Analytic Rényi Entropy Bounds for Device-Independent Cryptography

Thomas A. Hahn,[1] Aby Philip,[2] Ernest Y.-Z. Tan,[3] and Peter Brown[4]

[1] *The Center for Quantum Science and Technology, Department of Physics of Complex Systems, Weizmann Institute of Science, Rehovot, Israel*
[2] *Institute of Fundamental Technological Research, Polish Academy of Sciences, Pawińskiego 5B, 02-106 Warsaw, Poland*
[3] *University of Waterloo, Waterloo, Ontario N2L 3G1, Canada*
[4] *Télécom Paris, LTCI, Institut Polytechnique de Paris, Inria, 19 Place Marguerite Perey, 91120 Palaiseau, France*

(Dated: July 11, 2025)

Device-independent (DI) cryptography represents the highest level of security, enabling cryptographic primitives to be executed safely on uncharacterized devices. Moreover, with successful proof-of-concept demonstrations in randomness expansion, randomness amplification, and quantum key distribution, the field is steadily advancing toward commercial viability. Critical to this continued progression is the development of tighter finite-size security proofs. In this work, we provide a simple method to obtain tighter finite-size security proofs for protocols based on the CHSH game, which is the nonlocality test used in all of the proof-of-concept experiments. We achieve this by analytically solving key-rate optimization problems based on Rényi entropies, providing a simple method to obtain tighter finite-size key rates.

I. INTRODUCTION

Quantum theory exhibits certain correlations between distant agents that cannot be explained classically [1]. These so-called nonlocal correlations have far-reaching implications beyond fundamental physics. In particular, by observing certain nonlocal correlations, it is possible to make statements about the underlying quantum systems used to produce them. This leads to device-independent information processing, which guarantees the successful completion of information processing tasks without detailed characterization of the underlying hardware. Such protocols have already been demonstrated in practice, with recent implementations of DI randomness expansion [2, 3], amplification [4] and quantum key distribution (QKD) [5–7].

DIQKD protocols provide a method to generate shared secret key whilst relying on minimal assumptions. Here, one can leverage the fact that if two honest parties, Alice and Bob, observe a shared nonlocal distribution, then they can information-theoretically verify that their outputs are random from the perspective of an eavesdropper, Eve. There has been significant effort in designing protocols that are secure and efficient in practice [8, 9], and whilst proof-of-principle demonstrations of DIQKD have been achieved, the current achievable rates remain far from practical.

Protocols based on the CHSH inequality [10] have been studied extensively for DIQKD, owing to its simplicity both theoretically and experimentally. In particular, when considering security against collective attacks, [11] provides a tight analytical lower bound on the von Neumann entropy (and hence the asymptotic key-rate) in terms of the expected CHSH violation. Applying techniques like the Entropy Accumulation Theorem (EAT) [12–14], one can elevate these bounds to finite-size key rates against general adversaries. Recently, a new Rényi EAT (REAT) was proven in [15], which is based on the more general family of entropies known as Rényi entropies, and has the potential to yield tighter finite-size key-rates resulting in more practical protocols. However, in order to reap these benefits, tight and efficient methods for bounding Rényi entropies in a device-independent manner must be developed.

In this work, we derive a tight analytical relationship between the expected CHSH value and the amount of Rényi entropy in the output. This can be seen as a broad generalization of the expression derived in [11], which we recover as a special case. Crucially, our results unlock the potential of [15] and we demonstrate significantly improved finite-size key-rates for protocols based on the CHSH inequality. We further discuss how our results can be modified to include noisy preprocessing [16] and generalized to the asymmetric CHSH inequalities [17], further boosting the key rates and applicability of the technique. Overall, our work pushes DIQKD towards a new level of practicality.

II. ANALYTIC RÉNYI ENTROPY BOUNDS

We now present the main technical contribution of this work, which is a tight analytical relationship between the CHSH value and the accumulated Rényi entropy. To more precisely state the problem at hand, consider the following setup. We have two honest parties Alice and Bob, and an eavesdropper Eve. Alice and Bob each hold a device which can receive binary inputs X, Y and produce binary outputs A, B respectively. The behavior of the boxes may be modeled in the following way: a shared tripartite state $\rho_{Q_A Q_B E}$ is distributed to Alice, Bob and Eve; upon receiving the input $X = x$, Alice's box measures Q_A with a POVM $\{M_a^x\}_a$ and outputs the measurement outcome $A = a$; upon receiving the input $Y = y$,

Bob's box measures Q_B with the POVM $\{N_b^y\}_b$ and outputs the measurement outcome $B = b$. Given inputs $X = x$, $Y = y$, we can compute a post-measurement state $\rho_{ABE}^{xy} = \sum_{ab} |ab\rangle\langle ab|_{AB} \otimes \rho_E^{abxy}$, where

$$\rho_E^{abxy} = \mathrm{Tr}_{Q_A Q_B} \left[\rho_{Q_A Q_B E}(M_a^x \otimes N_b^y \otimes \mathbb{I}_E)\right] . \quad (1)$$

We call a tuple $(Q_A Q_B E, \rho_{Q_A Q_B E}, \{M_a^x\}, \{N_b^y\})$ of Hilbert spaces, a shared state, and POVMs a *quantum strategy*.

From the perspective of the honest parties, Alice and Bob, their devices are black boxes. As such, the exact Hilbert spaces, shared state, and POVMs used are unknown to them. On the other hand, we allow the eavesdropper Eve to have full control over the implementation of the devices, i.e., she can choose the quantum strategy. Despite the black-box nature of their devices, Alice and Bob are able to learn information about the correlations produced by their devices. In this work, we shall focus on protocols based on estimating the expected CHSH value of their devices, which is defined as

$$S = \sum_{abxy} (-1)^{xy+a+b} \mathrm{Tr}\left[\rho_{Q_A Q_B E} \left(M_a^x \otimes N_b^y \otimes \mathbb{1}\right)\right] . \quad (2)$$

For any $S > 2$, the correlations produced by the devices are nonlocal and can be used to produce device-independent randomness. Moreover, the maximal achievable value using quantum systems is $S = 2\sqrt{2}$ [18]. Given a quantum strategy $\Lambda = (Q_A Q_B E, \rho_{Q_A Q_B E}, \{M_{a|x}\}, \{N_{b|y}\})$ we denote its expected CHSH value, computed using Eq. (2), by $S_{\mathrm{CHSH}}(\Lambda)$.

In this section we are interested in solving the following optimization problem. Given a conditional entropy $\mathbb{H}$ (see Appendix A for the formal definitions) and an expected CHSH value $S = [2, 2\sqrt{2}]$, find the minimal value of $\mathbb{H}(A|X = 0, E)$ over all possible quantum strategies that have an expected CHSH value of S. In other words, we are interested in computing an $\mathbb{H}$ rate function for the CHSH Bell-inequality, as made precise in the following definition.

Definition 1 ($\mathbb{H}$ rate function for CHSH)**.** Let $\mathbb{H}$ be a conditional entropy and let $S \in [2, 2\sqrt{2}]$. We say that a function $f_{\mathbb{H}} : [2, 2\sqrt{2}] \to \mathbb{R}$ is a tight $\mathbb{H}$ rate function for the CHSH Bell inequality if

$$\begin{aligned} f_{\mathbb{H}}(S) \coloneqq \inf_{\Lambda} \quad & \mathbb{H}(A|X = 0, E) \\ \text{s.t.} \quad & S_{\mathrm{CHSH}}(\Lambda) = S , \end{aligned} \quad (3)$$

where the infimum is over all quantum strategies Λ.

We note that the case of the von Neumann entropy, H, was solved in [11] (see also [19]), where it was shown that

$$f_H(S) = 1 - h\left(\frac{1}{2} + \frac{1}{2}\sqrt{\frac{S^2}{4} - 1}\right) , \quad (4)$$

where $h(x) \coloneqq -x \log_2 x - (1-x)\log_2(1-x)$ is the binary entropy. Similarly, in [20] it was shown that for the min-entropy, $H_{\min}$, one has

$$f_{H_{\min}}(S) = 1 - \log\left(1 + \sqrt{2 - \frac{S^2}{4}}\right) . \quad (5)$$

Our main technical result is an exact analytical form of the rate functions for multiple major families of Rényi conditional entropy. In Theorem 2 below, we focus on presenting our results for two particular families of "sandwiched Rényi entropies" (see Appendix A or [21, 22] for full details), denoted as $\widetilde{H}_\alpha^\uparrow$ and $\widetilde{H}_\alpha^\downarrow$. We focus on these for now as they are the most relevant entropies for our finite-size analysis; however, we highlight that significant generalizations of this theorem are also possible. For instance, as we show in Appendix B, we can extend it to the family of asymmetric CHSH inequalities [17], modify $f_{\widetilde{H}_\alpha^\downarrow}(S)$ to include noisy preprocessing [16], and derive analogous results for the Petz-Rényi entropies [23]. The generalizations that include noisy preprocessing and the asymmetric CHSH inequalities have the potential to boost the achievable key-rates for DIQKD even further.

Theorem 2. *Let $\alpha > 1$ and $S \in [2, 2\sqrt{2}]$. Then we have*

$$f_{\widetilde{H}_\alpha^\uparrow}(S) = 1 + \frac{2\alpha - 1}{1 - \alpha} \log \phi_{\frac{\alpha}{2\alpha-1}}(S) , \quad (6)$$

$$f_{\widetilde{H}_\alpha^\downarrow}(S) = 1 + \frac{\alpha}{1 - \alpha} \log \phi_{\frac{1}{\alpha}}(S) , \quad (7)$$

where

$$\phi_\mu(S) = \left(\frac{1 - \sqrt{\frac{S^2}{4} - 1}}{2}\right)^\mu + \left(\frac{1 + \sqrt{\frac{S^2}{4} - 1}}{2}\right)^\mu . \quad (8)$$

We provide the proof in Appendix B, together with the generalizations. As expected, in the limit $\alpha \to 1$, both $f_{\widetilde{H}_\alpha^\uparrow}$ and $f_{\widetilde{H}_\alpha^\downarrow}$ converge to Eq. (4). Moreover, by setting $\alpha = 2$, one recovers the fact that $f_{\widetilde{H}_2^\downarrow}$ equals the rate function for the min-entropy in Eq. (5), a result that was first observed in [24].

Interestingly, the optimal strategy for Eve that achieves the infimum in Eq. (3) is the same for all $\alpha \geq 1$ and both entropy families. In particular, for an expected CHSH value, S, the optimal strategy for Eve is to program Alice and Bob's devices to measure the observables

$$A_0 = \sigma_z , \; A_1 = \sigma_x , \; B_{0,1} = \frac{\sigma_z \pm g_S \sigma_x}{\sqrt{1 + g_S^2}} , \quad (9)$$

where $g_S = \sqrt{\frac{S^2}{4} - 1}$, on the state

$$\sqrt{P_+} \, |\phi^+\rangle_{Q_A Q_B} |0\rangle_E + \sqrt{P_-} \, |\phi^-\rangle_{Q_A Q_B} |1\rangle_E , \quad (10)$$

for $|\phi^\pm\rangle = \frac{1}{\sqrt{2}}(|00\rangle \pm |11\rangle)$, and $P_\pm = \frac{1}{2}(1 \pm g_S)$.

III. IMPROVED FINITE-SIZE KEY RATES

A. Protocol Description

In addition to the CHSH set-up from the previous section, for DIQKD one generally allows Bob an extra measurement input, $Y = 2$, which he uses to generate secret key (his other settings are used to test the device). This additional measurement improves the key rates by reducing the cost of error correction, and due to non-signaling conditions, will not affect the validity of the rate functions described in the previous section, see e.g. [11]. This extended set-up is what will be considered here.

Each round of the DIQKD protocol is either a test or a key-generation round. During test rounds, which occur with some probability $\gamma \in [0, 1]$, each party chooses their measurement inputs X, Y uniformly at random from the set $\{0, 1\}$ and generates some measurement outputs A, B. During a subsequent public announcement step, Bob will announce a value $\bar{B}$ that is set equal to his output value B whenever it is a test round, which allows Alice to produce an additional test-data value for each round, $\bar{C}$, as follows:

$$\bar{C} = \begin{cases} 0, & \text{if } A \oplus B \neq X \cdot Y \\ 1, & \text{if } A \oplus B = X \cdot Y \ . \end{cases} \tag{11}$$

For single-round quantum strategies, the distribution of $\bar{C}$ can directly be related to the corresponding CHSH value, S. As such, the test data encodes the relevant information needed for applying Theorem 2. During generation rounds, Alice and Bob use the inputs $X = 0$ and $Y = 2$, respectively. No test data is produced and so Bob sets his public announcement $\bar{B}$ to some arbitrary value (say, $\bar{B} = 0$), while Alice sets $\bar{C}$ to a special symbol $\perp$.

After n rounds, both parties conduct two further classical postprocessing steps. First, they verify that the distribution of the test data lies within some predetermined set of probability distributions, S_Ω. By aborting the protocol whenever the observed distribution lies outside S_Ω, this step essentially ensures that Alice's and Bob's measurements produce (at least) weakly random bits, see e.g. [19, 25]. Afterwards, the protocol concludes with a post-processing step, which converts Alice's and Bob's raw measurement data, into (almost) ideal states according to a suitable DIQKD security definition [19, 25].

For clarity, we present an overall summary of this procedure as Protocol 1 below. More information regarding the classical post-processing steps can be found in Appendix D. Note that this protocol structure can also accommodate DI randomness expansion with only small changes, mostly in these classical post-processing steps; see e.g. [26] for such a description.

1 Cryptographic Protocol

1: For all rounds, $i \in \{1, \ldots, n\}$:
1.1. Alice and Bob generate a common random bit T_i, such that $\mathrm{P}(T_i = 0) = 1 - \gamma$ and $\mathrm{P}(T_i = 1) = \gamma$.
1.2. If $T_i = 0$, both parties will choose generation inputs $(X_i, Y_i) = (0, 2)$. If $T_i = 1$, both parties will choose test inputs $(X_i, Y_i) \in \{0, 1\}^2$ uniformly at random. They supply the inputs to the devices and obtain outputs (A_i, B_i).
2: Public announcements: Both parties announce all the input values (X_i, Y_i). Bob also announces some other registers $\bar{B}_i$ for $i \in [n]$ as follows. If $T_i = 0$, Bob announces $\bar{B}_i = 0$ and Alice sets $\bar{C}_i = \perp$. If $T_i = 1$, Bob announces $\bar{B}_i = B_i$, then Alice computes $\bar{C}_i$ according to the specified function of (A_i, B_i, X_i, Y_i).
3: Acceptance test: Alice checks if the observed frequency distribution $\mathrm{freq}_{\bar{c}_1^n}$ lies inside some predetermined set S_Ω, and aborts the protocol (via a public announcement) if it does not.
4: Classical postprocessing: Alice and Bob perform some additional classical operations such as error correction and privacy amplification (see Appendix D for details) to generate their final keys.

B. Rényi EAT and Key Rate

We provide a more detailed description of the finite-size security proof in Appendix D; here, we just outline the key steps of the proof. The global n-round state in the protocol after the public announcements is of the form $\rho_{A_1^n B_1^n \bar{C}_1^n X_1^n Y_1^n T_1^n \mathbf{E}}$, with $\mathbf{E}$ denoting quantum side-information Eve holds about the states in the devices (she also has access to the public announcements $X_1^n Y_1^n T_1^n$). Using the REAT, it was shown in [15, Lemmas 5.1 and 6.1] that as long as the set S_Ω is convex, the total accumulated Rényi entropy (conditioned on the acceptance test accepting, which we shall denote as the event Ω_{AT}) can be bounded by

$$\widetilde{H}_\alpha^\uparrow \left(A_1^n \bar{C}_1^n | X_1^n Y_1^n T_1^n \mathbf{E}\right)_{\rho_{|\Omega_{\mathrm{AT}}}} \geq n h_\alpha - \frac{\alpha}{\alpha - 1} \log \frac{1}{\Pr[\Omega_{\mathrm{AT}}]}, \tag{12}$$

where h_α is a quantity satisfying

$$h_\alpha \geq \inf_{\Lambda} \inf_{\mathbf{q} \in S_{\mathrm{acc}}} \frac{1}{\alpha - 1} D(\mathbf{q} \| \mathbf{p}_\Lambda) + q(\perp) \widetilde{H}_\alpha^\downarrow (A | X = 0, E) \ . \tag{13}$$

Here, the optimization takes place over all single-round quantum strategies Λ, and probability distributions $\mathbf{q}$ (on a single-round test-data register $\bar{C}$) within the acceptance set S_Ω. For each single-round strategy, $\widetilde{H}_\alpha^\downarrow (A | X = 0, E)$ refers to the corresponding Rényi entropy of the state produced from that strategy, and the $D(\mathbf{q} \| \mathbf{p}_\Lambda)$ term denotes the Kullback-Leibler (KL) divergence (see e.g. [28]) between $\mathbf{q}$ and the distribution $\mathbf{p}_\Lambda$ produced by that quantum strategy.

Qualitatively, the above optimization has an intu-

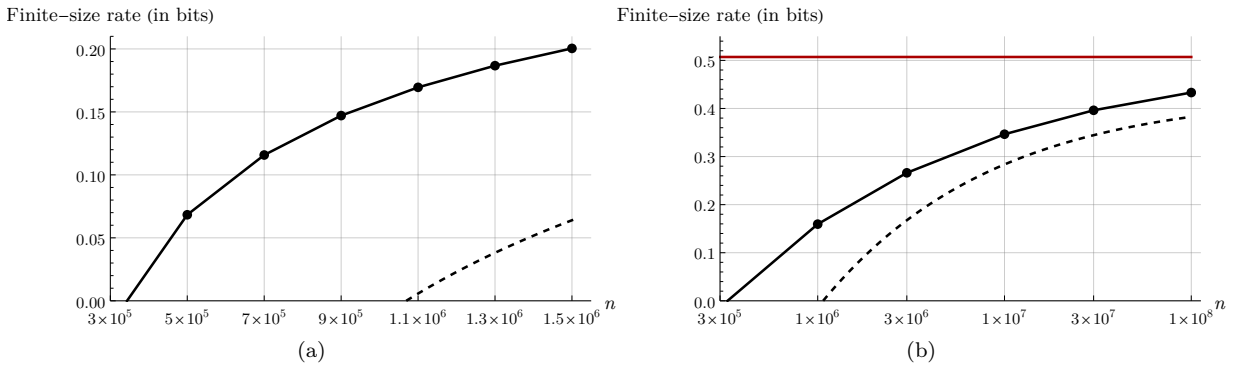

Figure 1. Plot of achievable finite-size key rates for the DIQKD experimental demonstration in [27], as a function of number of rounds n in two different ranges. The solid black curves show the results we obtain from our approach, the dashed black curves show the results previously computed in [27], and the solid red line displays the asymptotic rate (omitted in Fig. 1a due to the lower key rates in that plot). Note that we have kept the protocol and parameter choices (other than the final key length) nearly identical to the one in [27], except for minor changes and improvements we describe in Appendix D. We see there is a significant improvement in both the finite-size rate at the number of rounds used in that experiment ($n = 1.5 \times 10^6$), and the minimum n required to achieve nonzero finite-size rates.

itive informal interpretation, as follows. Observe that if the device behavior across the rounds were independent and identically distributed (IID), then in the asymptotic large-n limit, the best bound we could hope for on the global entropy $\widetilde{H}_\alpha^\uparrow\left(A_1^n\bar{C}_1^n|X_1^nY_1^nT_1^n\mathbf{E}\right)$ would be simply n times of the minimal single-round entropy $\widetilde{H}_\alpha^\downarrow(A|X=0,E)$ over all strategies Λ "compatible with" the accept condition S_Ω (more formally, such that the distribution $\mathbf{p}_\Lambda$ lies in S_Ω). The above optimization is similar in spirit to computing this minimal value, except that rather than the "hard" constraint of requiring $\mathbf{p}_\Lambda \in S_\Omega$, the KL divergence term serves to impose a "soft" version of this constraint, in that it acts as a "penalty" if $\mathbf{p}_\Lambda$ is far from S_Ω — see [15, Sec. 5.2] for more detailed exposition. We emphasize however that while this intuitive interpretation is informal, the bounds (12)–(13) constitute a *rigorous* lower bound on the global Rényi entropy, against general (non-IID) attacks.

In Appendix D, we explain how this bound can be used to compute finite-size key rates. Essentially, the LHS of Eq. (12) describes the Rényi entropy that Alice generates over all rounds of the protocol (including the test data), conditioned on the side-information registers $X_1^nY_1^nT_1^n\mathbf{E}$. Other side-information that Eve obtains, such as $\bar{B}_1^n$, can be accounted for separately by using appropriate chain rules; see Appendix D for details. The rate function derived in Theorem 2 can then be used to provide tight bounds on the $\widetilde{H}_\alpha^\downarrow(A|X=0,E)$ term in h_α, hence allowing us to compute finite-size key rates based on Rényi entropies.

In Fig. 1, we show the finite-size key rates obtained from this approach, as applied to the experimental parameters achieved in a DIQKD demonstration in [27]. We follow the parameters and implementation choices used in that work as closely as possible, apart from minor modifications we describe in Appendix D. (For Fig. 1a we also used exactly the same testing probability γ as in that work, whereas for Fig. 1b we optimized over the γ value; we discuss the details of this choice in that appendix as well.) We see that at the value $n = 1.5 \times 10^6$ used in that experiment, we improve the finite-size key rate by about a factor of 3. Similarly, we also reduce the minimum n required for nonzero finite-size key by nearly a factor of 3. Such improvements are critical in the context of practical demonstrations of DIQKD, as they significantly reduce the experimental requirements for a desired length of final key.

As a final remark, we note that in [29], a framework was developed to prove security for variable-length protocols, which do not simply make a binary accept/abort decision but rather adjust the length of the final key depending on the observed values. The key concept considered in their analysis is a "weighted" version of Rényi entropy; refer to e.g. [30] for further details. Our bound in Theorem 2 can also be applied to bound these weighted Rényi entropies, and would hence also be able to prove security for variable-length protocols, though we leave a detailed analysis for future work.

IV. CONCLUSION

This work represents an important step towards a practical implementation of DIQKD which represents the highest level of security, allowing for secret key generation using untrusted hardware. We leverage the Rényi Entropy Accumulation Theorem [15] and demonstrate that it yields significantly tighter finite-size key rates for DIQKD protocols based on the CHSH inequality. To do this, we derive tight analytical bounds on Rényi entropies in a device-independent manner, which in turn provide a tight relationship between the CHSH value and

the amount of Rényi entropy accumulated. Our results can be seen as a generalization of the expression derived in [11], which we recover as a special case. In Figure 1, we demonstrate the improvement in finite-size key rates for DIQKD protocols by comparing the rates from the DIQKD experimental demonstration in [27] to those that are achievable using our approach. In particular, we show that the work of [27] could triple their key-rates by using our technique, with no modifications to the experimental setup.

Our work prompts several pertinent questions towards the end goal of practical DI cryptography. Firstly, whilst we were also able to derive a tight analytical bound for the $\widetilde{H}_\alpha^\uparrow$ entropy, that bound is currently not applicable to improving finite-size analysis, as the current tools (e.g., [15]) use $\widetilde{H}_\alpha^\downarrow$ in their key rate expressions. Looking at the analogous results for device-dependent QKD [29–31], the key rates are actually computed in terms of $\widetilde{H}_\alpha^\uparrow$, providing tighter bounds on the finite-size rates. If we could develop similar results in the DI setting, we would then be able to use the bound $f_{\widetilde{H}_\alpha^\uparrow}(S)$ to obtain even higher finite-size key rates. In a second direction, the proofs of the analytical key rate formulae follow closely the work of [17] and are fairly generic in nature, hence it is likely that they could be extended to other families of Bell inequalities to obtain further analytical key rate formulae, providing simpler and tighter security proofs for protocols beyond CHSH. It may also be of some interest to see whether our techniques apply to the broader family of Rényi conditional entropies in [32] that unifies the cases considered in this work.

ACKNOWLEDGMENTS

TH acknowledges support from the Peter and Patricia Gruber Award and by the Air Force Office of Scientific Research under award number FA9550-22-1-0391. AP acknowledges support from the National Science Centre Poland (Grant No. 2022/46/E/ST2/00115). EYZT conducted research at the Institute for Quantum Computing, at the University of Waterloo, which is supported by Innovation, Science, and Economic Development Canada; support was also provided by NSERC under the Discovery Grants Program, Grant No. 341495. PB acknowledges support from the European union's Horizon Europe research and innovation programme under the project "Quantum Secure Networks Partnership" (QSNP, grant agreement No. 101114043).

[1] John Stewart Bell, "On the Einstein-Podolsky-Rosen paradox," Physics **1**, 195 (1964).
[2] Wen-Zhao Liu, Ming-Han Li, Sammy Ragy, Si-Ran Zhao, Bing Bai, Yang Liu, Peter J Brown, Jun Zhang, Roger Colbeck, Jingyun Fan, *et al.*, "Device-independent randomness expansion against quantum side information," Nature Physics , 1–4 (2021).
[3] Lynden K Shalm, Yanbao Zhang, Joshua C Bienfang, Collin Schlager, Martin J Stevens, Michael D Mazurek, Carlos Abellán, Waldimar Amaya, Morgan W Mitchell, Mohammad A Alhejji, *et al.*, "Device-independent randomness expansion with entangled photons," Nature Physics , 1–5 (2021).
[4] Anatoly Kulikov, Simon Storz, Josua D Schär, Martin Sandfuchs, Ramona Wolf, Florence Berterottière, Christoph Hellings, Renato Renner, and Andreas Wallraff, "Device-Independent Randomness Amplification," (2024), arXiv:2412.17931.
[5] Wei Zhang, Tim van Leent, Kai Redeker, Robert Garthoff, René Schwonnek, Florian Fertig, Sebastian Eppelt, Wenjamin Rosenfeld, Valerio Scarani, Charles C-W Lim, *et al.*, "A device-independent quantum key distribution system for distant users," Nature **607**, 687–691 (2022).
[6] David P Nadlinger, Peter Drmota, Bethan C Nichol, Gabriel Araneda, Dougal Main, Raghavendra Srinivas, David M Lucas, Christopher J Ballance, Kirill Ivanov, EY-Z Tan, *et al.*, "Experimental quantum key distribution certified by Bell's theorem," Nature **607**, 682–686 (2022).
[7] Wen-Zhao Liu, Yu-Zhe Zhang, Yi-Zheng Zhen, Ming-Han Li, Yang Liu, Jingyun Fan, Feihu Xu, Qiang Zhang, and Jian-Wei Pan, "Toward a Photonic Demonstration of Device-Independent Quantum Key Distribution," Physical Review Letters **129**, 050502 (2022).
[8] Víctor Zapatero, Tim van Leent, Rotem Arnon-Friedman, Wen-Zhao Liu, Qiang Zhang, Harald Weinfurter, and Marcos Curty, "Advances in device-independent quantum key distribution," npj quantum information **9**, 10 (2023).
[9] Ignatius W Primaatmaja, Koon Tong Goh, Ernest Y-Z Tan, John T-F Khoo, Shouvik Ghorai, and Charles C-W Lim, "Security of device-independent quantum key distribution protocols: a review," Quantum **7**, 932 (2023).
[10] John Clauser, Michael Horne, Abner Shimony, and R. Holt, "Proposed Experiment to Test Local Hidden-Variable Theories," Phys. Rev. Lett. **23**, 880–884 (1969).
[11] Stefano Pironio, Antonio Acín, Nicolas Brunner, Nicolas Gisin, Serge Massar, and Valerio Scarani, "Device-independent quantum key distribution secure against collective attacks," New Journal of Physics **11**, 045021 (2009).
[12] Frédéric Dupuis, Omar Fawzi, and Renato Renner, "Entropy Accumulation," Communications in Mathematical Physics **379**, 867–913 (2020).
[13] F. Dupuis and O. Fawzi, "Entropy accumulation with improved second-order term," IEEE Transactions on Information Theory , 1–1 (2019), 1805.11652.
[14] Tony Metger, Omar Fawzi, David Sutter, and Renato Renner, "Generalised entropy accumulation," in *2022 IEEE 63rd Annual Symposium on Foundations of Computer Science (FOCS)* (2022) pp. 844–850.
[15] Amir Arqand, Thomas A. Hahn, and Ernest Y.Z. Tan, "Generalized Rényi entropy accumulation theorem and

generalized quantum probability estimation," (2024), arXiv:2405.05912v3.
[16] M Ho, P Sekatski, EY-Z Tan, R Renner, J-D Bancal, and N Sangouard, "Noisy Preprocessing Facilitates a Photonic Realization of Device-Independent Quantum Key Distribution," Physical Review Letters **124**, 230502 (2020).
[17] Erik Woodhead, Antonio Acín, and Stefano Pironio, "Device-independent quantum key distribution with asymmetric CHSH inequalities," Quantum **5**, 443 (2021).
[18] Boris S Cirel'son, "Quantum generalizations of Bell's inequality," Letters in Mathematical Physics **4**, 93–100 (1980).
[19] Rotem Arnon-Friedman, Frédéric Dupuis, Omar Fawzi, Renato Renner, and Thomas Vidick, "Practical device-independent quantum cryptography via entropy accumulation," Nature Communications **9**, 459 (2018).
[20] Lluis Masanes, Stefano Pironio, and Antonio Acín, "Secure device-independent quantum key distribution with causally independent measurement devices," Nature communications **2**, 238 (2011), 1009.1567.
[21] Martin Müller-Lennert, Frédéric Dupuis, Oleg Szehr, Serge Fehr, and Marco Tomamichel, "On quantum Rényi entropies: A new generalization and some properties," Journal of Mathematical Physics **54** (2013), 1306.3142.
[22] Mark M Wilde, Andreas Winter, and Dong Yang, "Strong converse for the classical capacity of entanglement-breaking and Hadamard channels via a sandwiched Rényi relative entropy," Communications in Mathematical Physics **331**, 593–622 (2014).
[23] Dénes Petz, "Quasi-entropies for finite quantum systems," Reports on Mathematical Physics **23**, 57–65 (1986).
[24] G Murta, S B van Dam, J Ribeiro, R Hanson, and S Wehner, "Towards a realization of device-independent quantum key distribution," Quantum Science and Technology **4**, 035011 (2019).
[25] Ernest Y.-Z. Tan, Pavel Sekatski, Jean-Daniel Bancal, René Schwonnek, Renato Renner, Nicolas Sangouard, and Charles C.-W. Lim, "Improved DIQKD protocols with finite-size analysis," Quantum **6**, 880 (2022).
[26] Thomas A. Hahn, Ernest Y. Z. Tan, and Peter Brown, "Bounds on Petz-Rényi Divergences and their Applications for Device-Independent Cryptography," (2024), arXiv:2408.12313.
[27] D. P. Nadlinger, P. Drmota, B. C. Nichol, G. Araneda, D. Main, R. Srinivas, D. M. Lucas, C. J. Ballance, K. Ivanov, E. Y.-Z. Tan, P. Sekatski, R. L. Urbanke, R. Renner, N. Sangouard, and J.-D. Bancal, "Experimental quantum key distribution certified by Bell's theorem," Nature **607**, 682–686 (2022).
[28] Thomas M. Cover and Joy A. Thomas, *Elements of Information Theory 2nd Edition (Wiley Series in Telecommunications and Signal Processing)* (Wiley-Interscience, 2006).
[29] Thomas van Himbeeck and Peter Brown, "Tight and general finite-size security of quantum key distribution," (2025), in preparation.
[30] Omar Fawzi, Jan Kochanowski, Cambyse Rouzé, and Thomas Van Himbeeck, "Additivity and chain rules for quantum entropies via multi-index Schatten norms," (2025), arXiv:2502.01611.
[31] Amir Arqand and Ernest Y. Z. Tan, "Marginal-constrained entropy accumulation theorem," (2025), arXiv:2502.02563.
[32] Roberto Rubboli, Milad M. Goodarzi, and Marco Tomamichel, "Quantum Conditional Entropies," (2024), arXiv:2410.21976v1 [quant-ph].
[33] Pavel Sekatski, Jean-Daniel Bancal, Xavier Valcarce, Ernest Y.-Z. Tan, Renato Renner, and Nicolas Sangouard, "Device-independent quantum key distribution from generalized CHSH inequalities," Quantum **5**, 444 (2021).
[34] Marco Tomamichel, *Quantum Information Processing with Finite Resources* (Springer International Publishing, 2016).
[35] F.B. Hildebrand, *Introduction to Numerical Analysis* (McGraw-Hill, 1956).
[36] Andreas Bluhm, Ángela Capel, Paul Gondolf, and Antonio Pérez-Hernández, "Continuity of quantum entropic quantities via almost convexity," IEEE Transactions on Information Theory **69**, 5869–5901 (2023).
[37] Renato Renner, Nicolas Gisin, and Barbara Kraus, "Information-theoretic security proof for quantum-key-distribution protocols," Phys. Rev. A **72**, 012332 (2005).
[38] Wei Dai and Ted Krovetz, "VHASH Security," Cryptology ePrint Archive, Paper 2007/338 (2007).
[39] Christopher Portmann and Renato Renner, "Security in quantum cryptography," Reviews of Modern Physics **94**, 025008 (2022), 2102.00021.
[40] Frédéric Dupuis, "Privacy Amplification and Decoupling Without Smoothing," IEEE Transactions on Information Theory **69**, 7784–7792 (2023).
[41] Anindya De, Christopher Portmann, Thomas Vidick, and Renato Renner, "Trevisan's Extractor in the Presence of Quantum Side Information," SIAM Journal on Computing **41**, 915–940 (2012).
[42] Wolfgang Mauerer, Christopher Portmann, and Volkher B. Scholz, "A modular framework for randomness extraction based on Trevisan's construction," (2012), arXiv:1212.0520.

Appendix A: Notation and definitions

We begin by introducing the notation that we will be using. Quantum systems and their associated Hilbert spaces will often be denoted by capital letters, e.g. A. Given a space A, we denote the set of positive semidefinite operators acting on A by $\mathrm{Pos}(A)$. An operator $\rho \in \mathrm{Pos}(A)$ is called a *quantum state* if we have $\mathrm{Tr}[\rho] = 1$. The set of quantum states on A is denoted by $S_=(A)$. For two operators $\rho, \sigma \in \mathrm{Pos}(A)$ we write $\rho \ll \sigma$ if $\ker\sigma \subseteq \ker\rho$, where $\ker\tau \coloneqq \{|v\rangle : \tau|v\rangle = 0\}$. Further, we say ρ is orthogonal to σ, denoted by $\rho \perp \sigma$, if $\mathrm{Tr}[\rho\sigma] = 0$. The function log denotes the logarithm base 2. We conclude this section with formal definitions of the conditional entropies considered in this work.

Definition 3. Given any two positive semi-definite operators $\rho, \sigma \in \mathrm{Pos}(A)$ with $\mathrm{Tr}[\rho] > 0$, and $\alpha \in (1, \infty)$, the **sandwiched Rényi divergence** and **Petz-Rényi divergence** between ρ, σ are, respectively, given by:

$$\widetilde{D}_\alpha(\rho||\sigma) \coloneqq \begin{cases} \frac{1}{\alpha-1}\log\frac{\mathrm{Tr}\|\sigma^{\frac{1-\alpha}{2\alpha}}\rho\sigma^{\frac{1-\alpha}{2\alpha}}\|_\alpha^\alpha}{\mathrm{Tr}[\rho]} & \rho \ll \sigma \\ +\infty & \text{otherwise}, \end{cases} \tag{A1}$$

and

$$\bar{D}_\alpha(\rho||\sigma) \coloneqq \begin{cases} \frac{1}{\alpha-1}\log\frac{\mathrm{Tr}\left[\rho^\alpha\sigma^{1-\alpha}\right]}{\mathrm{Tr}[\rho]} & \rho \ll \sigma \\ +\infty & \text{otherwise}. \end{cases} \tag{A2}$$

These definitions are extended to $\alpha = 1$ and $\alpha = \infty$ by taking the respective limits.

Definition 4. For any bipartite, normalized state $\rho \in S_=(AB)$, and $\alpha \in [1, \infty]$, we define the following **conditional Rényi entropies**:

$$\widetilde{H}^\downarrow_\alpha(A|B)_\rho \coloneqq -\widetilde{D}_\alpha(\rho_{AB}||\mathbb{1}_A \otimes \rho_B) \tag{A3}$$

$$\widetilde{H}^\uparrow_\alpha(A|B)_\rho \coloneqq \sup_{\sigma\in S_=(B)} -\widetilde{D}_\alpha(\rho_{AB}||\mathbb{1}_A \otimes \sigma_B) \tag{A4}$$

$$\bar{H}^\downarrow_\alpha(A|B)_\rho \coloneqq -\bar{D}_\alpha(\rho_{AB}||\mathbb{1}_A \otimes \rho_B) \tag{A5}$$

$$\bar{H}^\uparrow_\alpha(A|B)_\rho \coloneqq \sup_{\sigma\in S_=(B)} -\bar{D}_\alpha(\rho_{AB}||\mathbb{1}_A \otimes \sigma_B)\,. \tag{A6}$$

It is these four conditional Rényi entropy families that we focus on in this work. For the Petz-Rényi entropies, i.e. the latter two expressions, we do not further consider $\alpha > 2$, as data-processing inequalities do not generally hold in this range. Also note that $\widetilde{H}^\uparrow_\infty$ is often referred to as the min-entropy $H_{\min}$ (some works instead refer to $\widetilde{H}^\downarrow_\infty$ as the min-entropy, though we shall not use this convention in this work).

Appendix B: Analytic Bounds and Proofs

Our main result, which encompasses the asymmetric CHSH score[1]

$$S_\beta = \sum_{abxy} (-1)^{xy+a+b}\beta^{1-x}\,\mathrm{Tr}\left[\rho_{Q_AQ_BE}\left(M_a^x \otimes N_b^y \otimes \mathbb{1}\right)\right] \tag{B1}$$

for all $\beta \in \mathbb{R}$ [17, 33], is given by the following theorem, which summarizes the results we obtain in the rest of this section. The result for the CHSH inequality is recovered by setting $\beta = 1$. We discuss how to include noisy preprocessing in Appendix C.

Theorem 5. *Let $|\beta| \geq 1$, $S_\beta \in \left[2|\beta|, 2\sqrt{1+\beta^2}\right]$, and $g_S = \sqrt{\frac{S_\beta^2}{4} - \beta^2}$. Then, we have*

$$f_{\widetilde{H}^\downarrow_\alpha}(S_\beta) = 1 + \frac{\alpha}{1-\alpha}\log\left[\left(\frac{1-g_S}{2}\right)^{\frac{1}{\alpha}} + \left(\frac{1+g_S}{2}\right)^{\frac{1}{\alpha}}\right] \tag{B2}$$

[1] We will refer to the expected value of a Bell-inequality as a "score", despite it not being formulated as a nonlocal game, as the term "value" is ambiguous in certain places.

$$f_{\widetilde{H}^{\uparrow}_{\alpha}}(S_\beta) = 1 + \left(\frac{2\alpha-1}{1-\alpha}\right)\log\left[\left(\frac{1-g_S}{2}\right)^{\frac{\alpha}{2\alpha-1}} + \left(\frac{1+g_S}{2}\right)^{\frac{\alpha}{2\alpha-1}}\right] \tag{B3}$$

for all $\alpha \in (1,\infty)$. *Similarly,*

$$f_{\bar{H}^{\downarrow}_{\alpha}}(S_\beta) = 1 + \frac{1}{1-\alpha}\log\left[\left(\frac{1-g_S}{2}\right)^{2-\alpha} + \left(\frac{1+g_S}{2}\right)^{2-\alpha}\right] \tag{B4}$$

$$f_{\bar{H}^{\uparrow}_{\alpha}}(S_\beta) = 1 + \frac{\alpha}{1-\alpha}\log\left[\left(\frac{1-g_S}{2}\right)^{\frac{1}{\alpha}} + \left(\frac{1+g_S}{2}\right)^{\frac{1}{\alpha}}\right] \tag{B5}$$

for all $\alpha \in (1,2)$.

Remark 6. *The above rate functions can be validly extended to* $\alpha = \infty$ *(for the first pair of formulas) and* $\alpha = 2$ *(for the second pair of formulas) by taking the limits* $\alpha \to \infty$ *and* $\alpha \nearrow 2$ *respectively, as we prove and discuss further in Sec. B 8. (Though for the case of* $f_{\bar{H}^{\uparrow}_2}$*, taking this limit is slightly unnecessary, in that one can directly substitute* $\alpha = 2$ *to obtain a well-defined expression which also matches the limiting value.) Note that the resulting formulas for* $f_{\widetilde{H}^{\downarrow}_{\infty}}$ *and* $f_{\bar{H}^{\downarrow}_2}$ *are discontinuous with respect to the CHSH score: more precisely, we have* $f_{\widetilde{H}^{\downarrow}_{\infty}}(S) = f_{\bar{H}^{\downarrow}_2}(S) = 0$ *for all* $S < 2\sqrt{2}$ *and* $f_{\widetilde{H}^{\downarrow}_{\infty}}(S) = f_{\bar{H}^{\downarrow}_2}(S) = 1$ *for* $S = 2\sqrt{2}$.

The same bounds hold for $\alpha < 1$, in suitable parameter ranges. However, this regime is not generally useful for QKD, so we omit it for brevity. We also note that this approach can be extended to $|\beta| \leq 1$ by instead replacing the expression for g_S with that in Eq. (B18) and then taking a concave envelope at an appropriate point in the formula, essentially the same as what was done in [17]. Note that taking this concave envelope is indeed necessary in this regime, as the formula resulting from only replacing the g_S expression does not generally satisfy the required convexity properties.

We note that the rate functions $f_{\widetilde{H}^{\downarrow}_{\alpha}}$ and $f_{\widetilde{H}^{\uparrow}_{\alpha}}$ are in fact related and we have $f_{\widetilde{H}^{\uparrow}_{\alpha}} = f_{\widetilde{H}^{\downarrow}_{2-1/\alpha}}$. From this, it immediately follows that our bounds satisfy $f_{\widetilde{H}^{\uparrow}_{\infty}} = f_{\widetilde{H}^{\downarrow}_2}$, which gives a reasoning for this special case that was observed in [24].

1. Qubit Reductions

This section contains several known results, which are necessary for future calculations. When both honest parties are restricted to two-input two-output projective measurements, Jordan's lemma, see e.g. [11], can be used to claim that it is sufficient to consider states and projective measurements of the form

$$\rho_{IQ_AQ_B} = \sum_i \Pr\left[I = i\right] |i\rangle\langle i|_I \otimes \rho^i_{Q_AQ_B} \tag{B6}$$

$$M^x_a = \sum_i |i\rangle\langle i|_I \otimes M^{i,x}_a \tag{B7}$$

$$N^y_b = \sum_i |i\rangle\langle i|_I \otimes N^{i,y}_b\,, \tag{B8}$$

where Q_A, Q_B are single-qubit Hilbert spaces on which the projective measurements $M^{i,x}_a, N^{i,y}_b$ act. Moreover, Eve's side-information consists of the classical register I, as well as the purification of the state $\rho^i_{Q_AQ_B}$, for any value $I = i$. We denote this *pure* tripartite state by $\rho^i_{Q_AQ_BE}$, and the post-measurement state is given by

$$\rho^{xy}_{IABE} = \sum_i \Pr\left[I = i\right] |i\rangle\langle i|_I \otimes \sum_{ab}\left[|ab\rangle\langle ab|_{AB} \otimes \rho_E^{iabxy}\right]. \tag{B9}$$

Using this qubit reduction, the asymmetric CHSH score S_β will similarly be expressed as a convex mixture over all i, i.e.

$$S_\beta = \sum_i \Pr\left[I = i\right] S^i_\beta \tag{B10}$$

$$S^i_\beta = \sum_{abxy} (-1)^{xy+a+b} \beta^{1-x} \operatorname{Tr}\left[\rho^i_{Q_A Q_B E}\left(M^{i,x}_a \otimes N^{i,y}_b \otimes \mathbb{1}\right)\right] . \tag{B11}$$

After applying the key-generation measurement, Alice's and Eve's joint post-measurement state is given by

$$\rho_{IAE} = \sum_i \Pr\left[I = i\right] |i\rangle\langle i|_I \otimes \sum_a \left[|a\rangle\langle a|_A \otimes \rho^{ia}_E\right] , \tag{B12}$$

where the measurement input $X = 0$ is kept implicit and $\rho^{ia}_E = \operatorname{Tr}_{Q_A}\left[\rho^i_{Q_A E}(M^{i,0}_a \otimes \mathbb{I}_E)\right]$.

Our goal is now to lower bound $\mathbb{H}(A|I = i, E)_\rho$ for each value $I = i$; or, in other words, to lower bound the entropy for states generated by qubit strategies. We note that by using the methods from [17], one can easily show the following (see Lemma 7 below for a general version of this property): each such state satisfies

$$\mathbb{H}(A|I = i, E)_\rho \geq \mathbb{H}(A|I = i, E)_\sigma , \tag{B13}$$

where

$$\sigma_{IAE} = \sum_i \Pr\left[I = i\right] |i\rangle\langle i|_I \otimes \sigma^i_{AE} \tag{B14}$$

$$\sigma^i_{AE} = \frac{1}{2} |0\rangle\langle 0| \otimes |\psi_=\rangle\langle\psi_=| + \frac{1}{2} |1\rangle\langle 1| \otimes |\psi_\neq\rangle\langle\psi_\neq| , \tag{B15}$$

for a pair of vectors $\{|\psi_=\rangle, |\psi_\neq\rangle\}$ that can be written in the following form (for some basis vectors $\{|0\rangle, |1\rangle\}$ of a two-dimensional subspace containing the span of $\{|\psi_=\rangle, |\psi_\neq\rangle\}$):

$$|\psi_=\rangle = |0\rangle \tag{B16}$$

$$|\psi_\neq\rangle = g^i_S |0\rangle + \sqrt{1 - g^{i2}_S} |1\rangle , \tag{B17}$$

where $g^i_S = \sqrt{\frac{S^{i2}_\beta}{4} - \beta^2}$ for $|\beta| \geq 1$. For $|\beta| \leq 1$, one instead uses $g^i_S = E_\beta(S^i_\beta)$, where

$$E_\beta = \begin{cases} \sqrt{\frac{S^{i2}_\beta}{4} - \beta^2}, & \text{if } |S^i_\beta| \geq 2\sqrt{1 + \beta^2 - \beta^4} \\ \sqrt{1 - \left(1 - \frac{1}{|\beta|}\sqrt{(1 - \beta^2)\left(\frac{S^{i2}_\beta}{4} - 1\right)}\right)^2}, & \text{if } |S^i_\beta| \leq 2\sqrt{1 + \beta^2 - \beta^4} . \end{cases} \tag{B18}$$

More generally, the above property of qubit strategies is an instance of the following lemma. Here, one should view $\mathbb{Q}$ as the function for which (depending on which case we are considering) either $\mathbb{H}^\downarrow(A|B) = \frac{1}{1-\alpha} \log \mathbb{Q}(A|B)$ or $\mathbb{H}^\uparrow(A|B) = \frac{\alpha}{1-\alpha} \log \mathbb{Q}(A|B)$ holds, using $\mathbb{H}$ to generically represent either Petz or sandwiched entropy.

Lemma 7. *Suppose $\mathbb{Q}(A|B) : S_=(AB) \to \mathbb{R}$ is a function satisfying*

1. *(Local unitary invariance): For any unitary V on A we have $\mathbb{Q}(A|B)_{\rho_{AB}} = \mathbb{Q}(A|B)_{V\rho_{AB}V^\dagger}$.*

2. *(Classical linearity): For any state $\rho_{ABC} = \sum_c \Pr[C = c]\rho^c_{AB} \otimes |c\rangle\langle c|_C$ classical on C, we have $\mathbb{Q}(A|BC)_{\rho_{ABC}} = \sum_c \Pr[C = c]\mathbb{Q}(A|B)_{\rho^c_{AB}}$.*

3. *(Data processing): For any $\rho_{ABC} \in S_=(ABC)$ we have $\mathbb{Q}(A|BC)_{\rho_{ABC}} \geq \mathbb{Q}(A|B)_{\rho_{AB}}$.*

Let $|\psi\rangle \in Q_A Q_B E$ with Q_A and Q_B being qubit systems, let $\{M_a\}_a$ be a rank-one projective measurement on Q_A and let

$$\rho_{AE} = \sum_a |a\rangle\langle a|_A \otimes \rho^a_E \tag{B19}$$

be the post-measurement state such that $g_S \geq 0$. Then there exists a state

$$\sigma_{AE} = \frac{1}{2} |0\rangle\langle 0| \otimes |\psi_=\rangle\langle\psi_=| + \frac{1}{2} |1\rangle\langle 1| \otimes |\psi_\neq\rangle\langle\psi_\neq| \tag{B20}$$

such that $|\langle\psi_{=}|\psi_{\neq}\rangle| \geq g_S$ *and*

$$\mathbb{Q}(A|E)_{\rho_{AE}} \leq \mathbb{Q}(A|E)_{\sigma_{AE}}\,, \tag{B21}$$

where g_S *is as defined above and, for any* $\beta \in \mathbb{R}$*, depends on the asymmetric CHSH score,* S_β*, achieved by the above system.*

Proof. The proof is detailed for the special case of the von Neumann entropy in [17]. We extend it to the more general setting considered here but we note that it remains almost exactly the same proof. Without loss of generality, let us assume that $\{M_a\}_a$ is a measurement in the computational basis.[2] Let

$$\rho_{AE} = \sum_a |a\rangle\langle a|_A \otimes \rho_E^a \tag{B22}$$

$$\rho'_{AE} = \sum_a |a\oplus 1\rangle\langle a\oplus 1|_A \otimes \rho_E^a\,. \tag{B23}$$

Due to local unitary invariance, we have that $\mathbb{Q}(A|E)_{\rho_{AE}} = \mathbb{Q}(A|E)_{\rho'_{AE}}$. Moreover, using classical linearity, it holds that

$$\mathbb{Q}(A|EF)_{\bar{\rho}_{AEF}} = \mathbb{Q}(A|E)_{\rho_{AE}}\,, \tag{B24}$$

where

$$\bar{\rho}_{AEF} = \frac{1}{2}\rho_{AE}\otimes|0\rangle\langle 0|_F + \frac{1}{2}\rho'_{AE}\otimes|1\rangle\langle 1|_F \tag{B25}$$

$$= \frac{1}{2}|0\rangle\langle 0|_A \otimes \left(\sum_a \rho_E^a \otimes |a\rangle\langle a|_F\right) + \frac{1}{2}|1\rangle\langle 1|_A \otimes \left(\sum_a \rho_E^{a\oplus 1}\otimes|a\rangle\langle a|_F\right)\,. \tag{B26}$$

Furthermore, the initial state can be written as

$$|\psi\rangle_{Q_A Q_B E} = |0\rangle_{Q_A}\otimes|\psi_0\rangle_{BE} + |1\rangle_{Q_A}\otimes|\psi_1\rangle_{BE}\,, \tag{B27}$$

and $|\psi_a\rangle_{BE}$ can be viewed as purifications of ρ_E^a. One potential extension of $\bar{\rho}_{AEF}$ is thus

$$\bar{\rho}_{ABEFF'} = \frac{1}{2}|0\rangle\langle 0|_A\otimes|\psi_{=}\rangle\langle\psi_{=}|_{BEFF'} + \frac{1}{2}|1\rangle\langle 1|_A\otimes|\psi_{\neq}\rangle\langle\psi_{\neq}|_{BEFF'}\,, \tag{B28}$$

where

$$|\psi_{=}\rangle = |\psi_0\rangle_{BE}\otimes|00\rangle_{FF'} + |\psi'_1\rangle_{BE}\otimes|11\rangle_{FF'} \tag{B29}$$

$$|\psi_{\neq}\rangle = |\psi'_1\rangle_{BE}\otimes|00\rangle_{FF'} + |\psi_0\rangle_{BE}\otimes|11\rangle_{FF'}\,, \tag{B30}$$

and $|\psi'_1\rangle_{BE} = (\sigma_X\otimes\mathbb{1})|\psi_1\rangle_{BE}$, where σ_X denotes the corresponding Pauli operator. Due to the data processing inequality, it holds that

$$\mathbb{Q}(A|E)_{\rho_{AE}} \leq \mathbb{Q}(A|BEFF')_{\bar{\rho}_{ABEFF'}}\,. \tag{B31}$$

We have thus identified a state for which the desired inequality holds. The states $\{|\psi_{=}\rangle, |\psi_{\neq}\rangle\}$ span (at most) a two-dimensional subspace, which can be embedded in the Hilbert space E, yielding the state in Eq. (B20). Moreover, by [17, Eqs. (73) and (95)] it must hold that $|\langle\psi_{=}|\psi_{\neq}\rangle| \geq g_S$. □

Remark 8. *Since* $\{|\psi_{=}\rangle, |\psi_{\neq}\rangle\}$ *span at most a two-dimensional subspace, they can always be written as*

$$|\psi_{=}\rangle = e^{i\phi}|0\rangle \tag{B32}$$

$$|\psi_{\neq}\rangle = |\langle\psi_{=}|\psi_{\neq}\rangle|\,|0\rangle + \sqrt{1-|\langle\psi_{=}|\psi_{\neq}\rangle|^2}\,|1\rangle\,. \tag{B33}$$

[2] Any two-output projective measurement on a qubit is related to it via a unitary, which we can implicitly apply a priori.

However, as can be seen in Eq. (B20), the angle ϕ disappears. As such, one can without loss of generality set it to $\phi = 0$. Moreover, in principle, the rate functions we derive in the following sections should depend on $|\langle\psi_=|\psi_\neq\rangle|$ rather than g_S. However, all our rate functions are monotonically increasing in the score (also in $|\langle\psi_=|\psi_\neq\rangle|$), and one can thus further lower bound the entropy by replacing $|\langle\psi_=|\psi_\neq\rangle|$ with g_S. This is a property that is also implicitly used in [17], and it is this feature that allows us to simply consider states of the form given by Eq. (B15). In our case, monotonicity is proven in Appendix B 6.

Remark 9. *The last two properties we wish to mention are as follows. Due to the relation between $\mathbb{Q}(A|E)$, and $\mathbb{H}(A|E)$, Eq. (B21) immediately implies Eq. (B13). Also, we do not consider degenerate measurements, which only have one potential outcome, in Lemma 7. The reason for this is simple, and also discussed in [17]: not only can such a measurement never be part of a CHSH set-up that violates the (asymmetric) CHSH inequality, but Alice's output would be deterministic. By choosing $\{|\psi_=\rangle, |\psi_\neq\rangle\}$ such that $|\langle\psi_=|\psi_\neq\rangle| = 0$, we have $\mathbb{H}(A|I=i,E)_\rho = \mathbb{H}(A|I=i,E)_\sigma = 0$. Eq. (B13) is thus trivially satisfied. Similarly, without loss of generality, we only consider set-ups for which g_S is well-defined. As an example, for $\beta = 1$, a well-defined g_S corresponds to achieving a CHSH score of $S \geq 2$. These other cases would not violate the desired (asymmetric) CHSH inequality. By instead choosing S_β such that $g_S = 0$, one achieves a higher score and we still bound any Rényi entropy by 0.*

2. Derivation of $f_{\widetilde{H}^\downarrow_\alpha}(S_\beta)$

Theorem 10. *Let $\alpha \in (1,\infty)$, $|\beta| \geq 1$, and $S_\beta \in \left[2|\beta|, 2\sqrt{1+\beta^2}\right]$. Then*

$$f_{\widetilde{H}^\downarrow_\alpha}(S_\beta) = 1 + \frac{\alpha}{1-\alpha}\log\left[\left(\frac{1-g_S}{2}\right)^{\frac{1}{\alpha}} + \left(\frac{1+g_S}{2}\right)^{\frac{1}{\alpha}}\right], \tag{B34}$$

where $g_S = \sqrt{\frac{S_\beta^2}{4} - \beta^2}$.

Proof. For any ρ_{IAE} of the form given by Eq. (B14), we first prove that Alice's measurement outcome after a key-generation measurement satisfies

$$\widetilde{H}^\downarrow_\alpha(A|X=0,IE)_\rho \geq 1 + \frac{\alpha}{1-\alpha}\log\left[\left(\frac{1-g_S}{2}\right)^{\frac{1}{\alpha}} + \left(\frac{1+g_S}{2}\right)^{\frac{1}{\alpha}}\right]. \tag{B35}$$

We now consider Alice's and Eve's bipartite state for some $I = i$. For any such σ_{AE} (we omit the index i for now) as in Eq. (B15), Eve's reduced density matrix is given by

$$\sigma_E = \frac{1-g_S}{2}|v_1\rangle\langle v_1| + \frac{1+g_S}{2}|v_2\rangle\langle v_2|, \tag{B36}$$

where

$$|v_1\rangle = -\sqrt{\frac{1-g_S}{2}}|0\rangle + \sqrt{\frac{1+g_S}{2}}|1\rangle \tag{B37}$$

$$|v_2\rangle = \sqrt{\frac{1+g_S}{2}}|0\rangle + \sqrt{\frac{1-g_S}{2}}|1\rangle. \tag{B38}$$

Plugging this directly into $\sigma_E^{\frac{1-\alpha}{2\alpha}}\sigma_{AE}\sigma_E^{\frac{1-\alpha}{2\alpha}}$ gives us that

$$\sigma_E^{\frac{1-\alpha}{2\alpha}}\sigma_{AE}\sigma_E^{\frac{1-\alpha}{2\alpha}} \tag{B39}$$

$$= \frac{1}{2}|0\rangle\langle 0|_A \otimes \left[\left(\frac{1-g_S}{2}\right)^{\frac{1}{\alpha}}|v_1\rangle\langle v_1|_E - \left(\frac{1-g_S^2}{4}\right)^{\frac{1}{2\alpha}}(|v_1\rangle\langle v_2|_E + |v_2\rangle\langle v_1|_E) + \left(\frac{1+g_S}{2}\right)^{\frac{1}{\alpha}}|v_2\rangle\langle v_2|_E\right]$$

$$+ \frac{1}{2}|1\rangle\langle 1|_A \otimes \left[\left(\frac{1-g_S}{2}\right)^{\frac{1}{\alpha}}|v_1\rangle\langle v_1|_E + \left(\frac{1-g_S^2}{4}\right)^{\frac{1}{2\alpha}}(|v_1\rangle\langle v_2|_E + |v_2\rangle\langle v_1|_E) + \left(\frac{1+g_S}{2}\right)^{\frac{1}{\alpha}}|v_2\rangle\langle v_2|_E\right] \tag{B40}$$

$$= \frac{1}{2}|0\rangle\langle 0|_A \otimes \left[\left(\left(\frac{1-g_S}{2}\right)^{\frac{1}{2\alpha}}|v_1\rangle_E - \left(\frac{1+g_S}{2}\right)^{\frac{1}{2\alpha}}|v_2\rangle_E\right)\left(\left(\frac{1-g_S}{2}\right)^{\frac{1}{2\alpha}}\langle v_1|_E - \left(\frac{1+g_S}{2}\right)^{\frac{1}{2\alpha}}\langle v_2|_E\right)\right]$$
$$+ \frac{1}{2}|1\rangle\langle 1|_A \otimes \left[\left(\left(\frac{1-g_S}{2}\right)^{\frac{1}{2\alpha}}|v_1\rangle_E + \left(\frac{1+g_S}{2}\right)^{\frac{1}{2\alpha}}|v_2\rangle_E\right)\left(\left(\frac{1-g_S}{2}\right)^{\frac{1}{2\alpha}}\langle v_1|_E + \left(\frac{1+g_S}{2}\right)^{\frac{1}{2\alpha}}\langle v_2|_E\right)\right]. \quad \text{(B41)}$$

This can alternatively be expressed as

$$\sigma_E^{\frac{1-\alpha}{2\alpha}}\sigma_{AE}\sigma_E^{\frac{1-\alpha}{2\alpha}} = \frac{1}{2}\left(\left(\frac{1-g_S}{2}\right)^{\frac{1}{\alpha}} + \left(\frac{1+g_S}{2}\right)^{\frac{1}{\alpha}}\right)|0\rangle\langle 0|_A \otimes |w_1\rangle\langle w_1|_E$$
$$+ \frac{1}{2}\left(\left(\frac{1-g_S}{2}\right)^{\frac{1}{\alpha}} + \left(\frac{1+g_S}{2}\right)^{\frac{1}{\alpha}}\right)|1\rangle\langle 1|_A \otimes |w_1'\rangle\langle w_1'|_E\,, \quad \text{(B42)}$$

where $\{|w_1\rangle, |w_1'\rangle\}$ are normalized vectors. It then directly follows from this that

$$\widetilde{H}_\alpha^\downarrow(A|E)_\sigma = \frac{1}{1-\alpha}\log\left[\mathrm{Tr}\left[\left(\sigma_E^{\frac{1-\alpha}{2\alpha}}\sigma_{AE}\sigma_E^{\frac{1-\alpha}{2\alpha}}\right)^\alpha\right]\right] \quad \text{(B43)}$$
$$= \frac{1}{1-\alpha}\log\frac{2\left[\left(\frac{1-g_S}{2}\right)^{\frac{1}{\alpha}} + \left(\frac{1+g_S}{2}\right)^{\frac{1}{\alpha}}\right]^\alpha}{2^\alpha} \quad \text{(B44)}$$
$$= \frac{1}{1-\alpha}\log\frac{\left[\left(\frac{1-g_S}{2}\right)^{\frac{1}{\alpha}} + \left(\frac{1+g_S}{2}\right)^{\frac{1}{\alpha}}\right]^\alpha}{2^{\alpha-1}} \quad \text{(B45)}$$
$$= 1 + \frac{\alpha}{1-\alpha}\log\left[\left(\frac{1-g_S}{2}\right)^{\frac{1}{\alpha}} + \left(\frac{1+g_S}{2}\right)^{\frac{1}{\alpha}}\right]. \quad \text{(B46)}$$

This concludes the calculations for the individual qubit block related to some index $I = i$. We now explicitly write the index for rest of the calculation. Let $h(S_\beta^i)$ denote the function

$$h(S_\beta^i) \coloneqq \left[\left(\frac{1-g_S^i}{2}\right)^{\frac{1}{\alpha}} + \left(\frac{1+g_S^i}{2}\right)^{\frac{1}{\alpha}}\right]^\alpha. \quad \text{(B47)}$$

Using both Eq. (B13) and [34, Prop. 5.1], it holds that

$$\widetilde{H}_\alpha^\downarrow(A|X=0,IE)_\rho \geq \frac{1}{1-\alpha}\log\left[\sum_i \Pr[I=i]\, 2^{(1-\alpha)\widetilde{H}_\alpha^\downarrow(A|I=i,E)_{\sigma^i}}\right] \quad \text{(B48)}$$
$$= 1 + \frac{1}{1-\alpha}\log\left[\sum_i \Pr[I=i]\, h(S_\beta^i)\right]. \quad \text{(B49)}$$

We show in Appendix B 6 that $h(S_\beta^i)$ is concave for $\alpha > 1$. The desired lower bound then follows from this property, together with the monotonicity of the logarithm, i.e.

$$\widetilde{H}_\alpha^\downarrow(A|X=0,IE)_\rho \geq 1 + \frac{1}{1-\alpha}\log\left[h(S_\beta)\right] \quad \text{(B50)}$$
$$= 1 + \frac{\alpha}{1-\alpha}\log\left[\left(\frac{1-g_S}{2}\right)^{\frac{1}{\alpha}} + \left(\frac{1+g_S}{2}\right)^{\frac{1}{\alpha}}\right], \quad \text{(B51)}$$

where $g_S = \sqrt{\frac{S_\beta^2}{4} - \beta^2}$ and $S_\beta = \Pr[I=i]\, S_\beta^i$. This inequality is saturated if Alice and Bob share the state

$$\sqrt{P_+}\,|\phi^+\rangle_{Q_AQ_B}|0\rangle_E + \sqrt{P_-}\,|\phi^-\rangle_{Q_AQ_B}|1\rangle_E\,, \quad \text{(B52)}$$

where $P_\pm = \frac{1}{2}(1 \pm g_S)$, and measure the observables $A_0 = \sigma_z$, $A_1 = \sigma_x$, $B_0 = \frac{\beta\sigma_z + g_S\sigma_x}{\sqrt{\beta^2+g_S^2}}$, $B_1 = \frac{\beta\sigma_z - g_S\sigma_x}{\sqrt{\beta^2+g_S^2}}$ (see Appendix B 7 for more details). The derived rate function is thus tight.

□

3. Derivation of $f_{\widetilde{H}^\uparrow_\alpha}(S_\beta)$

Theorem 11. *Let $\alpha \in (1,\infty)$, $|\beta| \geq 1$, and $S_\beta \in \left[2|\beta|, 2\sqrt{1+\beta^2}\right]$. Then*

$$f_{\widetilde{H}^\uparrow_\alpha}(S_\beta) = f_{\widetilde{H}^\downarrow_{2-\frac{1}{\alpha}}}(S_\beta)\,. \tag{B53}$$

Proof. For any ρ_{IAE} of the form given by Eq. (B14), we first consider Alice's and Eve's bipartite state for some $I = i$. Due to Eq. (A4), for any such σ_{AE} (we omit the index i for now) as in Eq. (B15), it must hold that

$$\widetilde{H}^\uparrow_\alpha(A|E)_\sigma \geq -\widetilde{D}_\alpha(\sigma_{AE}||\mathbb{1}_A \otimes \tau_E)\,, \tag{B54}$$

where we pick the following choice of state in the second argument:

$$\tau_E = \frac{\sigma_E^{\frac{\alpha}{2\alpha-1}}}{\mathrm{Tr}[\sigma_E^{\frac{\alpha}{2\alpha-1}}]}. \tag{B55}$$

For this state, however, one finds that

$$-\widetilde{D}_\alpha(\sigma_{AE}||\mathbb{1}_A \otimes \tau_E) = \frac{1}{1-\alpha}\log\left[\mathrm{Tr}\left[\left(\tau_E^{\frac{1-\alpha}{2\alpha}}\sigma_{AE}\tau_E^{\frac{1-\alpha}{2\alpha}}\right)^\alpha\right]\right] \tag{B56}$$

$$= \frac{1}{1-\alpha}\log\left[\frac{\mathrm{Tr}\left[\left(\sigma_E^{\frac{1-\alpha}{2(2\alpha-1)}}\sigma_{AE}\sigma_E^{\frac{1-\alpha}{2(2\alpha-1)}}\right)^\alpha\right]}{\mathrm{Tr}\left[\sigma_E^{\frac{\alpha}{2\alpha-1}}\right]^{1-\alpha}}\right] \tag{B57}$$

$$= \frac{1}{1-\alpha}\log\left[\frac{\mathrm{Tr}\left[\left(\sigma_E^{\frac{1-\alpha'}{2\alpha'}}\sigma_{AE}\sigma_E^{\frac{1-\alpha'}{2\alpha'}}\right)^\alpha\right]}{\mathrm{Tr}\left[\sigma_E^{\frac{1}{\alpha'}}\right]^{1-\alpha}}\right], \tag{B58}$$

where $\alpha' = 2 - \frac{1}{\alpha}$. Moreover, it holds that

$$\mathrm{Tr}\left[\left(\sigma_E^{\frac{1-\alpha'}{2\alpha'}}\sigma_{AE}\sigma_E^{\frac{1-\alpha'}{2\alpha'}}\right)^\alpha\right] = 2\left(\frac{\left(\frac{1-g_S}{2}\right)^{\frac{1}{\alpha'}} + \left(\frac{1+g_S}{2}\right)^{\frac{1}{\alpha'}}}{2}\right)^\alpha = \frac{1}{2^{\alpha-1}}\left(\left(\frac{1-g_S}{2}\right)^{\frac{1}{\alpha'}} + \left(\frac{1+g_S}{2}\right)^{\frac{1}{\alpha'}}\right)^\alpha \tag{B59}$$

$$\mathrm{Tr}\left[\sigma_E^{\frac{1}{\alpha'}}\right]^{1-\alpha} = \left(\left(\frac{1-g_S}{2}\right)^{\frac{1}{\alpha'}} + \left(\frac{1+g_S}{2}\right)^{\frac{1}{\alpha'}}\right)^{1-\alpha}, \tag{B60}$$

where the first equation directly follows from Eq. (B42) and the second equation is due to the decomposition $\sigma_E = \frac{1-g_S}{2}|v_1\rangle\langle v_1| + \frac{1+g_S}{2}|v_2\rangle\langle v_2|$, where $\{|v_1\rangle, |v_2\rangle\}$ are orthonormal vectors given by Eqs. (B37)–(B38). This then gives us that

$$\widetilde{H}^\uparrow_\alpha(A|E)_\sigma \geq \frac{1}{1-\alpha}\log\left[\frac{\mathrm{Tr}\left[\left(\sigma_E^{\frac{1-\alpha'}{2\alpha'}}\sigma_{AE}\sigma_E^{\frac{1-\alpha'}{2\alpha'}}\right)^\alpha\right]}{\mathrm{Tr}\left[\sigma_E^{\frac{1}{\alpha'}}\right]^{1-\alpha}}\right] \tag{B61}$$

$$= \frac{1}{1-\alpha}\log\left[2^{1-\alpha}\left(\left(\frac{1-g_S}{2}\right)^{\frac{1}{\alpha'}} + \left(\frac{1+g_S}{2}\right)^{\frac{1}{\alpha'}}\right)^{2\alpha-1}\right] \tag{B62}$$

$$= 1 + \frac{1}{1-\alpha} \log \left[\left(\left(\frac{1-g_S}{2} \right)^{\frac{1}{\alpha'}} + \left(\frac{1+g_S}{2} \right)^{\frac{1}{\alpha'}} \right)^{2\alpha-1} \right] \tag{B63}$$

$$= 1 + \frac{2\alpha-1}{1-\alpha} \log \left[\left(\frac{1-g_S}{2} \right)^{\frac{1}{\alpha'}} + \left(\frac{1+g_S}{2} \right)^{\frac{1}{\alpha'}} \right] \tag{B64}$$

$$= 1 + \frac{\alpha'}{1-\alpha'} \log \left[\left(\frac{1-g_S}{2} \right)^{\frac{1}{\alpha'}} + \left(\frac{1+g_S}{2} \right)^{\frac{1}{\alpha'}} \right] . \tag{B65}$$

This, however, is simply $f_{\widetilde{H}^{\downarrow}_{\alpha'}}(S_\beta)$. This concludes the calculations for the individual qubit block related to some index $I = i$. We now explicitly write the index for rest of the calculation. Let $h(S^i_\beta)$ denote the function

$$h(S^i_\beta) \coloneqq \left[\left(\frac{1-g^i_S}{2} \right)^{\frac{1}{\alpha'}} + \left(\frac{1+g^i_S}{2} \right)^{\frac{1}{\alpha'}} \right]^{\alpha'} . \tag{B66}$$

Using both Eq. (B13) and [34, Prop. 5.1], it holds that

$$\widetilde{H}^{\uparrow}_{\alpha}(A|X=0,IE)_\rho \geq \frac{\alpha}{1-\alpha} \log \left[\sum_i \Pr\left[I=i\right] 2^{\frac{(1-\alpha)}{\alpha} \widetilde{H}^{\uparrow}_{\alpha}(A|I=i,E)_{\sigma^i}} \right] \tag{B67}$$

$$= \frac{1}{1-\alpha'} \log \left[\sum_i \Pr\left[I=i\right] 2^{\left(1-\alpha'\right) \widetilde{H}^{\uparrow}_{\alpha}(A|I=i,E)_{\sigma^i}} \right] \tag{B68}$$

$$\geq 1 + \frac{1}{1-\alpha'} \log \left[\sum_i \Pr\left[I=i\right] h(S^i_\beta) \right] . \tag{B69}$$

We show in Appendix B 6 that $h(S^i_\beta)$ is concave for $\alpha' > 1$. The desired lower bound then follows from this property, together with the monotonicity of the logarithm, i.e.

$$\widetilde{H}^{\uparrow}_{\alpha}(A|X=0,IE)_\rho \geq 1 + \frac{1}{1-\alpha'} \log \left[h(S_\beta) \right] \tag{B70}$$

$$= 1 + \frac{\alpha'}{1-\alpha'} \log \left[\left(\frac{1-g_S}{2} \right)^{\frac{1}{\alpha'}} + \left(\frac{1+g_S}{2} \right)^{\frac{1}{\alpha'}} \right] \tag{B71}$$

$$= f_{\widetilde{H}^{\downarrow}_{\alpha'}}(S_\beta) , \tag{B72}$$

where $g_S = \sqrt{\frac{S^2_\beta}{4} - \beta^2}$ and $S_\beta = \Pr\left[I=i\right] S^i_\beta$. This inequality is saturated if Alice and Bob share the state

$$|\psi\rangle = \sqrt{P_+} \, |\phi^+\rangle_{Q_A Q_B} \, |0\rangle_E + \sqrt{P_-} \, |\phi^-\rangle_{Q_A Q_B} \, |1\rangle_E , \tag{B73}$$

where $P_\pm = \frac{1}{2}\left(1 \pm g_S\right)$, and measure the observables $A_0 = \sigma_z$, $A_1 = \sigma_x$, $B_0 = \frac{\beta\sigma_z + g_S\sigma_x}{\sqrt{\beta^2+g_S^2}}$, $B_1 = \frac{\beta\sigma_z - g_S\sigma_x}{\sqrt{\beta^2+g_S^2}}$. To see this, we first note it can be readily verified that this achieves the desired score, S_β. Moreover, after Alice applies measurement A_0, the classical output (stored in register A) satisfies

$$\widetilde{H}^{\uparrow}_{\alpha}\left(A|E\right)_\psi \leq \widetilde{H}^{\downarrow}_{2-\frac{1}{\alpha}}\left(A|E\right)_\psi , \tag{B74}$$

due to [34, Cor. 5.3]. The latter, however, is simply equal to $f_{\widetilde{H}^{\downarrow}_{2-\frac{1}{\alpha}}}(S_\beta)$ (see Appendix B 7 for more details). The derived upper and lower bounds are thus equal and the rate function must therefore be tight. □

4. Derivation of $f_{\bar{H}^{\downarrow}_{\alpha}}(S_\beta)$

Theorem 12. *Let $\alpha \in (1,2)$, $|\beta| \geq 1$, and $S_\beta \in \left[2|\beta|, 2\sqrt{1+\beta^2}\right]$. Then*

$$f_{\bar{H}^{\downarrow}_{\alpha}}(S_\beta) = 1 + \frac{1}{1-\alpha}\log\left[\left(\frac{1-g_S}{2}\right)^{2-\alpha} + \left(\frac{1+g_S}{2}\right)^{2-\alpha}\right], \tag{B75}$$

where $g_S = \sqrt{\frac{S_\beta^2}{4} - \beta^2}$.

Proof. For any ρ_{IAE} of the form given by Eq. (B14), we first prove that Alice's measurement outcome after a key-generation measurement satisfies

$$\bar{H}^{\downarrow}_{\alpha}(A|X=0,IE)_\rho \geq 1 + \frac{1}{1-\alpha}\log\left[\left(\frac{1-g_S}{2}\right)^{2-\alpha} + \left(\frac{1+g_S}{2}\right)^{2-\alpha}\right]. \tag{B76}$$

We now consider Alice's and Eve's bipartite state for some $I = i$. For any such σ_{AE} (we omit the index i for now) as in Eq. (B15), one can readily verify that

$$\sigma_{AE}^{\alpha} = \frac{1}{2^\alpha}|0\rangle\langle 0| \otimes |\psi_{=}\rangle\langle\psi_{=}| + \frac{1}{2^\alpha}|1\rangle\langle 1| \otimes |\psi_{\neq}\rangle\langle\psi_{\neq}| \tag{B77}$$

$$\sigma_{E}^{\frac{1-\alpha}{2}} = \left(\frac{1-g_S}{2}\right)^{\frac{1-\alpha}{2}}|v_1\rangle\langle v_1| + \left(\frac{1+g_S}{2}\right)^{\frac{1-\alpha}{2}}|v_2\rangle\langle v_2|, \tag{B78}$$

where $\{|\psi_{=}\rangle, |\psi_{\neq}\rangle\}$ are given by Eqs. (B16)–(B17) and $\{|v_1\rangle, |v_2\rangle\}$ by Eqs. (B37)–(B38). Using the fact that $\sqrt{1-g_S^2}\sqrt{\frac{1\pm g_S}{2}} = (1\pm g_S)\sqrt{\frac{1\mp g_S}{2}}$, we thus find that

$$\sigma_{E}^{\frac{1-\alpha}{2}}\sigma_{AE}^{\alpha}\sigma_{E}^{\frac{1-\alpha}{2}} = \frac{1}{2^\alpha}\left(\frac{1-g_S}{2}\right)^{2-\alpha}(|0\rangle\langle 0| + |1\rangle\langle 1|) \otimes |v_1\rangle\langle v_1| + \frac{1}{2^\alpha}\left(\frac{1+g_S}{2}\right)^{2-\alpha}(|0\rangle\langle 0| + |1\rangle\langle 1|) \otimes |v_2\rangle\langle v_2|. \tag{B79}$$

We note that one could alternatively have calculated $\sigma_{AE}^{\alpha}\sigma_{E}^{1-\alpha}$. Using this, we can express the Petz-Rényi entropy as

$$\bar{H}^{\downarrow}_{\alpha}(A|E)_\sigma = \frac{1}{1-\alpha}\log\left[\mathrm{Tr}\left[\sigma_{AE}^{\alpha}\sigma_{E}^{1-\alpha}\right]\right] \tag{B80}$$

$$= \frac{1}{1-\alpha}\log\left[\mathrm{Tr}\left[\sigma_{E}^{\frac{1-\alpha}{2}}\sigma_{AE}^{\alpha}\sigma_{E}^{\frac{1-\alpha}{2}}\right]\right] \tag{B81}$$

$$= \frac{1}{1-\alpha}\log\left[\frac{1}{2^{\alpha-1}}\left(\frac{1-g_S}{2}\right)^{2-\alpha} + \frac{1}{2^{\alpha-1}}\left(\frac{1+g_S}{2}\right)^{2-\alpha}\right] \tag{B82}$$

$$= 1 + \frac{1}{1-\alpha}\log\left[\left(\frac{1-g_S}{2}\right)^{2-\alpha} + \left(\frac{1+g_S}{2}\right)^{2-\alpha}\right] \tag{B83}$$

This concludes the calculations for the individual qubit block related to some index $I = i$. We now explicitly write the index for rest of the calculation. Let $h(S_\beta^i)$ denote the function

$$h(S_\beta^i) \coloneqq \left(\frac{1-g_S^i}{2}\right)^{2-\alpha} + \left(\frac{1+g_S^i}{2}\right)^{2-\alpha}. \tag{B84}$$

Using both Eq. (B13) and [34, Prop. 5.1], it holds that

$$\bar{H}^{\downarrow}_{\alpha}(A|X=0,IE)_\rho \geq \frac{1}{1-\alpha}\log\left[\sum_i \Pr[I=i]\, 2^{(1-\alpha)\bar{H}^{\downarrow}_{\alpha}(A|I=i,E)_{\sigma^i}}\right] \tag{B85}$$

$$= 1 + \frac{1}{1-\alpha}\log\left[\sum_i \Pr[I=i]\, h(S_\beta^i)\right]. \tag{B86}$$

We show in Appendix B 6 that $h(S_\beta^i)$ is concave for $\alpha > 1$. The desired lower bound then follows from this property, together with the monotonicity of the logarithm, i.e.

$$\bar{H}_\alpha^\downarrow(A|X=0,IE)_\rho \geq 1 + \frac{1}{1-\alpha}\log\left[h(S_\beta)\right] \tag{B87}$$

$$= 1 + \frac{1}{1-\alpha}\log\left[\left(\frac{1-g_S}{2}\right)^{2-\alpha} + \left(\frac{1+g_S}{2}\right)^{2-\alpha}\right], \tag{B88}$$

where $g_S = \sqrt{\frac{S_\beta^2}{4} - \beta^2}$ and $S_\beta = \Pr\left[I=i\right] S_\beta^i$. This inequality is saturated if Alice and Bob share the state

$$\sqrt{P_+}\left|\phi^+\right\rangle_{Q_A Q_B}\left|0\right\rangle_E + \sqrt{P_-}\left|\phi^-\right\rangle_{Q_A Q_B}\left|1\right\rangle_E, \tag{B89}$$

where $P_\pm = \frac{1}{2}\left(1 \pm g_S\right)$, and measure the observables $A_0 = \sigma_z$, $A_1 = \sigma_x$, $B_0 = \frac{\beta\sigma_z + g_S\sigma_x}{\sqrt{\beta^2+g_S^2}}$, $B_1 = \frac{\beta\sigma_z - g_S\sigma_x}{\sqrt{\beta^2+g_S^2}}$ (see Appendix B 7 for more details). The derived rate function is thus tight. □

5. Derivation of $f_{\bar{H}_\alpha^\uparrow}(S_\beta)$

Theorem 13. *Let $\alpha \in (1,2)$, $|\beta| \geq 1$, and $S_\beta \in \left[2|\beta|, 2\sqrt{1+\beta^2}\right]$. Then*

$$f_{\bar{H}_\alpha^\uparrow}(S_\beta) = 1 + \frac{\alpha}{1-\alpha}\log\left[\left(\frac{1-g_S}{2}\right)^{\frac{1}{\alpha}} + \left(\frac{1+g_S}{2}\right)^{\frac{1}{\alpha}}\right], \tag{B90}$$

where $g_S = \sqrt{\frac{S_\beta^2}{4} - \beta^2}$.

Proof. From [34, Lemma 5.1], we know that

$$\bar{H}_\alpha^\uparrow\left(A|E\right)_\sigma = \frac{\alpha}{1-\alpha}\log\left[\mathrm{Tr}\left[\mathrm{Tr}_A\left(\sigma_{AE}^\alpha\right)^{\frac{1}{\alpha}}\right]\right]. \tag{B91}$$

For any ρ_{IAE} of the form given by Eq. (B14), we first consider Alice's and Eve's bipartite state for some $I = i$. Recall that, for any σ_{AE} (we omit the index i for now) as in Eq. (B15), one can readily verify that

$$\sigma_{AE}^\alpha = \frac{1}{2^\alpha}\left|0\right\rangle\!\left\langle 0\right| \otimes \left|\psi_=\right\rangle\!\left\langle\psi_=\right| + \frac{1}{2^\alpha}\left|1\right\rangle\!\left\langle 1\right| \otimes \left|\psi_\neq\right\rangle\!\left\langle\psi_\neq\right| \tag{B92}$$

$$\mathrm{Tr}_A\left(\sigma_{AE}^\alpha\right) = \frac{1}{2^\alpha}\left|\psi_=\right\rangle\!\left\langle\psi_=\right| + \frac{1}{2^\alpha}\left|\psi_\neq\right\rangle\!\left\langle\psi_\neq\right| = \frac{1}{2^{\alpha-1}}\left(\frac{1-g_S}{2}\left|v_1\right\rangle\!\left\langle v_1\right| + \frac{1+g_S}{2}\left|v_2\right\rangle\!\left\langle v_2\right|\right), \tag{B93}$$

where $\{\left|\psi_=\right\rangle, \left|\psi_\neq\right\rangle\}$ are given by Eqs. (B16)–(B17) and $\{\left|v_1\right\rangle, \left|v_2\right\rangle\}$ by Eqs. (B37)–(B38).

$$\bar{H}_\alpha^\uparrow\left(A|E\right)_\sigma = \frac{\alpha}{1-\alpha}\log\left[\mathrm{Tr}\left[\mathrm{Tr}_A\left(\sigma_{AE}^\alpha\right)^{\frac{1}{\alpha}}\right]\right] \tag{B94}$$

$$= \frac{\alpha}{1-\alpha}\log\left[2^{\frac{1-\alpha}{\alpha}}\left(\left(\frac{1-g_S}{2}\right)^{\frac{1}{\alpha}} + \left(\frac{1+g_S}{2}\right)^{\frac{1}{\alpha}}\right)\right] \tag{B95}$$

$$= 1 + \frac{\alpha}{1-\alpha}\log\left[\left(\frac{1-g_S}{2}\right)^{\frac{1}{\alpha}} + \left(\frac{1+g_S}{2}\right)^{\frac{1}{\alpha}}\right]. \tag{B96}$$

This concludes the calculations for the individual qubit block related to some index $I = i$. We now explicitly write the index for rest of the calculation. Let $h(S_\beta^i)$ denote the function

$$h(S_\beta^i) \coloneqq \left(\frac{1-g_S^i}{2}\right)^{\frac{1}{\alpha}} + \left(\frac{1+g_S^i}{2}\right)^{\frac{1}{\alpha}}. \tag{B97}$$

Using both Eq. (B13) and [34, Prop. 5.1], it holds that

$$\bar{H}^{\uparrow}_{\alpha}(A|X=0,IE)_{\rho} \geq \frac{\alpha}{1-\alpha}\log\left[\sum_i \Pr\left[I=i\right] 2^{\frac{(1-\alpha)}{\alpha}\bar{H}^{\uparrow}_{\alpha}(A|I=i,E)_{\sigma^i}}\right] \tag{B98}$$

$$= 1 + \frac{\alpha}{1-\alpha}\log\left[\sum_i \Pr\left[I=i\right] h(S^i_{\beta})\right] . \tag{B99}$$

We show in Appendix B 6 that $h(S^i_{\beta})$ is concave for $\alpha > 1$. The desired lower bound then follows from this property, together with the monotonicity of the logarithm, i.e.

$$\bar{H}^{\uparrow}_{\alpha}(A|X=0,IE)_{\rho} \geq 1 + \frac{\alpha}{1-\alpha}\log\left[h(S_{\beta})\right] \tag{B100}$$

$$= 1 + \frac{\alpha}{1-\alpha}\log\left[\left(\frac{1-g_S}{2}\right)^{\frac{1}{\alpha}} + \left(\frac{1+g_S}{2}\right)^{\frac{1}{\alpha}}\right] , \tag{B101}$$

where $g_S = \sqrt{\frac{S^2_{\beta}}{4} - \beta^2}$ and $S_{\beta} = \Pr\left[I=i\right] S^i_{\beta}$. This inequality is saturated if Alice and Bob share the state

$$|\psi\rangle = \sqrt{P_+}\,|\phi^+\rangle_{Q_AQ_B}|0\rangle_E + \sqrt{P_-}\,|\phi^-\rangle_{Q_AQ_B}|1\rangle_E , \tag{B102}$$

where $P_{\pm} = \frac{1}{2}\left(1 \pm g_S\right)$, and measure the observables $A_0 = \sigma_z$, $A_1 = \sigma_x$, $B_0 = \frac{\beta\sigma_z + g_S\sigma_x}{\sqrt{\beta^2+g_S^2}}$, $B_1 = \frac{\beta\sigma_z - g_S\sigma_x}{\sqrt{\beta^2+g_S^2}}$ (see Appendix B 7 for more details). The derived rate function is thus tight. □

6. Monotonicity and Concavity Properties

In this section, we aim to prove the monotonicity and concavity of the following three functions for certain regions of $\alpha > 1$ and $|\beta| \geq 1$:

$$h_1(S_{\beta}) = \left[\left(\frac{1-g_S}{2}\right)^{\frac{1}{\alpha}} + \left(\frac{1+g_S}{2}\right)^{\frac{1}{\alpha}}\right]^{\alpha} \tag{B103}$$

$$h_2(S_{\beta}) = \left(\frac{1-g_S}{2}\right)^{2-\alpha} + \left(\frac{1+g_S}{2}\right)^{2-\alpha} \tag{B104}$$

$$h_3(S_{\beta}) = \left(\frac{1-g_S}{2}\right)^{\frac{1}{\alpha}} + \left(\frac{1+g_S}{2}\right)^{\frac{1}{\alpha}} , \tag{B105}$$

where $g_S = \sqrt{\frac{S^2_{\beta}}{4} - \beta^2}$. These functions can alternatively be expressed as

$$h_i(S_{\beta}) = f_i(\bar{g}(S_{\beta})) , \tag{B106}$$

where

$$f_1(x) = \left[\left(\frac{1-\sqrt{x}}{2}\right)^{\frac{1}{\alpha}} + \left(\frac{1+\sqrt{x}}{2}\right)^{\frac{1}{\alpha}}\right]^{\alpha} \tag{B107}$$

$$f_2(x) = \left(\frac{1-\sqrt{x}}{2}\right)^{2-\alpha} + \left(\frac{1+\sqrt{x}}{2}\right)^{2-\alpha} \tag{B108}$$

$$f_3(x) = \left(\frac{1-\sqrt{x}}{2}\right)^{\frac{1}{\alpha}} + \left(\frac{1+\sqrt{x}}{2}\right)^{\frac{1}{\alpha}} , \tag{B109}$$

and $\bar{g}(S_{\beta}) = \frac{S^2_{\beta}}{4} - \beta^2$. The first derivative of $h(S_{\beta})$ is given by $h_i'(S_{\beta}) = f_i'(\bar{g}(S_{\beta}))\cdot\bar{g}'(S_{\beta})$. Since $\bar{g}'(S_{\beta}) > 0$ in the regime $S_{\beta} \in \left[2|\beta|, 2\sqrt{1+\beta^2}\right]$, the monotonicity of $h_i(S)$ is purely determined by $f_i(x)$. We now show that $f_i'(x) \leq 0$

in the relevant regimes $\alpha > 1$. In particular this ensures that the rate functions are monotonically increasing, a property which is relevant for the discussion in Appendix B 1.

Proposition 14. *For all $x \in [0,1]$, the functions $f_1(x)$ and $f_3(x)$ are monotonically decreasing for $\alpha \in (1,\infty)$. Moreover, for all $x \in [0,1]$, the function $f_2(x)$ is monotonically decreasing for $\alpha \in (1,2)$.*

Proof. Note that because $f_1(x) = f_3(x)^\alpha$,

$$f_1'(x) = \alpha f_3(x)^{\alpha-1} f_3'(x)\,. \tag{B110}$$

However, because $f_3(x) \geq 0$, $f_1'(x)$ must have the same sign as $f_3'(x)$. We will now prove that $f_3'(x) \leq 0$ for all $x \in (0,1)$. The continuity of $f_3(x)$ then ensures that $f_3'(x)$ is monotonic over the entire regime $x \in [0,1]$. For $\alpha > 1$, it follows that

$$f_3'(x) = \frac{(1+\sqrt{x})^{\frac{1}{\alpha}-1} - (1-\sqrt{x})^{\frac{1}{\alpha}-1}}{2\alpha\sqrt{x}} \tag{B111}$$

$$\leq 0\,, \tag{B112}$$

thus ensuring monotonicity. For $f_2(x)$, one can either prove monotonicity by noting that it is equivalent to $f_3(x)$ after a suitable modification of the Rényi parameter α, or alternatively one can see it via

$$f_2'(x) = \frac{(2-\alpha)\left((1+\sqrt{x})^{1-\alpha} - (1-\sqrt{x})^{1-\alpha}\right)}{2\alpha\sqrt{x}} \tag{B113}$$

$$\leq 0\,, \tag{B114}$$

which holds for all $x \in (0,1)$. Again one can extend the result to $x \in [0,1]$ via a continuity argument. □

As was argued in [17], to prove concavity of $h_i(S_\beta)$, it is sufficient to prove that $f_i(x)$ is concave and monotonically decreasing and that $\bar{g}(S_\beta)$ is convex. Since $\bar{g}(S_\beta)$ is clearly convex and we have shown that all $f_i(x)$ are decreasing, it simply remains to show that each $f_i(x)$ is concave.

Proposition 15. *For all $x \in [0,1]$, the function*

$$f_3(x) = \left(1-\sqrt{x}\right)^{\frac{1}{\alpha}} + \left(1+\sqrt{x}\right)^{\frac{1}{\alpha}} \tag{B115}$$

is concave for $\alpha \in (1,\infty)$.

Proof. Due to continuity arguments, it is sufficient to prove it for all $x \in (0,1)$. The second derivative of this function is given by

$$f_3''(x) = \frac{(1-\sqrt{x})^{\frac{1}{\alpha}-1}}{4\alpha x^{\frac{3}{2}}} - \frac{(1+\sqrt{x})^{\frac{1}{\alpha}-1}}{4\alpha x^{\frac{3}{2}}} - \frac{(\alpha-1)\left((1+\sqrt{x})^{\frac{1}{\alpha}-2} + (1-\sqrt{x})^{\frac{1}{\alpha}-2}\right)}{4\alpha^2 x} \tag{B116}$$

$$= \frac{1}{4\alpha x^{\frac{3}{2}}}\left(\left(1-\sqrt{x}\right)^{\frac{1}{\alpha}-2}\left(1-\sqrt{x}-\frac{\alpha-1}{\alpha}\sqrt{x}\right) - \left(1+\sqrt{x}\right)^{\frac{1}{\alpha}-2}\left(1+\sqrt{x}+\frac{\alpha-1}{\alpha}\sqrt{x}\right)\right) \tag{B117}$$

$$= \frac{1}{4\alpha x^{\frac{3}{2}}}\left(\left(1-\sqrt{x}\right)^{\frac{1}{\alpha}-2}\left(1-\left(2-\frac{1}{\alpha}\right)\sqrt{x}\right) - \left(1+\sqrt{x}\right)^{\frac{1}{\alpha}-2}\left(1+\left(2-\frac{1}{\alpha}\right)\sqrt{x}\right)\right). \tag{B118}$$

Note that, for any $\alpha > 1$ and $y \in [0,1)$,

$$(1-y)^{2-\frac{1}{\alpha}} \geq 1 - \left(2-\frac{1}{\alpha}\right)y + \frac{(\alpha-1)(2\alpha-1)}{2\alpha^2}y^2 \tag{B119}$$

$$(1+y)^{2-\frac{1}{\alpha}} \leq 1 - \left(2+\frac{1}{\alpha}\right)y + \frac{(\alpha-1)(2\alpha-1)}{2\alpha^2}y^2\,. \tag{B120}$$

The right-hand-side are simply Taylor approximations of the left-hand-side. These inequalities follow from the fact that the error arising from these Taylor approximations are proportional to the third derivatives of the left-hand-side

at some value $\xi \in [0, y]$; see e.g. [35, Eq. (1.3.2)]. The third derivative, i.e.

$$\frac{d^3}{dy^3}(1-y)^{2-\frac{1}{\alpha}} = \frac{(2\alpha-1)(\alpha-1)}{\alpha^3}(1-y)^{-1-\frac{1}{\alpha}} \tag{B121}$$

$$\frac{d^3}{dy^3}(1+y)^{2-\frac{1}{\alpha}} = -\frac{(2\alpha-1)(\alpha-1)}{\alpha^3}(1+y)^{-1-\frac{1}{\alpha}}\ , \tag{B122}$$

are positive and negative, respectively, for all $y \in [0,1)$, thus ensuring that the bounds hold. Using this and setting $y = \sqrt{x}$, we find that for all $\alpha > 1$,

$$\begin{aligned} f_3''(x) \leq{}& \frac{1}{4\alpha x^{\frac{3}{2}}}\left(1-\sqrt{x}\right)^{\frac{1}{\alpha}-2}\left(\left(1-\sqrt{x}\right)^{2-\frac{1}{\alpha}} - \frac{(\alpha-1)(2\alpha-1)}{2\alpha^2}x\right) \\ &- \frac{1}{4\alpha x^{\frac{3}{2}}}\left(1+\sqrt{x}\right)^{\frac{1}{\alpha}-2}\left(\left(1+\sqrt{x}\right)^{2-\frac{1}{\alpha}} - \frac{(\alpha-1)(2\alpha-1)}{2\alpha^2}x\right) && \text{(B123)} \\ ={}& \frac{1}{4\alpha x^{\frac{3}{2}}}\left(\left(1-\left(1-\sqrt{x}\right)^{\frac{1}{\alpha}-2}\frac{(\alpha-1)(2\alpha-1)}{2\alpha^2}x\right) - \left(1-\left(1+\sqrt{x}\right)^{\frac{1}{\alpha}-2}\frac{(\alpha-1)(2\alpha-1)}{2\alpha^2}x\right)\right) && \text{(B124)} \\ ={}& \frac{1}{4\alpha x^{\frac{3}{2}}}\frac{(\alpha-1)(2\alpha-1)}{2\alpha^2}x\left(\left(1+\sqrt{x}\right)^{\frac{1}{\alpha}-2} - \left(1-\sqrt{x}\right)^{\frac{1}{\alpha}-2}\right) && \text{(B125)} \\ \leq{}& 0\,. && \text{(B126)} \end{aligned}$$

This concludes the concavity proof for $\alpha > 1$.

□

Proposition 16. *For all $x \in [0,1]$, the function*

$$f_1(x) = \left(\left(1-\sqrt{x}\right)^{\frac{1}{\alpha}} + \left(1+\sqrt{x}\right)^{\frac{1}{\alpha}}\right)^{\alpha} \tag{B127}$$

is concave for $\alpha \in (1,\infty)$.

Proof. Due to continuity arguments, it is sufficient to prove it for all $x \in (0,1)$. The second derivative of this function is given by

$$\begin{aligned} f_1''(x) = \frac{k(x)}{(1-x)^2}\Big(&(1-x)^{\frac{1}{\alpha}}\left((4-2\alpha)\sqrt{x} - 2\alpha x^{\frac{3}{2}}\right) - \alpha\left(1-\sqrt{x}\right)^{\frac{2}{\alpha}}\left(-1-\sqrt{x}+x+x^{\frac{3}{2}}\right) \\ &-\alpha\left(1+\sqrt{x}\right)^{\frac{2}{\alpha}}\left(1-\sqrt{x}-x+x^{\frac{3}{2}}\right)\Big)\ , \end{aligned} \tag{B128}$$

where $k(x) = \dfrac{\left(\left(1-\sqrt{x}\right)^{\frac{1}{\alpha}}+\left(1+\sqrt{x}\right)^{\frac{1}{\alpha}}\right)^{\alpha-2}}{4\alpha x^{\frac{3}{2}}} \geq 0$. This can be simplified as follows.

$$\begin{aligned} f_1''(x) &= \frac{k(x)}{(1-x)^2}\left((1-x)^{\frac{1}{\alpha}}\left((4-2\alpha)\sqrt{x} - 2\alpha x^{\frac{3}{2}}\right) + \alpha\left(1-\sqrt{x}\right)^{\frac{2}{\alpha}}\left(1+\sqrt{x}\right)(1-x) - \alpha\left(1+\sqrt{x}\right)^{\frac{2}{\alpha}}\left(1-\sqrt{x}\right)(1-x)\right) \\ &= \frac{k(x)}{(1-x)^2}\left((1-x)^{\frac{1}{\alpha}}\left((4-2\alpha)\sqrt{x} - 2\alpha x^{\frac{3}{2}}\right) + \alpha(1-x)^2\left(\left(1-\sqrt{x}\right)^{\frac{2}{\alpha}-1} - \left(1+\sqrt{x}\right)^{\frac{2}{\alpha}-1}\right)\right) && \text{(B129)} \\ &= k(x)\left((1-x)^{\frac{1}{\alpha}-2}\left((4-2\alpha)\sqrt{x} - 2\alpha x^{\frac{3}{2}}\right) + \alpha\left(\left(1-\sqrt{x}\right)^{\frac{2}{\alpha}-1} - \left(1+\sqrt{x}\right)^{\frac{2}{\alpha}-1}\right)\right) && \text{(B130)} \end{aligned}$$

For the regime $\alpha \geq 2$, we find that for any $y \in [0,1)$

$$\alpha\left(1-y^2\right)^{2-\frac{1}{\alpha}}\left((1+y)^{\frac{2}{\alpha}-1} - (1-y)^{\frac{2}{\alpha}-1}\right) \geq (4-2\alpha)\,y\,. \tag{B131}$$

Note that equality holds when $y = 0$. The inequality then follows from the fact that

$$\frac{d}{dy}\left(\alpha\left(1-y^2\right)^{2-\frac{1}{\alpha}}\left((1+y)^{\frac{2}{\alpha}-1} - (1-y)^{\frac{2}{\alpha}-1}\right) - (4-2\alpha)\,y\right) \tag{B132}$$

$$= \left(1-y^2\right)^{1-\frac{1}{\alpha}} \left((3\alpha y-\alpha+2)\,(1-y)^{\frac{2}{\alpha}-1} - (3\alpha y+\alpha-2)\,(1+y)^{\frac{2}{\alpha}-1}\right) - (4-2\alpha) \tag{B133}$$

$$\geq \left(1-y^2\right)^{1-\frac{1}{\alpha}} \left((2-\alpha)\,(1-y)^{\frac{2}{\alpha}-1} + (2-\alpha)\,(1+y)^{\frac{2}{\alpha}-1}\right) - (4-2\alpha) \tag{B134}$$

$$= (2-\alpha)\left(1-y^2\right)^{1-\frac{1}{\alpha}} \left((1-y)^{\frac{2}{\alpha}-1} + (1+y)^{\frac{2}{\alpha}-1}\right) - (4-2\alpha) \tag{B135}$$

$$\geq (2-\alpha)\left(1-y^2\right)^{1-\frac{2}{\alpha}} \left((1-y)^{\frac{2}{\alpha}-1} + (1+y)^{\frac{2}{\alpha}-1}\right) - (4-2\alpha) \tag{B136}$$

$$= (2-\alpha)\left((1+y)^{1-\frac{2}{\alpha}} + (1-y)^{1-\frac{2}{\alpha}}\right) - (4-2\alpha) \tag{B137}$$

$$\geq 2\,(2-\alpha) - (4-2\alpha) \tag{B138}$$

$$= 0\,. \tag{B139}$$

The second-to-last line holds with equality for $y = 0$. For $y > 0$, the inequality holds because $(1+y)^{1-\frac{2}{\alpha}} + (1-y)^{1-\frac{2}{\alpha}} \geq 0$ and

$$\frac{d}{dy}\left((1+y)^{1-\frac{2}{\alpha}} + (1-y)^{1-\frac{2}{\alpha}}\right) = -\frac{(\alpha-2)\left((1-y)^{-\frac{2}{\alpha}} - (1+y)^{-\frac{2}{\alpha}}\right)}{\alpha} \tag{B140}$$

$$\leq 0\,. \tag{B141}$$

Using Eq. (B131) and setting $y = \sqrt{x}$, it follows that

$$f_1''(x) \leq -2\alpha x^{\frac{3}{2}} k(x)\,(1-x)^{1/\alpha-2} \tag{B142}$$

$$\leq 0\,. \tag{B143}$$

For $1 < \alpha \leq 2$, we instead use the inequality

$$\alpha\left(1-y^2\right)^{2-\frac{1}{\alpha}} \left((1+y)^{\frac{2}{\alpha}-1} - (1-y)^{\frac{2}{\alpha}-1}\right) \geq (4-2\alpha)\,y + \frac{2\,(\alpha-2)\left(3\alpha^2+2\alpha-2\right)}{3\alpha^2} y^3\,. \tag{B144}$$

The right-hand-side should be viewed as a third-order Taylor approximation, and the bound holds because the error arising from this Taylor approximation is proportional to the fourth derivative at some $\xi \in [0, y]$, see [35, Eq. (1.3.2)], and

$$\frac{d^4}{dy^4}\left(\alpha\left(1-y^2\right)^{2-\frac{1}{\alpha}} \left((1+y)^{\frac{2}{\alpha}-1} - (1-y)^{\frac{2}{\alpha}-1}\right)\right) \tag{B145}$$

$$= \frac{16\,(\alpha-1)\,(\alpha+1)\,(2\alpha-1)\left(1-y^2\right)^{-\frac{1}{\alpha}-2} \left((1+y)^{\frac{2}{\alpha}-1} - (1-y)^{\frac{2}{\alpha}-1}\right)}{a^4} \tag{B146}$$

$$\geq 0\,. \tag{B147}$$

It then holds that

$$f_1''(x) \leq x^{\frac{3}{2}} k(x)\,(1-x)^{1/\alpha-2} \left(\frac{2\,(2-\alpha)\left(3\alpha^2+2\alpha-2\right)}{3\alpha^2} - 2\alpha\right) \leq 0\,. \tag{B148}$$

The last inequality follows from the fact that

$$2\,(2-\alpha)\left(3\alpha^2+2\alpha-2\right) = 6\alpha^3 \tag{B149}$$

has the three solutions $\alpha \in \{-1, \frac{2}{3}, 1\}$. The fraction

$$\left(\frac{2\,(2-\alpha)\left(3\alpha^2+2\alpha-2\right)}{3\alpha^2} - 2\alpha\right) \tag{B150}$$

must thus have the same sign for all $1 < \alpha \leq 2$. Explicitly verifying the sign for one such α then suffices to prove the claim. This concludes the proof that $f_1(x)$ is concave for all $\alpha > 1$. □

Proposition 17. *For all $x \in [0,1]$, the function*

$$f_2(x) = \left(1-\sqrt{x}\right)^{2-\alpha} + \left(1+\sqrt{x}\right)^{2-\alpha} \tag{B151}$$

is concave for $\alpha \in (1,2)$.

Proof. For this function, one can either prove concavity by noting that this statement is equivalent to Proposition 15 after a suitable modification of the Rényi parameter α, or alternatively one can prove it as follows. Due to continuity arguments, it is sufficient to prove it for all $x \in (0,1)$. The second derivative is given by

$$f_2''(x) = \frac{2-\alpha}{4x^{\frac{3}{2}}}\left[\left(1-\alpha\sqrt{x}\right)\left(1-\sqrt{x}\right)^{-\alpha} - \left(1+\alpha\sqrt{x}\right)\left(1+\sqrt{x}\right)^{-\alpha}\right] . \tag{B152}$$

Note that, for any $\alpha \in (1,2)$ and $y \in [0,1)$,

$$(1-y)^\alpha \geq 1 - \alpha y + \frac{\alpha(\alpha+1)}{2}y^2 \tag{B153}$$

$$(1+y)^\alpha \leq 1 + \alpha y + \frac{\alpha(\alpha+1)}{2}y^2 . \tag{B154}$$

The right-hand-side are simply Taylor approximations of the left-hand-side, and the bound holds because the errors arising from such Taylor approximations are proportional to the third derivative at some $\xi \in [0,y]$, see [35, Eq. (1.3.2)], and

$$\frac{d^3}{dy^3}(1-y)^\alpha = \alpha(\alpha-1)(2-\alpha)(1+y)^{\alpha-3} \tag{B155}$$

$$\frac{d^3}{dy^3}(1+y)^\alpha = -\alpha(\alpha-1)(2-\alpha)(1-y)^{\alpha-3} , \tag{B156}$$

which are positive and negative, respectively, for all $y \in [0,1)$. Setting $y = \sqrt{x}$ and using the inequalities from Eqs. (B154)–(B153), then yields

$$f_2''(x) \leq \frac{(2-\alpha)\alpha(\alpha+1)}{8\sqrt{x}(1-x)^\alpha}\left[\left(1-\sqrt{x}\right)^\alpha - \left(1+\sqrt{x}\right)^\alpha\right] \tag{B157}$$

$$\leq 0 \tag{B158}$$

This concludes the proof that $f_2(x)$ is concave for all $\alpha \in (1,2)$. □

7. Tightness of Rate Bounds

It can be shown that all inequalities are tight by considering the following attack which saturates the bound. First, it is easy to verify that measuring that state

$$\sqrt{P_+}\,|\phi^+\rangle_{Q_AQ_B}\,|0\rangle_E + \sqrt{P_-}\,|\phi^-\rangle_{Q_AQ_B}\,|1\rangle_E , \tag{B159}$$

where $P_\pm = \frac{1}{2}(1 \pm g_S)$ and $g_S = \sqrt{\frac{S_\beta^2}{4} - \beta^2}$, in the observables $A_0 = \sigma_z$, $A_1 = \sigma_x$, $B_0 = \frac{\beta\sigma_z + g_S\sigma_x}{\sqrt{\beta^2+g_S^2}}$, $B_1 = \frac{\beta\sigma_z - g_S\sigma_x}{\sqrt{\beta^2+g_S^2}}$ achieves a score of S_β. Moreover, if Alice measures this state in the observable A_0, one finds that the post-measurement state is given by

$$\rho_{AE} = \frac{1}{2}|0\rangle\langle 0|_A \otimes |\psi_0\rangle\langle\psi_0| + \frac{1}{2}|1\rangle\langle 1|_A \otimes |\psi_1\rangle\langle\psi_1| , \tag{B160}$$

where

$$|\psi_0\rangle = \sqrt{P_+}\,|0\rangle + \sqrt{P_-}\,|1\rangle \tag{B161}$$

$$|\psi_1\rangle = \sqrt{P_+}\,|0\rangle - \sqrt{P_-}\,|1\rangle . \tag{B162}$$

In the basis

$$|0'\rangle := |\psi_0\rangle = \sqrt{P_+}\,|0\rangle + \sqrt{P_-}\,|1\rangle \tag{B163}$$
$$|1'\rangle := \sqrt{P_-}\,|0\rangle - \sqrt{P_+}\,|1\rangle\,, \tag{B164}$$

one finds that

$$|\psi_0\rangle = |0'\rangle \tag{B165}$$
$$|\psi_1\rangle = |\,\langle\psi_0|\psi_1\rangle\,|\,|0'\rangle + \sqrt{1-|\,\langle\psi_0|\psi_1\rangle\,|^2}\,|1'\rangle\,, \tag{B166}$$

where

$$|\,\langle\psi_0|\psi_1\rangle\,| = P_+ - P_- \tag{B167}$$
$$= g_S\,. \tag{B168}$$

This post-measurement state thus has the same form as those from Eq. (B15). However, it is precisely on these states that we prove explicit bounds on $\mathbb{H}\,(A|E)$. Whenever we provide an exact expression for $\mathbb{H}\,(A|E)$ for these states, the same proof must also hold for Eq. (B159). The only key difference is that the score of Eq. (B159) is already S_β. We note that although one can technically calculate $\widetilde{H}^{\uparrow}_{\alpha}(A|E)$ for such states, we use a slight variation of this argument to prove tightness in Appendix B 3.

The proofs for deriving tight rate functions require that $\mathbb{Q}$ satisfies the properties mentioned in Lemma 7 and that the bounds we derive for $\mathbb{Q}$ are monotonically decreasing in g_S. These two conditions are necessary to reduce the analysis to states of the form of Eq. (B15). For $\alpha > 1$, concavity is necessary to derive an expression which no longer depends on the index I. Whenever these properties hold and one derives an exact expression for $\mathbb{H}\,(A|E)$ for states of the form of Eq. (B15), then the bounds on the corresponding rate function must be tight. We additionally note that if the bound on $\mathbb{Q}$ were not concave for $\alpha > 1$, but satisfied the properties of Lemma 7 and monotonicity, then one can still achieve tight bounds; however they would depend on the bound's concave envelope. We discuss and explicitly use this property in Appendix C, when incorporating noisy preprocessing into the analysis.

8. Discontinuity Behavior for Edge Cases

In this subsection, we justify the claim that the Theorem 5 bounds can be extended to the right-endpoints of the α ranges by taking the respective limits from below. As briefly mentioned previously, some of the resulting formulas are discontinuous with respect to S_β (for $|\beta| \geq 1$). Specifically, $f_{\widetilde{H}^{\downarrow}_{\infty}}(S_\beta)$ and $f_{\bar{H}^{\downarrow}_{2}}(S_\beta)$ have a discontinuity at $S_\beta = 2\sqrt{1+\beta^2}$, taking the value 1 at that point and 0 elsewhere. We further note that these discontinuities are a genuine property of $\widetilde{H}^{\downarrow}_{\infty}(A|E)_\rho$ and $\bar{H}^{\downarrow}_{2}(A|E)_\rho$, respectively, and not a result of the methods used to obtain these bounds — we shall show this by constructing a family of states saturating these discontinuous bounds.

To prove the claim, we simply note that for all of the Rényi entropies $\mathbb{H}_\alpha$ in Definition 4, for any $\alpha^\star \in (1,\infty]$ we have $\mathbb{H}_{\alpha^\star} = \lim_{\alpha\nearrow\alpha^\star}\mathbb{H}_\alpha = \inf_{\alpha\in(1,\alpha^\star)}\mathbb{H}_\alpha$ because they are monotone decreasing with respect to α. Given this, we can freely interchange the infimum over α with the infimum over quantum strategies in the definition of the rate functions, from which we can conclude the desired claim $f_{\mathbb{H}_{\alpha^\star}}(S_\beta) = \inf_{\alpha\in(1,\alpha^\star)} f_{\mathbb{H}_\alpha}(S_\beta) = \lim_{\alpha\nearrow\alpha^\star} f_{\mathbb{H}_\alpha}(S_\beta)$ (the second equality holds because $f_{\mathbb{H}_\alpha}$ immediately inherits the monotonicity in α from $\mathbb{H}_\alpha$).

In fact, for the cases $f_{\widetilde{H}^{\downarrow}_{\infty}}$ and $f_{\bar{H}^{\downarrow}_{2}}$, we can instead prove this via a more direct analysis of the optimal attack for each S_β, which also shows that the discontinuities are a genuine feature of the bounds. Specifically, to calculate $\widetilde{H}^{\downarrow}_{\alpha}(A|E)_\rho$ for the optimal attack from Eq. (B160), we need to first calculate $\rho_E^{\frac{1-\alpha}{2\alpha}}\,\rho_{AE}\rho_E^{\frac{1-\alpha}{2\alpha}}$. From Eq. (B42), we see that as $\alpha\to\infty$ for $S_\beta < 2\sqrt{1+\beta^2}$,

$$\rho_E^{\frac{1-\alpha}{2\alpha}}\,\rho_{AE}\rho_E^{\frac{1-\alpha}{2\alpha}} \to |0\rangle\langle 0|_A \otimes |w_1\rangle\langle w_1|_E + |1\rangle\langle 1|_A \otimes |w_1'\rangle\langle w_1'|_E\,, \tag{B169}$$

where $\{|w_1\rangle\,, |w_1'\rangle\}$ are normalized vectors. Thus,

$$\lim_{\alpha\to\infty}\widetilde{H}^{\downarrow}_{\alpha}(A|E)_\rho = \lim_{\alpha\to\infty}\frac{1}{1-\alpha}\cdot\lim_{\alpha\to\infty}\log\left[\mathrm{Tr}\left[\left(\sigma_E^{\frac{1-\alpha}{2\alpha}}\,\sigma_{AE}\sigma_E^{\frac{1-\alpha}{2\alpha}}\right)^\alpha\right]\right] \tag{B170}$$
$$= \lim_{\alpha\to\infty}\frac{1}{1-\alpha}\log 2 \tag{B171}$$

$$= 0\,, \tag{B172}$$

where the first lines follows from the fact that both limits are finite and well-defined. Hence, we get that $\widetilde{H}^{\downarrow}_{\infty}(A|E)_{\rho} = 0$ is indeed an achievable bound. Conversely, for the maximal score, one can lower bound $\widetilde{H}^{\downarrow}_{\infty}(A|X=0,IE)_{\rho}$ as follows. Recall that

$$\bar{H}^{\downarrow}_{\alpha}(A|X=0,IE)_{\rho} \geq \frac{1}{1-\alpha}\log\left[\sum_i \Pr\left[I=i\right] 2^{(1-\alpha)\bar{H}^{\downarrow}_{\alpha}(A|I=i,E)_{\sigma^i}}\right] \tag{B173}$$

were the states σ^i_{AE} are defined in Eq. (B15). Each σ^i_{AE} has to achieve the maximum score and therefore must be of the form

$$\sigma^i_{AE} = \frac{1}{2}\left|0\right\rangle\!\left\langle 0\right|_A \otimes \left|0\right\rangle\!\left\langle 0\right|_E + \frac{1}{2}\left|1\right\rangle\!\left\langle 1\right|_A \otimes \left|0\right\rangle\!\left\langle 0\right|_E\,. \tag{B174}$$

Using

$$\left(\sigma^i_E\right)^{\frac{1-\alpha}{2\alpha}} \sigma^i_{AE} \left(\sigma^i_E\right)^{\frac{1-\alpha}{2\alpha}} = \frac{1}{2}\left|0\right\rangle\!\left\langle 0\right|_A \otimes \left|0\right\rangle\!\left\langle 0\right|_E + \frac{1}{2}\left|1\right\rangle\!\left\langle 1\right|_A \otimes \left|0\right\rangle\!\left\langle 0\right|_E\,, \tag{B175}$$

we get that

$$\lim_{\alpha\to\infty} \widetilde{H}^{\downarrow}_{\alpha}(A|E)_{\sigma^i} = \lim_{\alpha\to\infty} \frac{1}{1-\alpha}\log\left[\mathrm{Tr}\left[\left(\left(\sigma^i_E\right)^{\frac{1-\alpha}{2\alpha}} \sigma^i_{AE} \left(\sigma^i_E\right)^{\frac{1-\alpha}{2\alpha}}\right)^{\alpha}\right]\right] \tag{B176}$$

$$= \lim_{\alpha\to\infty} \frac{1}{1-\alpha}\log 2^{1-\alpha} \tag{B177}$$

$$= \lim_{\alpha\to\infty} 1 \tag{B178}$$

$$= 1\,. \tag{B179}$$

It follows from this that $\widetilde{H}^{\downarrow}_{\infty}(A|X=0,IE)_{\rho} = 1$ as well.

To calculate $\bar{H}^{\downarrow}_{\alpha}(A|E)_{\rho}$ for the optimal attack from Eq. (B160), we need to calculate $\rho_E^{\frac{1-\alpha}{2}} \rho^{\alpha}_{AE} \rho_E^{\frac{1-\alpha}{2}}$. From Eq. (B80), we see that as $\alpha \nearrow 2$ for $S_{\beta} < 2\sqrt{1+\beta^2}$,

$$\rho_E^{\frac{1-\alpha}{2}} \rho^{\alpha}_{AE} \rho_E^{\frac{1-\alpha}{2}} \to \frac{1}{4}\left(\left|0\right\rangle\!\left\langle 0\right| + \left|1\right\rangle\!\left\langle 1\right|\right) \otimes \left|v_1\right\rangle\!\left\langle v_1\right| + \frac{1}{4}\left(\left|0\right\rangle\!\left\langle 0\right| + \left|1\right\rangle\!\left\langle 1\right|\right) \otimes \left|v_2\right\rangle\!\left\langle v_2\right|\,. \tag{B180}$$

where $\{\left|v_1\right\rangle, \left|w_2\right\rangle\}$ are normalized vectors. Thus,

$$\lim_{\alpha\nearrow 2} \bar{H}^{\downarrow}_{\alpha}(A|E)_{\rho} = \lim_{\alpha\nearrow 2} \frac{1}{1-\alpha} \cdot \lim_{\alpha\nearrow 2} \log\left[\mathrm{Tr}\left[\rho_E^{\frac{1-\alpha}{2}} \rho^{\alpha}_{AE} \rho_E^{\frac{1-\alpha}{2}}\right]\right] \tag{B181}$$

$$= \lim_{\alpha\nearrow 2} \frac{1}{1-\alpha}\log 1 \tag{B182}$$

$$= 0\,. \tag{B183}$$

Conversely, if we witness a maximal score, then we again simply need to consider the state from Eq. (B174). For this state, we find that

$$\left(\sigma^i_E\right)^{\frac{1-\alpha}{2}} \left(\sigma^i_{AE}\right)^{\alpha} \left(\sigma^i_E\right)^{\frac{1-\alpha}{2}} = \frac{1}{2^{\alpha}}\left|0\right\rangle\!\left\langle 0\right|_A \otimes \left|0\right\rangle\!\left\langle 0\right|_E + \frac{1}{2^{\alpha}}\left|1\right\rangle\!\left\langle 1\right|_A \otimes \left|0\right\rangle\!\left\langle 0\right|_E\,, \tag{B184}$$

and

$$\lim_{\alpha\nearrow 2} \bar{H}^{\downarrow}_{\alpha}(A|E)_{\sigma^i} = \lim_{\alpha\nearrow 2} \frac{1}{1-\alpha}\log\left[\mathrm{Tr}\left[\left(\sigma^i_E\right)^{\frac{1-\alpha}{2}} \left(\sigma^i_{AE}\right)^{\alpha} \left(\sigma^i_E\right)^{\frac{1-\alpha}{2}}\right]\right] \tag{B185}$$

$$= \lim_{\alpha\nearrow 2} \frac{1}{1-\alpha}\log 2^{1-\alpha} \tag{B186}$$

$$= 1 \tag{B187}$$

It again follows that $\bar{H}^{\downarrow}_{\infty}(A|X=0,IE)_{\rho}=1$.

From the calculations presented above, it is evident that the discontinuities of $f_{\widetilde{H}^{\downarrow}_{\infty}}(S_\beta)$ and $f_{\bar{H}^{\downarrow}_{2}}(S_\beta)$ at $S_\beta = 2\sqrt{1+\beta^2}$ arise due to the edge-case behaviors of $\widetilde{H}^{\downarrow}_{\infty}(A|E)_{\rho}$ and $\bar{H}^{\downarrow}_{2}(A|E)_{\rho}$. Hence, our bounds are tight even at these discontinuities. Such behavior can also be found in the Belavkin-Staszewski conditional entropy which is discontinuous on states that are not full-rank [36]. One can then show that this implies it has the same rate function as $\widetilde{H}^{\downarrow}_{\infty}$ and $\bar{H}^{\downarrow}_{2}$. Finally, we note that the functions $f_{\widetilde{H}^{\uparrow}_{\alpha}}(S_\beta)$ and $f_{\bar{H}^{\uparrow}_{\alpha}}(S_\beta)$ have no discontinuities for $\alpha > 1$ (within their respective domains of validity, i.e. $\alpha \leq 2$ for the latter).

9. Alternative Concavity Proofs

Proposition 18. *For all $x \in [0,1]$, the function*

$$f_3(x) = \left(1-\sqrt{x}\right)^{\frac{1}{\alpha}} + \left(1+\sqrt{x}\right)^{\frac{1}{\alpha}} \tag{B188}$$

is concave for $\alpha \in (1,2)$.

Proof. Due to continuity arguments, it is sufficient to consider $x \in (0,1)$. To prove that Eq. (B188) is concave, we need to calculate its second derivative. The second derivative of Eq. (B188) is as follows:

$$\frac{1}{4\alpha x^{\frac{3}{2}}}\left(-\left(1+\sqrt{x}\right)^{\frac{1}{\alpha}-2}\left[1+\left(2-\frac{1}{\alpha}\right)\sqrt{x}\right] + \left(1-\sqrt{x}\right)^{\frac{1}{\alpha}-2}\left[1-\left(2-\frac{1}{\alpha}\right)\sqrt{x}\right]\right). \tag{B189}$$

To show that Eq. (B188) is concave, we need to show that Eq. (B189) is non-positive. The first term in the above expression is negative. The second term is negative when $x > x'$, where x' is the root of $d(x) = 1-(2-1/\alpha)\sqrt{x}$ between $x=0$ and $x=1$. We can see that the x' exists by observing that $d(x)$ switches signs and is a decreasing between $x=0$ and $x=1$. Let us consider the case when $\alpha > 1$ and $x < x'$. For concavity of Eq. (B188),

$$\frac{1}{4\alpha x^{\frac{3}{2}}}\left(-\left(1+\sqrt{x}\right)^{\frac{1}{\alpha}-2}\left[1+\left(2-\frac{1}{\alpha}\right)\sqrt{x}\right] + \left(1-\sqrt{x}\right)^{\frac{1}{\alpha}-2}\left[1-\left(2-\frac{1}{\alpha}\right)\sqrt{x}\right]\right) \leq 0 \tag{B190}$$

$$\implies \left(1+\sqrt{x}\right)^{\frac{1}{\alpha}-2}\left[1+\left(2-\frac{1}{\alpha}\right)\sqrt{x}\right] \geq \left(1-\sqrt{x}\right)^{\frac{1}{\alpha}-2}\left[1-\left(2-\frac{1}{\alpha}\right)\sqrt{x}\right] \tag{B191}$$

$$\implies \left(\frac{1+(2-1/\alpha)\sqrt{x}}{1-(2-1/\alpha)\sqrt{x}}\right)\left(\frac{1-\sqrt{x}}{1+\sqrt{x}}\right)^{2-\frac{1}{\alpha}} \geq 1\,. \tag{B192}$$

Note that as $x \to 0$, the left-hand-side tends to 1. Also when $x \to x'$, the left-hand-side tends to ∞. The above inequality holds for $x < x'$ if, additionally,

$$\left(\frac{1+(2-1/\alpha)\sqrt{x}}{1-(2-1/\alpha)\sqrt{x}}\right)\left(\frac{1-\sqrt{x}}{1+\sqrt{x}}\right)^{2-\frac{1}{\alpha}} \tag{B193}$$

is a monotone function. To show this, let $t \coloneqq \sqrt{x}$, where $t \in [0,1)$. Then

$$\left(\frac{1+(2-1/\alpha)\sqrt{x}}{1-(2-1/\alpha)\sqrt{x}}\right)\left(\frac{1-\sqrt{x}}{1+\sqrt{x}}\right)^{2-\frac{1}{\alpha}} = \left(\frac{1+(2-1/\alpha)\,t}{1-(2-1/\alpha)\,t}\right)\left(\frac{1-t}{1+t}\right)^{2-\frac{1}{\alpha}}, \tag{B194}$$

and the first derivative of Eq. (B193) is as follows:

$$\begin{aligned}&\frac{d}{dt}\left[\left(\frac{1+(2-1/\alpha)\,t}{1-(2-1/\alpha)\,t}\right)\left(\frac{1-t}{1+t}\right)^{2-\frac{1}{\alpha}}\right] \\ &= \frac{[(2-1/\alpha)\,(1-t)^{2-\frac{1}{\alpha}} - (1+(2-1/\alpha)\,t)(2-1/\alpha)(1-t)^{1-\frac{1}{\alpha}}](1-(2-1/\alpha)\,t)(1+t)^{2-\frac{1}{\alpha}})}{(1-(2-1/\alpha)\,t)^2(1+t)^{4-\frac{2}{\alpha}}}\end{aligned} \tag{B195}$$

$$+\frac{[(2-1/\alpha)\,(1+t)^{2-\frac{1}{\alpha}}-(1-(2-1/\alpha)\,t)(2-1/\alpha)(1+t)^{1-\frac{1}{\alpha}}](1+(2-1/\alpha)\,t)(1-t)^{2-\frac{1}{\alpha}})}{(1-(2-1/\alpha)\,t)^2(1+t)^{4-\frac{2}{\alpha}}} \tag{B196}$$

$$=\frac{2\,(2-1/\alpha)\,(1-t^2)^{2-\frac{1}{\alpha}}-2(1-(2-1/\alpha)^2\,t^2)(2-1/\alpha)(1-t^2)^{1-\frac{1}{\alpha}}}{(1-(2-1/\alpha)\,t)^2(1+t)^{4-\frac{2}{\alpha}}} \tag{B197}$$

$$=\left(4-\frac{2}{\alpha}\right)(1-t^2)^{1-\frac{1}{\alpha}}\frac{((2-1/\alpha)^2-1)t^2}{(1-(2-1/\alpha)\,t)^2(1+t)^{4-\frac{2}{\alpha}}}\geq 0\,, \tag{B198}$$

when $t\in[0,\sqrt{x'})$. Hence, Eq. (B193) is monotone when $x\in[0,x')$. Therefore, we have shown that Eq. (B189) is non-positive and thus Eq. (B188) is concave. □

Proposition 19. *For all $x\in[0,1]$, the function*

$$f_2(x)=\left(1-\sqrt{x}\right)^{2-\alpha}+\left(1+\sqrt{x}\right)^{2-\alpha} \tag{B199}$$

is concave for $\alpha\in(1,2)$.

Proof. To prove that Eq. (B199) is concave, we need to calculate its second derivative. Due to continuity arguments, it is sufficient to consider $x\in[0,1)$. The second derivative is given by

$$\frac{2-\alpha}{4x^{\frac{3}{2}}}\left[-\left(1+\alpha\sqrt{x}\right)\left(1+\sqrt{x}\right)^{-\alpha}+\left(1-\alpha\sqrt{x}\right)\left(1-\sqrt{x}\right)^{-\alpha}\right]. \tag{B200}$$

To show that Eq. (B199) is concave, we need to show that Eq. (B200) is non-positive. The first term in the above expression is negative. The second term is negative when $x>x'$, where x' is the root of $h(x)=1-\alpha\sqrt{x}$ between $x=0$ and $x=1$. We can see that the x' exists by observing that $h(x)$ switches signs and is a decreasing between $x=0$ and $x=1$. Let us consider the case when, $\alpha>1$ and $x<x'$

$$\frac{2-\alpha}{4x^{\frac{3}{2}}}\left[-\left(1+\alpha\sqrt{x}\right)\left(1+\sqrt{x}\right)^{-\alpha}+\left(1-\alpha\sqrt{x}\right)\left(1-\sqrt{x}\right)^{-\alpha}\right]\leq 0 \tag{B201}$$

$$\implies \left(1+\alpha\sqrt{x}\right)\left(1+\sqrt{x}\right)^{-\alpha}\geq\left(1-\alpha\sqrt{x}\right)\left(1-\sqrt{x}\right)^{-\alpha} \tag{B202}$$

$$\implies \frac{(1+\alpha\sqrt{x})\,(1-\sqrt{x})^{\alpha}}{(1-\alpha\sqrt{x})\,(1+\sqrt{x})^{\alpha}}\geq 1\,. \tag{B203}$$

Note that as $x\to 0$, the left-hand-side tends to 1. Also when $x\to x'$, the left-hand-side tends to ∞. The above inequality holds for $x<x'$ if, additionally,

$$\frac{(1+\alpha\sqrt{x})\,(1-\sqrt{x})^{\alpha}}{(1-\alpha\sqrt{x})\,(1+\sqrt{x})^{\alpha}} \tag{B204}$$

is a monotone function. To show this, let $t:=\sqrt{x}$, where $t\in[0,1)$. Then

$$\frac{(1+\alpha\sqrt{x})\,(1-\sqrt{x})^{\alpha}}{(1-\alpha\sqrt{x})\,(1+\sqrt{x})^{\alpha}}=\frac{(1+\alpha t)\,(1-t)^{\alpha}}{(1-\alpha t)\,(1+t)^{\alpha}}\,, \tag{B205}$$

and the first derivative of Eq. (B204) is as follows:

$$\frac{d}{dt}\left[\frac{(1+\alpha t)\,(1-t)^{\alpha}}{(1-\alpha t)\,(1+t)^{\alpha}}\right] \tag{B206}$$

$$=\frac{[\alpha(1-t)^{\alpha}-\alpha(1+\alpha t)(1-t)^{\alpha-1}]((1-\alpha t)(1+t)^{\alpha})}{(1-\alpha t)^2(1+t)^{2\alpha}}+\frac{[\alpha(1+t)^{\alpha}-\alpha(1-\alpha t)(1+t)^{\alpha-1}]((1+\alpha t)(1-t)^{\alpha})}{(1-\alpha t)^2(1+t)^{2\alpha}}$$

$$=\frac{2\alpha(1-t^2)^{\alpha}+2\alpha(1-t^2)^{\alpha-1}(\alpha^2t^2-1)}{(1-\alpha t)^2(1+t)^{2\alpha}}=2\alpha(1-t^2)^{\alpha-1}\frac{(\alpha^2-1)t^2}{(1-\alpha t)^2(1+t)^{2\alpha}}\geq 0\,. \tag{B207}$$

when $t\in[0,\sqrt{x'})$. Hence, Eq. (B204) is monotone when $x\in[0,x')$. Therefore, we have shown that Eq. (B200) is non-positive and thus Eq. (B199) is concave. □

Appendix C: Incorporating Noisy Preprocessing

Noisy preprocessing is a technique to boost key rates by injecting randomness into Alice's raw string during generation rounds [37]. The intuition is that in certain settings this can decrease Eve's information about Alice's raw string more than it decreases Bob's information, overall leading to an increase in the key rate. It has already been shown for the CHSH-based DIQKD protocol that this can lower the minimal detection efficiencies [16]. To implement noisy preprocessing, Alice simply flips her output bit with some probability q in generation rounds, using private local randomness. After incorporating such a step, the resulting classical-quantum state ρ_{IAE} is no longer given by Eq. (B14); rather it is of the form

$$\rho_{IAE} = \sum_i \Pr[I=i] \, |i\rangle\langle i|_I \otimes \sum_a \left[|a\rangle\langle a|_A \otimes \left((1-q)\rho_E^{ia} + q\rho_E^{i(a\oplus 1)} \right) \right] . \tag{C1}$$

Following the same arguments as those used to prove Lemma 7, to lower bound $\mathbb{H}(A|I,E)_\rho$ it suffices to consider states of the form

$$\sigma_{IAE} = \sum_i \Pr[I=i] \, |i\rangle\langle i|_I \otimes \sigma_{AE}^i \tag{C2}$$

$$\sigma_{AE}^i = \frac{1}{2} |0\rangle\langle 0| \otimes \left((1-q) |\psi_=\rangle\langle\psi_=| + q |\psi_{\neq}\rangle\langle\psi_{\neq}| \right) + \frac{1}{2} |1\rangle\langle 1| \otimes \left(q |\psi_=\rangle\langle\psi_=| + (1-q) |\psi_{\neq}\rangle\langle\psi_{\neq}| \right) , \tag{C3}$$

such that $|\langle\psi_=|\psi_{\neq}\rangle| \geq g_S$. In the following theorem, we will use the notation that for $g_S = \sqrt{\frac{S_\beta^2}{4} - \beta^2}$, we define the following quantities:

$$N_1 = \sqrt{\left(g_S - \sqrt{g_S^2 + (1-2q)^2 (1-g_S^2)} \right)^2 + (1-2q)^2 (1-g_S^2)} \tag{C4}$$

$$N_2 = \sqrt{\left(g_S + \sqrt{g_S^2 + (1-2q)^2 (1-g_S^2)} \right)^2 + (1-2q)^2 (1-g_S^2)} . \tag{C5}$$

Theorem 20. *Let $|\beta| \geq 1$, $S_\beta \in \left[2|\beta|, 2\sqrt{1+\beta^2}\right]$, and $g_S = \sqrt{\frac{S_\beta^2}{4} - \beta^2}$. Then, for any $q \in [0,1]$, we have (writing $\hat{h}$ to denote the concave envelope of an arbitrary function $h : \left[2|\beta|, 2\sqrt{1+\beta^2}\right] \to \mathbb{R}$):*

$$f_{\widetilde{H}_\alpha^\downarrow}(S_\beta) = 1 + \frac{1}{1-\alpha} \log \left[\hat{h}_{\widetilde{H}_\alpha^\downarrow}(S_\beta) \right] \tag{C6}$$

for all $\alpha \in (1,\infty)$, where

$$h_{\widetilde{H}_\alpha^\downarrow}(S_\beta) = \frac{1}{2^{\alpha+1}} \left[\left((1-g_S)^{\frac{1}{\alpha}} + (1+g_S)^{\frac{1}{\alpha}} + \sqrt{\left((1-g_S)^{\frac{1}{\alpha}} + (1+g_S)^{\frac{1}{\alpha}} \right)^2 - 16 (1-g_S^2)^{\frac{1}{\alpha}} (q-q^2)} \right)^\alpha \right.$$
$$\left. + \left((1-g_S)^{\frac{1}{\alpha}} + (1+g_S)^{\frac{1}{\alpha}} - \sqrt{\left((1-g_S)^{\frac{1}{\alpha}} + (1+g_S)^{\frac{1}{\alpha}} \right)^2 - 16 (1-g_S^2)^{\frac{1}{\alpha}} (q-q^2)} \right)^\alpha \right] . \tag{C7}$$

Similarly,

$$f_{\bar{H}_\alpha^\downarrow}(S_\beta) = 1 + \frac{1}{1-\alpha} \log \left[\hat{h}_{\bar{H}_\alpha^\downarrow}(S_\beta) \right] \tag{C8}$$

$$f_{\bar{H}_\alpha^\uparrow}(S_\beta) = 1 + \frac{\alpha}{1-\alpha} \log \left[\hat{h}_{\bar{H}_\alpha^\uparrow}(S_\beta) \right] \tag{C9}$$

for all $\alpha \in (1,2]$, where

$$h_{\bar{H}^{\downarrow}_{\alpha}}(S_\beta) = \sum_{j=1}^{2} \frac{1}{2} \left(1 + (-1)^{j+1} \sqrt{g_S^2 + (2q-1)^2(1-g_S^2)}\right)^{\alpha} (1-g_S)^{1-\alpha} \left(\frac{g_S + (-1)^j \sqrt{g_S^2 + (2q-1)^2(1-g_S^2)}}{N_j}\right)^2 + \sum_{j=1}^{2} \frac{1}{2} \left(1 + (-1)^{j+1} \sqrt{g_S^2 + (2q-1)^2(1-g_S^2)}\right)^{\alpha} (1+g_S)^{1-\alpha} \left(\frac{(2q-1)\sqrt{1-g_S^2}}{N_j}\right)^2 \tag{C10}$$

and

$$h_{\bar{H}^{\uparrow}_{\alpha}}(S_\beta) = \left[\sum_{j=1}^{2} \left(\frac{1 + (-1)^{j+1} \sqrt{g_S^2 + (2q-1)^2(1-g_S^2)}}{2}\right)^{\alpha} \left(1 - \left(\frac{(2q-1)\sqrt{1-g_S^2}}{N_j}\right)^2\right)\right]^{\frac{1}{\alpha}} + \left[\sum_{j=1}^{2} \left(\frac{1 + (-1)^{j+1} \sqrt{g_S^2 + (2q-1)^2(1-g_S^2)}}{2}\right)^{\alpha} \left(\frac{(2q-1)\sqrt{1-g_S^2}}{N_j}\right)^2\right]^{\frac{1}{\alpha}} . \tag{C11}$$

A proof of Theorem 20 can be found in Appendix C 1. Note that the final bounds $f_{\mathbb{H}}$ are presented in terms of the concave envelopes $\hat{h}_{\mathbb{H}}$ of the functions $h_{\mathbb{H}}$ — in order to avoid having to take this concave envelope, it would suffice to show that all $h_{\mathbb{H}}$ are concave. We note that, up to numerical precision, these functions indeed appear to be concave. We leave a rigorous proof of concavity for future work, and highlight that a consequence of this concavity would be that the equality $f_{\widetilde{H}^{\downarrow}_{\alpha}} = f_{\bar{H}^{\uparrow}_{\alpha}}$ does not generally hold for every $q \in [0,1]$ and $\alpha \in (1,2]$.

1. Proof of Theorem 20

Proof. For any σ_{IAE} of the form given by Eq. (C2), we first consider Alice's and Eve's bipartite state for some $I = i$. For any σ_{AE} as in Eq. (C3), Eve's reduced density matrix is given by

$$\sigma_E = \frac{1-g_x}{2} |v_1\rangle\langle v_1| + \frac{1+g_x}{2} |v_2\rangle\langle v_2| \ , \tag{C12}$$

where $g_x = |\langle \psi_= | \psi_{\neq} \rangle|$ and[3]

$$|v_1\rangle = -\sqrt{\frac{1-g_x}{2}} |0\rangle + \sqrt{\frac{1+g_x}{2}} |1\rangle \tag{C13}$$

$$|v_2\rangle = \sqrt{\frac{1+g_x}{2}} |0\rangle + \sqrt{\frac{1-g_x}{2}} |1\rangle \ . \tag{C14}$$

We first derive an exact expression for $\widetilde{H}^{\downarrow}_{\alpha}(A|E)_\sigma$. Plugging the above expressions directly into $\sigma_E^{\frac{1-\alpha}{2\alpha}} \sigma_{AE} \sigma_E^{\frac{1-\alpha}{2\alpha}}$ gives us that

$$\sigma_E^{\frac{1-\alpha}{2\alpha}} \sigma_{AE} \sigma_E^{\frac{1-\alpha}{2\alpha}} \tag{C15}$$

$$= \frac{1}{2} |0\rangle\langle 0|_A \otimes \left[\left(\frac{1-g_x}{2}\right)^{\frac{1}{\alpha}} |v_1\rangle\langle v_1|_E + \left(\frac{1-g_x^2}{4}\right)^{\frac{1}{2\alpha}} (q - \bar{q}) \left(|v_1\rangle\langle v_2|_E + |v_2\rangle\langle v_1|_E\right) + \left(\frac{1+g_x}{2}\right)^{\frac{1}{\alpha}} |v_2\rangle\langle v_2|_E\right] + \frac{1}{2} |1\rangle\langle 1|_A \otimes \left[\left(\frac{1-g_x}{2}\right)^{\frac{1}{\alpha}} |v_1\rangle\langle v_1|_E + \left(\frac{1-g_x^2}{4}\right)^{\frac{1}{2\alpha}} (\bar{q} - q) \left(|v_1\rangle\langle v_2|_E + |v_2\rangle\langle v_1|_E\right) + \left(\frac{1+g_x}{2}\right)^{\frac{1}{\alpha}} |v_2\rangle\langle v_2|_E\right] , \tag{C16}$$

[3] Note that we work with g_x, rather than g_S, as we do not explicitly prove the fact that $h_{\mathbb{H}}(S_\beta)$ is monotonically decreasing. Rather, at an appropriate point, we use the fact that its concave envelope is monotonically decreasing.

where $\bar{q} = 1 - q$. This expression is diagonalizable, and it is easily verifiable that

$$\sigma_E^{\frac{1-\alpha}{2\alpha}} \sigma_{AE} \sigma_E^{\frac{1-\alpha}{2\alpha}} = 2^{-1-\frac{1}{\alpha}} \left(|0\rangle\langle 0|_A \otimes \left[\lambda_1 |w_1\rangle\langle w_1|_E + \lambda_2 |w_2\rangle\langle w_2|_E \right] + |1\rangle\langle 1|_A \otimes \left[\lambda_1' |w_1'\rangle\langle w_1'|_E + \lambda_2' |w_2'\rangle\langle w_2'|_E \right] \right) , \tag{C17}$$

where $\{|w_1\rangle, |w_2\rangle\}$ and $\{|w_1'\rangle, |w_2'\rangle\}$ are pairwise orthonormal vectors, and

$$\lambda_{1,2} = \frac{(1-g_x)^{\frac{1}{\alpha}} + (1+g_x)^{\frac{1}{\alpha}}}{2} \pm \sqrt{\left(\frac{(1-g_x)^{\frac{1}{\alpha}} + (1+g_x)^{\frac{1}{\alpha}}}{2} \right)^2 - (1-g_x^2)^{\frac{1}{\alpha}} \left(4q - 4q^2\right)} \tag{C18}$$

$$\lambda_{1,2}' = \frac{(1-g_x)^{\frac{1}{\alpha}} + (1+g_x)^{\frac{1}{\alpha}}}{2} \pm \sqrt{\left(\frac{(1-g_x)^{\frac{1}{\alpha}} + (1+g_x)^{\frac{1}{\alpha}}}{2} \right)^2 - (1-g_x^2)^{\frac{1}{\alpha}} \left(4\bar{q} - 4\bar{q}^2\right)} . \tag{C19}$$

Note that because $q - q^2 = \bar{q} - \bar{q}^2$, it must also hold that $\lambda_{1,2} = \lambda_{1,2}'$. It then directly follows that

$$\widetilde{H}_\alpha^\downarrow(A|E)_\sigma = \frac{1}{1-\alpha} \log \left[\mathrm{Tr} \left[\left(\sigma_E^{\frac{1-\alpha}{2\alpha}} \sigma_{AE} \sigma_E^{\frac{1-\alpha}{2\alpha}} \right)^\alpha \right] \right] \tag{C20}$$

$$= \frac{1}{1-\alpha} \log \frac{\left[\lambda_1^\alpha + \lambda_2^\alpha + \lambda_1'^\alpha + \lambda_2'^\alpha\right]}{2^{\alpha+1}} \tag{C21}$$

$$= \frac{1}{1-\alpha} \log \frac{\left[2\lambda_1^\alpha + 2\lambda_2^\alpha\right]}{2^{\alpha+1}} \tag{C22}$$

$$= \frac{1}{1-\alpha} \log \frac{\left[\lambda_1^\alpha + \lambda_2^\alpha\right]}{2^{\alpha}} . \tag{C23}$$

This thus yields

$$\widetilde{H}_\alpha^\downarrow(A|E)_\sigma = 1 + \frac{1}{1-\alpha} \log \left[\frac{1}{2^{\alpha+1}} \left[\left((1-g_x)^{\frac{1}{\alpha}} + (1+g_x)^{\frac{1}{\alpha}} + \sqrt{\left((1-g_x)^{\frac{1}{\alpha}} + (1+g_x)^{\frac{1}{\alpha}} \right)^2 - 16 \left(1-g_x^2\right)^{\frac{1}{\alpha}} \left(q - q^2\right)} \right)^\alpha \right.$$
$$\left. + \left((1-g_x)^{\frac{1}{\alpha}} + (1+g_x)^{\frac{1}{\alpha}} - \sqrt{\left((1-g_x)^{\frac{1}{\alpha}} + (1+g_x)^{\frac{1}{\alpha}} \right)^2 - 16 \left(1-g_x^2\right)^{\frac{1}{\alpha}} \left(q - q^2\right)} \right)^\alpha \right] \right] . \tag{C24}$$

Up to the use of the concave envelope and the difference between g_x and g_S, this is the expression from Eq. (C6). Next, we consider $\bar{H}_\alpha^\downarrow(A|E)_\sigma$. For any σ_{AE} as in Eq. (C3), it holds that

$$\sigma_{AE}^\alpha = |0\rangle\langle 0| \otimes \left[\left(\frac{1 + \sqrt{g_x^2 + (q-\bar{q})^2 (1-g_x^2)}}{4} \right)^\alpha |u_1\rangle\langle u_1| + \left(\frac{1 - \sqrt{g_x^2 + (q-\bar{q})^2 (1-g_x^2)}}{4} \right)^\alpha |u_2\rangle\langle u_2| \right]$$
$$+ |1\rangle\langle 1| \otimes \left[\left(\frac{1 + \sqrt{g_x^2 + (q-\bar{q})^2 (1-g_x^2)}}{4} \right)^\alpha |u_1'\rangle\langle u_1'| + \left(\frac{1 - \sqrt{g_x^2 + (q-\bar{q})^2 (1-g_x^2)}}{4} \right)^\alpha |u_2'\rangle\langle u_2'| \right] , \tag{C25}$$

where

$$|u_1\rangle = \frac{g_x - \sqrt{g_x^2 + (q-\bar{q})^2 (1-g_x^2)}}{N_1} |v_1\rangle - \frac{(q-\bar{q})\sqrt{1-g_x^2}}{N_1} |v_2\rangle \tag{C26}$$

$$|u_2\rangle = \frac{g_x + \sqrt{g_x^2 + (q-\bar{q})^2 (1-g_x^2)}}{N_2} |v_1\rangle - \frac{(q-\bar{q})\sqrt{1-g_x^2}}{N_2} |v_2\rangle \tag{C27}$$

$$|u_1'\rangle = \frac{g_x - \sqrt{g_x^2 + (q-\bar{q})^2 (1-g_x^2)}}{N_1} |v_1\rangle + \frac{(q-\bar{q})\sqrt{1-g_x^2}}{N_1} |v_2\rangle \tag{C28}$$

$$|u_2'\rangle = \frac{g_x + \sqrt{g_x^2 + (q-\bar{q})^2 (1-g_x^2)}}{N_2} |v_1\rangle + \frac{(q-\bar{q})\sqrt{1-g_x^2}}{N_2} |v_2\rangle . \tag{C29}$$

It then holds that

$$\bar{H}_\alpha^\downarrow(A|E)_\sigma = \frac{1}{1-\alpha}\log\left[\mathrm{Tr}\left[\sigma_{AE}^\alpha \sigma_E^{1-\alpha}\right]\right] \tag{C30}$$

$$= \frac{1}{1-\alpha}\log\left[\sum_{j=1}^{2} 2\left(\frac{1+(-1)^{j+1}\sqrt{g_x^2+(q-\bar{q})^2(1-g_x^2)}}{4}\right)^\alpha \left(\frac{1-g_x}{2}\right)^{1-\alpha}\left(\frac{g_x+(-1)^j\sqrt{g_x^2+(q-\bar{q})^2(1-g_x^2)}}{N_j}\right)^2\right.$$

$$\left.+\sum_{j=1}^{2} 2\left(\frac{1+(-1)^{j+1}\sqrt{g_x^2+(q-\bar{q})^2(1-g_x^2)}}{4}\right)^\alpha \left(\frac{1+g_x}{2}\right)^{1-\alpha}\left(\frac{(q-\bar{q})\sqrt{1-g_x^2}}{N_j}\right)^2\right]. \tag{C31}$$

Slightly rewriting this equality, we then attain that this is equal to

$$1+\frac{1}{1-\alpha}\log\left[\sum_{j=1}^{2}\frac{1}{2}\left(1+(-1)^{j+1}\sqrt{g_x^2+(2q-1)^2(1-g_x^2)}\right)^\alpha (1-g_x)^{1-\alpha}\left(\frac{g_x+(-1)^j\sqrt{g_x^2+(2q-1)^2(1-g_x^2)}}{N_j}\right)^2\right.$$

$$\left.+\sum_{j=1}^{2}\frac{1}{2}\left(1+(-1)^{j+1}\sqrt{g_x^2+(2q-1)^2(1-g_x^2)}\right)^\alpha (1+g_x)^{1-\alpha}\left(\frac{(2q-1)\sqrt{1-g_x^2}}{N_j}\right)^2\right]. \tag{C32}$$

To calculate $\bar{H}_\alpha^\uparrow(A|E)_\sigma$, we again use [34, Lemma 5.1], i.e.

$$\bar{H}_\alpha^\uparrow(A|E)_\sigma = \frac{\alpha}{1-\alpha}\log\left[\mathrm{Tr}\left[\mathrm{Tr}_A\left(\sigma_{AE}^\alpha\right)^{\frac{1}{\alpha}}\right]\right]. \tag{C33}$$

A consequence of the above calculations is that

$$\mathrm{Tr}_A[\sigma_{AE}^\alpha] = \sum_{j=1}^{2}\left(\frac{1+(-1)^{j+1}\sqrt{g_x^2+(q-\bar{q})^2(1-g_x^2)}}{4}\right)^\alpha \left(|u_j\rangle\langle u_j| + |u_j'\rangle\langle u_j'|\right) \tag{C34}$$

$$= \sum_{j=1}^{2} 2\left(\frac{1+(-1)^{j+1}\sqrt{g_x^2+(q-\bar{q})^2(1-g_x^2)}}{4}\right)^\alpha$$

$$\cdot\left[\left(\frac{g_x+(-1)^j\sqrt{g_x^2+(q-\bar{q})^2(1-g_x^2)}}{N_j}\right)^2 |v_1\rangle\langle v_1| + \left(\frac{(q-\bar{q})\sqrt{1-g_x^2}}{N_j}\right)^2 |v_2\rangle\langle v_2|\right] \tag{C35}$$

$$= \sum_{j=1}^{2} 2\left(\frac{1+(-1)^{j+1}\sqrt{g_x^2+(q-\bar{q})^2(1-g_x^2)}}{4}\right)^\alpha$$

$$\cdot\left[\left(1-\left(\frac{(q-\bar{q})\sqrt{1-g_x^2}}{N_j}\right)^2\right)|v_1\rangle\langle v_1| + \left(\frac{(q-\bar{q})\sqrt{1-g_x^2}}{N_j}\right)^2 |v_2\rangle\langle v_2|\right]. \tag{C36}$$

The eigenvalues of $\mathrm{Tr}_A[\sigma_{AE}^\alpha]$ are given by

$$\mu_1 = 2\sum_{j=1}^{2}\left(\frac{1+(-1)^{j+1}\sqrt{g_x^2+(q-\bar{q})^2(1-g_x^2)}}{4}\right)^\alpha \left(1-\left(\frac{(q-\bar{q})\sqrt{1-g_x^2}}{N_j}\right)^2\right) \tag{C37}$$

$$\mu_2 = 2\sum_{j=1}^{2}\left(\frac{1+(-1)^{j+1}\sqrt{g_x^2+(q-\bar{q})^2(1-g_x^2)}}{4}\right)^\alpha \left(\frac{(q-\bar{q})\sqrt{1-g_x^2}}{N_j}\right)^2, \tag{C38}$$

and therefore

$$\bar{H}_\alpha^\uparrow(A|E)_\sigma = 1 + \frac{\alpha}{1-\alpha}\log\left[\left[\sum_{j=1}^{2}\left(\frac{1+(-1)^{j+1}\sqrt{g_x^2+(2q-1)^2(1-g_x^2)}}{2}\right)^\alpha\left(1-\left(\frac{(2q-1)\sqrt{1-g_x^2}}{N_j}\right)^2\right)\right]^{\frac{1}{\alpha}} + \left[\sum_{j=1}^{2}\left(\frac{1+(-1)^{j+1}\sqrt{g_x^2+(2q-1)^2(1-g_x^2)}}{2}\right)^\alpha\left(\frac{(2q-1)\sqrt{1-g_x^2}}{N_j}\right)^2\right]^{\frac{1}{\alpha}}\right]. \tag{C39}$$

Again, up to the use of the concave envelope and the difference between g_x and g_S, these expressions for the Petz-Rényi entropies are equal to Eqs. (C8) – (C9). Let us now reintroduce the index $I = i$. We prove in Appendix C 2 that all three functions $\hat{h}_{\mathbb{H}}$ are monotonically decreasing. By upper bounding $h_{\mathbb{H}}$ with $\hat{h}_{\mathbb{H}}$ and then using the fact that, due to the monotonicity of $\hat{h}_{\mathbb{H}}$, one can provide a subsequent upper bound by replacing g_x with g_S, the following must be true. For any ρ_{IAE} of the form given by Eq. (C1), it must hold that

$$\widetilde{H}_\alpha^\downarrow(A|X=0,IE)_\rho \geq \frac{1}{1-\alpha}\log\left[\sum_i \Pr[I=i]\, 2^{(1-\alpha)\widetilde{H}_\alpha^\downarrow(A|I=i,E)_{\sigma^i}}\right] \tag{C40}$$

$$\geq 1 + \frac{1}{1-\alpha}\log\left[\sum_i \Pr[I=i]\,\hat{h}_{\widetilde{H}_\alpha^\downarrow}(S_\beta^i)\right] \tag{C41}$$

$$\geq 1 + \frac{1}{1-\alpha}\log\left[\hat{h}_{\widetilde{H}_\alpha^\downarrow}(S_\beta)\right], \tag{C42}$$

where we use [34, Prop. 5.1] in the first inequality. To prove that this bound is tight, first note that, by construction, for any $S_\beta \in \left[2|\beta|, 2\sqrt{1+\beta^2}\right]$, there must exist a set of $\{S_\beta^i\}_i$ and a distribution defined by the elements $\{p_i\}_i$ such that

$$\sum_i p_i S_\beta^i = S_\beta \tag{C43}$$

$$\sum_i p_i h_{\widetilde{H}_\alpha^\downarrow}\left(S_\beta^i\right) = \hat{h}_{\widetilde{H}_\alpha^\downarrow}\left(S_\beta\right). \tag{C44}$$

For any such S_β^i, the explicit attack from Appendix B 7 produces a post-measurement state, σ_{AE}^i, of the form as in Eq. (C3). However, for such states, we have explicitly calculated the Rényi entropy, and

$$\widetilde{H}_\alpha^\downarrow(A|X=0,E)_{\sigma^i} = 1 + \frac{1}{1-\alpha}\log\left(h_{\widetilde{H}_\alpha^\downarrow}\left(S_\beta^i\right)\right). \tag{C45}$$

If, with probability p_i, Eve constructs the attack from Appendix B 7 that achieves a score of S_β^i and stores the index $I = i$ on a classical register, then she can generate the post-measurement state $\sigma_{IAE} = \sum_i p_i\,|i\rangle\langle i|_I \otimes \sigma_{AE}^i$. For this state,

$$\widetilde{H}_\alpha^\downarrow(A|X=0,IE)_\sigma = 1 + \frac{1}{1-\alpha}\log\left(\sum_i p_i h_{\widetilde{H}_\alpha^\downarrow}(S_\beta^i)\right) \tag{C46}$$

$$= 1 + \frac{1}{1-\alpha}\log\left(\hat{h}_{\widetilde{H}_\alpha^\downarrow}\left(S_\beta\right)\right), \tag{C47}$$

due to [34, Prop. 5.1] (or else see the classical linearity property in Lemma 7). This attack thus saturates our bounds. Using the same arguments, it must also hold that

$$\bar{H}_\alpha^\downarrow(A|X=0,IE)_\rho = 1 + \frac{1}{1-\alpha}\log\left[\hat{h}_{\bar{H}_\alpha^\downarrow}(S_\beta)\right] \tag{C48}$$

$$\bar{H}_\alpha^\uparrow(A|X=0,IE)_\rho = 1 + \frac{\alpha}{1-\alpha}\log\left[\hat{h}_{\bar{H}_\alpha^\uparrow}(S_\beta)\right]. \tag{C49}$$

□

2. Monotonicity Properties

In this section, we show that all three functions $\hat{h}_{\mathbb{H}}(S_\beta)$ described in Theorem 20 are monotonically decreasing in the score, S_β (for α, β in the appropriate ranges).

Proposition 21. *For all $\alpha \in (1,\infty)$ and $q \in [0,1]$, the function $\hat{h}_{\widetilde{H}^{\downarrow}_{\alpha}}(S_\beta)$ is monotonically decreasing in the interval $S_\beta \in \left[2|\beta|, 2\sqrt{1+\beta^2}\right]$. Similarly, for all $\alpha \in (1,2)$ and $q \in [0,1]$, the functions $\hat{h}_{\bar{H}^{\downarrow}_{\alpha}}(S_\beta)$ and $\hat{h}_{\bar{H}^{\uparrow}_{\alpha}}(S_\beta)$ are monotonically decreasing in the interval $S_\beta \in \left[2|\beta|, 2\sqrt{1+\beta^2}\right]$.*

Proof. Intuitively, the monotonicity will be a consequence of the fact that $\hat{h}_{\mathbb{H}}(S_\beta)$ is concave and non-increasing at the point $S_\beta = 2|\beta|$. More concretely, for any $h_{\mathbb{H}}(S_\beta)$, we are evaluating (up to a multiplicative constant) the corresponding quantity $\mathbb{Q}(A|E)$ for a classical-quantum state of the form in Eq. (C3) such that $|\langle\psi_{=}|\psi_{\neq}\rangle| = g_S$. For these classical-quantum states, we shall show that

$$h_{\mathbb{H}}(S_\beta) \leq h_{\mathbb{H}}(2|\beta|)\,. \tag{C50}$$

In other words, at $g_S = 0$ (i.e. $S_\beta = 2|\beta|$), Eve only needs to guess the noisy preprocessing bit; for all other $g_S > 0$, the conditional entropy, $\mathbb{H}(A|E)$, cannot be lower than that value. To prove this, let A' contain Alice's output bit before applying noisy preprocessing, let Q contain the noisy preprocessing bit that is added to A', and let A contain the bit Alice stores after incorporating noisy preprocessing. It must hold that

$$\mathbb{H}(Q) = \mathbb{H}(Q|A'E) = \mathbb{H}(AQ|A'E) = \mathbb{H}(A|A'E) \leq \mathbb{H}(A|E)\,. \tag{C51}$$

The first inequality holds, as Q is sampled independently and not related to the registers A' and E. The second equality holds, because A can be constructed deterministically from A' and Q, in the sense of [12, Lemma B.7]. The third equality holds for the same reason, i.e. Q can be constructed deterministically from A' and A. The last inequality holds due to data processing [34, Corollary 5.1].

For any $\mathbb{H}$, the constant function $l(S_\beta) = h_{\mathbb{H}}(2|\beta|)$ is an upper bound on $h_{\mathbb{H}}(S_\beta)$ that is (trivially) concave, and tight at the endpoint $S_\beta = 2|\beta|$; this implies that the concave envelope of $h_{\mathbb{H}}(S_\beta)$ also satisfies

$$\hat{h}_{\mathbb{H}}(2|\beta|) = h_{\mathbb{H}}(2|\beta|)\,. \tag{C52}$$

We are now ready to prove monotonicity. For any $a, b \in \left[2|\beta|, 2\sqrt{1+\beta^2}\right]$ such that $a \leq b$, it must hold due to concavity that

$$\hat{h}_{\mathbb{H}}(a) \geq \frac{b-a}{b-2|\beta|}\hat{h}_{\mathbb{H}}(2|\beta|) + \frac{a-2|\beta|}{b-2|\beta|}\hat{h}_{\mathbb{H}}(b) = \frac{b-a}{b-2|\beta|}l(b) + \frac{a-2|\beta|}{b-2|\beta|}\hat{h}_{\mathbb{H}}(b)\,. \tag{C53}$$

Moreover, as $l(S_\beta)$ is concave, it must also be an upper bound on $\hat{h}_{\mathbb{H}}(S_\beta)$. It then follows that

$$\hat{h}_{\mathbb{H}}(a) \geq \frac{b-a}{b-2|\beta|}\hat{h}_{\mathbb{H}}(b) + \frac{a-2|\beta|}{b-2|\beta|}\hat{h}_{\mathbb{H}}(b) = \hat{h}_{\mathbb{H}}(b)\,. \tag{C54}$$

This proves that all three functions are monotonically decreasing. □

Appendix D: Finite-size key rates

We first elaborate on the classical postprocessing used in the last steps. For a DIQKD protocol, this consists of the following procedures:

1. Error correction and error verification: In error correction, Alice sends a string $\mathbf{L}_{\mathrm{EC}}$ (of some fixed length ℓ_{EC}) to Bob, who uses it to produce a guess for Alice's string A_1^n. Then in error verification, Alice draws some choice of hash function H_{EV} from a δ-almost-universal hash family [38] (with fixed output length ℓ_{EV}), then applies it to A_1^n and sends the resulting value $\mathbf{L}_{\mathrm{EV}}$ to Bob, along with the choice of hash function H_{EV}. Bob then computes the corresponding hash of his guess and aborts if it does not match.

2. Privacy amplification: Alice applies a privacy amplification procedure to A_1^n to produce a final key of length ℓ_{key}, and Bob does the same to his guess for A_1^n.

When designing a protocol for the finite-size regime, there are two critical "overall" parameters that should be considered. We briefly outline them here, deferring the details to e.g. [27, 39].

1. The **completeness** parameter ϵ_{com}: this is an upper bound on the probability that the honest behavior aborts. Since this protocol might abort during either the acceptance test or the error verification step, it is convenient to construct upper bounds $\epsilon_{\text{com}}^{\text{AT}}$ and $\epsilon_{\text{com}}^{\text{EV}}$ on the abort probabilities in each of those two steps respectively, after which one can validly take $\epsilon_{\text{com}} = \epsilon_{\text{com}}^{\text{AT}} + \epsilon_{\text{com}}^{\text{EV}}$ due to the union bound. Here we design the protocol such that $\epsilon_{\text{com}}^{\text{AT}} = 10^{-3}$ and $\epsilon_{\text{com}}^{\text{EV}} \leq 10^{-3}$ in an essentially similar fashion to [27][4] (we elaborate on this in Sec. D 1 below). We emphasize that this parameter does not affect the security properties of the protocol in any way, which are instead quantified by the next parameter.

2. The **soundness** parameter ϵ_{sound}: informally, this quantifies the "security" of the final key; refer to [27, 39] for a rigorous definition. As discussed in those works, to analyze this parameter it suffices to separately consider a **correctness** parameter ϵ_{corr} and a **secrecy** parameter ϵ_{secret}, then set $\epsilon_{\text{sound}} = \epsilon_{\text{corr}} + \epsilon_{\text{secret}}$. Basically, ϵ_{corr} is an upper bound on the probability that the final keys do not match and the protocol accepts, while roughly speaking ϵ_{secret} quantifies how well Alice's final key is decoupled from Eve; again, see [27, 39] for details. Following [27], we design the protocol such that $\epsilon_{\text{sound}} = 10^{-10}$, by setting $\epsilon_{\text{corr}} = 2^{-61}$ and $\epsilon_{\text{secret}} = \epsilon_{\text{sound}} - \epsilon_{\text{corr}}$.

We now discuss the details of our protocol in terms of the above parameters. For the testing probability γ, in Fig. 1a we followed the value $\gamma = 13/256$ used in [27] for all data points, whereas in Fig. 1b we optimized over γ in units of $1/256$ (as was done in [27] so that the test/generation decision could be straightforwardly determined by drawing 8 uniformly random bits). We did so because we found that for the range of n values in the former, optimizing the choice of γ only improved the finite-size key rates by less than 0.002. In contrast, for the larger n values considered in the latter, we found that it was important to optimize over the choice of γ to obtain better finite-size key rates — this is due to some subtle limitations we discuss in Appendix D 4 later.

We emphasize that apart from the above point regarding γ, our protocol only differs from the protocol in [27] in terms of using a slightly different accept condition (see Remark 22), a technical point in privacy amplification (see Remark 23), and having Bob directly announce the values $\bar{B}_1^n$ for Alice to compute $\bar{C}_1^n$ (which slightly simplifies the analysis without sacrificing key rate; see Remark 24).

1. Completeness

To discuss completeness, we need to specify some honest behavior for the devices. We suppose that the honest behavior is IID, and each round produces some distribution $\mathbf{q}_{\text{hon}}$ on the register $\bar{C}_j$ for that round. For our protocol, this distribution would have the form

$$q_{\text{hon}}(0) = \gamma(1-\omega_{\text{hon}}), \quad q_{\text{hon}}(1) = \gamma\omega_{\text{hon}}, \quad q_{\text{hon}}(\perp) = 1-\gamma, \tag{D1}$$

where ω_{hon} is the expected CHSH winning probability of the honest behavior in test rounds. Furthermore, let $Q_{\text{hon}}^{\text{err}}$ denote the probability of Alice and Bob getting different outcomes in generation rounds. Following [27], we set

$$\omega_{\text{hon}} = 0.83, \quad Q_{\text{hon}}^{\text{err}} = 0.018, \tag{D2}$$

where the ω_{hon} value corresponds to the expected CHSH "correlator" score of $S = 2.64$ used in that work (as a somewhat conservative estimate of the device performance in that experiment).

For a given $\epsilon_{\text{com}}^{\text{AT}}$, we need to choose the set S_{acc} in the accept condition such that the probability of the honest behavior yielding a frequency distribution outside S_{acc} is at most $\epsilon_{\text{com}}^{\text{AT}}$. We shall focus on S_{acc} of the following form: for each value $\bar{c} \in \{0, 1, \perp\}$ we take some values $\delta_{\bar{c}}^{\text{low}}, \delta_{\bar{c}}^{\text{upp}} > 0$, and set S_{acc} to be the set of distributions $\mathbf{q}$ satisfying

$$\forall \bar{c} \in \{0,1,\perp\}, \quad q_{\text{hon}}(\bar{c}) - \delta_{\bar{c}}^{\text{low}} \leq q(\bar{c}) \leq q_{\text{hon}}(\bar{c}) + \delta_{\bar{c}}^{\text{upp}}. \tag{D3}$$

For S_{acc} of this form, to achieve some desired $\epsilon_{\text{com}}^{\text{AT}}$, it suffices to choose the values $\delta_{\bar{c}}^{\text{low}}, \delta_{\bar{c}}^{\text{upp}}$ such that for the honest behavior we have

$$\forall \bar{c} \in \{0,1,\perp\}, \quad \Pr[\text{freq}_{\bar{C}_1^n}(\bar{c}) < q_{\text{hon}}(\bar{c}) - \delta_{\bar{c}}^{\text{low}}] \leq \frac{\epsilon_{\text{com}}^{\text{AT}}}{6} \quad \text{and} \quad \Pr[\text{freq}_{\bar{C}_1^n}(\bar{c}) > q_{\text{hon}}(\bar{c}) + \delta_{\bar{c}}^{\text{upp}}] \leq \frac{\epsilon_{\text{com}}^{\text{AT}}}{6}, \tag{D4}$$

[4] While the final completeness parameter reported in that work was $\epsilon_{\text{com}} = 10^{-2}$, that was a somewhat conservative estimate.

since by the union bound, the probability of violating one or more of the inequalities is upper bounded by the sum of the individual probabilities of violating each one. (It would of course be possible to "distribute" $\epsilon_{\text{com}}^{\text{AT}}$ in some other fashion across the terms; however, we found heuristically that this appears to give better performance as compared to e.g. distributing it such that all the values $\delta_{\bar{c}}^{\text{low}}, \delta_{\bar{c}}^{\text{upp}}$ are equal.) Since the honest behavior is IID, the probabilities in (D4) can be written in terms of the CDF of a binomial distribution, which allows us to use the inverse CDF (available in most computational software) to solve for $\delta_{\bar{c}}^{\text{low}}, \delta_{\bar{c}}^{\text{upp}}$ in terms of $\epsilon_{\text{com}}^{\text{AT}}$.

Remark 22. *This choice of accept condition differs slightly from [27], which used an accept condition with only a one-sided bound on $q(1)$. We chose to use the form presented here because in some cases it seems to improve the key rates from the REAT (albeit usually only by a small amount), and also because a lower bound on $q(\perp)$ in the accept condition is needed to apply an improved chain rule we use later (in the first line of Eq. (D9) below).*

Furthermore, the accept condition in [27] was based on a 3-standard-deviation "tolerance", rather than exactly computing the CDF of a binomial distribution to achieve a desired $\epsilon_{\text{com}}^{\text{AT}}$. For this work we choose to conservatively match this by setting $\epsilon_{\text{com}}^{\text{AT}} = 10^{-3}$, since a (one-sided) 3-standard-deviation fluctuation in a normal distribution occurs with probability $1.35 \times 10^{-3} > 10^{-3}$.

As for $\epsilon_{\text{com}}^{\text{EV}}$, we first observe that error verification can only abort if Bob's guess for A_1^n is wrong, thus any upper bound on the probability of the latter (under the honest behavior) is a valid choice of $\epsilon_{\text{com}}^{\text{EV}}$. We then note that a specialized error correction procedure was developed in [27] with the following properties: for an IID honest behavior of the described form, an error-correction string of length

$$\ell_{\text{EC}} = n\left((1-\gamma)h_{\text{bin}}(Q_{\text{hon}}^{\text{err}}) + \gamma h_{\text{bin}}(1-\omega_{\text{hon}})\right) + 50\sqrt{n} \tag{D5}$$

suffices to ensure that Bob's guess is correct with probability over 99.9%, as estimated by simulations. (While this value is a somewhat heuristic estimate, recall that it only affects the probability that the honest behavior aborts, not any of the security properties of the protocol.) Hence performing error correction according to this procedure suffices to heuristically ensure $\epsilon_{\text{com}}^{\text{EV}} \leq 10^{-3}$.

2. Correctness

In [27], error verification was performed using a δ-almost-universal hash with $\delta = 2^{-61}$ and $\ell_{\text{EV}} = 64$ (under the condition that the message length in bits is at most $2^{64} \approx 10^{19}$, which is indeed the case here). We leave this aspect entirely unchanged, which suffices to ensure a correctness parameter of $\epsilon_{\text{corr}} = 2^{-61}$ as proven in [27].

3. Secrecy

This is the part of our analysis that differs the most from [27], in that apart from improving the entropy accumulation bound, we simplify or improve a number of other steps in the analysis. We shall show that to achieve a desired secrecy parameter ϵ_{secret}, it suffices to take the length of the final key to be

$$\ell_{\text{key}} = nh_\alpha - n\left(\gamma + \delta_\perp^{\text{low}}\right) - \ell_{\text{EC}} - \ell_{\text{EV}} - \frac{\alpha}{\alpha-1}\log\frac{1}{\epsilon_{\text{secret}}} + 2, \tag{D6}$$

where h_α is computed in terms of the S_{acc} choice defined in Sec. D 1, while ℓ_{EC} and ℓ_{EV} are as described in (D5) and Sec. D 2 respectively. Note that to evaluate h_α, we used generic heuristic numerical methods rather than a convex solver that returns explicit dual bounds, because our bounds on the Rényi entropy do not fall within the standard disciplined-convex-programming ruleset for such solvers. However, as the optimization is convex with respect to each of the individual variables, every local minimum is a global minimum and hence we believe it is unlikely that the resulting value we obtain for h_α is a significant overestimate of its true value.

Remark 23. *In order for the following analysis to hold, we currently require an implementation difference between the protocol described here and the protocol in [27], in that privacy amplification would have to be performed using 2-universal hashing [40] rather than Trevisan's extractor [27, 41, 42] as used in that work. This is because a Rényi privacy amplification theorem has currently only been proven for the former, not the latter. However, it seems likely that it should be possible to obtain such a result for the latter, and it would be a useful question to investigate in future work.*

Let Ω_{AT} denote the event that the protocol accepts during the acceptance test, and let Ω_{EV} denote the event that it accepts during error verification (so the event of the protocol accepting overall is $\Omega_{\mathrm{AT}} \wedge \Omega_{\mathrm{EV}}$. By applying the REAT (specifically [15, Lemma 5.1 with Lemma 6.1]), the state conditioned on Ω_{AT} satisfies

$$\widetilde{H}^{\uparrow}_{\alpha}\left(A_1^n \bar{C}_1^n | X_1^n Y_1^n T_1^n \mathbf{E}\right)_{\rho_{|\Omega_{\mathrm{AT}}}} \geq n h_\alpha - \frac{\alpha}{\alpha-1} \log \frac{1}{\Pr[\Omega_{\mathrm{AT}}]} \,. \tag{D7}$$

However, in order to apply the relevant privacy amplification theorem from [40], we would need to account for various other publicly announced registers, such as $\bar{B}_1^n$, $\mathbf{L}_{\mathrm{EC}}$, $\mathbf{L}_{\mathrm{EV}}$ and H_{EV}, as well as further conditioning on Ω_{EV}. (In fact the REAT by itself technically allows us to directly condition on $\Omega_{\mathrm{AT}} \wedge \Omega_{\mathrm{EV}}$ instead; however, doing so in this proof would obstruct the last line in (D8) later where we need to "factor off" H_{EV}.) Furthermore, since Alice performs privacy amplification only on A_1^n, the relevant Rényi entropy should only have A_1^n (not $\bar{C}_1^n$) on the left side of the conditioning.

To address these points, we first handle the conditioning on Ω_{EV}, and remove the error-correction and error-verification registers from the conditioning:

$$\begin{aligned}
&\widetilde{H}^{\uparrow}_{\alpha}\left(A_1^n | \bar{B}_1^n \mathbf{L}_{\mathrm{EC}} \mathbf{L}_{\mathrm{EV}} H_{\mathrm{EV}} X_1^n Y_1^n T_1^n \mathbf{E}\right)_{\rho_{|\Omega_{\mathrm{AT}} \wedge \Omega_{\mathrm{EV}}}} \\
\geq &\widetilde{H}^{\uparrow}_{\alpha}\left(A_1^n | \bar{B}_1^n \mathbf{L}_{\mathrm{EC}} \mathbf{L}_{\mathrm{EV}} H_{\mathrm{EV}} X_1^n Y_1^n T_1^n \mathbf{E}\right)_{\rho_{|\Omega_{\mathrm{AT}}}} - \frac{\alpha}{\alpha-1} \log \frac{1}{\Pr[\Omega_{\mathrm{EV}} | \Omega_{\mathrm{AT}}]} \\
\geq &\widetilde{H}^{\uparrow}_{\alpha}\left(A_1^n | \bar{B}_1^n H_{\mathrm{EV}} X_1^n Y_1^n T_1^n \mathbf{E}\right)_{\rho_{|\Omega_{\mathrm{AT}}}} - \ell_{\mathrm{EC}} - \ell_{\mathrm{EV}} - \frac{\alpha}{\alpha-1} \log \frac{1}{\Pr[\Omega_{\mathrm{EV}} | \Omega_{\mathrm{AT}}]} \\
= &\widetilde{H}^{\uparrow}_{\alpha}\left(A_1^n | \bar{B}_1^n X_1^n Y_1^n T_1^n \mathbf{E}\right)_{\rho_{|\Omega_{\mathrm{AT}}}} - \ell_{\mathrm{EC}} - \ell_{\mathrm{EV}} - \frac{\alpha}{\alpha-1} \log \frac{1}{\Pr[\Omega_{\mathrm{EV}} | \Omega_{\mathrm{AT}}]} \,.
\end{aligned} \tag{D8}$$

where the second line is [12, Lemma B.5], the third line is a standard chain rule[5] for classical conditioning registers, and the fourth line holds because $\rho_{A_1^n \bar{B}_1^n H_{\mathrm{EV}} X_1^n Y_1^n T_1^n \mathbf{E}}$ can be viewed as the state immediately after the choice of hash function H_{EV} was drawn, in which case H_{EV} is independent of all other registers (even conditioned on Ω_{AT}) due to how it was generated.

Next, we relate this to $\widetilde{H}^{\uparrow}_{\alpha}\left(A_1^n \bar{C}_1^n | X_1^n Y_1^n T_1^n \mathbf{E}\right)_{\rho_{|\Omega_{\mathrm{AT}}}}$ following the approach in [15]: observe that

$$\begin{aligned}
\widetilde{H}^{\uparrow}_{\alpha}\left(A_1^n | \bar{B}_1^n X_1^n Y_1^n T_1^n \mathbf{E}\right)_{\rho_{|\Omega_{\mathrm{AT}}}} &\geq \widetilde{H}^{\uparrow}_{\alpha}\left(A_1^n \bar{B}_1^n | X_1^n Y_1^n T_1^n \mathbf{E}\right)_{\rho_{|\Omega_{\mathrm{AT}}}} - n(\gamma + \delta_{\perp}^{\mathrm{low}}) \log \dim(\bar{B}_i) \\
&= \widetilde{H}^{\uparrow}_{\alpha}\left(A_1^n \bar{B}_1^n \bar{C}_1^n | X_1^n Y_1^n T_1^n \mathbf{E}\right)_{\rho_{|\Omega_{\mathrm{AT}}}} - n(\gamma + \delta_{\perp}^{\mathrm{low}}) \log \dim(\bar{B}_i) \\
&\geq \widetilde{H}^{\uparrow}_{\alpha}\left(A_1^n \bar{C}_1^n | X_1^n Y_1^n T_1^n \mathbf{E}\right)_{\rho_{|\Omega_{\mathrm{AT}}}} - n(\gamma + \delta_{\perp}^{\mathrm{low}}) \log \dim(\bar{B}_i) \,,
\end{aligned} \tag{D9}$$

where the first line is proven in [15, Remark 8.1] (noting that the number of test rounds conditioned on Ω_{AT} is at most $\gamma + \delta_{\perp}^{\mathrm{low}}$), the second line holds because $\bar{C}_1^n$ can be "projectively reconstructed" from $A_1^n B_1^n X_1^n Y_1^n T_1^n$ in the sense described in [12, Lemma B.7], and the last line holds because classical registers have non-negative contributions to entropy [34, Lemma 5.3] (this last step is not necessary in general, but we employ it here since our analytic bounds are only for the entropy of Alice's output, not Bob's). Putting all the above bounds together, we conclude (since $\dim(\bar{B}_i) = 2$ and $\Pr[\Omega_{\mathrm{AT}}] \Pr[\Omega_{\mathrm{EV}} | \Omega_{\mathrm{AT}}] = \Pr[\Omega_{\mathrm{EV}} \wedge \Omega_{\mathrm{AT}}]$):

$$\widetilde{H}^{\uparrow}_{\alpha}\left(A_1^n | \bar{B}_1^n \mathbf{L}_{\mathrm{EC}} \mathbf{L}_{\mathrm{EV}} X_1^n Y_1^n T_1^n \mathbf{E}\right)_{\rho_{|\Omega_{\mathrm{AT}} \wedge \Omega_{\mathrm{EV}}}} \geq n h_\alpha - n(\gamma + \delta_{\perp}^{\mathrm{low}}) - \ell_{\mathrm{EC}} - \ell_{\mathrm{EV}} - \frac{\alpha}{\alpha-1} \log \frac{1}{\Pr[\Omega_{\mathrm{EV}} \wedge \Omega_{\mathrm{AT}}]} \,. \tag{D10}$$

With this, we note that if we let K_A denote Alice's final key and $\mathbf{E}_{\mathrm{fin}}$ denote Eve's final side-information after privacy amplification, then we have

$$\begin{aligned}
&\Pr[\Omega_{\mathrm{EV}} \wedge \Omega_{\mathrm{AT}}] \, d\left(\rho_{K_A \mathbf{E}_{\mathrm{fin}} | \Omega_{\mathrm{EV}} \wedge \Omega_{\mathrm{AT}}}, \frac{\mathbb{I}_{K_A}}{|K_A|} \otimes \rho_{\mathbf{E}_{\mathrm{fin}} | \Omega_{\mathrm{EV}} \wedge \Omega_{\mathrm{AT}}}\right) \\
\leq &\Pr[\Omega_{\mathrm{EV}} \wedge \Omega_{\mathrm{AT}}] 2^{\frac{2}{\alpha} - 2} 2^{\frac{\alpha-1}{\alpha}\left(\ell_{\mathrm{key}} - \widetilde{H}^{\uparrow}_{\alpha}\left(A_1^n | \bar{B}_1^n \mathbf{L}_{\mathrm{EC}} \mathbf{L}_{\mathrm{EV}} H_{\mathrm{EV}} X_1^n Y_1^n T_1^n \mathbf{E}\right)_{\rho_{|\Omega_{\mathrm{EV}} \wedge \Omega_{\mathrm{AT}}}}\right)}
\end{aligned}$$

[5] Specifically, [21, Prop. 8] together with the fact that classical registers have non-negative contributions to entropy [34, Lemma 5.3].

$$
\begin{aligned}
&= \Pr[\Omega_{\mathrm{EV}} \wedge \Omega_{\mathrm{AT}}] 2^{\frac{\alpha-1}{\alpha}\left(\ell_{\mathrm{key}} - \widetilde{H}^{\uparrow}_{\alpha}\left(A_1^n|\bar{B}_1^n \mathbf{L}_{\mathrm{EC}} \mathbf{L}_{\mathrm{EV}} H_{\mathrm{EV}} X_1^n Y_1^n T_1^n \mathbf{E}\right)_{\rho|\Omega_{\mathrm{EV}}\wedge\Omega_{\mathrm{AT}}} - 2\right)} \\
&\leq \Pr[\Omega_{\mathrm{EV}} \wedge \Omega_{\mathrm{AT}}] 2^{\log \frac{1}{\Pr[\Omega_{\mathrm{EV}}\wedge\Omega_{\mathrm{AT}}]} - \log \frac{1}{\epsilon_{\mathrm{secret}}}} \\
&= \epsilon_{\mathrm{secret}} \,,
\end{aligned} \tag{D11}
$$

where the second line is the Rényi privacy amplification theorem in [40, Theorem 9 with Lemma 7] (noting that the 1-norm distance and trace distance differ by a factor of 1/2), the third line simply regroups the exponents, and the fourth line follows from combining (D6) with (D10). This fulfills the definition of $\epsilon_{\mathrm{secret}}$-secrecy as described in e.g. [19, 25].

Remark 24. *Comparing the above analysis to that in [27], apart from the main change of the former being based on Rényi entropies*[6] *(which also simplified some points regarding event conditioning), the other notable difference is the chain rule used above to obtain the first line of* (D9)*. We believe it is likely to yield better results than the chain rules used in [27] to handle the $\bar{B}_1^n$ registers. Moreover, this chain rule also places $\bar{B}_1^n$ directly in the conditioning registers, which allowed us to modify the protocol such that Bob publicly announces the $\bar{B}_1^n$ registers for Alice to compute $\bar{C}_1^n$ — this simplifies the analysis compared to [27], where instead Bob computes $\bar{C}_1^n$ using his guess for Alice's A_1^n registers, and extra steps had to be taken in that proof to accommodate the possibility that his guess could be wrong.*

4. Possible modifications

Finally, we make some informal comments regarding some potential for slightly sharpening the above analysis. Namely, the way we computed the lower bound on h_α is slightly suboptimal, in that we were effectively "sacrificing" entropy contributions from test rounds (in that comparing to the REAT statement in [15, Lemma 5.1], we have handled the entropy contributions from terms with $\bar{c} \neq \perp$ by trivially lower bounding them with zero). In contrast, the earlier security proofs in e.g. [19, 27] using previous EAT versions (based on von Neumann entropy in single rounds) were able to incorporate the entropy contributions from those rounds, due to certain properties of von Neumann entropy that did not carry over to the REAT. Due to this, we found that at the larger n values studied in Fig. 1b (where the rates are closer to the asymptotic values), the REAT-based approach here gives worse rates than those in [27] when using the same value of γ — to obtain the improved rates in that figure, we instead had to optimize the choice of γ, using a smaller value that "sacrifices" less of the key rate to the test-round component.

From a theoretical point of view, one way to overcome this drawback would be to use [15, Theorem 5.1] rather than [15, Lemma 5.1], in that it "retains" entropy contributions from test rounds. However, the former involves some slightly elaborate Rényi divergence terms that do not seem straightforward to analyze using the methods in this work. It would be an interesting task for future work to generalize these methods to handle those Rényi divergence terms. Another prospect could be to note that [15, Lemma 5.1] itself technically involves Rényi entropy terms conditioned on each value of $\bar{c}$ that could be individually analyzed (here we have basically only retained the $\bar{c} = \perp$ term); however, that approach seems less promising for future use, because those terms would all be zero if the protocol is one where $\bar{c}$ contains the full input-output values in test rounds.

Alternatively, from a practical perspective it might seem expedient to address this by having Alice perform privacy amplification on only the generation-round data (since the test-round entropy is anyway "sacrificed"), in which case she would not need to include test-round data in the error correction information (as implicitly accounted for in (D5)) — this would reduce the value of ℓ_{EC} accordingly and also make that step practically easier to implement. However, due to the structure of the above proof (mainly the use of the "projective reconstruction" property from [12, Lemma B.7] in the second line of (D9), which requires Alice's test-round outputs to appear at some point in the entropy terms), it does not seem entirely straightforward to "directly" get a bound for the final entropy of Alice's generation-round registers only.

We observe that technically, one way to obtain a bound would be as follows: define registers $\bar{A}_i$ that are equal to A_i in test rounds and set to 0 in generation rounds, and define registers $\hat{A}_i$ that are equal to A_i in generation rounds and set to 0 in test rounds. Then

$$
\begin{aligned}
\widetilde{H}^{\uparrow}_{\alpha}\left(\hat{A}_1^n|\bar{B}_1^n X_1^n Y_1^n T_1^n \mathbf{E}\right)_{\rho|\Omega_{\mathrm{AT}}} &\geq \widetilde{H}^{\uparrow}_{\alpha}\left(\hat{A}_1^n|\bar{A}_1^n \bar{B}_1^n X_1^n Y_1^n T_1^n \mathbf{E}\right)_{\rho|\Omega_{\mathrm{AT}}} \\
&\geq \widetilde{H}^{\uparrow}_{\alpha}\left(\hat{A}_1^n \bar{A}_1^n \bar{B}_1^n|X_1^n Y_1^n T_1^n \mathbf{E}\right)_{\rho|\Omega_{\mathrm{AT}}} - n(\gamma + \delta_{\perp}^{\mathrm{low}}) \log \dim(\bar{A}_i \bar{B}_i)
\end{aligned}
$$

[6] Refer to [26, Fig. 1] for an analysis of the effect of making only this change, without any of the other improvements we employ here.

$$= \widetilde{H}^{\uparrow}_{\alpha}\left(A_1^n \bar{B}_1^n | X_1^n Y_1^n T_1^n \mathbf{E}\right)_{\rho_{|\Omega_{\mathrm{AT}}}} - n(\gamma + \delta_{\perp}^{\mathrm{low}}) \log \dim(\bar{A}_i \bar{B}_i)\,, \tag{D12}$$

where the second line is again via [15, Remark 8.1], and the third line is again via [12, Lemma B.7], observing that A_1^n can be "projectively reconstructed" from $\hat{A}_1^n \bar{A}_1^n T_1^n$ (and vice versa, $\hat{A}_1^n \bar{A}_1^n$ can also be "projectively reconstructed" from $A_1^n T_1^n$).

With this we can continue on exactly the same way as in the second and third lines of (D9). However, observe that this results in $2n\left(\gamma + \delta_{\perp}^{\mathrm{low}}\right)$ in place of $n\left(\gamma + \delta_{\perp}^{\mathrm{low}}\right)$ in the final key length formula. In order to obtain an overall benefit from this approach, we would need a more detailed analysis of how much the $O(\sqrt{n})$ term in ℓ_{EC} (Eq. (D5)) can be improved by not having to include the test-round data, which is a somewhat more specialized coding-theory question that we shall not consider within this work. (Furthermore, we informally note that this proposed approach still ends up reducing the analysis to $\widetilde{H}^{\uparrow}_{\alpha}\left(A_1^n \bar{B}_1^n | X_1^n Y_1^n T_1^n \mathbf{E}\right)_{\rho_{|\Omega_{\mathrm{AT}}}}$, which again includes Alice's test-round data and hence does not really "exploit" the fact that our single-round analysis excludes the entropy contributions from those terms.)