho^{-1}, \quad (\text{F31})$$

$$\nabla_{\sigma\sigma}^2 f(h, \rho, \sigma) = H_\sigma \mapsto \frac{1}{z^2} \langle \nabla_\sigma \hat{\Psi}_\alpha(\rho, \sigma), H_\sigma \rangle \nabla_\sigma \hat{\Psi}_\alpha(\rho, \sigma) + \frac{s_\alpha}{z} \nabla_{\sigma\sigma}^2 \hat{\Psi}_\alpha(\rho, \sigma) H_\sigma + \sigma^{-1} H_\sigma \sigma^{-1}, \quad (\text{F32})$$

$$\nabla_{\rho\sigma}^2 f(h, \rho, \sigma) = H_\sigma \mapsto \frac{1}{z^2} \langle \nabla_\sigma \hat{\Psi}_\alpha(\rho, \sigma), H_\sigma \rangle \nabla_\rho \hat{\Psi}_\alpha(\rho, \sigma) + \frac{s_\alpha}{z} \nabla_{\rho\sigma}^2 \hat{\Psi}_\alpha(\rho, \sigma) H_\sigma, \quad (\text{F33})$$

$$\nabla_{\sigma\rho}^2 f(h, \rho, \sigma) = H_\rho \mapsto \frac{1}{z^2} \langle \nabla_\rho \hat{\Psi}_\alpha(\rho, \sigma), H_\rho \rangle \nabla_\sigma \hat{\Psi}_\alpha(\rho, \sigma) + \frac{s_\alpha}{z} \nabla_{\sigma\rho}^2 \hat{\Psi}_\alpha(\rho, \sigma) H_\rho, \quad (\text{F34})$$

where $z = h - s_\alpha \hat{\Psi}_\alpha(\rho, \sigma)$.

Similarly to the derivation of the gradient, we obtain $D_{GG} \hat{\Psi}_\alpha(G, Z)$ by differentiating (F20) over G :

$$D_{GG}^2 \hat{\Psi}_\alpha(G, Z)[H_{G_1}, H_{G_2}] = \langle H_{G_1}, Z_S^\dagger D g'(Z_S^\dagger G Z_S^\dagger)[Z_S^\dagger H_{G_2} Z_S^\dagger] Z_S^\dagger \rangle \quad (\text{F35})$$

We obtain $D_{ZZ}^2 \hat{\Psi}_\alpha(G, Z)$ by differentiating (F21) over Z with the product rule (F4):

$$D_{ZZ}^2 \hat{\Psi}_\alpha(G, Z)[H_{Z_1}, H_{Z_2}] = \langle D^2 h(Z)[H_{Z_1}, H_{Z_2}], SG^{\frac{1}{2}} g'(G^{\frac{1}{2}} Z_S G^{\frac{1}{2}}) G^{\frac{1}{2}} S^\dagger \rangle \\ + \langle Dh(Z)[H_{Z_1}], SG^{\frac{1}{2}} Dg'(G^{\frac{1}{2}} Z_S G^{\frac{1}{2}})[G^{\frac{1}{2}} S^\dagger Dh(Z)[H_{Z_2}] SG^{\frac{1}{2}}] G^{\frac{1}{2}} S^\dagger \rangle. \quad (\text{F36})$$

Using He's Lemma 5.1 [18], we have that

$$D_Z \hat{\Psi}_\alpha(G, Z)[H_Z] = \langle Dh(Z)[H_Z], SG^{\frac{1}{2}} g'(G^{\frac{1}{2}} Z_S G^{\frac{1}{2}}) G^{\frac{1}{2}} S^\dagger \rangle = \langle Dh(Z)[H_Z], SZ_S^{-\frac{1}{2}} \tilde{g}(Z_S^{\frac{1}{2}} G Z_S^{\frac{1}{2}}) Z_S^{-\frac{1}{2}} S^\dagger \rangle, \quad (\text{F37})$$

where $\tilde{g}(x) = xg'(x)$. With this second expression we can obtain D_{ZG}^2 easily by differentiating it over G :

$$D_{ZG}^2 \hat{\Psi}_\alpha(G, Z)[H_{Z_1}, H_{G_2}] = \langle Dh(Z)[H_{Z_1}], SZ_S^{-\frac{1}{2}} D\tilde{g}(Z_S^{\frac{1}{2}} G Z_S^{\frac{1}{2}})[Z_S^{\frac{1}{2}} H_{G_2} Z_S^{\frac{1}{2}}] Z_S^{-\frac{1}{2}} S^\dagger \rangle. \quad (\text{F38})$$

Now we can substitute $G = \hat{\mathcal{G}}(\rho)$, $Z = \hat{\mathcal{Z}}(\sigma)$ and the vector directions $H_{G_i} = \hat{\mathcal{G}}(H_{\rho_i})$ and $H_{Z_i} = \hat{\mathcal{Z}}(H_{\sigma_i})$. To obtain the product of the Hessian times a vector we need to isolate the first vector directions, H_{ρ_1} and H_{σ_1} :

$$\nabla_{\rho\rho}^2 \hat{\Psi}_\alpha(\rho, \sigma) H_{\rho_2} = \hat{\mathcal{G}}^\dagger(Z_S^{\frac{1}{2}} Dg'(Z_S^{\frac{1}{2}} G Z_S^{\frac{1}{2}})[Z_S^{\frac{1}{2}} \hat{\mathcal{G}}(H_{\rho_2}) Z_S^{\frac{1}{2}}] Z_S^{\frac{1}{2}}), \quad (\text{F39})$$

$$\nabla_{\sigma\rho}^2 \hat{\Psi}_\alpha(\rho, \sigma) H_{\rho_2} = \hat{\mathcal{Z}}^\dagger(Dh(Z)[SZ_S^{-\frac{1}{2}} D\tilde{g}(Z_S^{\frac{1}{2}} G Z_S^{\frac{1}{2}})[Z_S^{\frac{1}{2}} \hat{\mathcal{G}}(H_{\rho_2}) Z_S^{\frac{1}{2}}] Z_S^{-\frac{1}{2}} S^\dagger]), \quad (\text{F40})$$

$$\nabla_{\rho\sigma}^2 \hat{\Psi}_\alpha(\rho, \sigma) H_{\sigma_2} = \hat{\mathcal{G}}^\dagger(Z_S^{\frac{1}{2}} D\tilde{g}(Z_S^{\frac{1}{2}} G Z_S^{\frac{1}{2}})[Z_S^{-\frac{1}{2}} S^\dagger Dh(Z)[\hat{\mathcal{Z}}(H_{\sigma_2})] SZ_S^{-\frac{1}{2}}] Z_S^{\frac{1}{2}}), \quad (\text{F41})$$

$$\nabla_{\sigma\sigma}^2 \hat{\Psi}_\alpha(\rho, \sigma) H_{\sigma_2} = \hat{\mathcal{Z}}^\dagger(D^2 h(Z)[\hat{\mathcal{Z}}(H_{\sigma_2}), SG^{\frac{1}{2}} g'(G^{\frac{1}{2}} Z_S G^{\frac{1}{2}}) G^{\frac{1}{2}} S^\dagger]) \\ + \hat{\mathcal{Z}}^\dagger(Dh(Z)[SG^{\frac{1}{2}} Dg'(G^{\frac{1}{2}} Z_S G^{\frac{1}{2}})[G^{\frac{1}{2}} S^\dagger Dh(Z)[\hat{\mathcal{Z}}(H_{\sigma_2})] SG^{\frac{1}{2}}] G^{\frac{1}{2}} S^\dagger]). \quad (\text{F42})$$

And we obtain the Hessian, $\nabla^2 \hat{\Psi}_\alpha$, by vectorizing (see Appendix A of [19]) the previous equations and thus isolating the remaining vector directions, H_{ρ_2} and H_{σ_2} .

c. Third derivatives

The third derivatives of the barrier (33) are

$$\nabla_{hhh}^3 f(h, \rho, \sigma) = -\frac{2}{z^3}, \quad (\text{F43})$$

$$\nabla_{hh\rho}^3 f(h, \rho, \sigma) = \frac{2s_\alpha}{z^3} \nabla_\rho \hat{\Psi}_\alpha(\rho, \sigma), \quad (\text{F44})$$

$$\nabla_{hh\sigma}^3 f(h, \rho, \sigma) = \frac{2s_\alpha}{z^3} \nabla_\sigma \hat{\Psi}_\alpha(\rho, \sigma), \quad (\text{F45})$$

$$\nabla_{h\rho\rho}^3 f(h, \rho, \sigma) = H_\rho \mapsto -\frac{2s_\alpha^2}{z^3} \langle \nabla_\rho \hat{\Psi}_\alpha(\rho, \sigma), H_\rho \rangle \nabla_\rho \hat{\Psi}_\alpha(\rho, \sigma) - \frac{s_\alpha}{z^2} \nabla_{\rho\rho}^2 \hat{\Psi}_\alpha(\rho, \sigma) H_\rho, \quad (\text{F46})$$

$$\nabla_{h\sigma\sigma}^3 f(h, \rho, \sigma) = H_\sigma \mapsto -\frac{2s_\alpha^2}{z^3} \langle \nabla_\sigma \hat{\Psi}_\alpha(\rho, \sigma), H_\sigma \rangle \nabla_\sigma \hat{\Psi}_\alpha(\rho, \sigma) - \frac{s_\alpha}{z^2} \nabla_{\sigma\sigma}^2 \hat{\Psi}_\alpha(\rho, \sigma) H_\sigma, \quad (\text{F47})$$

$$\nabla_{h\sigma\rho}^3 f(h, \rho, \sigma) = H_\rho \mapsto -\frac{2s_\alpha^2}{z^3} \langle \nabla_\rho \hat{\Psi}_\alpha(\rho, \sigma), H_\rho \rangle \nabla_\sigma \hat{\Psi}_\alpha(\rho, \sigma) - \frac{s_\alpha}{z^2} \nabla_{\sigma\rho}^2 \hat{\Psi}_\alpha(\rho, \sigma) H_\rho, \quad (\text{F48})$$

$$\begin{aligned}
\nabla_{\rho\rho\rho}^3 f(h, \rho, \sigma) &= (H_\rho, H_\rho) \mapsto \frac{2s_\alpha^3}{z^3} \langle \nabla_\rho \hat{\Psi}_\alpha(\rho, \sigma), H_\rho \rangle^2 \nabla_\rho \hat{\Psi}_\alpha(\rho, \sigma) \\
&+ \frac{s_\alpha^2}{z^2} \langle \nabla_{\rho\rho}^2 \hat{\Psi}_\alpha H_\rho, H_\rho \rangle \nabla_\rho \hat{\Psi}_\alpha(\rho, \sigma) + \frac{2s_\alpha^2}{z^2} \langle \nabla_\rho \hat{\Psi}_\alpha, H_\rho \rangle \nabla_{\rho\rho}^2 \hat{\Psi}_\alpha(\rho, \sigma) H_\rho \\
&+ \frac{s_\alpha}{z} \nabla_{\rho\rho\rho}^3 \hat{\Psi}_\alpha(\rho, \sigma) [\cdot, H_\rho, H_\rho] - 2\rho^{-1} H_\rho \rho^{-1} H_\rho \rho^{-1},
\end{aligned} \tag{F49}$$

$$\begin{aligned}
\nabla_{\sigma\sigma\sigma}^3 f(h, \rho, \sigma) &= (H_\sigma, H_\sigma) \mapsto \frac{2s_\alpha^3}{z^3} \langle \nabla_\sigma \hat{\Psi}_\alpha(\rho, \sigma), H_\sigma \rangle^2 \nabla_\sigma \hat{\Psi}_\alpha(\rho, \sigma) \\
&+ \frac{s_\alpha^2}{z^2} \langle \nabla_{\sigma\sigma}^2 \hat{\Psi}_\alpha H_\sigma, H_\sigma \rangle \nabla_\sigma \hat{\Psi}_\alpha(\rho, \sigma) + \frac{2s_\alpha^2}{z^2} \langle \nabla_\sigma \hat{\Psi}_\alpha, H_\sigma \rangle \nabla_{\sigma\sigma}^2 \hat{\Psi}_\alpha(\rho, \sigma) H_\sigma \\
&+ \frac{s_\alpha}{z} \nabla_{\sigma\sigma\sigma}^3 \hat{\Psi}_\alpha(\rho, \sigma) [\cdot, H_\sigma, H_\sigma] - 2\sigma^{-1} H_\sigma \sigma^{-1} H_\sigma \sigma^{-1},
\end{aligned} \tag{F50}$$

$$\begin{aligned}
\nabla_{\rho\rho\sigma}^3 f(h, \rho, \sigma) &= (H_\rho, H_\sigma) \mapsto \frac{2s_\alpha^3}{z^3} \langle \nabla_\rho \hat{\Psi}_\alpha(\rho, \sigma), H_\rho \rangle \langle \nabla_\sigma \hat{\Psi}_\alpha(\rho, \sigma), H_\sigma \rangle \nabla_\rho \hat{\Psi}_\alpha(\rho, \sigma) \\
&+ \frac{s_\alpha^2}{z^2} \langle \nabla_{\rho\sigma}^2 \hat{\Psi}_\alpha H_\sigma, H_\rho \rangle \nabla_\rho \hat{\Psi}_\alpha(\rho, \sigma) + \frac{s_\alpha^2}{z^2} \langle \nabla_\rho \hat{\Psi}_\alpha, H_\rho \rangle \nabla_{\rho\sigma}^2 \hat{\Psi}_\alpha(\rho, \sigma) H_\sigma \\
&\times \frac{s_\alpha^2}{z^2} \langle \nabla_\sigma \hat{\Psi}_\alpha, H_\sigma \rangle \nabla_{\rho\rho}^2 \hat{\Psi}_\alpha(\rho, \sigma) H_\rho + \frac{s_\alpha}{z} \nabla_{\rho\rho\sigma}^3 \hat{\Psi}_\alpha(\rho, \sigma) [\cdot, H_\rho, H_\sigma],
\end{aligned} \tag{F51}$$

where $z = h - s_\alpha \hat{\Psi}_\alpha(\rho)$. The remaining derivatives are completely analogous. Here we simplified the third derivative $\nabla_{\rho\rho\rho}^3 f$ (and similarly for other terms) by applying it twice to the same vector because that is what the conic solver uses.

To obtain $\nabla^3 \hat{\Psi}_\alpha(\rho, \sigma)$ we first differentiate the second derivatives of $\hat{\Psi}_\alpha(G, Z)$ and then substitute G and Z . We obtain $D_{GGG} \hat{\Psi}_\alpha(G, Z)$ by differentiating (F35) over G :

$$D_{GGG}^3 \hat{\Psi}_\alpha[H_{G_1}, H_{G_2}, H_{G_3}] = \langle H_{G_1}, Z_S^{\frac{1}{2}} D^2 g'(Z_S^{\frac{1}{2}} G Z_S^{\frac{1}{2}}) [Z_S^{\frac{1}{2}} H_{G_2} Z_S^{\frac{1}{2}}, Z_S^{\frac{1}{2}} H_{G_3} Z_S^{\frac{1}{2}}] Z_S^{\frac{1}{2}} \rangle. \tag{F52}$$

We obtain $D_{ZGG}^3 \hat{\Psi}_\alpha(G, Z)$ by differentiating (F38) over G :

$$\begin{aligned}
D_{ZGG}^3 \hat{\Psi}_\alpha[H_{Z_1}, H_{G_2}, H_{G_3}] &= \langle Dh(Z)[H_{Z_1}], SZ_S^{-\frac{1}{2}} D^2 \tilde{g}(Z_S^{\frac{1}{2}} G Z_S^{\frac{1}{2}}) [Z_S^{\frac{1}{2}} H_{G_2} Z_S^{\frac{1}{2}}, Z_S^{\frac{1}{2}} H_{G_3} Z_S^{\frac{1}{2}}] Z_S^{-\frac{1}{2}} S^\dagger \rangle \\
&= \langle H_{Z_1}, Dh(Z)[SZ_S^{-\frac{1}{2}} D^2 \tilde{g}(Z_S^{\frac{1}{2}} G Z_S^{\frac{1}{2}}) [Z_S^{\frac{1}{2}} H_{G_2} Z_S^{\frac{1}{2}}, Z_S^{\frac{1}{2}} H_{G_3} Z_S^{\frac{1}{2}}] Z_S^{-\frac{1}{2}} S^\dagger] \rangle.
\end{aligned} \tag{F53}$$

Because the third derivatives are continuous, $D_{GZG}^3 \hat{\Psi}_\alpha = D_{ZGG}^3 \hat{\Psi}_\alpha$. In order to solve the first vector direction and obtain $\nabla_{GZG}^3 \hat{\Psi}_\alpha[\cdot, H_{Z_2}, H_{G_3}]$, it is convenient to reorganize $D_{GZG}^3 \hat{\Psi}_\alpha$:

$$D_{GZG}^3 \hat{\Psi}_\alpha(G, Z)[H_{G_1}, H_{Z_2}, H_{G_3}] = \langle H_{G_1}, Z_S^{\frac{1}{2}} D^2 \tilde{g}(Z_S^{\frac{1}{2}} G Z_S^{\frac{1}{2}}) [Z_S^{-\frac{1}{2}} S^\dagger Dh(Z)[H_{Z_2}] SZ_S^{-\frac{1}{2}}, Z_S^{\frac{1}{2}} H_{G_3} Z_S^{\frac{1}{2}}] Z_S^{\frac{1}{2}} \rangle, \tag{F54}$$

and analogously for $D_{GGZ}^3 \hat{\Psi}_\alpha$.

To obtain $D_{GZZ}^3 \hat{\Psi}_\alpha$, we follow [[18], lemma 5.1] and use the expression

$$\langle A, Xf(XY^2X)X \rangle = \langle A, Y^{-1}(YX^2Y)f(YX^2Y)Y^{-1} \rangle. \tag{F55}$$

With this equation, we can rewrite

$$D_G \hat{\Psi}_\alpha(G, Z)[H_G] = \langle H_G, Z_S^{\frac{1}{2}} g'(Z_S^{\frac{1}{2}} G Z_S^{\frac{1}{2}}) Z_S^{\frac{1}{2}} \rangle = \langle H_G, G^{-\frac{1}{2}} \tilde{g}(G^{\frac{1}{2}} Z_S G^{\frac{1}{2}}) G^{-\frac{1}{2}} \rangle, \tag{F56}$$

where $\tilde{g}(x) = xg'(x)$, and differentiate twice with respect to Z :

$$D_{GZZ}\hat{\Psi}_\alpha[H_{G_1}, H_{Z_2}, H_{Z_3}] = \langle G^{-\frac{1}{2}}H_{G_1}G^{-\frac{1}{2}}, D\tilde{g}(G^{\frac{1}{2}}Z_S G^{\frac{1}{2}})[G^{\frac{1}{2}}S^\dagger D^2 h[H_{Z_2}, H_{Z_3}]SG^{\frac{1}{2}}] \rangle \\ + \langle G^{-\frac{1}{2}}H_{G_1}G^{-\frac{1}{2}}, D^2\tilde{g}(G^{\frac{1}{2}}Z_S G^{\frac{1}{2}})[G^{\frac{1}{2}}S^\dagger Dh[H_{Z_2}]SG^{\frac{1}{2}}, G^{\frac{1}{2}}S^\dagger Dh[H_{Z_3}]SG^{\frac{1}{2}}] \rangle. \quad (F57)$$

Again, to obtain $\nabla_{ZGZ}^3\hat{\Psi}_\alpha[\cdot, H_{G_2}, H_{Z_3}]$ we simply reorganize $D_{GZZ}\hat{\Psi}_\alpha$:

$$D_{ZGZ}\hat{\Psi}_\alpha[H_{Z_1}, H_{G_2}, H_{Z_3}] = \langle H_{Z_1}, D^2 h[SG^{\frac{1}{2}}D\tilde{g}(G^{\frac{1}{2}}Z_S G^{\frac{1}{2}})[G^{-\frac{1}{2}}H_{G_2}G^{-\frac{1}{2}}]G^{\frac{1}{2}}S^\dagger, H_{Z_3}] \rangle \\ + \langle H_{Z_1}, Dh[SG^{\frac{1}{2}}D^2\tilde{g}(G^{\frac{1}{2}}Z_S G^{\frac{1}{2}})[G^{-\frac{1}{2}}H_{G_2}G^{-\frac{1}{2}}, G^{\frac{1}{2}}S^\dagger Dh[H_{Z_3}]SG^{\frac{1}{2}}]G^{\frac{1}{2}}S^\dagger] \rangle. \quad (F58)$$

And analogously for $D_{ZZG}\hat{\Psi}_\alpha$.

Finally, we differentiate Eq. (F36) with respect to Z using the product rule (F4) and use the self-adjointness of the spectral derivatives (F13) to obtain

$$D_{ZZZ}\hat{\Psi}_\alpha[H_{Z_1}, H_{Z_2}, H_{Z_3}] = \langle H_{Z_1}, D^3 h(Z)[SG^{\frac{1}{2}}g'(G^{\frac{1}{2}}Z_S G^{\frac{1}{2}})G^{\frac{1}{2}}S^\dagger, H_{Z_2}, H_{Z_3}] \rangle \\ + \langle H_{Z_1}, D^2 h(Z)[H_{Z_2}, SG^{\frac{1}{2}}Dg'(G^{\frac{1}{2}}Z_S G^{\frac{1}{2}})[G^{\frac{1}{2}}S^\dagger Dh(Z)[H_{Z_3}]SG^{\frac{1}{2}}]G^{\frac{1}{2}}S^\dagger] \rangle \\ + \langle H_{Z_1}, D^2 h(Z)[SG^{\frac{1}{2}}Dg'(G^{\frac{1}{2}}Z_S G^{\frac{1}{2}})[G^{\frac{1}{2}}S^\dagger Dh(Z)[H_{Z_2}]SG^{\frac{1}{2}}]G^{\frac{1}{2}}S^\dagger, H_{Z_3}] \rangle \\ + \langle H_{Z_1}, Dh(Z)[SG^{\frac{1}{2}}Dg'(G^{\frac{1}{2}}Z_S G^{\frac{1}{2}})[G^{\frac{1}{2}}S^\dagger D^2 h(Z)[H_{Z_2}, H_{Z_3}]SG^{\frac{1}{2}}]G^{\frac{1}{2}}S^\dagger] \rangle \\ + \langle H_{Z_1}, Dh(Z)[SG^{\frac{1}{2}}D^2 g'(G^{\frac{1}{2}}Z_S G^{\frac{1}{2}})[G^{\frac{1}{2}}S^\dagger Dh(Z)[H_{Z_2}]SG^{\frac{1}{2}}, G^{\frac{1}{2}}S^\dagger Dh(Z)[H_{Z_3}]SG^{\frac{1}{2}}]G^{\frac{1}{2}}S^\dagger] \rangle. \quad (F59)$$

Now we can simply substitute $G = \hat{G}(\rho)$, $Z = \hat{Z}(\sigma)$ and the vector directions $H_{G_i} = \hat{G}(H_{\rho_i})$ and $H_{Z_i} = \hat{Z}(H_{\sigma_i})$. We then solve the first vector direction to recover the third derivatives $\nabla^3\hat{\Psi}_\alpha(\rho, \sigma)$ applied to two vectors (actually, twice the same vector), which is what the conic solver needs.

3. FastRényiQKD cone

a. Gradient

The gradient of the barrier (34) is

$$\nabla_h f(h, \rho) = -\frac{1}{z}, \quad (F60)$$

$$\nabla_\rho f(h, \rho) = \frac{s_\alpha}{z} \nabla_\rho \hat{\Psi}_\alpha(\rho) - \rho^{-1}, \quad (F61)$$

where $z = h - s_\alpha \hat{\Psi}_\alpha(\rho)$ and $\hat{\Psi}_\alpha$ is given by Eq. (30).

The gradient of $\hat{\Psi}_\alpha(\rho)$ can be obtained from the gradient of $\hat{\Psi}_\alpha(\rho, \sigma)$ (computed in F 2 a) just by composing $\hat{\Psi}_\alpha(\rho, \sigma)$ with the projection $(\rho, \sigma) \mapsto (\rho, \rho)$ and replacing $\hat{Z}$ with $\hat{Z}\hat{G}$:

$$\nabla_\rho \hat{\Psi}_\alpha(\rho) = \hat{G}^\dagger(Z_S^{\frac{1}{2}}g'(Z_S^{\frac{1}{2}}GZ_S^{\frac{1}{2}})Z_S^{\frac{1}{2}}) + \widehat{Z\hat{G}}^\dagger(Dh(Z)[SG^{\frac{1}{2}}g'(G^{\frac{1}{2}}Z_S G^{\frac{1}{2}})G^{\frac{1}{2}}S^\dagger]) \\ = \hat{G}^\dagger(Z_S^{\frac{1}{2}}Ug'(\Lambda^2)U^\dagger Z_S^{\frac{1}{2}}) + \widehat{Z\hat{G}}^\dagger(Dh(Z)[SG^{\frac{1}{2}}Vg'(\Lambda^2)V^\dagger G^{\frac{1}{2}}S^\dagger]),$$

where in the last equality we used the singular value decomposition of $Z_S^{\frac{1}{2}}G^{\frac{1}{2}} = U\Lambda V^\dagger$.

b. Hessian

The Hessian of the barrier (34) is

$$\nabla_{hh}^2 f(h, \rho) = \frac{1}{z^2}, \quad (F62)$$

$$\nabla_{h\rho}^2 f(h, \rho) = -\frac{s_\alpha}{z^2} \nabla_\rho \hat{\Psi}_\alpha(\rho), \quad (\text{F63})$$

$$\nabla_{\rho\rho}^2 f(h, \rho) = H \mapsto \frac{s_\alpha^2}{z^2} \langle \nabla_\rho \hat{\Psi}_\alpha(\rho), H \rangle \nabla_\rho \hat{\Psi}_\alpha(\rho) + \frac{s_\alpha}{z} \nabla_{\rho\rho}^2 \hat{\Psi}_\alpha(\rho) H + \rho^{-1} H \rho^{-1}, \quad (\text{F64})$$

where $z = h - s_\alpha \hat{\Psi}_\alpha(\rho)$.

To obtain the product of the Hessian times a vector H , we can simply compose $\hat{\Psi}_\alpha(\rho, \sigma)$ with the projection $(\rho, \sigma) \mapsto (\rho, \rho)$, replace $\hat{\mathcal{Z}}$ with $\hat{\mathcal{Z}}\hat{\mathcal{G}}$, and recycle the equations for the Hessian of $\hat{\Psi}_\alpha(\rho, \sigma)$, Eqs. (F39)–(F42):

$$\nabla^2 \hat{\Psi}_\alpha(\rho) H = \nabla_{\rho\rho}^2 \hat{\Psi}_\alpha(\rho, \rho) H + \nabla_{\sigma\rho}^2 \hat{\Psi}_\alpha(\rho, \rho) H + \nabla_{\rho\sigma}^2 \hat{\Psi}_\alpha(\rho, \rho) H + \nabla_{\sigma\sigma}^2 \hat{\Psi}_\alpha(\rho, \rho) H. \quad (\text{F65})$$

To obtain the Hessian, we can vectorize $\nabla^2 \hat{\Psi}_\alpha(\rho) H$ and isolate the vector direction H .

c. Third derivatives

Finally, the conic solver needs the third derivatives of the barrier (34) are

$$\nabla_{hhh}^3 f(h, \rho) = -\frac{2}{z^3}, \quad (\text{F66})$$

$$\nabla_{hh\rho}^3 f(h, \rho) = \frac{2s_\alpha}{z^3} \nabla_\rho \hat{\Psi}_\alpha(\rho), \quad (\text{F67})$$

$$\nabla_{h\rho\rho}^3 f(h, \rho) = H \mapsto -\frac{2s_\alpha^2}{z^3} \langle \nabla_\rho \hat{\Psi}_\alpha(\rho), H \rangle \nabla_\rho \hat{\Psi}_\alpha(\rho) - \frac{s_\alpha}{z^2} \nabla_{\rho\rho}^2 \hat{\Psi}_\alpha(\rho) H, \quad (\text{F68})$$

$$\begin{aligned} \nabla_{\rho\rho\rho}^3 f(h, \rho) = (H_1, H_2) \mapsto & \frac{2s_\alpha^3}{z^3} \langle \nabla_\rho \hat{\Psi}_\alpha(\rho), H_1 \rangle \langle \nabla_\rho \hat{\Psi}_\alpha(\rho), H_2 \rangle \nabla_\rho \hat{\Psi}_\alpha(\rho) \\ & + \frac{s_\alpha^2}{z^2} \langle \nabla_{\rho\rho}^2 \hat{\Psi}_\alpha H_2, H_1 \rangle \nabla_\rho \hat{\Psi}_\alpha(\rho) + \frac{s_\alpha^2}{z^2} \langle \nabla_\rho \hat{\Psi}_\alpha, H_1 \rangle \nabla_{\rho\rho}^2 \hat{\Psi}_\alpha(\rho) H_2 \\ & + \frac{s_\alpha^2}{z^2} \langle \nabla_\rho \hat{\Psi}_\alpha, H_2 \rangle \nabla_{\rho\rho}^2 \hat{\Psi}_\alpha(\rho) H_1 + \frac{s_\alpha}{z} \nabla_{\rho\rho\rho}^3 \hat{\Psi}_\alpha(\rho) [\cdot, H_1, H_2] \\ & - \rho^{-1} H_2 \rho^{-1} H_1 \rho^{-1} - \rho^{-1} H_1 \rho^{-1} H_2 \rho^{-1}, \end{aligned} \quad (\text{F69})$$

where $z = h - s_\alpha \hat{\Psi}_\alpha(\rho)$. Since the conic solver will use $\nabla^3 f$ applied twice to the same vector direction H , we can simplify the last derivative:

$$\begin{aligned} \nabla_{\rho\rho\rho}^3 f(h, \rho) [\cdot, H, H] = & \frac{2s_\alpha^3}{z^3} \langle \nabla_\rho \hat{\Psi}_\alpha(\rho), H \rangle^2 \nabla_\rho \hat{\Psi}_\alpha(\rho) \\ & + \frac{s_\alpha^2}{z^2} \langle \nabla_{\rho\rho}^2 \hat{\Psi}_\alpha H, H \rangle \nabla_\rho \hat{\Psi}_\alpha(\rho) + 2 \frac{s_\alpha^2}{z^2} \langle \nabla_\rho \hat{\Psi}_\alpha, H \rangle \nabla_{\rho\rho}^2 \hat{\Psi}_\alpha(\rho) H \\ & + \frac{s_\alpha}{z} \nabla_{\rho\rho\rho}^3 \hat{\Psi}_\alpha(\rho) [\cdot, H, H] - 2 \rho^{-1} H \rho^{-1} H \rho^{-1}. \end{aligned}$$

Finally, as in the previous sections, by composing $\hat{\Psi}_\alpha(\rho, \sigma)$ with the projection $(\rho, \sigma) \mapsto (\rho, \rho)$ and taking $H := H_\sigma = H_\rho$, we can recycle the third derivatives of Sec. F 2 c:

$$\begin{aligned} \nabla_{\rho\rho\rho}^3 \hat{\Psi}_\alpha(\rho) [\cdot, H, H] = & \nabla_{\rho\rho\rho}^3 \hat{\Psi}_\alpha(\rho, \rho) [\cdot, H, H] + \nabla_{\sigma\rho\rho}^3 \hat{\Psi}_\alpha(\rho, \rho) [\cdot, H, H] \\ & + \nabla_{\rho\sigma\rho}^3 \hat{\Psi}_\alpha(\rho, \rho) [\cdot, H, H] + \nabla_{\rho\rho\sigma}^3 \hat{\Psi}_\alpha(\rho, \rho) [\cdot, H, H] \\ & + \nabla_{\sigma\sigma\rho}^3 \hat{\Psi}_\alpha(\rho, \rho) [\cdot, H, H] + \nabla_{\sigma\rho\sigma}^3 \hat{\Psi}_\alpha(\rho, \rho) [\cdot, H, H] \\ & + \nabla_{\rho\sigma\sigma}^3 \hat{\Psi}_\alpha(\rho, \rho) [\cdot, H, H] + \nabla_{\sigma\sigma\sigma}^3 \hat{\Psi}_\alpha(\rho, \rho) [\cdot, H, H]. \end{aligned} \quad (\text{F70})$$

APPENDIX G: STARTING POINT

1. RényiQKD cone

The starting point for the RényiQKD cone (31) is given by the solution of

$$(h, \rho, \sigma) = -\nabla f(h, \rho, \sigma), \quad (\text{G1})$$

where $f(h, \rho, \sigma)$ is the barrier function (33). Recalling the derivatives from Appendix F 2 a, we obtain the following system of non-linear equations:

$$h = \frac{1}{z}, \quad (\text{G2a})$$

$$\rho = \frac{1}{z} \nabla_\rho z + \rho^{-1}, \quad (\text{G2b})$$

$$\sigma = \frac{1}{z} \nabla_\sigma z + \sigma^{-1}, \quad (\text{G2c})$$

where $z = h - s_\alpha \hat{\Psi}_\alpha$.

If d_ρ and d_σ are the dimensions of ρ and σ , we can plug in the ansatz $\rho = \gamma \mathbb{1}_{d_\rho}$, $\sigma = \delta \mathbb{1}_{d_\sigma}$. In the particular case where $\hat{\mathcal{G}}(\mathbb{1}_{d_\rho}) = \hat{\mathcal{G}}^\dagger(\mathbb{1}_{d_\rho}) = \mathbb{1}_{d_\rho}$ and $\hat{\mathcal{Z}}(\mathbb{1}_{d_\sigma}) = \hat{\mathcal{Z}}^\dagger(\mathbb{1}_{d_\sigma}) = \mathbb{1}_{d_\sigma}$, Eqs. (F24) and (F25) give us

$$\nabla_\rho \hat{\Psi}_\alpha(\rho, \sigma) = \alpha \left(\frac{\gamma}{\delta} \right)^{\alpha-1} \mathbb{1}_{d_\rho} \quad (\text{G3})$$

$$\nabla_\sigma \hat{\Psi}_\alpha(\rho, \sigma) = (1 - \alpha) \left(\frac{\gamma}{\delta} \right)^\alpha \mathbb{1}_{d_\sigma} \quad (\text{G4})$$

In this particular case we get $\hat{\Psi}_\alpha = d_\rho ((\gamma/\delta)^\alpha \delta)$, so we substitute that in Eqs. (G2), obtaining

$$h^2 - s_\alpha h d_\rho \left(\frac{\gamma}{\delta} \right)^\alpha \delta - 1 = 0 \quad (\text{G5a})$$

$$-s_\alpha h \alpha \left(\frac{\gamma}{\delta} \right)^{\alpha-1} = \gamma - \frac{1}{\gamma} \quad (\text{G5b})$$

$$-s_\alpha h (1 - \alpha) \left(\frac{\gamma}{\delta} \right)^\alpha = \delta - \frac{1}{\delta} \quad (\text{G5c})$$

We eliminate h by combining (G5a) and (G5b) and also by combining (G5b) and (G5c), obtaining

$$(\gamma^2 - 1)^2 \gamma^{-2\alpha} (\delta^2)^{\alpha-1} + d_\rho \alpha (\gamma^2 - 1) - \alpha^2 = 0 \quad (\text{G6})$$

$$\frac{\gamma^2 - 1}{\alpha} = \frac{\delta^2 - 1}{1 - \alpha} \quad (\text{G7})$$

We eliminate δ^2 from Eq. (G6) using (G7), obtaining an equation on a single variable that can then be solved

numerically by Newton's method. We substitute back in Eq. (G7) to obtain δ , and in Eq. (G2a) to obtain h .

In general, computing the initial point for an arbitrary choice of $\hat{\mathcal{G}}, \hat{\mathcal{Z}}$ is not feasible. Note, however, that this strategy always produces a valid starting point, since $\gamma \mathbb{1}_{d_\rho} > 0, \delta \mathbb{1}_{d_\sigma} > 0$ and Eq. (G2a) implies that $h > s_\alpha \hat{\Psi}_\alpha$, and therefore $(h, \gamma \mathbb{1}_{d_\rho}, \delta \mathbb{1}_{d_\sigma})$ belongs to the interior of the cone.

If the solver is going to be run repeatedly for a fixed choice of $\hat{\mathcal{G}}, \hat{\mathcal{Z}}$, one can compute numerically the central point for these maps and supply the result to the solver.

2. FastRényiQKD cone

The starting point for the FastRényiQKD cone (32) is given by the solution of

$$(h, \rho) = -\nabla f(h, \rho), \quad (\text{G8})$$

where $f(h, \rho)$ is the barrier (34). Substituting the derivatives from Appendix F 3 a we obtain the equations

$$h = \frac{1}{z}, \quad (\text{G9})$$

$$\rho = \frac{1}{z} \nabla_\rho z + \rho^{-1}, \quad (\text{G10})$$

where $z = h - s_\alpha \hat{\Psi}_\alpha$.

The first equation, together with the requirement that $h \geq s_\alpha \hat{\Psi}_\alpha$, implies

$$h = \frac{s_\alpha \hat{\Psi}_\alpha}{2} + \sqrt{1 + \frac{\hat{\Psi}_\alpha^2}{4}} \quad (\text{G11})$$

Using the ansatz $\rho = \gamma \mathbb{1}$ and in the particular case that $\hat{\mathcal{G}}(\mathbb{1}) = \hat{\mathcal{G}}^\dagger(\mathbb{1}) = \hat{\mathcal{Z}}(\mathbb{1}) = \hat{\mathcal{Z}}^\dagger(\mathbb{1}) = \mathbb{1}$, Eq. (F62) gives us

$$\nabla_\rho \Psi(\rho) = \alpha \mathbb{1} + (1 - \alpha) \mathbb{1} = \mathbb{1}. \quad (\text{G12})$$

Equation (G10) then boils down to

$$-s_\alpha h = \gamma - \frac{1}{\gamma} \quad (\text{G13})$$

In this particular case, we get $\hat{\Psi}_\alpha = \gamma d$, so we solve

$$-\frac{\gamma d}{2} - s_\alpha \sqrt{1 + \frac{\gamma^2 d^2}{4}} = \gamma - \frac{1}{\gamma} \quad (\text{G14})$$

to obtain

$$\gamma = \sqrt{\frac{d+3}{2d+2} - \frac{s_\alpha}{2}} \sqrt{1 + \frac{4}{(d+1)^2}}. \quad (\text{G15})$$

As with the RényiQKD cone, computing the initial point for an arbitrary choice of $\hat{\mathcal{G}}, \widehat{\mathcal{ZG}}$ is not feasible. Note, however, that choosing $\rho = \gamma\mathbb{1}$ and setting h to be given by Eq. (G11) is always a valid starting point, since $\gamma\mathbb{1} > 0$ and $\frac{s_\alpha \hat{\Psi}_\alpha}{2} + \sqrt{1 + \frac{\hat{\Psi}_\alpha^2}{4}} > s_\alpha \hat{\Psi}_\alpha$, and thus it belongs to the interior of the cone.

-
- [1] C. H. Bennett and G. Brassard, *Quantum cryptography: Public key distribution and coin tossing*, *Theor. Comput. Sci.* **560**, 7 (1984) (reprint).
- [2] A. K. Ekert, *Quantum cryptography based on Bell's theorem*, *Phys. Rev. Lett.* **67**, 661 (1991).
- [3] R. Renner, *Security of quantum key distribution*, *Int. J. Quantum Inform.* **06**, 1 (2008).
- [4] C. Portmann and R. Renner, *Security in quantum cryptography*, *Rev. Mod. Phys.* **94**, 025008 (2022).
- [5] E. Y.-Z. Tan, P. Sekatski, J.-D. Bancal, R. Schwonnek, R. Renner, N. Sangouard, and C. C.-W. Lim, *Improved DIQKD protocols with finite-size analysis*, *Quantum* **6**, 880 (2022).
- [6] S. Bäuml, C. Pascual-García, V. Wright, O. Fawzi, and A. Acín, *Security of discrete-modulated continuous-variable quantum key distribution*, *Quantum* **8**, 1418 (2024).
- [7] L. Kamin, J. Burniston, and E. Y. Z. Tan, *Rényi security framework against coherent attacks applied to decoy-state QKD*, *arXiv:2504.12248*.
- [8] F. Dupuis, O. Fawzi, and R. Renner, *Entropy accumulation*, *Commun. Math. Phys.* **379**, 867 (2020).
- [9] F. Dupuis and O. Fawzi, *Entropy accumulation with improved second-order term*, *IEEE Trans. Inf. Theory* **65**, 7596 (2019).
- [10] A. Arqand and E. Y. Z. Tan, *Marginal-constrained entropy accumulation theorem*, *arXiv:2502.02563*.
- [11] T. Metger and R. Renner, *Security of quantum key distribution from generalised entropy accumulation*, *Nat. Commun.* **14**, 5272 (2023).
- [12] A. Arqand, T. A. Hahn, and E. Y.-Z. Tan, *Generalized Rényi entropy accumulation theorem and generalized quantum probability estimation*, *Phys. Rev. X* **15**, 041013 (2025).
- [13] R. R. B. Chung, N. H. Y. Ng, and Y. Cai, *Generalized numerical framework for improved finite-sized key rates with Rényi entropy*, *Phys. Rev. A* **112**, 012612 (2025).
- [14] M. Frank and P. Wolfe, *An algorithm for quadratic programming*, *Naval Res. Logist. Q.* **3**, 95 (1956).
- [15] A. Winick, N. Lütkenhaus, and P. J. Coles, *The many faces of degeneracy in conic optimization*, *Quantum* **2**, 77 (2018).
- [16] D. Drusvyatskiy and H. Wolkowicz, *The many faces of degeneracy in conic optimization*, *Found. Trends Optim.* **3**, 77 (2017).
- [17] H. Hu, J. Im, J. Lin, N. Lütkenhaus, and H. Wolkowicz, *Robust interior point method for quantum key distribution rate computation*, *Quantum* **6**, 792 (2022).
- [18] K. He, J. Saunderson, and H. Fawzi, *Operator convexity along lines, self-concordance, and sandwiched Rényi entropies*, *arXiv:2502.05627*.
- [19] A. G. Lorente, P. V. Parellada, M. Castillo-Celeita, and M. Araújo, *Quantum key distribution rates from non-symmetric conic optimization*, *Quantum* **9**, 1657 (2025).
- [20] K. He, J. Saunderson, and H. Fawzi, *Mathematical programming computation*, *arXiv:2407.00241*.
- [21] F. Dupuis, *Privacy amplification and decoupling without smoothing*, *IEEE Trans. Inf. Theory* **69**, 7784 (2023).
- [22] C. H. Bennett, G. Brassard, and N. D. Mermin, *Quantum cryptography without Bell's theorem*, *Phys. Rev. Lett.* **68**, 557 (1992).
- [23] N. J. Cerf, M. Bourennane, A. Karlsson, and N. Gisin, *Security of quantum key distribution using d-level systems*, *Phys. Rev. Lett.* **88**, 127902 (2002).
- [24] L. Sheridan and V. Scarani, *Security proof for quantum key distribution using qudit systems*, *Phys. Rev. A* **82**, 030301(R) (2010).
- [25] Y. Zhang, Y. Bian, Z. Li, S. Yu, and H. Guo, *Continuous-variable quantum key distribution system: Past, present, and future*, *Appl. Phys. Rev.* **11**, 011318 (2024).
- [26] C. Pascual-García, S. Bäuml, M. Araújo, R. Liss, and A. Acín, *Improved finite-size key rates for discrete-modulated continuous variable quantum key distribution under coherent attacks*, *Phys. Rev. A* **111**, 022610 (2025).
- [27] J. Bezanson, A. Edelman, S. Karpinski, and V. B. Shah, *Julia: A fresh approach to numerical computing*, *SIAM Rev.* **59**, 65 (2017).
- [28] C. Coey, L. Kapelevich, and J. P. Vielma, *Solving natural conic formulations with Hypatia.jl*, *INFORMS J. Comput.* **34**, 2686 (2022).
- [29] C. Coey, L. Kapelevich, and J. P. Vielma, *Performance enhancements for a generic conic interior point algorithm*, *Math. Program. Comput.* **15**, 53 (2023).
- [30] A. Skajaa and Y. Ye, *A homogeneous interior-point algorithm for nonsymmetric convex conic optimization*, *Math. Program.* **150**, 391 (2015).
- [31] D. Papp and S. Yıldız, *On "A homogeneous interior-point algorithm for non-symmetric convex conic optimization"*, *arXiv:1712.00492*.
- [32] M. Lubin, O. Dowson, J. Dias Garcia, J. Huchette, B. Legat, and J. P. Vielma, *JuMP 1.0: Recent improvements to a modeling language for mathematical optimization*, *Math. Program. Comput.* **15**, 581 (2023).
- [33] M. Pereira, G. Currás-Lorenzo, and M. Araújo, *Optimal key rates for quantum key distribution with partial source characterization*, *arXiv:2510.13085*.
- [34] In the case of continuous variable QKD, it should be understood that $S^{(n)}$ and $C^{(n)}$ are actually held by Bob due to the reverse nature of information reconciliation for continuous variable QKD.
- [35] Such as a virtual tomography, for implementations as in [6].
- [36] M. Curty, M. Lewenstein, and N. Lütkenhaus, *Entanglement as a precondition for secure quantum key distribution*, *Phys. Rev. Lett.* **92**, 217903 (2004).
- [37] A. Ferenczi and N. Lütkenhaus, *Symmetries in quantum key distribution and the connection between optimal attacks and optimal cloning*, *Phys. Rev. A* **85**, 052310 (2012).

- [38] D. Tupkary, E. Y.-Z. Tan, and N. Lütkenhaus, *Security proof for variable-length quantum key distribution*, *Phys. Rev. Res.* **6**, 023002 (2024).
- [39] J. Bretagnolle and C. Huber, *Lois empiriques et distance de Prokhorov*, *Sém. Probab. Strasbourg* **12**, 332 (1978).
- [40] A. W. van der Vaart and J. A. Wellner, *Weak Convergence and Empirical Processes: With Applications to Statistics* (Springer International Publishing, New York, 2023).
- [41] Information reconciliation is exclusively performed through the classical channel, which cannot be attacked by Eve beyond wiretapping. Therefore, it can be assumed that its implementation is always honest.
- [42] J. M. Renes and R. Renner, *One-shot classical data compression with quantum side information and the distillation of common randomness or secret keys*, *IEEE Trans. Inf. Theory* **58**, 1985 (2012). [
- [43] D. Slepian and J. Wolf, *Noiseless coding of correlated information sources*, *IEEE Trans. Inf. Theory* **19**, 471 (1973).
- [44] K. He, J. Saunderson, and H. Fawzi, *QICS: Quantum information conic solver*, [arXiv:2410.17803](https://arxiv.org/abs/2410.17803).
- [45] R. L. Frank and E. H. Lieb, *Monotonicity of a relative Rényi entropy*, *J. Math. Phys.* **54**, 122201 (2013).
- [46] One might hope to be able to use this cone directly for optimizing (28) or (30), but this is not possible, as they are not expressible in terms of $\Psi_\alpha(\rho, \sigma)$. Analogously, the von Neumann QKD cone is not expressible in terms of the relative entropy cone [19]
- [47] F. Hiai, *Concavity of certain matrix trace and norm functions*, *Linear Algebra Appl.* **439**, 1568 (2013).
- [48] ConicQKD.jl: Implementation of convex cones for quantum key distribution (2025).
- [49] This cone is better known as the exponential cone, formulated in terms of the exponential function, but both formulations are equivalent.
- [50] We note, in particular, that the protocol outline in [7] is compatible with [[10], Corollary 4.2] as explained in [[7], Sec. IV.B].
- [51] N. Walenta, A. Burg, D. Caselunghe, J. Constantin, N. Gisin, O. Guinnard, R. Houlmann, P. Junod, B. Korzh, N. Kulesza, M. Legré, C. W. Lim, T. Lunghi, L. Monat, C. Portmann, M. Soucarros, R. T. Thew, P. Trinkler, G. Trollet, F. Vannel, and H. Zbinden, *A fast and versatile quantum key distribution system with hardware key distillation and wavelength multiplexing*, *New J. Phys.* **16**, 013047 (2014).
- [52] M. Araújo, M. Huber, M. Navascués, M. Pivoluska, and A. Tavakoli, *Quantum key distribution rates from semidefinite programming*, *Quantum* **7**, 1019 (2023).
- [53] A. Leverrier and P. Grangier, *Continuous-variable quantum-key-distribution protocols with a non-Gaussian modulation*, *Phys. Rev. A* **83**, 042312 (2011).
- [54] S. Ghorai, P. Grangier, E. Diamanti, and A. Leverrier, *Asymptotic security of continuous-variable quantum key distribution with a discrete modulation*, *Phys. Rev. X* **9**, 021059 (2019).
- [55] J. Lin, T. Upadhyaya, and N. Lütkenhaus, *Asymptotic security analysis of discrete-modulated continuous-variable quantum key distribution*, *Phys. Rev. X* **9**, 041064 (2019).
- [56] Similarly, the entropy $H(S|Y)$ in (19) should be taken with S as Bob's secret key after the public announcements, and Y as Alice's.
- [57] F. Kanitschar, I. George, J. Lin, T. Upadhyaya, and N. Lütkenhaus, *Finite-size security for discrete-modulated continuous-variable quantum key distribution protocols*, *PRX Quantum* **4**, 040306 (2023).
- [58] This is not true when post selection is included, such as in Ref. [57].
- [59] I. W. Primaatmaja, W. Y. Kon, and C. Lim, *Discrete-modulated continuous-variable quantum key distribution secure against general attacks*, [arXiv:2409.02630](https://arxiv.org/abs/2409.02630).
- [60] L. Kamin, J. Burniston, and E. Y.-Z. Tan, *Rényi security framework against coherent attacks applied to decoy-state QKD—GitHub repository*, <https://github.com/Optical-Quantum-Communication-Theory/Renyi-security-framework> (2025).
- [61] M. Navarro, A. Acín, and C. Pascual-García, *Effective discrete-modulated continuous variable QKD under general attacks*, [arXiv:2606.20346](https://arxiv.org/abs/2606.20346).
- [62] *Finite-size quantum key distribution rates from Rényi entropies using conic optimization—GitHub repository*, GitHub repository, <https://github.com/MarianaNvrt/Renyi-ConicQKD> (2025).
- [63] J. Burniston *et al.*, *Open QKD security: version 2.1.0*, Zenodo, [10.5281/zenodo.14262569](https://zenodo.org/record/14262569) (2024).
- [64] M. Tomamichel, M. Berta, and M. Hayashi, *Relating different quantum generalizations of the conditional Rényi entropy*, *J. Math. Phys. (N.Y.)* **55**, 082206 (2014).
- [65] Ref. [[13], Eqs. (35) and (40)] also uses a further reduction $\psi_{RAB} = G(\psi_{AB})$ in the minimization. However, this assumption is not mathematically sound, as it unnecessarily restricts the optimization to a subset of the valid state space.
- [66] M. Müller-Lennert, F. Dupuis, O. Szechr, S. Fehr, and M. Tomamichel, *On quantum Rényi entropies: A new generalization and some properties*, *J. Math. Phys. (N.Y.)* **54**, 122203 (2013).
- [67] M. Tomamichel, *Quantum Information Processing with Finite Resources* (Springer International Publishing, New York, 2016).
- [68] M. Berta, F. Furrer, and V. B. Scholz, *The smooth entropy formalism for von Neumann algebras*, *J. Math. Phys. (N.Y.)* **57**, 015213 (2016).
- [69] Y. Nesterov, *Lectures on Convex Optimization* (Springer, New York, 2018).
- [70] Y. Nesterov and A. Nemirovskii, *Interior-Point Polynomial Algorithms in Convex Programming* (SIAM Society for Industrial and Applied Mathematics, Philadelphia, USA, 1994).
- [71] R. A. Horn and C. R. Johnson, *Matrix Analysis* (Cambridge University Press, Cambridge, 2013).
- [72] F. Zhang, *Matrix Theory, Basic Results and Techniques* (Springer, New York, 2011).
- [73] F. Hiai, *Concavity of certain matrix trace functions*, *Taiwanese J. Math.* **5**, 535 (2001).
- [74] N. J. Higham, *Functions of Matrices: Theory and Computation* (SIAM Society for Industrial and Applied Mathematics, Philadelphia, USA, 2008).
- [75] F. Hiai and D. Petz, *Introduction to Matrix Analysis and Applications* (Springer, New York, 2014).