Fully device-independent conference key agreement

Ribeiro, Jérémy; Murta Guimarães, Gláucia; Wehner, Stephanie

DOI
10.1103/PhysRevA.97.022307

Publication date
2018

Document Version
Final published version

Published in
Physical Review A

Fully device-independent conference key agreement

Jérémy Ribeiro, Gláucia Murta, and Stephanie Wehner
QuTech, Delft University of Technology, Lorentzweg 1, NL-2628 CJ Delft, The Netherlands

(Received 23 August 2017; published 6 February 2018)

We present a security analysis of conference key agreement (CKA) in the most adversarial model of device independence (DI). Our protocol can be implemented by any experimental setup that is capable of performing Bell tests [specifically, the Mermin-Ardehali-Belinskii-Klyshko (MABK) inequality], and security can in principle be obtained for any violation of the MABK inequality that detects genuine multipartite entanglement among the N parties involved in the protocol. As our main tool, we derive a direct physical connection between the N-partite MABK inequality and the Clauser-Horne-Shimony-Holt (CHSH) inequality, showing that certain violations of the MABK inequality correspond to a violation of the CHSH inequality between one of the parties and the other $N-1$. We compare the asymptotic key rate for device-independent conference key agreement (DICKA) to the case where the parties use $N-1$ device-independent quantum key distribution protocols in order to generate a common key. We show that for some regime of noise the DICKA protocol leads to better rates.

DOI: 10.1103/PhysRevA.97.022307

Quantum communication allows cryptographic security that is provably impossible to obtain using any classical means. Probably the most famous example of a quantum advantage is quantum key distribution (QKD) [1,2], which allows two parties, Alice and Bob, to exchange an encryption key whose security is guaranteed even if the adversary has an arbitrarily powerful quantum computer. What's more, properties of entanglement lead to the remarkable feature that security is sometimes possible even if the quantum devices used to execute the protocol are largely untrusted—specifically, the notion of *device-independent* (DI) security [3–5] model quantum devices as black boxes in which we may only choose measurement settings and observe measurement outcomes. Yet, the quantum state and measurements employed by such boxes are unknown, and may even be prepared arbitrarily by the adversary.

Significant efforts have been undertaken to establish the security of device-independent QKD [5–11], leading to ever more sophisticated security proofs. Initial proofs assumed a simple model in which the devices act independently and identically (i.i.d.) in each round of the protocol. This significantly simplifies the security analysis since the underlying properties of the devices may first be estimated by gaining statistical confidence from the observation of the measurement outcomes in the tested rounds. The main challenge overcome by the more recent security proofs [8–11] was to establish security even if the devices behave arbitrarily from one round to the next, including having an arbitrary memory of the past that they might use to thwart the efforts of Alice and Bob. Assuming that the devices carry at least some memory of past interactions is an extremely realistic assumption due to technical limitations, even if Alice and Bob prepare their own trusted, but imperfect, devices, highlighting the extreme importance of such analyses for the implementation of device-independent QKD. In contrast, relatively little is known about device independence outside the realm of QKD [12–16].

Conference key agreement [17–19] (CKA or N-CKA) is the task of distributing a secret key among N parties. In order to achieve this goal, one could make use of $N-1$ individual QKD protocols to distribute $N-1$ different keys between one of the parties (Alice) and the others ($\mathrm{Bob}_1,\ldots,\mathrm{Bob}_{N-1}$), followed by Alice using these keys to encrypt a common key to all the participants. However, the existence of genuine multipartite quantum correlations can bring some advantage to multipartite tasks, and, as shown in Ref. [19], exploring properties of genuine multipartite entanglement can lead to protocols with better performance for conference key agreement.

Here we present a security analysis of conference key agreement in the most adversarial model of device independence. Our protocol can be implemented using any experimental setup that is capable of testing the Mermin-Ardehali-Belinskii-Klyshko (MABK) [20–22] inequality. The MABK inequalities have previously been used for randomness amplification and certification [23,24], and for self-testing N-partite systems [25,26]. Our proof is based on a physical insight linking the N-partite MABK inequality and the Clauser-Horne-Shimony-Holt (CHSH) inequality [27]. Specifically, we show that a violation of the MABK inequality detecting genuine multipartite entanglement, among the N parties involved in the protocol, corresponds to a violation of the CHSH inequality defined between one party and the other $N-1$. We also compare the asymptotic rates obtained for device-independent conference key agreement (DICKA) with the implementation of $N-1$ device-independent quantum key distribution (DIQKD), and show that for some regime of noise it is advantageous to perform DICKA. The manuscript is organized as follows: In the next section we present the protocol and state the security definitions for conference key agreement. Then we sketch the security proof of our DICKA protocol and present the MABK-CHSH correspondence. We finish with a comparison of the asymptotic key rates. An expanded and detailed derivation of the security proof and the noise model for the asymptotic key rates are presented in the Appendix.

I. THE PROTOCOL

For a device-independent implementation of CKA, we consider a protocol with N parties: Alice who possesses one device with two inputs $\{0,1\}$, and $\text{Bob}_1,\dots,\text{Bob}_{N-1}$ who each possess a device with three inputs $\{0,1,2\}$, every input with two outputs. During the protocol, Alice and the Bobs randomly choose some rounds to test for the violation of the MABK inequality. They abort the protocol if the frequency of rounds where they win the MABK game do not reach a specified threshold δ. We also consider that Alice has a source for generation of the states, which is independent of her measurement device.

Protocol 1 (DICKA):

1. *For every round $i \in [n]$ do:*
   (a) *Alice uses her source to produce and distribute an N-partite state, $\rho_{A_i B_{(1\dots N-1),i}}$, shared among herself and the $N-1$ Bobs.*
   (b) *Alice randomly picks T_i, s.t. $P(T_i = 1) = \mu$, and publicly communicates it to all the Bobs.*
   (c) *If $T_i = 0$ Alice and the Bobs choose $(X_i, Y_{(1\dots N-1),i}) = (0,2,\dots,2)$, and if $T_i = 1$ they all choose $X_i, Y_{(1\dots N-1),i} \in_R \{0,1\}$ uniformly at random.*
   (d) *Alice and the Bobs input the previously chosen values in their respective device and record the outputs as $A'_i, B'_{(1\dots N-1),i}$.*
2. *They all communicate publicly the list of bases $X_1^n Y_{(1\dots N-1)}{}_1^n$ they used.*
3. ***Error correction.** Alice and the Bobs apply an error correction protocol. We call O_A the classical information that Alice sends to the Bobs. For the purpose of parameter estimation, the Bobs also send some error correction information for the bits produced during the test rounds ($T_i = 1$); we denote $O_{(k)}$ the error correction information sent by* Bob_k. *If the error correction protocol aborts for at least one Bob then they abort the protocol. If it does not abort they obtain the raw keys $\tilde{K}_A = A', \tilde{K}_{B_{(1\dots N-1)}}$.*
4. ***Parameter estimation.** If $T_i = 1$, Alice uses A'_i and her guess on $B'_{(1\dots N-1),i}$ to set $C_i = 1$ if they have won the N-partite MABK game, and she sets $C_i = 0$ if they have lost it. If $T_i = 0$, she sets $C_i = \perp$. She aborts if $\sum_i C_i < \delta \cdot \sum_i T_i$, where $\delta \in\,]p_{\min}, p_{\max}[$.*
5. ***Privacy amplification.** Alice and the Bobs apply a privacy amplification protocol to create final keys $K_A, K_{B_{(1\dots N-1)}}$. We denote S the classical information publicly sent by Alice during this step.*

Security definitions For completeness, before stating our main result, which establishes the secret key length of Protocol 1, we first formalize what it means for a DICKA protocol to be secure. As for QKD [28,29] the security of conference key agreement [19] can be split into two terms: *correctness* and *secrecy*. Correctness is a statement about how sure we are that the N parties share identical keys, and secrecy is a statement about how much information the adversary can have about Alice's key.

Definition 1. (correctness and secrecy) A DICKA protocol is ϵ_{corr}-correct if Alice's and Bobs' keys, $K_A, K_{B_{(1)}}, \dots, K_{B_{(N-1)}}$, are all identical with probability at least $1-\epsilon_{\text{corr}}$. And it is ϵ_{sec}-secret, if Alice's key K_A is ϵ_{sec}-close to a key that Eve is ignorant about. This condition can be formalized as

$$p_{\hat{\Omega}} \left\| \rho_{K_A E|\hat{\Omega}} - \frac{\mathbb{1}_A}{2^l} \otimes \rho_{E|\hat{\Omega}} \right\|_{\text{tr}} \leqslant \epsilon_{\text{sec}},$$

where $\|\cdot\|_{tr}$ denotes the trace norm, l is the key length, $\hat{\Omega}$ is the event of the protocol not aborting, and $p_{\hat{\Omega}}$ is the probability for $\hat{\Omega}$.

If a protocol is ϵ_{corr}-correct and ϵ_{sec}-secret then it is ϵ^s-correct-and-secret for any $\epsilon^s \geqslant \epsilon_{\text{corr}} + \epsilon_{\text{sec}}$.

So in general when we say that a CKA (or a QKD) protocol is ϵ^s secure, we mean that for any possible physical implementation of the protocol, either it aborts with probability higher than $1-\epsilon^s$ or it is ϵ^s-correct-and-secret, according to Definition 1 (see Appendix, Sec. 2 b).

A combination of Definition 1 and the leftover hashing lemma [28] relates the length of a secret key, that can be obtained from a particular protocol, with the smooth min-entropy of Alice's raw key A' conditioned on Eve's information (see [28] for a detailed derivation of this statement): An ϵ_{sec}-secret key of size,

$$l = H_{\min}^{\epsilon}(A'|E) - 2\log_2 \frac{1}{\epsilon_{\text{PA}}}, \tag{1}$$

can be obtained, for $\epsilon_{\text{sec}} > 2\epsilon + \epsilon_{\text{PA}}$. The conditional smooth min-entropy is defined as $H_{\min}^{\epsilon}(A|E)_\rho := \sup_{\sigma\in\mathcal{B}(\rho)} H_{\min}(A|E)_\sigma$, with the supremum taken over all positive semidefinite operators ϵ-close to ρ in the purified distance (see [30]). For A a classical register and σ a quantum state, $H_{\min}(A|E)_\sigma$ represents the maximum probability with which Eve can guess the value of A if they share the state σ. In general $H_{\min}(A|E)_\sigma := \sup_{\tau_E} \sup\{\lambda : \sigma_{\text{AE}} \leqslant 2^{-\lambda} \mathbb{1}_A \otimes \tau_E\}$, where the supremum is taken over all quantum states τ_E.

Definition 1 was proved to be a criteria for composable security for QKD in the device-dependent scenario [29]. However, it is important to note that for the DI case it is not known whether such a criterion is enough for composable security. Indeed, Ref. [31] suggests that this is not the case if the same devices are used for generation of a subsequent key, since this new key can leak information about the first key. Following Ref. [11] we chose to adopt these definitions as the security criteria for DICKA.

Our main result establishes the length of a secure key that can be obtained from Protocol 1.

Theorem 1. Protocol 1 generates an ϵ^s-correct-and-secret key, with $\epsilon^s \leqslant \epsilon_{\rm PA} + 2(N-1)\epsilon'_{\rm EC} + 2\epsilon + \epsilon_{\rm EA}$, of length:

$$\begin{aligned} l = & \max_{p_{\min}\leqslant\delta_{\rm opt}\leqslant p_{\max}} [(f(\delta,\delta_{\rm opt}) - \mu)\cdot n - \tilde{v}\sqrt{n}] \\ & + 3\log_2(1-\sqrt{1-(\epsilon/4)^2}) - 2\log_2\left(\epsilon_{\rm PA}^{-1}\right) \\ & - \mathrm{leak}_{\rm EC}(O_A) - \sum_{k=1}^{N-1}\mathrm{leak}_{\rm EC}(O_{(k)}), \end{aligned} \tag{2}$$

where $\epsilon'_{\rm EC}$ is an error parameter of the error correction protocol, $\epsilon_{\rm PA}$ is the privacy amplification error probability, $\epsilon_{\rm EA}$ is a chosen security parameter for the protocol, and ϵ is a smoothing parameter. δ is the specified threshold bellow which the protocol aborts. The function $f(\cdot,\delta_{\rm opt})$ is the tangent of $\hat{f}(\cdot)$ [see Eq. (10)] in the point $\delta_{\rm opt}$, where $\delta_{\rm opt} \in]p_{\min},p_{\max}[$ is a parameter to be optimized. $\tilde{v} = 2(\log_2(13) + (\hat{f}'(p_{\rm opt})/\mu + 1))\sqrt{1-2\log_2(\epsilon\cdot\epsilon_{\rm EA})} + 2\log_2(7)\sqrt{-\log_2(\epsilon_{\rm EA}^2(1-\sqrt{1-(\epsilon/4)^2}))}$. And the leakages due to error correction, $\mathrm{leak}_{\rm EC}$, can be estimated according to a particular implementation of the protocol.

The security proof of Protocol 1 consists of two main steps: We first use the recently developed Entropy Accumulation Theorem [32] to split the overall entropy of Alice's string, produced during the protocol, into a sum of the entropy produced on each round of the protocol. Then we develop a new method to bound the entropy produced in one round by a function of the violation of the N-partite MABK inequality, which generalizes the bound for the bipartite case derived in [5,6]. In the following section we sketch the steps of the proof of Theorem 1. An expanded and detailed derivation of this result is presented in the Appendix.

II. SECURITY ANALYSIS

Step 1: Breaking the entropy round by round with the entropy accumulation theorem (EAT). To prove the security of Protocol 1 we need to lower bound the smooth min-entropy of the string produced by Alice's device conditioned on all the information Eve obtains during the protocol (evaluated on the output state of Protocol 1 given the event $\hat{\Omega}$ of not aborting),

$$H_{\min}^{\epsilon}\left(A'^n_1\middle|X_1^nY_{(1\ldots N-1)}{}_1^nT_1^nO_AO_{(1\ldots N-1)}E\right)_{\rho_{|\hat{\Omega}}}, \tag{3}$$

where E denotes Eve's quantum side information and all the other registers have been defined in Protocol 1. We can treat the error correction information $O_AO_{(1\ldots N-1)}$ that is communicated between Alice and the Bobs as a leakage:

$$\begin{aligned} (3) \geqslant & H_{\min}^{\epsilon}\left(A'^n_1\middle|X_1^nY_{(1\ldots N-1)}{}_1^nT_1^nE\right)_{\rho_{|\hat{\Omega}}} \\ & - \mathrm{leak}_{\rm EC}(O_A) - \sum_{k=1}^{N-1}\mathrm{leak}_{\rm EC}(O_{(k)}). \end{aligned} \tag{4}$$

This relation follows from the properties of the smooth min-entropy (see [33], Lemma 6.8).

Now, in order to bound the term $H_{\min}^{\epsilon}(A'^n_1|X_1^nY_{(1\ldots N-1)}{}_1^nT_1^nE)_{\rho_{|\hat{\Omega}}}$, we use the entropy accumulation theorem [32]. The EAT has already been used to prove security of device-independent QKD [11]. This theorem permits one to lower bound the above entropy by a sum of Von Neumann entropies evaluated on each round i. More precisely,

$$H_{\min}^{\epsilon}\left(A'^n_1\middle|X_1^nY_{(1\ldots N-1)}{}_1^nT_1^nE\right)_{\rho_{|\hat{\Omega}}} \geqslant nt - v\sqrt{n}, \tag{5}$$

where v is a prefactor independent of the number of rounds and t is a lower bound (for every round i) on the Von Neumann entropy $H(A'_i|X_1^iY_{(1\ldots N-1)}{}_1^iA'^{i-1}_1T_1^iE)_{\mathcal{M}_i(\sigma)}$ for all initial states σ that would achieve a Bell violation larger than the chosen threshold δ (see Appendix, Sec. 1 d). The EAT then reduces the security proof in the most adversarial scenario to the estimation of t.

Step 2: Bounding the entropy by a function of the Bell violation. We now proceed to lower bound t for Protocol 1, i.e., we find a lower bound on the Von Neumann entropy $H(A'_i|X_1^iY_{(1\ldots N-1)}{}_1^iA'^{i-1}_1T_1^iE)_{\mathcal{M}_i(\sigma)}$ as a function of the violation of the MABK inequality for N parties. The MABK inequalities [20–22] are N-partite Bell inequalities that reduce to the CHSH inequality for $N=2$. In order to see that, we first define the CHSH function $F_{\rm CHSH}$ that takes four operators A_0,A_1,B_0,B_1 as

$$\begin{aligned} & F_{\rm CHSH}(A_0,A_1,B_0,B_1) \\ & \quad := A_0\otimes B_0 + A_0\otimes B_1 + A_1\otimes B_0 - A_1\otimes B_1. \end{aligned} \tag{6}$$

Note that this is the standard Bell operator used to test the CHSH inequality.

Now we can define by recursion the MABK inequalities for N parties $\mathrm{Paul}_1,\ldots,\mathrm{Paul}_N$.

Definition 2 (MABK inequalities). Let P_0^i,P_1^i be the two binary observables ($P_k^{i\dagger} = P_k^i$ & $P_k^{i\,2} \leqslant \mathbb{1}_{\mathcal{P}_i}$) of Paul_i, $\forall i \in [N]$. Then we define by recursion,

$$MK_2 := \tfrac{1}{2}F_{\rm CHSH}\left(P_0^1,P_1^1,P_0^2,P_1^2\right), \tag{7}$$

$$MK_N := \frac{F_{\rm CHSH}\left(MK_{N-1},\overline{MK}_{N-1},P_0^N,P_1^N\right)}{2}, \tag{8}$$

where $\overline{MK}_l$ is the operator obtained from MK_l by replacing P_k^i by P_{1-k}^i, $\forall i\in[l],\forall k\in\{0,1\}$.

The N-partite MABK inequalities, $\forall N\geqslant 2, m\in[N]$, are

$$\mathcal{MK}_N := |\mathrm{tr}(MK_N\,\rho_{\mathcal{P}_{(1\ldots N)}})| \leqslant 2^{\frac{m-1}{2}}. \tag{9}$$

The bound for $m=1$ gives the classical value of the N-MABK inequality, and $m=N$ gives an upper bound (tight) on what can be achieved with quantum mechanics. For $1<m<N$, it was shown in [34] that $2^{(m-1)/2}$ is the maximal value that can be achieved by $(N-m+1)$-separable states. In particular, the violation of the inequality for $m=N-1$ witnesses the existence of genuine N-partite entanglement [34,35]. Note that MK_2 is the normalized CHSH operator, as it corresponds to an expression with classical value 1.

Now we are ready to state the result that constitutes our main tool.

Theorem 2 (MABK-CHSH correspondence). An N-partite MABK inequality with a violation $\mathcal{MK}_N > 2^{(m-1)/2}$, for $m = N-1$, can be reinterpreted as a CHSH inequality for a bipartite splitting consisting of one party on one side and the $N-1$ other parties on the other side, achieving a violation of $\mathcal{MK}_2 = \mathcal{MK}_N/2^{(N-2)/2} > 1$.

Proof. To see that we replace the operators MK_{N-1} and $\overline{MK}_{N-1}$ in Eq. (8) by their renormalized versions: $MK_{N-1}/2^{(N-2)/2}$ and $\overline{MK}_{N-1}/2^{(N-2)/2}$. Now, note that these renormalized operators can be seen as observables (they are Hermitian and their square is smaller than $\mathbb{1}$). Therefore, the N-partite MABK inequality (9) divided by $2^{(N-2)/2}$ corresponds to a CHSH inequality MK_2 between Paul_N and the $N-1$ other Pauls. ■

According to Theorem 2, if Alice and the $N-1$ Bobs of Protocol 1 violate the N-partite MABK inequality for $m = N-1$ (i.e., the value that certifies genuine N-partite entanglement), it is equivalent to Alice playing a CHSH game with the $N-1$ Bobs and achieving a violation. Therefore we can use the main result of Ref. [5] (which is a lower bound on the entropy of Alice's bit conditioned on Eve's information, as a function of the CHSH violation) to show that the function defined as

$$\hat{f}(p_w) := \left(1-\frac{\mu}{2}\right)\left(1-h\left(\frac{1}{2}+\frac{1}{2}\sqrt{\left(\frac{\mathcal{MK}_N(p_w)}{2^{(N-2)/2}}\right)^2-1}\right)\right), \tag{10}$$

lower bounds $H := H(A_i'|X_1^i Y_{(1\ldots N-1)}{}_1^i A_1'^{i-1} T_1^i E)_{\mathcal{M}_i(\sigma)}$. Here $h(\cdot)$ is the binary entropy, μ is the testing probability defined in Protocol 1, and the MABK value relates with the probability of winning the MABK game by

$$\mathcal{MK}_N(p_w) = 2^{\lfloor\frac{N}{2}\rfloor+1}\left(2^{N-2\lfloor\frac{N}{2}\rfloor}p_w - \frac{1}{2}\right). \tag{11}$$

As the protocol aborts when the observed violation is smaller than $\mathcal{MK}_N(\delta)$, where δ is the threshold specified in Protocol 1, we have

$$H \geqslant \hat{f}(\delta). \tag{12}$$

And note that, since $\hat{f}$ is a convex function of δ, its tangent in any point is also a lower bound on H, which defines t for Protocol 1 [see Appendix, Sec. 2 b for a detailed derivation of of Eq. (10)].

III. ASYMPTOTIC KEY RATE AND COMPARISON WITH DIQKD-BASED PROTOCOL

We remark that bipartite QKD has of course been studied in the device-independent setting [11], but as we are going to see in Fig. 1, a conference key agreement protocol can be beneficial for certain regimes of noise.

Combining Eqs. (3), (4), and (5) we get a lower bound on the length of secret key we can obtain with Protocol 1, which, when divided by the number of rounds n, gives us a lower bound on the secret key rate.

In order to calculate the secret key rate, we also need to estimate the leakages due to error correction, Eq. (4), and for that we need to specify the model for an honest implementation. Modeling the noise on the distributed state as a depolarizing noise we get

$$\text{leak}_{\text{EC}}(O_A) \leqslant [(1-\mu)h(Q)+\mu]n + O(\sqrt{n}), \tag{13}$$

and

$$\text{leak}_{\text{EC}}(O_{(k)}) \leqslant \mu n + O(\sqrt{n}), \tag{14}$$

where Q is the *quantum bit error rate* (QBER) between Alice and one of the Bobs. A detailed calculation of the leakage for this particular noise model is presented in the Appendix, Sec. 3.

Using this estimation of the leakage in the bounds for the entropy (3), and by taking $\mu \to 0$, s.t. $\mu\sqrt{n} \to \infty$, we get the asymptotic key rate for Protocol 1:

$$r^\infty_{N\text{-DICKA}} = 1 - h\left(\tfrac{1}{2}+\tfrac{1}{2}\sqrt{2(1-2Q)^N-1}\right) - h(Q). \tag{15}$$

We compare the above rate with the one we would have if Alice was performing $N-1$ DIQKD protocols in order to establish a common key with all the Bobs [11]:

$$r^\infty_{(N-1)\times\text{DIQKD}} = \frac{1-h\left(\frac{1}{2}+\frac{1}{2}\sqrt{2(1-2Q)^2-1}\right)-h(Q)}{N-1}. \tag{16}$$

Because when Alice runs $N-1$ DIQKD protocols she needs n rounds for each of the $N-1$ Bobs, the key rate $r^\infty_{(N-1)\times\text{DIQKD}}$ gets a factor of $\frac{1}{N-1}$. Note that here we consider that the cost for locally producing an N-partite GHZ state is comparable to the cost of producing EPR pairs. An analysis taking into account these costs for particular implementations will lead to a more fair comparison.

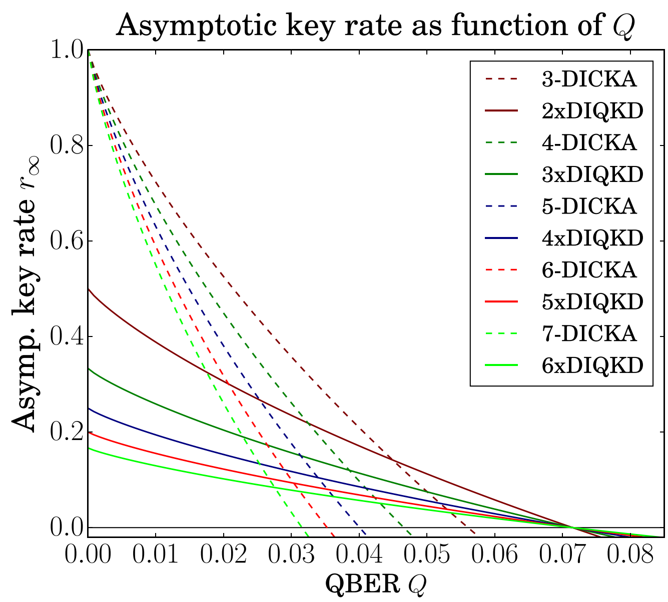

FIG. 1. Asymptotic key rate for N-DICKA (dashed lines), and for the distribution of a secret key between N parties through $N-1$ DIQKD protocols (solid lines), when each qubit experiences independent bit errors measured at a bit error rate (QBER) Q. From top to bottom, the lines correspond to $N = \{3,4,5,6,7\}$. We observe that for the low noise regime it is advantageous to use DICKA instead of $(N-1)\times$ DIQKD [11]. In general, the comparison between the two methods depends on the cost and noisiness of producing GHZ states over pairwise EPR pairs.

A comparison of these key rates is given in Fig. 1, where we see that in some regimes of noise, it can be advantageous to use the N-partite DICKA Protocol 1 instead of $N-1$-independent DIQKD protocols.

IV. CONCLUSION

We presented the first security proof for a fully device-independent implementation of conference key agreement. We have shown that, in principle, security can be achieved for any violation of the MABK inequality that detects genuine multipartite entanglement. It remains an open point whether the protocol can be extended in such a way that for violations of the MABK inequality that do not certify genuine N-partite entanglement we can still guarantee security.

We have compared the asymptotic key rates achieved with the DICKA protocol versus $N-1$ implementations of DIQKD, modeling the quantum channel connecting the parties as depolarizing channels. For implementations where the cost of local generation of GHZ states and EPR pairs is comparable, we show that it is advantageous to use DICKA for low noise regimes. A careful analysis that takes into account the costs of generation of the states is still needed for particular implementations.

We remark that proving advantage for a small number of parties already leads to better protocols for networks. Indeed, instead of using DIQKD as a building block for an N-DICKA protocol (for large N), one can use k-DICKA protocols, upon availability of k-GHZ states for $k=3,4$ or 5.

Finally, we also remark that our DICKA protocol can be adapted for other multipartite Bell inequalities. However, in general, finding good lower bounds on Eve's information about Alice's output as a function of the Bell violation is a difficult task. The MABK-CHSH correspondence proved in Theorem 2 represents an advance in this direction. Further exploration of this technique can lead to useful relations between other Bell inequalities.

ACKNOWLEDGMENTS

We thank Johannes Borregaard, Mikhail D. Lukin, and Valentina Caprara Vivoli for helpful discussions. This work was supported by STW Netherlands, NWO VIDI, and ERC Starting Grant.

APPENDIX

Here we expand in detail upon the security proof of the DICKA protocol presented in the main text, Protocol 1. A more detailed version of Protocol 1 is given in this Appendix in Protocol 2.

The Appendix is organized as follows: In Sec. 1 we introduce some background. We start by introducing the notation and some definitions which are going to be used in the main proofs. Then we present the entropy accumulation theorem, which constitutes an important tool of our security proof. We finish discussing the set of hypotheses contained in the device-independent model. In Sec. 2, we state the DICKA protocol and present the detailed security proof. In Sec. 3 we present the noise model to compare the asymptotic key rate of the DICKA protocol to the case where the parties perform $N-1$-independent DIQKD protocols in order to generate a common key.

1. Preliminaries

a. Notation

We denote $\mathcal{H}_A$ the Hilbert space of the system A with dimension $|A|$ and $\mathcal{H}_{AB}:=\mathcal{H}_A\otimes\mathcal{H}_B$ the Hilbert space of the composite system, with $\otimes$ the tensor product. By $\mathcal{L}(\mathcal{H})$, $\mathcal{S}_a(\mathcal{H})$, $\mathcal{P}(\mathcal{H})$, and $\mathcal{S}(\mathcal{H})$ we mean the set of linear, self-adjoint, positive semidefinite, and (quantum) density operators on $\mathcal{H}$, respectively. For two operators $A,B\in\mathcal{S}_a(\mathcal{H})$, $A\geqslant B$ means $(A-B)\in\mathcal{P}(\mathcal{H})$. For $M\in\mathcal{L}(\mathcal{H})$, we denote $|M|:=\sqrt{M^\dagger M}$, and the Schatten p-norm $\|M\|_p:=\mathrm{tr}(|M|^p)^{1/p}$ for $p\in[1,\infty[$, and $\|M\|_\infty$ is the largest singular value of M. For $M\in\mathcal{P}(\mathcal{H})$, M^{-1} is the generalized inverse of M, meaning that the relation $MM^{-1}M=M$ holds. If $\rho_{AB}\in\mathcal{S}(\mathcal{H}_{AB})$ then we denote $\rho_A:=\mathrm{tr}_B(\rho_{AB})$ and $\rho_B:=\mathrm{tr}_A(\rho_{AB})$ to be the respective reduced states. We use $[n]$ as a shorthand for $\{1,\ldots,n\}$. If we deal with a system composed with N subsystems within a round i of a protocol we denote $A_{(k\ldots l),i}$ for $A_{(k),i},\ldots A_{(l),i}$ $(k,l\in[N]: k\leqslant l)$, where $A_{(k),i}$ is the k^{th} subsystem of the round i. If we deal with a system composed of n subsystems across the n rounds of a protocol we denote A_k^l for $A_k,\ldots,A_l$ $(k,l\in[n]: k\leqslant l)$. Therefore $A_{(k\ldots l)_m^o}$ is a short for $A_{(k\ldots l),m},\ldots A_{(k\ldots l),o}$ $(k,l\in[N],m,o\in[n]: k\leqslant l, m\leqslant o)$.

For classical-quantum states (or cq states),

$$\rho_{XA}:=\sum_{x\in\mathcal{X}}p_x|x\rangle\langle x|_X\otimes\rho_{A|x},$$

where $\{p_x\}$ is a probability distribution on the alphabet $\mathcal{X}$ of X. We define a cq state $\rho_{XA|\Omega}$ conditioned on an event $\Omega\subset\mathcal{X}$ as

$$\rho_{XA|\Omega}:=\frac{1}{p_\Omega}\sum_{x\in\Omega}p_x|x\rangle\langle x|_X\otimes\rho_{A|x},\quad\text{where}\quad p_\Omega:=\sum_{x\in\Omega}p_x. \tag{A1}$$

We will denote by CPTP maps the linear maps that are completely positive and trace preserving.

Let $\mathcal{C}$ be an alphabet, and $C_1,\ldots,C_n$ be n random variables on this alphabet. We call $\mathsf{freq}(C_1^n)$ the vector whose components labeled by $c\in\mathcal{C}$ are the frequencies of the symbol c:

$$\mathsf{freq}(C_1^n)_c:=\frac{|\{i: C_i=c\}|}{n}.$$

b. Entropies

Throughout this work we will make use the smooth min-(max-) entropy. To define them we first define the min- and max-entropies [33].

Definition 3. If ρ_{AB} is a bipartite state and $\epsilon\in]0,1[$, we define the min- and max-entropies as

$$H_{\min}(A|B)_\rho:=-\log_2\left(\inf_{\sigma_B}\|\rho_{AB}^{\frac{1}{2}}\sigma_B^{-\frac{1}{2}}\|_\infty^2\right), \tag{A2}$$

$$H_{\max}(A|B)_\rho:=\log_2\left(\sup_{\sigma_B}\|\rho_{AB}^{\frac{1}{2}}\sigma_B^{-\frac{1}{2}}\|_1^2\right), \tag{A3}$$

where the infimum and the supremum are taken over all states $\sigma_B \in \mathcal{S}(\mathcal{B})$. Their smooth versions are defined as

$$H^{\epsilon}_{\min}(A|B)_{\rho} := \sup_{\hat{\rho}_{AB}} H_{\min}(A|B)_{\hat{\rho}}, \tag{A4}$$

$$H^{\epsilon}_{\max}(A|B)_{\rho} := \inf_{\hat{\rho}_{AB}} H_{\max}(A|B)_{\hat{\rho}}, \tag{A5}$$

where the supremum and infimum are over all operators $\hat{\rho}_{AB} \in \mathcal{P}(\mathcal{H}_{AB})$ in a ϵ ball (in the purified distance) centered in ρ_{AB}. Moreover if A is classical, the optimization can be restricted to an ϵ ball in $\mathcal{S}(\mathcal{H}_{AB})$.

c. Markov condition

The technique we are going to use for the security analysis of our DICKA protocol strongly relies on the fact that some variables satisfy the so-called *Markov condition*.

Definition 4 (Markov condition). Let ρ_{ABC} be a state in $\mathcal{S}(\mathcal{H}_{ABC})$. We say that ρ_{ABC} satisfies the Markov condition $A \leftrightarrow B \leftrightarrow C$ if and only if

$$I(A:C|B)_{\rho} = 0, \tag{A6}$$

where $I(A:C|B)_{\rho}$ is the mutual information between A and C conditioned on B for the state ρ_{ABC}.

This condition becomes trivial when A,B and C are independent random variables. For more details on the definition of the *Markov condition* see [32], Sec. 2.2 and Appendix C.

d. The entropy accumulation theorem

The security proof of our DICKA protocol makes use of a very powerful tool called the entropy accumulation theorem, recently introduced in [32]. The EAT relates the smooth min-(max-) entropy of N subsystems to the Von Neumann entropy of each subsystem. In this section we recall some necessary definitions from [32] and state the EAT.

The entropy accumulation theorem applies to states of the form,

$$\rho_{C_1^n A_1^n B_1^n E} := (\mathrm{tr}_{R_n} \circ \mathcal{M}_n \circ \ldots \circ \mathcal{M}_1 \otimes \mathbb{1}_E)(\rho_{R_0 E}), \tag{A7}$$

for some initial state $\rho_{R_0E} \in \mathcal{S}(\mathcal{H}_{R_0E})$ and $\forall i \in [n]$, $\mathcal{M}_i$ is an *EAT channel* defined as follows.

Definition 5 (EAT channels (from [11])). For $i \in [n]$ we call $\mathcal{M}_i$ an *EAT channel* if $\mathcal{M}_i$ is a CPTP map from R_{n-1} to $C_iA_iB_iR_i$ such that $\forall i \in [n]$.

(1) A_i,B_i,C_i are finite dimensional systems, C_i is classical, and R_i is an arbitrary quantum system.

(2) For any state $\sigma_{R_{i-1}R}$, where R is isomorphic to R_{i-1}, the output state $\sigma_{R_iA_iB_iC_iR} := (\mathcal{M}_i \otimes \mathbb{1}_R)\sigma_{R_{i-1}R}$ is such that the classical register C_i can be measured from $\sigma_{A_iB_i}$.

(3) Any state defined as in (A7) satisfies the following Markov conditions,

$$\forall i \in [n],\ A_1^{i-1} \leftrightarrow B_1^{i-1}E \leftrightarrow B_i. \tag{A8}$$

To state the EAT we also need the notion of min- and max-tradeoff functions. Let $\mathbb{P}(\mathcal{C})$ be the set of distributions on the alphabet $\mathcal{C}$ of C_i. For any $q \in \mathbb{P}(\mathcal{C})$ we define the set of states,

$$\Sigma_i(q) := \{\sigma_{C_iA_iB_iR_iR} = (\mathcal{M}_i \otimes \mathbb{1}_R)(\sigma_{R_{i-1}R}) : \sigma_{R_{i-1}R} \in \mathcal{S}(\mathcal{H}_{R_{i-1}R}) \quad \sigma_{C_i} = q\}. \tag{A9}$$

Definition 6. A real function f on $\mathbb{P}(\mathcal{C})$ is called a *min-tradeoff* function for a map $\mathcal{M}_i$ if

$$f_i(q) \leqslant \inf_{\sigma \in \Sigma_i(q)} H(A_i|B_iR)_{\sigma}, \tag{A10}$$

and *max-tradeoff* function for a map $\mathcal{M}_i$ if

$$f_i(q) \geqslant \sup_{\sigma \in \Sigma_i(q)} H(A_i|B_iR)_{\sigma}. \tag{A11}$$

If $\Sigma_i(q) = \varnothing$, the infimum is taken to be $+\infty$ and the supremum $-\infty$.

We can now state the EAT.

Theorem 3 (EAT from [32], Theorem 4.4).

Let $\mathcal{M}_1,\ldots,\mathcal{M}_n$ be an EAT channel and $\rho_{C_1^nA_1^nB_1^nE}$ be a state as defined in (A7), let $h \in \mathbb{R}$, f be an affine min-tradeoff function for all the maps $\mathcal{M}_i$,$i \in [n]$, and $\epsilon \in]0,1[$. For any event $\Omega \subset \mathcal{C}^n$ such that $f(\mathsf{freq}(C_1^n)) \geqslant h$,

$$H^{\epsilon}_{\min}\big(A_1^n\big|B_1^nE\big)_{\rho_{|\Omega}} \geqslant nh - v\sqrt{n}, \tag{A12}$$

where $v = 2(\log_2(1+2d_A) + \lceil \|\nabla f\|_{\infty} \rceil)\sqrt{1 - 2\log_2(\epsilon \cdot p_{\Omega})}$, where d_A is the maximum dimension of the system A_i. On the other hand we have

$$H^{\epsilon}_{\max}\big(A_1^n\big|B_1^nE\big)_{\rho_{|\Omega}} \leqslant n\tilde{h} + v\sqrt{n}, \tag{A13}$$

where we replace f by an affine *max-tradeoff function* $\tilde{f}$, such that the event Ω implies $\tilde{h} \geqslant \tilde{f}(\mathrm{freq}(C_1^n))$.

e. Device-independent assumptions

When dealing with cryptographic tasks it is important to be precise under which assumptions a protocol is proven secure. If an assumption is not satisfied in a particular implementation, the entire security of the protocol may be compromised. The device-independent framework allows one to relax many strong assumptions about the underlying system and devices, however, some assumptions (without which we can probably not achieve any security) are still present and it is important to make them explicit. In the following we state the assumptions present in our model, which constitutes the standard set of assumptions made in all device independent protocols. This minimal set of assumptions is crucial for security in the device-independent framework, as a relaxation of any of them compromises the security of the protocol.

Assumptions 1. Our DICKA protocol considers N parties, namely Alice, Bob$_1$, ..., Bob$_{N-1}$, and the eavesdropper, Eve. They satisfy the following assumptions.

(1) Each party is in a laboratory which is isolated from the outside (in particular from Eve). As a consequence no nonintended information can go in or out of the labs.

(2) Each party holds a trusted random number generator (RNG).

(3) All classical communications between the parties are assumed to be authenticated, and all classical operations are assumed to be trusted.

(4) Each party has a measurement device in their laboratory in which they can input classical information and which outputs 0 or 1. The measurement devices are otherwise arbitrary, and therefore could be prepared by Eve.

(5) Alice has a source that produces some N-partite quantum state $\rho_{A_iB_{(1\ldots N-1),i}}$ in the round i. We allow Eve to hold

the purification of $\rho_{A_1^n B_{(1\ldots N-1)}{}_1^n}$ (the state between Alice and the Bobs for the n rounds of the protocol) and we denote the pure global state $\rho_{A_1^n B_{(1\ldots N-1)}{}_1^n E}$. This source is also assumed to be arbitrary, and therefore we can assume that it is prepared by Eve.

(6) We will assume that Alice's source and her measurement device are independent (*e.g.,* Alice can isolate the source from the measurement device). Therefore there is no nonintended communication between the source and her measurement device.

Point 6 of Assumption 1 is usually not explicitly stated in previous works on device-independent QKD, however, we remark that this assumption is also present in all previous protocols. Indeed Assumption 6 is important to guarantee that no extra information about the outcomes of Alice's device is leaked to Eve (since Alice and Bob are in isolated labs), apart from what she can learn from the purifying system in her possession and the classical communication intentionally leaked during the protocol. Previous protocols usually assume that an external source is responsible for producing the states. However, note that in order to distribute the states to Alice and Bob's devices one needs a quantum channel connecting the external source with their labs, and similarly it is assumed that no information from the devices is leaked through this quantum channel. An alternative approach is to assume that the full state for the n rounds of the protocol is already shared between the two parties at the very beginning of the protocol (and any quantum channel connecting the source and the devices is disconnected once the protocol starts).

However, this is an unrealistic assumption, since an implementation of such a protocol would require quantum memory to last for the entire duration of the protocol. For that reason, here we chose NOT to assume that the state is already shared among all the parties, and Assumption 6 prevents the simple attack described in [36], Appendix C, where the outcome of round i is leaked throughout the state transmitted to Bob in the next rounds.

2. From self-testing to device-independent conference key agreement

The Clauser-Horne-Shimony-Holt inequality [27] has been successfully used to prove security of DIQKD [11] in the most adversarial scenario, where only a minimal set of assumptions (similar to Assumption 1) is required. The main point of using the CHSH inequality for cryptographic protocols is due to its self-testing properties, which allows one to derive properties about the devices used during the protocol. Therefore, in order to prove the security of DICKA it is very natural to think of an N-partite XOR game (or an equivalent Bell inequality) to self-test the N parties. It has recently been proven that the family of Mermin-Ardehali-Belinskii-Klyshko inequalities can self-test devices with a rigidity statement for the maximal violation [26] of the inequalities. This family of inequalities are a simple generalization of the well-known bipartite CHSH inequality to N parties. In this section we first relate the MABK inequalities to the CHSH inequality, and then we use this result to prove security of a DICKA protocol.

a. Mermin-Ardehali-Belinskii-Klyshko inequalities vs Clauser-Horne-Shimony-Holt inequality

MABK inequalities [20–22] are Bell inequalities for N parties ($N \geqslant 2$) that reduce to the CHSH inequality for $N = 2$. In this section we will show that for any $N > 2$ it is possible to reinterpret an N-partite MABK inequality as a CHSH inequality. More precisely if N parties, say Alice and $N-1$ Bobs, are involved in an N-partite MABK experiment, we can reinterpret this experiment as a bipartite CHSH experiment between Alice on one side and all the Bobs together on the other side. Before we formalize this argument, we will recall the definitions of CHSH and MABK inequalities. We first define the CHSH function F_{CHSH} that takes four operators A_0, A_1, B_0, B_1 as

$$F_{\mathrm{CHSH}}(A_0,A_1,B_0,B_1) := A_0 \otimes B_0 + A_0 \otimes B_1 + A_1 \otimes B_0 - A_1 \otimes B_1. \quad \text{(A14)}$$

This allows us to define the CHSH inequality.

Definition 7 (CHSH inequality). Let $A_0, A_1 \in \mathcal{S}_a(\mathcal{H}_A)$ be the binary observables corresponding to the two measurements applied by Alice during the CHSH experiment, and $B_0, B_1 \in \mathcal{S}_a(\mathcal{H}_B)$ the ones that Bob applies. Therefore we have $A_{0,1}^2 \leqslant \mathbb{1}_A$ and $B_{0,1}^2 \leqslant \mathbb{1}_B$. The CHSH inequality can be written as

$$S_2 := |\mathrm{tr}(F_{\mathrm{CHSH}}(A_0,A_1,B_0,B_1)\rho_{AB})| \leqslant 2, \quad \text{(A15)}$$

where S_2 is called the CHSH value and $\rho_{AB} \in \mathcal{S}(\mathcal{H}_{AB})$ is the state that Alice and Bob share.

Note that if Alice and Bob violate the CHSH inequality, meaning that $S_2 > 2$, then Alice and Bob use a nonclassical strategy i.e., ρ_{AB} is an entangled state, A_0 does not commute with A_1, and similarly B_0 does not commute with B_1.

One way to generalize the CHSH inequality to inequalities between N parties, say $\mathrm{Paul}_1, \ldots, \mathrm{Paul}_N$, is the following.

Definition 8 (MABK inequality). Let $P_0^i, P_1^i \in \mathcal{S}_a(\mathcal{H}_{\mathcal{P}^i})$ be the two binary observables ($\forall k \in \{0,1\}$, ${P_k^i}^\dagger = P_k^i$ & ${P_k^i}^2 \leqslant \mathbb{1}_{\mathcal{P}_i}$) for Paul_i, $\forall i \in [n]$. Then the N-partite MABK operator MK_N is defined by recursion as follows,

$$MK_2 := \tfrac{1}{2} F_{\mathrm{CHSH}}\big(P_0^1, P_1^1, P_0^2, P_1^2\big), \quad \text{(A16)}$$

$$MK_N := \tfrac{1}{2} F_{\mathrm{CHSH}}\big(MK_{N-1}, \overline{MK}_{N-1}, P_0^N, P_1^N\big). \quad \text{(A17)}$$

The N-partite MABK inequalities are then defined as

$$\forall N \geqslant 2,\ \mathcal{MK}_N := |\mathrm{tr}(MK_N\, \rho_{\mathcal{P}_{(1\ldots N)}})| \leqslant 2^{\frac{m-1}{2}},\ m \in [N], \quad \text{(A18)}$$

where $\mathcal{MK}_N$ is called the MABK value, $\overline{MK}_l, l \geqslant 2$ is the operator obtained from MK_l by replacing P_k^i by P_{1-k}^i, $\forall i \in [l], \forall k \in \{0,1\}$, and m is the largest number of parties that are entangled in the N-partite state $\rho_{\mathcal{P}_{(1\ldots N)}}$.

The MABK inequalities are such that a violation of the inequalities for $m = 1$ proves that at least two parties are entangled: The violation of the inequalities for $m = N-1$ proves genuine N-partite entanglement, and the case where $m = N$ gives an upper bound (tight) on what is achievable by quantum mechanics.

In order to show the reinterpretation of a MABK experiment into a CHSH experiment, we will define a rescaled version of the operator MK_N, namely $R_N := 2^{-\frac{N-1}{2}} MK_N$. One can show using the recursion relation (A17) that

$$\forall N \geqslant 3,\ R_N = \frac{1}{2\sqrt{2}} R_{N-1} \otimes \left(P_0^N + P_1^N\right) + \frac{1}{2\sqrt{2}} \overline{R}_{N-1} \otimes \left(P_0^N - P_1^N\right). \tag{A19}$$

One can also check that $\forall\, l \in [N]$, R_l are Hermitian operators and $R_l{}^2 \leqslant \mathbb{1}_{\mathcal{P}_{(1\ldots l)}}$.

Let us now consider a MABK experiment where the N parties $\text{Paul}_1,\ldots,\text{Paul}_N$ violate the MABK inequality for $m = N-1$, namely they achieve $\mathcal{MK}_N > 2^{\frac{N-2}{2}}$. We will show in the following lemma that this can be interpreted as a CHSH experiment between Alice and Bob, Alice being $\text{Paul}_1,\ldots,\text{Paul}_{N-1}$ and Bob being Paul_N, where they achieve a CHSH value of $S_2 = 2\sqrt{2} \times 2^{-\frac{N-1}{2}} \mathcal{MK}_N$.

Lemma 1. A MABK experiment between N parties achieving a MABK value $\mathcal{MK}_N > 2^{\frac{N-2}{2}}$ can be seen as a CHSH experiment between any of the N parties on one side and the $N-1$ other parties on the other side achieving a CHSH value of $S_2 = 2^{-\frac{N-1}{2}+\frac{3}{2}} \times \mathcal{MK}_N > 2$.

Proof. Let us write the MABK value for the MABK experiment,

$$\mathcal{MK}_N := |\mathrm{tr}(MK_N\, \rho_{\mathcal{P}_{(1\ldots N)}})| \tag{A20}$$

$$= |\mathrm{tr}(2^{\frac{N-1}{2}} R_N\, \rho_{\mathcal{P}_{(1\ldots N)}})| \tag{A21}$$

$$= 2^{\frac{N-1}{2}-\frac{3}{2}} \left|\mathrm{tr}\left(\left\{R_{N-1} \otimes \left(P_0^N + P_1^N\right) + \overline{R}_{N-1} \otimes \left(P_0^N - P_1^N\right)\right\} \rho_{\mathcal{P}_{(1\ldots N)}}\right)\right|, \tag{A22}$$

where we used in the first equality the definition of R_N and for the second equality the recursion relation (A19). Let us call $A_0^{\mathcal{P}_{(1\ldots N-1)}} := R_{N-1}$, $A_1^{\mathcal{P}_{(1\ldots N-1)}} := \overline{R}_{N-1}$, $B_0^{\mathcal{P}_{(N)}} := P_0^N$, and $B_1^{\mathcal{P}_{(N)}} := P_1^N$. Plugging it into Eq. (A22) gives us

$$\mathcal{MK}_N = 2^{\frac{N-1}{2}-\frac{3}{2}} \left|\mathrm{tr}\left(\left\{A_0^{\mathcal{P}_{(1\ldots N-1)}} \otimes B_0^{\mathcal{P}_{(N)}} + A_0^{\mathcal{P}_{(1\ldots N-1)}} \otimes B_1^{\mathcal{P}_{(N)}} + A_1^{\mathcal{P}_{(1\ldots N-1)}} \otimes B_0^{\mathcal{P}_{(N)}} - A_1^{\mathcal{P}_{(1\ldots N-1)}} \otimes B_1^{\mathcal{P}_{(N)}}\right\} \rho_{\mathcal{P}_{(1\ldots N-1)},\mathcal{P}_{(N)}}\right)\right| \tag{A23}$$

$$= 2^{\frac{N-1}{2}-\frac{3}{2}} \times \left|\mathrm{tr}\left(F_{\text{CHSH}}\left(A_0^{\mathcal{P}_{(1\ldots N-1)}}, A_1^{\mathcal{P}_{(1\ldots N-1)}}, B_0^{\mathcal{P}_{(N)}}, B_1^{\mathcal{P}_{(N)}}\right) \times \rho_{\mathcal{P}_{(1\ldots N-1)},\mathcal{P}_{(N)}}\right)\right| \tag{A24}$$

$$=: 2^{\frac{N-1}{2}-\frac{3}{2}} \times S_2, \text{ by definition of } S_2, \tag{A25}$$

where $F_{\text{CHSH}}(A_0^{\mathcal{P}_{(1\ldots N-1)}}, A_1^{\mathcal{P}_{(1\ldots N-1)}}, B_0^{\mathcal{P}_{(N)}}, B_1^{\mathcal{P}_{(N)}})$ is the CHSH operator between the parties $\{\text{Paul}_1,\ldots,\text{Paul}_{N-1}\}$ together and Paul_N. Note that here we have split the N parties into Paul_N on one side and $\{\text{Paul}_1,\ldots,\text{Paul}_{N-1}\}$ on the other side, but by symmetry of the MABK inequality we can exchange Paul_N with any $\text{Paul}_i, i \in [N-1]$, which proves the statement. ■

Remark 1. Since no bipartite bound entangled state can violate the CHSH inequality [37], Lemma 1 implies that for any finite dimensional N-partite state that permits one to violate the MABK inequalities for $m = N-1$ (see Def. 8), there exists at least N splits of the N parties into two groups given by Lemma 1 such that the bipartite state between these two groups is distillable, which is a similar result as in Refs. [38,39].

To each of the MABK inequalities we can associate an XOR game [40]. Indeed we can write the N-MABK operator as

$$MK_N = 2^{-2\lfloor\frac{N}{2}\rfloor} \sum_{x\in\{0,1\}^N} (-1)^{f(x)} \bigotimes_{i\in[N]} P^i_{x_i},$$

and the MABK value as, (A26)

$$\mathcal{MK}_N = 2^{-2\lfloor\frac{N}{2}\rfloor} \cdot \left|\sum_{x\in\{0,1\}^N} (-1)^{f(x)} \left\langle \bigotimes_{i\in[N]} P^i_{x_i} \right\rangle\right|, \tag{A27}$$

where $x_i \in \{0,1\}$ is the i^{th} bit of x, $f : \{0,1\}^N \mapsto \{0,1,\perp\}$ is a function, and we adopt the convention that $(-1)^{\perp} = 0$. One can note that f can take the value $\perp$ only when N is odd, as a consequence of the fact that for N odd, half of the terms $x \in \{0,1\}^N$ do not appear in the inequality.

We can now define an XOR game between N parties $\text{Paul}_1,\ldots,\text{Paul}_N$, where we ask to all the Paul_k ($k \in [N]$) the question $x_{(k)} \in \{0,1\}$ uniformly at random and independently of the questions $x_{(1,\ldots,k-1,k+1,\ldots,N)}$ asked to the others. Each Paul will reply $a_{(k)} \in \{0,1\}$. They can agree on a strategy (that might be quantum) before the game but they are assumed not to communicate during the game. They win if $w_{\text{MABK}}(a_{(1\ldots N)},x_{(1\ldots N)}) = 1$, where $w_{\text{MABK}}(a_{(1\ldots N)},x_{(1\ldots N)})$ is the function $\{0,1\}^{2N} \mapsto \{0,1\}$ defined as

$$w_{\text{MABK}}(a_{(1\ldots N)},x_{(1\ldots N)}) = \begin{cases} 1 & \text{if } \bigoplus_i^N a_{(i)} = f(x_{(1\ldots N)}) \\ 0 & \text{otherwise} \end{cases}, \tag{A28}$$

where f is the function defined in the previous equation by the N-partite MABK operator MK_N. Note that when $f(x_{(1\ldots N)}) = \perp$ we always have $w_{\text{MABK}}(a_{(1\ldots N)},x_{(1\ldots N)}) = 0$.

We now relate the probability of winning the N-MABK game to the the N-partite MABK value $\mathcal{MK}_N$.

Lemma 2. Let $\text{Paul}_1,\ldots,\text{Paul}_N$ be N parties playing an N-MABK game with a quantum strategy given by their observables $P_0^1,\ldots P_0^N$ for the question 0, $P_1^1,\ldots,P_1^N$ for the question 1, and the N-partite state $\rho_{\mathcal{P}_{(1\ldots N)}}$. The probability p_w that they win the game is

$$p_w = 2^{2\lfloor\frac{N}{2}\rfloor-N}\left[\frac{1}{2} \pm \frac{2^{-\lfloor\frac{N}{2}\rfloor}\mathcal{MK}_N}{2}\right], \tag{A29}$$

where $\pm$ corresponds to the sign of $\mathrm{tr}(MK_N\rho)$, with MK_N being the MABK operator defined by Pauls' observables and $\mathcal{MK}_N$ being the corresponding MABK value. For $\mathcal{MK}_N \in]2^{\frac{N-2}{2}}, 2^{\frac{N-1}{2}}[$, and when $\mathrm{tr}(MK_N\,\rho) \geqslant 0$, we have $p_w \in]p_{\min}, p_{\max}[$, where $p_{\min} := 2^{2\lfloor N/2\rfloor-N-1} + 2^{\lfloor N/2\rfloor-N/2-2}$ and $p_{\max} := 2^{2\lfloor N/2\rfloor-N-1} + 2^{\lfloor N/2\rfloor-N/2-3/2}$.

Proof. By definition of p_w we have

$$p_w := \sum_{x_{(1\ldots N)}\in\{0,1\}^N} P(x_{(1\ldots N)})P\left(\bigoplus_i a_i = f(x_{(1\ldots N)})\middle| x_{(1\ldots N)}\right). \tag{A30}$$

Here $x_{(1\ldots N)}$ is chosen uniformly at random so $P(x_{(1\ldots N)}) = 2^{-N}$. Also we can split the above sum according to the three possible values that f can take which gives us

$$p_w = 2^{-N} \times \left[\sum_{\substack{x_{(1\ldots N)}:\\ f(x_{(1\ldots N)})=0}} P\left(\bigoplus_i a_i = 0\middle| x_{(1\ldots N)}\right) + \sum_{\substack{x_{(1\ldots N)}:\\ f(x_{(1\ldots N)})=1}} P\left(\bigoplus_i a_i = 1\middle| x_{(1\ldots N)}\right) + \sum_{\substack{x_{(1\ldots N)}:\\ f(x_{(1\ldots N)})=\perp}} \underbrace{P\left(\bigoplus_i a_i = \perp\middle| x_{(1\ldots N)}\right)}_{=0}\right]. \tag{A31}$$

We can rewrite the above conditional probabilities in terms of the average of the observable $P^1_{x_1}\otimes\ldots\otimes P^N_{x_N}$ as

$$P\left(\bigoplus_i a_i = 0\middle| x_{(1\ldots N)}\right) = \frac{1+\langle P^1_{x_1}\otimes\ldots\otimes P^N_{x_N}\rangle}{2} \quad \text{and} \quad P\left(\bigoplus_i a_i = 1\middle| x_{(1\ldots N)}\right) = \frac{1-\langle P^1_{x_1}\otimes\ldots\otimes P^N_{x_N}\rangle}{2}. \tag{A32}$$

Plugging it into Eq. (A31) we get

$$p_w = 2^{-N}\left[\sum_{\substack{x_{(1\ldots N)}:\\ f(x_{(1\ldots N)})=0}} \frac{1+\langle P^1_{x_1}\otimes\ldots\otimes P^N_{x_N}\rangle}{2} + \sum_{\substack{x_{(1\ldots N)}:\\ f(x_{(1\ldots N)})=1}} \frac{1-\langle P^1_{x_1}\otimes\ldots\otimes P^N_{x_N}\rangle}{2}\right] \tag{A33}$$

$$= 2^{-N} \underbrace{\sum_{\substack{x_{(1\ldots N)}:\\ f(x_{(1\ldots N)})\neq\perp}} \frac{1}{2}}_{2^{2\lfloor N/2\rfloor}\cdot 1/2} + \; 2^{-N}\underbrace{\sum_{\substack{x_{(1\ldots N)}:\\ f(x_{(1\ldots N)})\neq\perp}} (-1)^{f(x_{(1\ldots N)})}\cdot\frac{\langle P^1_{x_1}\otimes\ldots\otimes P^N_{x_N}\rangle}{2}}_{=\pm 2^{\lfloor N/2\rfloor}\mathcal{MK}_N/2,\ \text{see eq. (A27)}} \tag{A34}$$

$$= 2^{2\lfloor\frac{N}{2}\rfloor - N}\left[\frac{1}{2} \pm \frac{2^{-\lfloor\frac{N}{2}\rfloor}\mathcal{MK}_N}{2}\right]. \tag{A35}$$

In the second line we have $\sum_{\substack{x:\\ f(x)\neq\perp}} \frac{1}{2} = \frac{1}{2}2^{2\lfloor N/2\rfloor}$ because when N is odd only half of the term $x\in\{0,1\}^N$ are present in the inequality. ∎

b. Device-independent conference key agreement

We now present a DICKA protocol and prove its security in two steps. We first use the recently developed entropy accumulation theorem [32] to split the overall entropy of Alice's string produced during the protocol, into a sum of entropy produced on each round of the protocol. Then we use the relation between the MABK inequalities and the CHSH inequality, derived in the previous section, to bound the entropy produced in one round by a function of the violation of the N-partite MABK inequality, which generalize the bounds found for the bipartite case in [6].

The protocol

Before we describe our DICKA protocol let us first state the security definitions for DICKA. We follow the definitions given in [11] for DIQKD and generalize it to the multipartite case.

Definition 9. (Correctness) We will call a DICKA protocol $\epsilon_{\rm corr}$-correct for an implementation, if Alice's and Bobs' keys, $K_A, K_{B_{(1)}},\ldots,K_{B_{(N-1)}}$, are all identical with probability at least $1-\epsilon_{\rm corr}$.

Definition 10. (Secrecy) We say that a DICKA protocol is $\epsilon_{\rm sec}$-secret for an implementation, if conditioned on not aborting Alice's key K_A is $\epsilon_{\rm sec}$-close to a key that Eve is ignorant about. More formally for a key of length l, we want

$$p_{\hat\Omega}\left\|\rho_{K_AE|\hat\Omega} - \frac{\mathbb{1}_A}{2^l}\otimes\rho_{E|\hat\Omega}\right\|_{\rm tr} \leqslant \epsilon_{\rm sec},$$

where $\hat\Omega$ is the event of the protocol not aborting, and $p_{\hat\Omega}$ is the probability for $\hat\Omega$.

Note that if a protocol is $\epsilon_{\rm corr}$-correct and $\epsilon_{\rm sec}$-secret then it is ϵ^s-correct-and-secret for $\epsilon^s \geqslant \epsilon_{\rm corr}+\epsilon_{\rm sec}$.

Definition 11 (Security). A DICKA protocol is called $(\epsilon^s,\epsilon^c,l)$-secure if the following.

(1) (Soundness) For any implementation of the protocol, either it aborts with probability greater than $1-\epsilon^s$ or it is ϵ^s-correct-and-secret.

(2) (Completeness) There exists an honest implementation of the protocol such that the probability of aborting the protocol is less than ϵ^c, that is, $1 - p_{\hat{\Omega}} \leqslant \epsilon^c$.

We remark again that Definition 11 was proven to be a criteria for composable security for quantum key distribution in the device-dependent scenario [29]. However, for the device-independent case it is not known whether such a criteria is enough for composable security. Indeed, Ref. [31] suggests that this is not the case if the same devices are used for generation of a subsequent key since this new key can leak information about the first key. Following Ref. [11] we chose to adopt Definition 11 as the security criteria for DICKA.

We now prove that the DICKA protocol presented in the main text, under Assumption 1, satisfies the above definitions of security. For completeness we restate the protocol here.

Protocol 2 (More detailed version of Protocol 1): *The protocol runs as follows for N parties.*

1. *For every round $i \in [n]$ do*
   (a) *Alice uses her source to produce and distribute an N-partite state, $\rho_{A_i B_{(1\ldots N-1),i}}$, shared among herself and the $N-1$ Bobs.*
   (b) *Alice randomly picks T_i, s.t. $P(T_i = 1) = \mu$, and publicly communicates it to all the Bobs.*
   (c) *If $T_i = 0$ Alice and the Bobs choose $(X_i, Y_{(1\ldots N-1),i}) = (0,2,\ldots,2)$, and if $T_i = 1$ they all choose $X_i, Y_{(1\ldots N-1),i} \in_R \{0,1\}$ uniformly at random.*
   (d) *Alice and the Bobs input the value they chose previously in their respective device and record the output as $A'_i, B'_{(1\ldots N-1),i}$.*
2. *They all communicate publicly the list of bases $X_1^n Y_{(1\ldots N-1)}{}_1^n$ they used.*
3. ***Error correction:** Alice and the Bobs apply an error correction protocol. Here we chose a protocol based on universal hashing [41,42]. If the error correction protocol aborts for at least one Bob then they abort the protocol. If it does not abort they obtain the raw keys $\tilde{K}_A, \tilde{K}_{B_{(1\ldots N-1)}}$. We call O_A the classical information that Alice has sent to the Bobs during the error correction protocol. Also the Bobs will send some error correction information but only for the bits produced during the testing rounds ($T_i = 1$), for the purpose of parameter estimation. We call Alice's guess on Bobs' strings $G_{(1\ldots N-1)}$, and we denote $O_{(k)}$ the error correction information sent by* Bob$_k$.
4. ***Parameter estimation:** For all the rounds i such that $T_i = 1$, Alice uses A'_i and her guess on $B'_{(1\ldots N-1),i}$ to set $C_i = 1$ if they have won the N-partite MABK game in the round i, she sets $C_i = 0$ if they have lost it, and finally she sets $C_i = \perp$ for the rounds i where $T_i = 0$. She aborts if $\sum_i C_i < \delta \cdot \sum_i T_i$, where $\delta \in]p_{\min}, p_{\max}[$.*
5. ***Privacy amplification:** Alice and the Bobs apply a privacy amplification protocol (namely the universal hashing described in [43]) to create final keys $K_A, K_{B_{(1\ldots N-1)}}$. We call S the classical information that Alice sent to the Bobs during the privacy amplification protocol.*

Note that the above Protocol 2 is very similar to the DIQKD protocol given in [11], the difference being that since N parties are present here we use a shared N-partite GHZ state, instead of EPR pairs, and we have to add error corrections. Indeed we have an error correction protocol that permits all the parties to get the same raw key. But since we have N parties involved in the protocol, at least one of the parties needs to know all the other parties' outputs for the testing rounds (when $T_i = 1$) in order to estimate, in the parameter estimation phase, how many times they succeed in the MABK game. For simplicity of the analysis we choose, in Protocol 2, to communicate this information through error correction protocols.

In the ideal scenario (when there is no noise and no interference of Eve) the state $\rho_{A_1^n B_{(1\ldots N-1)}{}_1^n}$ produced corresponds to n copies of the N-partite GHZ state, N-GHZ state, distributed across the N parties, and Alice and the Bobs measure the following observables.

(1) Alice's observable for $X_i = 0$ is σ_z and for $X_i = 1$ it is σ_x.

(2) For the Bobs, they have the observable σ_z for $Y_{(k),i} = 2$, and for $Y_{(k),i} \in \{0,1\}$ they have observables that are defined by a strategy that maximally violates the N-MABK inequality when the measurements are performed on a N-GHZ state [21]. In particular, for each party the observable for $Y_{(k),i} = 0$ and the one for $Y_{(k),i} = 1$ must be maximally incompatible [26].

In the next sections we are going to present the detailed proof of the following main result.

Theorem 4. Let $\epsilon_{\mathrm{EC}}, \epsilon'_{\mathrm{EC}} \in]0,1[$ be the two error parameters of the error correction protocol as described in Sec. 2 b, $\epsilon_{\mathrm{PA}} \in]0,1[$ be the privacy amplification error probability, $\epsilon_{\mathrm{EA}} \in]0,1[$ be a chosen security parameter for Protocol 2, and $\epsilon \in]0,1[$ be a smoothing parameter. Protocol 2 is $(\epsilon^s, \epsilon^c, l)$-secure according to Definition 11, with $\epsilon^s \leqslant \epsilon_{\mathrm{PA}} + 2(N-1)\epsilon'_{\mathrm{EC}} + 2\epsilon + \epsilon_{\mathrm{EA}}, \epsilon^c \leqslant (N-1)(2\epsilon_{\mathrm{EC}} + \epsilon'_{\mathrm{EC}}) + (1 - \mu(1 - \exp[-2(p_{\exp} - \delta)^2]))^n$, and

$$
\begin{aligned}
l = {} & \max_{p_{\min} \leqslant \frac{p_{\mathrm{opt}}}{\mu} \leqslant p_{\max}} ((f(\hat{q}, p_{\mathrm{opt}}) - \mu)n - \tilde{v}\sqrt{n}) \\
& + 3\log_2(1 - \sqrt{1 - (\epsilon/4)^2}) - 2\log_2\left(\epsilon_{\mathrm{PA}}^{-1}\right) - \mathrm{leak}_{\mathrm{EC}}(O_A) \\
& - \sum_{k=1}^{N-1} \mathrm{leak}_{\mathrm{EC}}(O_{(k)}), \qquad \text{(A36)}
\end{aligned}
$$

where $\tilde{v} = 2(\log_2(13) + (\hat{f}'(p_{\rm opt})+1))\sqrt{1-2\log_2(\epsilon\epsilon_{\rm EA})} + 2\log_2(7)\sqrt{-\log_2(\epsilon_{\rm EA}^2(1-\sqrt{1-(\epsilon/4)^2}))}$, $p_{\rm opt} \in]\mu p_{\min},\mu p_{\max}[$ ($p_{\min},p_{\max}$ are defined in Lemma 2) is a parameter to be optimized: More precisely $p_{\rm opt}$ is the unique point where the tangent function $f(\cdot,p_{\rm opt})$ to the function $\hat{f}(\cdot)$ (see Lemma 5) is such that $f(p_{\rm opt},p_{\rm opt}) = \hat{f}(p_{\rm opt})$ [by convexity of $\hat{f}$ we have $\forall x \in [0,1]$ $f(x,p_{\rm opt}) \leqslant \hat{f}(x)$]. Finally $p_{\rm exp}$ is the expected winning probability to win a single round of the MABK game for an honest implementation, $\delta \in]p_{\min},p_{\max}[$ is the threshold defined in Protocol 2, and $\hat{q}$ is the vector $(\mu\delta,\mu-\mu\delta,1-\mu)^t$.

Correctness

The correctness of Protocol 2 comes from the first part of the error correction protocol used by the parties, where Alice sends information to the Bobs so that they generate the raw keys $\tilde{K}_A,\tilde{K}_{B_{(1\dots N-1)}}$. We want here an error correction protocol that uses only communication from Alice to the Bobs and that minimizes the amount of communication needed. Therefore we are going to use an error correction protocol as the one described in [41,42]. The idea of this error correction code is that Alice chooses a hash function and sends to the Bobs the chosen function and the hashed value of her bits. We denote this communication O_A. Then each Bob$_k$ will individually use O_A and his own prior knowledge $B_{(k)}{}_1^n X_A{}_1^n Y_{(1\dots N-1)}{}_1^n T_1^n$ to guess Alice's string. Each of the Bobs can fail to produce a guess, so if one of them fails the protocol aborts. In an honest implementation of the protocol, the probability that one particular Bob, say Bob$_k$ ($k \in [N-1]$), aborts is upper bounded by $\epsilon_{\rm EC}$. Therefore the probability that at least one of them aborts in an honest implementation is at most $(N-1)\epsilon_{\rm EC}$. If for $k \in [N-1]$ Bob$_k$ does not abort we then have $P(\tilde{K}_A \neq \tilde{K}_{B_{(k)}}) \leqslant \epsilon'_{\rm EC}$. Therefore if none of the Bobs abort we have

$$
\begin{aligned}
&P(\tilde{K}_A = \tilde{K}_{B_{(1)}} = \dots = \tilde{K}_{B_{(N-1)}})\\
&\quad = 1 - P(\tilde{K}_A \neq \tilde{K}_{B_{(1)}} \text{ OR} \dots \text{OR } \tilde{K}_A \neq \tilde{K}_{B_{(N-1)}})\\
&\quad \geqslant 1-(N-1)\epsilon'_{\rm EC} \geqslant 1-\epsilon_{\rm corr},
\end{aligned}
$$

where we take $\epsilon_{\rm corr} \geqslant (N-1)\epsilon'_{\rm EC}$, which proves the following lemma.

Lemma 3. The Protocol 2 is $\epsilon_{\rm corr}$-correct, for any $\epsilon_{\rm corr} \geqslant (N-1)\epsilon'_{\rm EC}$, where $\epsilon'_{\rm EC}$ is such that if $\forall k \in [N-1]$ Bob$_k$ does not abort the error correction protocol then $P(\tilde{K}_A \neq \tilde{K}_{B_{(k)}}) \leqslant \epsilon'_{\rm EC}$.

Completeness

We call an *honest implementation* of the protocol, an implementation where the measurement devices used act in the same way in all the rounds of the protocol, the state used for the n rounds is of the form $\rho_{AB_{(1\dots N-1)}}^{\otimes n}$ [the measurements and the state are then said to be identically and independently distributed (i.i.d.)], and such that for one single round, the probability of winning the N-partite MABK game is $p_{\rm exp} \in]p_{\min},p_{\max}]$.

Lemma 4. For any parameter $\delta \in]p_{\min},p_{\max}[$, Protocol 2 is ϵ^c-complete, for

$$
\begin{aligned}
\epsilon^c &\leqslant (N-1)(2\epsilon_{\rm EC}+\epsilon'_{\rm EC})\\
&\quad + (1-\mu(1-\exp[-2(p_{\rm exp}-\delta)^2]))^n, \qquad \text{(A37)}
\end{aligned}
$$

where $p_{\rm exp} > \delta$, δ is a threshold and $p_{\min},p_{\max}$ are defined in Lemma 2.

Proof. Protocol 2 can abort at two moments: It can abort during the error correction or during the parameter estimation. For the error correction step, the protocol aborts if one of the Bobs aborts while trying to guess Alice's string, or if Alice aborts while guessing Bobs' testing bits. We are assuming that the Bobs use the same error correction protocol in order to send information about their outputs in the test rounds so that Alice can make her guess. Therefore the overall probability of aborting during the error correction protocol is then bounded by $2(N-1)\epsilon_{\rm EC}$ for an honest implementation. The probability of aborting during the parameter estimation part (conditioned on not aborting the error correction step) is given by

$$
\begin{aligned}
P_{\rm PE}(\text{abort}) &= P(G_{(1\dots N-1)}\text{ is correct})P\left(\sum_i C_i < \delta\sum_i T_i \middle| G_{(1\dots N-1)}\text{ is correct}\right)\\
&\quad + P(\exists k: G_{(k)}\text{is wrong})P\left(\sum_i C_i < \delta\sum_i T_i \middle| \exists k: G_{(k)}\text{ is wrong}\right), \qquad \text{(A38)}
\end{aligned}
$$

where $G_{(k)}$ is Alice's guess for Bob$_k$'s testing round bits. It is said to be correct when the string $G_{(k)} = B'_{(k),I}$ for $I := \{i \in [n] : T_i = 1\}$. By bounding $P(G_{(1\dots N-1)}$ is correct) by 1, $P(\exists k : G_{(k)}$is wrong) by $(N-1)\epsilon'_{\rm EC}$, and $P(\sum_i C_i < \delta\cdot\sum_i T_i|\exists k : G_{(k)}$is wrong) by 1, we get

$$
P_{\rm PE}(\text{abort}) \leqslant \sum_{j=0}^n P\left(\sum_i T_i = j\right)P\left(\sum_i C_i < \delta j \middle| \sum_i T_i = j \;\&\; \forall k \tilde{K}_A = \tilde{K}_{B_{(k)}}\right) + (N-1)\epsilon'_{\rm EC}. \qquad \text{(A39)}
$$

Let us consider an honest implementation such that $p_{\rm exp} > \delta$; we can then rewrite (A39) as

$$
P_{\rm PE}(\text{abort}) \leqslant \sum_{j=0}^n P\left(\sum_i T_i = j\right)P\left(\sum_i C_i < (p_{\rm exp}-(p_{\rm exp}-\delta))j \middle| \sum_i T_i = j\; G_{(1\dots N-1)}\text{ is correct}\right) + (N-1)\epsilon'_{\rm EC}. \qquad \text{(A40)}
$$

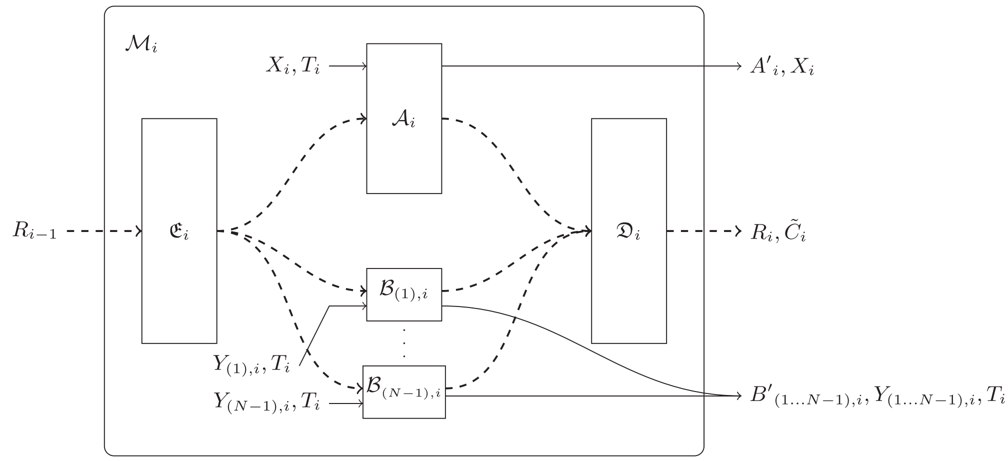

FIG. 2. Description of the map $\mathcal{M}_i$. This map describes the round i of the first step of Protocol 2. T_i is chosen at random such that $P(T_i = 1) = \mu$. $X_i \in \{0,1\}$ represents the "basis" in which Alice's device, represented by the CPTP map $\mathcal{A}_i$, measures its input to get the output $A'_i \in \{0,1\}$. $X_i = 0$ when $T_i = 0$ and $X_i \in_R \{0,1\}$ otherwise. $Y_{(k),i} \in \{0,1,2\}$ represents the "basis" in which Bob$_k$'s device, represented by the CPTP map $\mathcal{B}_{(k),i}$, measures its input to get the output $B'_{(k),i} \in \{0,1\}$. If $T_i = 0$ we have $Y_{(k),i} = 2$, else we have $Y_{(k),i} \in_R \{0,1\}$. If $T_i = 0$ then $\tilde{C}_i = \perp$, else $\tilde{C}_i = w_{\text{MABK}}(A'_i, B'_{(1\ldots N-1),i}, X_i, Y_{(1\ldots N-1),i})$.

Note that the expectation value $\mathbb{E}(C_i) = p_{\text{exp}}$ and because an honest implementation is i.i.d. we can use Hoeffding inequalities to bound $P(\sum_i C_i < (p_{\text{exp}} - (p_{\text{exp}} - \delta))j \,|\, \sum_i T_i = j \;\&\; G_{(1\ldots N-1)} \text{ is correct}) < \exp(-2(p_{\text{exp}} - \delta)^2 j)$. Moreover the i.i.d. random variables T_i follow a Bernoulli distribution with $P(T_i = 1) = \mu$. Plugging all of this into Eq. (A40) gives us

$$P_{\text{PE}}(\text{abort}) \leqslant \sum_{j=0}^{n} \binom{n}{j} (1-\mu)^{n-j} \mu^j \times \exp(-2(p_{\text{exp}} - \delta)^2 j) + (N-1)\epsilon'_{\text{EC}} \tag{A41}$$

$$= \sum_{j=0}^{n} \binom{n}{j} (1-\mu)^{n-j} (\mu \times \exp(-2(p_{\text{exp}} - \delta)^2))^j + (N-1)\epsilon'_{\text{EC}} \tag{A42}$$

$$= (1 - \mu(1 - \exp[-2(p_{\text{exp}} - \delta)^2]))^n + (N-1)\epsilon'_{\text{EC}}, \tag{A43}$$

where the last equality comes from the binomial theorem. ■

Soundness

In order to complete the security proof of Protocol 2, it remains to prove secrecy. Let $\hat{\Omega}'$ be the event that Protocol 2 does not abort and that the error correction step is successful. The Leftover Hashing Lemma [28], Corollary 5.6.1 states that the secrecy of the final key, after a privacy amplification protocol using a family of two-universal hashing functions, depends on the amount of smooth min-entropy of the state before privacy amplification conditioned on the event $\hat{\Omega}'$.

Theorem 5 (Leftover Hashing Lemma [28]): Let $\mathcal{F}$ be a family of two-universal hashing functions from $\{0,1\}^n \to \{0,1\}^l$, such that $F(A_1^n) = K_A$ for $F \in \mathcal{F}$, then it holds that

$$\left\| \rho_{K_A E|\hat{\Omega}'} - \frac{\mathbb{1}_A}{2^l} \otimes \rho_{E|\hat{\Omega}'} \right\|_{\text{tr}} \leqslant 2\epsilon + 2^{-\frac{1}{2}(H_{\min}^{\epsilon}(A_1^n|E)_{\rho_{|\hat{\Omega}'}} - l)}. \tag{A44}$$

According to Theorem 5, in order to prove the secrecy of Protocol 2 we need to lower bound the smooth min-entropy $H_{\min}^{\epsilon}(A'^n_1|X_1^n Y_{(1\ldots N-1)}{}_1^n T_1^n O\, O_{(1\ldots N-1)} E)_{\rho_{|\hat{\Omega}'}}$. The proof goes in the following steps: In Lemma 6, we introduce an error correction map and bound the entropy $H_{\min}^{\epsilon}(A'^n_1|X_1^n Y_{(1\ldots N-1)}{}_1^n T_1^n E)$ for the state after the action of the error correction map, conditioned on the event that a particular violation is observed and the error correction protocol is successful. In Lemma 7, we relate the state generated by Protocol 2 conditioned on the event that the error correction protocols were successful to the state artificially introduced in Lemma 6, and we estimate $H_{\min}^{\epsilon}(A'^n_1|X_1^n Y_{(1\ldots N-1)}{}_1^n T_1^n O\, O_{(1\ldots N-1)} E)$, taking into account the information leaked during the error correction protocol. Finally, in Lemma 8, we combine the previous results proving the soundness of Protocol 2.

To bound the smooth min-entropy we will use the EAT. Indeed, before the error correction part, Protocol 2 can be described by a composition of EAT channels that we will call $\mathcal{M}_1, \ldots, \mathcal{M}_n$ (see Fig. 2).

In order to apply the EAT we need to find a min-tradeoff function for the maps $\mathcal{M}_i$ defined by Fig. 2. i.e., we need to find a function f such that

$$f(q) \leqslant \inf_{\sigma \in \Sigma_i(q)} H(A'_i \tilde{C}_i | X_i Y_{(1\ldots N-1),i} T_i R)_\sigma, \tag{A45}$$

for

$$\Sigma_i(q) := \{\sigma_{\tilde{C}_i A'_i B'_{(1\ldots N-1),i} X_i Y_{(1\ldots N-1),i} T_i R_i R} = (\mathcal{M}_i \otimes \mathbb{1}_R)(\sigma_{R_{i-1}R}) : \sigma_{R_{i-1}R} \in \mathcal{S}(\mathcal{H}_{R_{i-1}R}) \;\&\; \sigma_{\tilde{C}_i} = q\},$$

where $\Sigma_i(q)$ is the set of states that can be generated by the action of the channel $\mathcal{M}_i \otimes \mathbb{1}_R$ on an arbitrary state and such that the classical register $\tilde{C}_i$ has distribution q.

Lemma 5. The real function defined as

$$\hat{f}(x) := \left(1 - \frac{\mu}{2}\right) \left(1 - h\left(\frac{1}{2}+\frac{1}{2}\sqrt{2^{4+2\lfloor \frac{N}{2}\rfloor - N}\left(2^{N-2\lfloor \frac{N}{2}\rfloor}\frac{x}{\mu} - \frac{1}{2}\right)^2 - 1}\right)\right), \tag{A46}$$

is a min-tradeoff function for the EAT channels $\mathcal{M}_i$ defined by Fig. 2. Here μ is the testing probability of Protocol 2, N is the number of parties in Protocol 2, and $h(x)$ is the binary entropy: $h(x) = -x\log_2(x) - (1-x)\log_2(1-x)$.

We define the affine function $f(\,\cdot\,, p_{\rm opt})$ over the probability distribution $\mathbb{P}(\{1,0,\perp\})$ as $\forall\ q = (q(1),q(0),q(\perp))^t \in \mathbb{P}(\{1,0,\perp\})$,

$$f(q,p_{\rm opt}) := \hat{f}'(p_{\rm opt})q(1) + \hat{f}(p_{\rm opt}) - \hat{f}'(p_{\rm opt})p_{\rm opt}, \tag{A47}$$

where $p_{\rm opt} \in]\mu p_{\rm min}, \mu p_{\rm max}[$.

Note that, for simplicity, in the main text we have defined $\hat{f}$ and f as functions of the observed winning probability. In order to make the argument more rigorous and general, here $f(\cdot,p_{\rm opt})$ is a function that takes as input the vector of frequencies $q = (q(1),q(0),q(\perp))^t$.

Proof. Let us take a state $\sigma_{\tilde{C}_i A'_i B'_{(1\ldots N-1),i} X_i Y_{(1\ldots N-1),i} T_i R_i R} \in \Sigma_i(q)$. Then we define the state,

$$\sigma'_{\tilde{C}_i A''_i B''_{(1),i} B'_{(2\ldots N-1),i} X_i Y_{(1\ldots N-1),i} T_i F_i R_i R}, \tag{A48}$$

to be the state we obtain from $\sigma_{\tilde{C}_i A'_i B'_{(1\ldots N-1),i} X_i Y_{(1\ldots N-1),i} T_i R_i R}$ by replacing A'_i by $A''_i := A'_i \oplus F_i$ and $B'_{(1),i}$ by $B''_{(1),i} := B'_{(1),i} \oplus F_i$ where F_i is a bit that is chosen uniformly at random. None of the other registers are changed, in particular, note that we still have $\sigma'_{\tilde{C}_i} = q$, where the value of $\tilde{C}_i$ can be determined by the registers A''_i, $B''_{(1),i}$, and $B'_{(2\ldots N-1),i}$. Moreover, since F_i is completely independent of the other variables and given the definition of A''_i, it is easy to check that

$$H(A'_i\tilde{C}_i|X_iY_{(1\ldots N-1),i}T_iR)_\sigma = H(A''_i\tilde{C}_i|F_iX_iY_{(1\ldots N-1),i}T_iR)_{\sigma'}. \tag{A49}$$

Using the chain rule,

$$H(A''_i\tilde{C}_i|F_iX_iY_{(1\ldots N-1),i}T_iR)_{\sigma'} \geqslant H(A''_i|F_iX_iY_{(1\ldots N-1),i}T_iR)_{\sigma'}, \tag{A50}$$

and since $P(X_i=0) = 1 - \frac{\mu}{2}$,

$$H(A''_i|F_iX_iY_{(1\ldots N-1),i}T_iR)_{\sigma'} \geqslant \left(1-\frac{\mu}{2}\right)H(A''_i|F_iY_{(1\ldots N-1),i}T_iR,X_i=0)_{\sigma'}. \tag{A51}$$

Given that for $X_i = 0$ Alice's measurement is independent of $Y_{(1\ldots N-1),i}$ and T_i we have

$$H(A''_i|F_iY_{(1\ldots N-1),i}T_iR,X_i=0)_{\sigma'} = H(A''_i|F_iR,X_i=0)_{\sigma'}. \tag{A52}$$

Using the definition of the conditional Von Neumann entropy we can write

$$H(A''_i|F_iR,X_i=0)_{\sigma'} = H(A''_iF_iR|X_i=0)_{\sigma'} - H(F_iR|X_i=0)_{\sigma'} \tag{A53}$$

$$= H(A''_i|X_i=0)_{\sigma'} + \underbrace{H(F_iR|X_i=0)_{\sigma'} - H(F_iR|X_i=0)_{\sigma'}}_{=-\chi(A''_i:F_iR|X=0)_{\sigma'}} \tag{A54}$$

$$= 1 - \chi(A''_i : F_iR|X_i=0)_{\sigma'}, \tag{A55}$$

where $\chi(A''_i : F_iR|X_i=0)$ is the Holevo quantity, and the last equality comes from the definition of A''_i being a uniform variable (for any value of X_i).

Since A_i and F_i are independent (even conditioned on X_i), we get

$$\chi(A_i : F_iR|X_i=0)_{\sigma'} = H(F_iR|X_i=0)_{\sigma'} - H(F_iR|A_i,X_i=0)_{\sigma'} \tag{A56}$$

$$= H(R|F_i,X_i=0)_{\sigma'} - H(R|F_iA_i,X_i=0)_{\sigma'} + \underbrace{H(R|F_i,X_i)_{\sigma'} - H(F_i|A_iF_i,X_i=0)_{\sigma'}}_{=0} \tag{A57}$$

$$= \chi(A_i : R|F_i,X_i=0)_{\sigma'}. \tag{A58}$$

For any state leading to a CHSH violation of $S_2 \in [2,2\sqrt{2}]$, Ref. [6], Sec. 2.3 gives a tight upper bound on $\chi(A_i : R|F_i,X_i=0)$:

$$\chi(A_i : R|F_i,X_i=0) \leqslant h\left(\frac{1}{2}+\frac{1}{2}\sqrt{\frac{{S_2}^2}{4}-1}\right), \tag{A59}$$

where $h(x) = -x\log_2(x) - (1-x)\log_2(1-x)$.

However, in our DICKA protocol the parties are using a MABK game to test their devices. But as we have seen, according to Lemma 1, any state leading to a MABK value $\mathcal{MK}_N > 2^{\frac{N-2}{2}}$, can be reinterpreted as a bipartite state leading to a CHSH value of $S_2 = 2^{-\frac{N-1}{2}+\frac{3}{2}}\mathcal{MK}_N > 2$. Plugging it into Eq. (A59) gives

$$\chi(A_i : R|F_i, X_i = 0) \leqslant h\left(\frac{1}{2} + \frac{1}{2}\sqrt{\frac{2^{4-N}\mathcal{MK}_N{}^2}{4} - 1}\right). \tag{A60}$$

Assuming that $\mathrm{tr}(MK_N\,\rho) \geqslant 0$ in Lemma 2, and using the result of this lemma, the MABK value can be rewritten as the probability of winning the MABK game:

$$p_w = 2^{2\lfloor\frac{N}{2}\rfloor - N}\left[\frac{1}{2} + \frac{2^{-\lfloor\frac{N}{2}\rfloor}\mathcal{MK}_N}{2}\right] \tag{A61}$$

$$\Leftrightarrow \mathcal{MK}_N = 2^{\lfloor\frac{N}{2}\rfloor + 1}\left(2^{N-2\lfloor\frac{N}{2}\rfloor}\cdot p_w - \frac{1}{2}\right), \tag{A62}$$

hence we get

$$\chi(A_i : R|F_i, X_i = 0) \leqslant h\left(\frac{1}{2} + \frac{1}{2}\sqrt{\frac{2^{6+2\lfloor\frac{N}{2}\rfloor - N}\cdot\left(2^{N-2\lfloor\frac{N}{2}\rfloor}\cdot p_w - \frac{1}{2}\right)^2}{4} - 1}\right). \tag{A63}$$

Combining all together we have

$$H(A_i'\tilde{C}_i|X_iY_{(1\ldots N-1),i}T_iR)_\sigma \geqslant \left(1 - \frac{\mu}{2}\right)\left(1 - h\left(\frac{1}{2} + \frac{1}{2}\sqrt{\frac{2^{6+2\lfloor\frac{N}{2}\rfloor - N}\cdot\left(2^{N-2\lfloor\frac{N}{2}\rfloor}\cdot p_w - \frac{1}{2}\right)^2}{4} - 1}\right)\right). \tag{A64}$$

Note that p_w can be expressed in terms of the probability distribution $q = (q(1),q(0),q(\perp))^t$ (where ^{t} is the transpose) as $p_w = \frac{q(1)}{1-q(\perp)}$. And because in our case the definition of the maps $\mathcal{M}_i$ implies $1 - q(1) = \mu$ we have $p_w = \frac{q(1)}{\mu}$. Therefore the function,

$$\bar{f}(q) = \hat{f}(q(1)) = \left(1 - \frac{\mu}{2}\right)\left(1 - h\left(\frac{1}{2} + \frac{1}{2}\sqrt{2^{4+2\lfloor\frac{N}{2}\rfloor - N}\cdot\left(2^{N-2\lfloor\frac{N}{2}\rfloor}\cdot\frac{q(1)}{\mu} - \frac{1}{2}\right)^2 - 1}\right)\right), \tag{A65}$$

is a min-tradeoff function, and $\hat{f}$ is a differentiable convex increasing function of one variable. To find an affine min-tradeoff function f we take a tangent to $\hat{f}$ for some value $p_{\mathrm{opt}}(n,\delta) \in]\mu\cdot p_{\min}, \mu\cdot p_{\max}[$ to be chosen, where μ and δ are defined in the Protocol 2, which gives us

$$f(q,p_{\mathrm{opt}}) := \hat{f}'(p_{\mathrm{opt}})q(1) + \hat{f}(p_{\mathrm{opt}}) - \hat{f}'(p_{\mathrm{opt}})p_{\mathrm{opt}}. \tag{A66}$$

■

In the following Lemma we show that the state $\tilde{\rho}$ created by applying a sequence of n CPTP maps of the form described by Fig. 2 on some initial state [when conditioned on the event of having (statistically) high enough Bell violation] possesses a linear amount of entropy.

Lemma 6. Let $\mathcal{M}_{\mathrm{EC}}$ be the CPTP map $A'^n_1 B'_{(1\ldots N-1)}{}^n_1 \mapsto A'^n_1 B'_{(1\ldots N-1)}{}^n_1 K_{B_{(1\ldots N-1)}} G_{(1\ldots N-1)}$ that models the error correction protocols, applied during Step 3 of Protocol 2, which produce the raw keys $K_{B_{(1\ldots N-1)}}$ and the guess $G_{(1\ldots N-1)}$. For $i \in [n]$ let $\mathcal{M}_i$ be the CPTP map from R_{i-1} to $A'_i B'_{(1\ldots N-1)_i}\tilde{C}_i X_i Y_{(1\ldots N-1),i} T_i R_i$ defined in the Fig. 2. Let Ω be the event $\{\sum_j \tilde{C}_j \geqslant \delta\cdot\sum_j T_j$ for $\delta \in]p_{\min}, p_{\max}[$ **and** all the error correction protocols were successful, meaning that $\forall k$, $A'^n_1 = K_{B_{(k)}}$ and Alice guess $G_{(1\ldots N-1)}$ is correct$\}$. We define the state,

$$\begin{aligned}&\tilde{\rho}_{A'^n_1\tilde{C}^n_1 B'_{(1\ldots N-1)}{}^n_1 X^n_1 Y_{(1\ldots N-1)}{}^n_1 T^n_1 E}\\ &:= (\mathrm{tr}_{R_n}\circ\mathcal{M}_n\circ\ldots\circ\mathcal{M}_1\otimes\mathbb{1}_E)(\rho_{R_0E}),\end{aligned} \tag{A67}$$

where $R_0 = A^n_1 B_{(1\ldots N-1)}{}^n_1$, and ρ_{R_0E} is the state shared between Alice, the Bobs, and Eve (produced by Alice's source) across the n rounds of the Protocol 2 before they apply any measurement. Then we have for any $\epsilon \in]0,1[$,

$$H^\epsilon_{\min}\left(A'^n_1|X^n_1 Y_{(1\ldots N-1)}{}^n_1 T^n_1 E\right)_{\mathcal{M}_{\mathrm{EC}}(\tilde{\rho})_{|\Omega}}(f(\hat{q},p_{\mathrm{opt}}) - \mu)n - \tilde{v}\sqrt{n} + 3\log_2(1 - \sqrt{1-(\epsilon/4)^2}), \tag{A68}$$

where $\tilde{v} = 2(\log_2(13) + (\hat{f}'(p_{\mathrm{opt}}) + 1))\sqrt{1 - 2\log_2(\epsilon p_\Omega)} + 2\log_2(7)\sqrt{-\log_2(p^2_\Omega(1 - \sqrt{1-(\epsilon/4)^2}))}$, and $\hat{q} = (\delta\mu, \mu - \delta\mu, 1-\mu)^t \in \mathbb{P}(\{1,0,\perp\})$.

Proof. Note that $\tilde{\rho}_{|\Omega} := \mathrm{tr}_{K_{B_{(1\ldots N-1)}}G_{(1\ldots N-1)}}(\mathcal{M}_{\mathrm{EC}}(\tilde{\rho})_{|\Omega})$, therefore $H^\epsilon_{\min}(A'^n_1|X^n_1 Y_{(1\ldots N-1)}{}^n_1 T^n_1 E)_{\mathcal{M}_{\mathrm{EC}}(\tilde{\rho})_{|\Omega}} = H^\epsilon_{\min}(A'^n_1|X^n_1 Y_{(1\ldots N-1)}{}^n_1 T^n_1 E)_{\tilde{\rho}_{|\Omega}}$.

The maps $\mathcal{M}_1,\ldots,\mathcal{M}_n$ are EAT channels with the following Markov conditions,

$$\forall i \in [n], A'^{i-1}_1 \tilde{C}^{i-1}_1 \leftrightarrow X^{i-1}_1 Y_{(1\ldots N-1)}{}^{i-1}_1 T^{i-1}_1 E \leftrightarrow X^i_1 Y_{(1\ldots N-1)}{}^i_1 T^i_1. \tag{A69}$$

Indeed for any round $i \in [n]$ the variables $X_i Y_{(1\ldots N-1),i} T_i$ are chosen independently of any other round $j \neq i$. We have proven that the function $f(\,\cdot\,,p_{\text{opt}})$ is a min-tradeoff function for the maps $\mathcal{M}_1,\ldots,\mathcal{M}_n$. We can therefore use the EAT to bound $H^{\epsilon}_{\min}(A'^n_1\tilde{C}^n_1|X^n_1 Y_{(1\ldots N-1)}{}^n_1 T^n_1 E)_{\tilde{\rho}_{|\Omega}}$:

$$H^{\epsilon}_{\min}\big(A'^n_1\tilde{C}^n_1|X^n_1 Y_{(1\ldots N-1)}{}^n_1 T^n_1 E\big)_{\tilde{\rho}_{|\Omega}} \geqslant nf(\hat{q},p_{\text{opt}}) - c\sqrt{n}, \tag{A70}$$

where $\hat{q} = (\mu\delta,\mu-\mu\delta,1-\mu)$, $c = 2(\log_2(13) + \lceil \hat{f}'(p_{\text{opt}})\rceil)\sqrt{1-2\log_2(\epsilon p_\Omega)}$, and p_Ω is the probability of the event Ω. This is true because $f(q,p_{\text{opt}})$ is an increasing function of $q(1)$, so for any event that implies $\sum_j \tilde{C}_j \geqslant \delta \sum_j T_j$ we have that $f(\mathsf{freq}(\tilde{C}^n_1),p_{\text{opt}}) \geqslant f(\hat{q},p_{\text{opt}})$, in particular $\Omega \Rightarrow f(\mathsf{freq}(\tilde{C}^n_1),p_{\text{opt}}) \geqslant f(\hat{q},p_{\text{opt}})$. Note that because $\forall x \in \mathbb{R}$, $\lceil x\rceil \leqslant x+1$ we can upper bound $\lceil \hat{f}'(p_{\text{opt}})\rceil$ by $\hat{f}'(p_{\text{opt}})+1$ and then take $c = 2(\log_2(13) + (\hat{f}'(p_{\text{opt}})+1))\sqrt{1-2\log_2(\epsilon\cdot p_\Omega)}$.

Using [33], Eq. (6.57) we can relate $H^{\epsilon}_{\min}(A'^n_1\tilde{C}^n_1|X^n_1 Y_{(1\ldots N-1)}{}^n_1 T^n_1 E)_{\tilde{\rho}_{|\Omega}}$ to $H^{\epsilon}_{\min}(A'^n_1|X^n_1 Y_{(1\ldots N-1)}{}^n_1 T^n_1 E)_{\tilde{\rho}_{|\Omega}}$:

$$\begin{aligned} &H^{\epsilon}_{\min}\big(A'^n_1\big|X^n_1 Y_{(1\ldots N-1)}{}^n_1 T^n_1 E\big)_{\tilde{\rho}_{|\Omega}} \\ &\geqslant H^{\frac{\epsilon}{4}}_{\min}\big(A'^n_1\tilde{C}^n_1\big|X^n_1 Y_{(1\ldots N-1)}{}^n_1 T^n_1 E\big)_{\tilde{\rho}_{|\Omega}} - H^{\frac{\epsilon}{4}}_{\max}\big(\tilde{C}^n_1\big|X^n_1 Y_{(1\ldots N-1)}{}^n_1 T^n_1 E\big)_{\tilde{\rho}_{|\Omega}} + 3\log_2(1-\sqrt{1-(\epsilon/4)^2}). \end{aligned} \tag{A71}$$

We now need to upper bound $H^{\frac{\epsilon}{4}}_{\max}(\tilde{C}^n_1|X^n_1 Y_{(1\ldots N-1)}{}^n_1 T^n_1 E)_{\tilde{\rho}_{|\Omega}}$. First we note that

$$H^{\frac{\epsilon}{4}}_{\max}\big(\tilde{C}^n_1\big|X^n_1 Y_{(1\ldots N-1)}{}^n_1 T^n_1 E\big)_{\tilde{\rho}_{|\Omega}} \leqslant H^{\frac{\epsilon}{4}}_{\max}\big(\tilde{C}^n_1\big|T^n_1 E\big)_{\tilde{\rho}_{|\Omega}}.$$

To upper bound $H^{\frac{\epsilon}{4}}_{\max}(\tilde{C}^n_1|T^n_1 E)_{\tilde{\rho}_{|\Omega}}$ we will use [36], Lemma 28. Indeed $H^{\frac{\epsilon}{4}}_{\max}(\tilde{C}^n_1|T^n_1 E)_{\tilde{\rho}_{|\Omega}}$ can be bounded exactly in the same as in [36], Lemma 28, and leads to

$$\begin{aligned} H^{\epsilon/4}_{\max}\big(\tilde{C}^n_1\big|T^n_1 E\big)_{\tilde{\rho}_{|\Omega}} &\leqslant \mu n + n(\alpha-1)\log_2^2(7) + \frac{\alpha}{\alpha-1}\log_2\left(\frac{1}{p_\Omega}\right) - \frac{\log_2(1-\sqrt{1-(\epsilon/4)^2})}{\alpha-1} && \text{(A72)}\\ &\leqslant \mu n + n(\alpha-1)\log_2^2(7) - \frac{\log_2(p_\Omega^2(1-\sqrt{1-(\epsilon/4)^2}))}{\alpha-1}, && \text{(A73)} \end{aligned}$$

for $\alpha \in]1,2]$.

Taking $\alpha = 1 + \sqrt{\frac{-\log_2(p_\Omega^2(1-\sqrt{1-(\epsilon/4)^2}))}{n\log_2^2(7)}}$ gives us

$$H^{\epsilon/4}_{\max}\big(\tilde{C}^n_1\big|T^n_1 E\big)_{\tilde{\rho}_{|\Omega}} \leqslant \mu n + 2\sqrt{n}\log_2(7)\sqrt{-\log_2\big(p_\Omega^2(1-\sqrt{1-(\epsilon/4)^2})\big)}. \tag{A74}$$

Putting Eqs. (A70),(A71), and (A74) together gives us

$$H^{\epsilon}_{\min}\big(A'^n_1\big|X^n_1 Y_{(1\ldots N-1)}{}^n_1 T^n_1 E\big)_{\tilde{\rho}_{|\Omega}} \geqslant (f(\hat{q},p_{\text{opt}})-\mu)n - \tilde{v}\sqrt{n} + 3\log_2(1-\sqrt{1-(\epsilon/4)^2}), \tag{A75}$$

where $\tilde{v} = 2(\log_2(13) + (\hat{f}'(p_{\text{opt}})+1))\sqrt{1-2\log_2(\epsilon p_\Omega)} + 2\log_2(7)\sqrt{-\log_2(p_\Omega^2(1-\sqrt{1-(\epsilon/4)^2}))}$.

Since $H^{\epsilon}_{\min}(A'^n_1|X^n_1 Y_{(1\ldots N-1)}{}^n_1 T^n_1 E)_{\mathcal{M}_{\text{EC}}(\tilde{\rho})_{|\Omega}} = H^{\epsilon}_{\min}(A'^n_1|X^n_1 Y_{(1\ldots N-1)}{}^n_1 T^n_1 E)_{\tilde{\rho}_{|\Omega}}$, we have

$$H^{\epsilon}_{\min}\big(A'^n_1\big|X^n_1 Y_{(1\ldots N-1)}{}^n_1 T^n_1 E\big)_{\mathcal{M}_{\text{EC}}(\tilde{\rho})_{|\Omega}} \geqslant (f(\hat{q},p_{\text{opt}})-\mu)n - \tilde{v}\sqrt{n} + 3\log_2(1-\sqrt{1-(\epsilon/4)^2}). \tag{A76}$$

This bound holds for any $p_{\text{opt}} \in]\mu p_{\min},\mu p_{\max}[$. ■

In the following Lemma we link the result of the previous lemma to the real state ρ generated by Protocol 2. Indeed, in the real state, the "Bell violation" is not estimated directly, but via the error corrections that might fail with some small probability. We show that the real state of the protocol, when conditioned on the event that *Protocol 2 does not abort* **and** *the error corrections were successful*, possesses a linear amount on entropy.

Lemma 7. Let us call $\hat{\Omega}$ the event of not aborting Protocol 2 and $\hat{\Omega}'$ the event $\hat{\Omega}$ **and** all the error correction protocols were successful, meaning that $\forall k \in [N-1]$, $K_{B_{(k)}} = A'^n_1$, and Alice's guess $G_{(1\ldots N-1)}$ is correct. Then, for any $\epsilon_{\text{EA}},\epsilon'_{\text{EC}},\epsilon \in]0,1[$, Protocol 2 either aborts with a probability $1-P(\hat{\Omega})\geqslant 1-(1-2(N-1)\epsilon'_{\mathrm{EC}})\epsilon_{\mathrm{EA}}$ ($\Leftrightarrow P(\hat{\Omega}')\leqslant\epsilon_{\mathrm{EA}}$) or

$$\begin{aligned} &H^{\epsilon}_{\min}\big(A'^{n}_{1}\big|X^{n}_{1}Y_{(1\ldots N-1)}{}^{n}_{1}T^{n}_{1}O_{A}O_{(1\ldots N-1)}E\big)_{\rho_{|\hat{\Omega}'}}\\ &\geqslant \max_{p_{\min}\leqslant\frac{p_{\mathrm{opt}}}{\mu}\leqslant p_{\max}} n\left((f(\hat{q},p_{\mathrm{opt}})-\mu)-\frac{2(\log_2(13)+(\hat{f}'(p_{\mathrm{opt}})+1))\sqrt{1-2\log_2(\epsilon\epsilon_{\mathrm{EA}})}}{\sqrt{n}}\right)\\ &\quad-\sqrt{n}\Big(2\log_2(7)\sqrt{-\log_2\ \big(\epsilon^2_{\mathrm{EA}}(1-\sqrt{1-(\epsilon/4)^2})\big)}\Big)+3\log_2(1-\sqrt{1-(\epsilon/4)^2})-\mathrm{leak}_{\mathrm{EC}}(O_A)-\sum_{k=1}^{N-1}\mathrm{leak}_{\mathrm{EC}}(O_{(k)}), \end{aligned} \tag{A77}$$

where $\hat{q}=(\mu\delta,\mu-\mu\delta,1-\mu)^t$.

Proof. Using the chain rule [33], Lemma 6.8 we get

$$H^{\epsilon}_{\min}\big(A'^{n}_{1}\big|X^{n}_{1}Y_{(1\ldots N-1)}{}^{n}_{1}T^{n}_{1}O_{A}O_{(1\ldots N-1)}E\big)_{\rho_{|\hat{\Omega}'}}\geqslant H^{\epsilon}_{\min}\big(A'^{n}_{1}\big|X^{n}_{1}Y_{(1\ldots N-1)}{}^{n}_{1}T^{n}_{1}E\big)_{\rho_{|\hat{\Omega}'}}-\mathrm{leak}_{\mathrm{EC}}(O_A)-\sum_{k=1}^{N-1}\mathrm{leak}_{\mathrm{EC}}(O_{(k)}), \tag{A78}$$

where $\mathrm{leak}_{\mathrm{EC}}(O_A)$ is the leakage due to the error correction protocol (when the Bobs try to guess Alice's bits) and $\mathrm{leak}_{\mathrm{EC}}(O_{(k)})$ is the leakage due to error correction (when Alice tries to guess Bob_k's test rounds bits). These leakages will be estimated in Sec. 3.

We now need to bound $H^{\epsilon}_{\min}(A'^{n}_{1}|X^{n}_{1}Y_{(1\ldots N-1)}{}^{n}_{1}T^{n}_{1}E)_{\rho_{|\hat{\Omega}'}}$. Note that the reduced state on $A'^{n}_{1}X^{n}_{1}Y_{(1\ldots N-1)}{}^{n}_{1}T^{n}_{1}E$ of the global state at the end of Protocol 2 conditioned on the event $\hat{\Omega}'$ of not aborting *and* all the error correction protocol were successful, is equal to the state $\mathcal{M}_{\mathrm{EC}}(\tilde{\rho}_{A'^{n}_{1}X^{n}_{1}Y_{(1\ldots N-1)}{}^{n}_{1}T^{n}_{1}E})_{|\Omega}$, therefore using Lemma 6 we get

$$H^{\epsilon}_{\min}\big(A'^{n}_{1}\big|X^{n}_{1}Y_{(1\ldots N-1)}{}^{n}_{1}T^{n}_{1}E\big)_{\rho_{|\hat{\Omega}'}}\geqslant(f(\hat{q},p_{\mathrm{opt}})-\mu)n-\tilde{v}\sqrt{n}+3\log_2(1-\sqrt{1-(\epsilon/4)^2}), \tag{A79}$$

where $\tilde{v}=2(\log_2(13)+(\hat{f}'(p_{\mathrm{opt}})+1))\sqrt{1-2\log_2(\epsilon\cdot p_{\hat{\Omega}'})}+2\log_2(7)\sqrt{-\log_2(p^2_{\hat{\Omega}'}(1-\sqrt{1-(\epsilon/4)^2}))}$. ■

The following Lemma concludes on the soundness of Protocol 2. To do so we need to relate the event $\hat{\Omega}$ that Protocol 2 does not abort, with the event $\hat{\Omega}'$ that Protocol 2 does not abort **and** that the error corrections are successful.

Lemma 8. For any implementation of Protocol 2, either the protocol aborts with a probability greater than $1-\epsilon_{\mathrm{EA}}$ or it is $((N-1)\epsilon_{\mathrm{EC}}+\epsilon_{\mathrm{PA}}+\epsilon)$-correct-and-secret while producing keys of length l defined in Eq. (A.36).

Proof. Let $\hat{\Omega}$ be the event of not aborting in Protocol 2, and $\hat{\Omega}'$ the event $\hat{\Omega}$ **and** all the error correction protocols were successful. According to Lemma 7 we are into one of the two following cases:

The protocol aborts with a probability $1-P(\hat{\Omega})\geqslant 1-(1-2(N-1)\epsilon'_{\mathrm{EC}})\epsilon_{\mathrm{EA}}$. This is equivalent to $P(\hat{\Omega}')\leqslant\epsilon_{\mathrm{EA}}$ and implies that $1-P(\hat{\Omega})\geqslant 1-\epsilon_{\mathrm{EA}}$.

The aborting probability is $1-P(\hat{\Omega})\leqslant 1-\epsilon_{\mathrm{EA}}$ (which implies that $P(\hat{\Omega}')\geqslant\epsilon_{\mathrm{EA}}$) and the smooth min-entropy of the final state conditioned on $\hat{\Omega}'$ is bounded as in Eq. (A79). Conditioned on $\hat{\Omega}$ there are two cases.

(1) The error correction step failed. This happens with probability at most $2(N-1)\epsilon'_{\mathrm{EC}}$.

(2) The error correction were successful and then all the keys agree. We have then the event $\hat{\Omega}'$. Therefore according to Lemma 7 the entropy is high enough to produce keys of length l such that

$$\left\|\rho_{K_AE|\hat{\Omega}}-\frac{\mathbb{1}_A}{2^l}\otimes\rho_{E|\hat{\Omega}}\right\|_{\mathrm{tr}}\leqslant\epsilon_{\mathrm{PA}}+2\epsilon, \tag{A80}$$

where ϵ_{PA} is the privacy amplification error probability and ϵ is the smoothing parameter. ■

By combining the two above cases we have that Protocol 2 is $(\epsilon_{\mathrm{PA}}+2(N-1)\epsilon'_{\mathrm{EC}}+2\epsilon)$-correct-and-secret.

3. Asymptotic key rate analysis

In this section we evaluate the asymptotic key rate of the DICKA Protocol 2 and compare it to the case where the parties perform $N-1$ DIQKD protocols in order to establish a common key. In implementations where the efficiency of generation of GHZ states is comparable to the efficiency of the generation of EPR pairs a common key using a DICKA protocol can be, in principle, established in a much smaller number of rounds, however, one needs to analyze how the QBER and the leakages in the error correction protocol affects the key generation.

To analyze the key rate we need to evaluate the length l of the final key produced by Protocol 2, Eq. (A.36), and compute the rate $r:=\frac{l}{\#\mathrm{rounds}}$. To achieve this, we need to estimate the leakage due to the error correction step. We use in our analysis an error correction protocol based on universal hashing [41,42]. The size of the leakage is taken to be the amount of correction information needed if the implementation were honest, for some abort probability of the error correction protocol of at most ϵ_{EC}, and such that the guess (when not aborting) is correct with probability at least $1-\epsilon'_{\mathrm{EC}}$. For a given honest implementation, this leakage can be bounded as follows [42]:

$$\begin{aligned}\mathrm{leak}(\mathrm{O_A})\leqslant&\max_{k\in[N-1]}H_0^{\bar{\epsilon}_{\mathrm{EC}}}\big(A^n_1\big|B'_{(k)}{}^{n}_{1}X^n_1Y_{(1\ldots N-1)}T^n_1\big)\\&+\log_2(\epsilon'^{-1}_{\mathrm{EC}}),\end{aligned} \tag{A81}$$

$$\mathrm{leak}(\mathrm{O}_{(k)}) \leqslant H_0^{\tilde{\epsilon}_{\mathrm{EC}}}(B'_{(k),I}|A'^n_1 X^n_1 Y_{(1\ldots N-1)}T^n_1) + \log_2(\epsilon'^{-1}_{\mathrm{EC}}), \tag{A82}$$

for $\epsilon_{\mathrm{EC}} = \tilde{\epsilon}_{\mathrm{EC}} + \epsilon'_{\mathrm{EC}}$, $I := \{i \in [n] : T_i = 1\}$ and where $H_0^{\tilde{\epsilon}_{\mathrm{EC}}}$ is evaluated on the state produced by the honest implementation. If it turns out that the implementation is not the expected one then the protocol will just abort with a higher probability but the security is not affected.

We will consider here one particular honest implementation to evaluate the leakage. Then we will compare it to what we would get using $N-1$ device-independent quantum key distribution ($(N-1)\times$ DIQKD) protocols to distribute the key to the N parties. For the key rate of the latter we will use the recent and most general analysis given in [11]. Of course the following calculations can be adapted to other implementations.

Lemma 9 (Asymptotic key rate). There exists an implementation of Protocol 2 in which the achieved asymptotic key rate is given by

$$r_{N-\mathrm{CKA},\infty} = 1 - h\left(\tfrac{1}{2} + \tfrac{1}{2}\sqrt{2(1-2Q)^N - 1}\right) - h(Q), \tag{A83}$$

where Q is the QBER between Alice and each of the Bobs.

Proof. In the following analysis we chose an i.i.d. honest implementation scenario where we assume that the channel between Alice and each of the Bobs is a depolarizing channel:

$$\mathcal{D}(\rho) = (1-p_{\mathrm{dep}})\rho + p_{\mathrm{dep}}\frac{\mathbb{1}}{2}, \tag{A84}$$

for $p_{\mathrm{dep}} \in]0,1[$. We will also apply this channel to model the noise on Alice's side. The state that is produced by Alice's source is supposed to be an N-GHZ state denoted $\mathrm{GHZ}_N := |\mathrm{GHZ}_N\rangle\langle\mathrm{GHZ}_N|$, where $|\mathrm{GHZ}_N\rangle := \frac{|0\rangle^{\otimes N}+|1\rangle^{\otimes N}}{\sqrt{2}}$. Therefore the state shared between Alice and the Bobs in one round is $\rho_{AB_{(1\ldots N-1)}} := \mathcal{D}^{\otimes N}(\mathrm{GHZ}_N)$. The QBER between Alice and each of the Bobs can then be expressed as $Q = \frac{2p_{\mathrm{dep}} - p_{\mathrm{dep}}^2}{2}$ ($\Leftrightarrow p_{\mathrm{dep}} = 1 - \sqrt{1-2Q}$) and the expected winning probability of the MABK game is given by

$$p_{\mathrm{exp}} = 2^{2\lfloor N/2\rfloor - N}\left[\frac{1}{2} + (1-p_{\mathrm{dep}})^N \frac{2^{-\lfloor N/2\rfloor}2^{(N-1)/2}}{2}\right].$$

We can bound H_0 by $H_{\max}$ [44], Lemma 18 as

$$\begin{aligned} &H_0^{\tilde{\epsilon}_{\mathrm{EC}}}\left(A^n_1\middle|B'^{\,n}_{(k)1} X^n_1 Y_{(1\ldots N-1)}T^n_1\right) \\ &\quad\leqslant H_{\max}^{\tilde{\epsilon}_{\mathrm{EC}}/2}\left(A^n_1\middle|B'^{\,n}_{(k)1} X^n_1 Y_{(1\ldots N-1)}T^n_1\right) \\ &\qquad + \log_2\left(8/\tilde{\epsilon}^2_{\mathrm{EC}} + 2/(2-\tilde{\epsilon}_{\mathrm{EC}})\right). \end{aligned} \tag{A85}$$

Using the nonasymptotic version of the asymptotic equipartition theorem [45], Theorem 9 we get

$$\begin{aligned} &H_{\max}^{\tilde{\epsilon}_{\mathrm{EC}}/2}\left(A^n_1\middle|B'^{\,n}_{(k)1} X^n_1 Y_{(1\ldots N-1)}T^n_1\right) \\ &\quad\leqslant nH(A'_i|B'_{(1\ldots N-1),i}X_iY_{(1\ldots N-1),i}T_i) + \sqrt{n}\Delta(\tilde{\epsilon}_{\mathrm{EC}}), \end{aligned} \tag{A86}$$

where $\Delta(\tilde{\epsilon}_{\mathrm{EC}}) := 4\log_2\left(2\sqrt{2^{H_{\max}(A'_i|B'_{(1\ldots N-1),i}X_iY_{(1\ldots N-1),i}T_i)}} + 1\right)\cdot \sqrt{2\log_2(8/\tilde{\epsilon}^2_{\mathrm{EC}})}$. We can now upper bound the entropy for honest implementation of Protocol 2 as

$$\begin{aligned} &H(A'_i|B'_{(1\ldots N-1),i}X_iY_{(1\ldots N-1),i}T_i) \\ &\quad= (1-\mu)H(A'_i|B'_{(1\ldots N-1),i}X_iY_{(1\ldots N-1),i},T_i=0) \\ &\qquad+ \mu\underbrace{H(A'_i|B'_{(1\ldots N-1),i}X_iY_{(1\ldots N-1),i},T_i=1)}_{\leqslant 1} \end{aligned} \tag{A87}$$

$$\leqslant (1-\mu)h(Q) + \mu, \tag{A88}$$

and $H_{\max}(A'_i|B'_{(1\ldots N-1),i}X_iY_{(1\ldots N-1),i}T_i) \leqslant 1$. This gives us an upper bound on $\mathrm{leak}(O_A)$:

$$\begin{aligned} \mathrm{leak}(O_A) \leqslant\, & n((1-\mu)h(Q)+\mu) + \sqrt{n}4\log_2(2\sqrt{2}+1) \\ & \sqrt{2\log_2\left(8/\tilde{\epsilon}^2_{\mathrm{EC}}\right)} + \log_2\left(8/\tilde{\epsilon}^2_{\mathrm{EC}} + 2/(2-\tilde{\epsilon}_{\mathrm{EC}})\right). \end{aligned} \tag{A89}$$

Using the same reasoning, we get

$$\begin{aligned} \mathrm{leak}(O_{(k)}) \leqslant\, & n\mu + \sqrt{n}4\log_2(2\sqrt{2}+1)\sqrt{2\log_2\left(8/\tilde{\epsilon}^2_{\mathrm{EC}}\right)} \\ & + \log_2\left(8/\tilde{\epsilon}^2_{\mathrm{EC}} + 2/(2-\tilde{\epsilon}_{\mathrm{EC}})\right). \end{aligned} \tag{A90}$$

Putting this into Eq. (A36) we get

$$\begin{aligned} l =\, & (f(\hat{q},p_{\mathrm{opt}}) - (1-\mu)h(Q) - (N+1)\mu)n \\ & - \hat{v}\sqrt{n} + 3\log_2(1-\sqrt{1-(\epsilon/4)^2}) - \log_2\left(\epsilon_{\mathrm{PA}}^{-1}\right) \\ & - N\log_2\left(8/\tilde{\epsilon}^2_{\mathrm{EC}} + 2/(2-\tilde{\epsilon}_{\mathrm{EC}})\right), \end{aligned} \tag{A91}$$

where $\hat{v} = \tilde{v} + N\cdot 4\log_2(2\sqrt{2}+1)\cdot\sqrt{2\log_2(8/\tilde{\epsilon}^2_{\mathrm{EC}})}$, and $\tilde{v}$ is defined in Theorem 4.

Note that in the asymptotic regime $n\to\infty$ we can take the threshold δ to be $\delta = p_{\mathrm{exp}}$, and the optimal p_{opt} will be $p_{\mathrm{opt}} = \mu\delta = \mu p_{\mathrm{exp}}$. Also for the asymptotic analysis we chose $\mu = n^{-1/10}$. Therefore the asymptotic rate $r_\infty := \lim_{n\to\infty}\frac{l}{\#\mathrm{rounds}}$ becomes

$$\begin{aligned} r_{N-\mathrm{CKA},\infty} &= \hat{f}(\mu p_{\mathrm{exp}}) - h(Q) \\ &= \hat{f}(\mu(1-\sqrt{1-2Q})) - h(Q) \\ &= 1 - h\left(\frac{1}{2} + \frac{1}{2}\sqrt{2(1-2Q)^N - 1}\right) - h(Q). \end{aligned} \tag{A92}$$

■

We then compare it to the asymptotic rate we would get if in order to distribute a key to N parties, Alice were to use a DIQKD protocol for each of the Bobs. To get the asymptotic rate for the $(N-1)$ DIQKD protocols, we use the analysis given in [11]. In their DIQKD protocol they consider an honest implementation where the state is a depolarized EPR pair $(1-\nu)\Phi_{AB} + \nu\frac{\mathbb{1}}{2}$. If we say that, for each Bob, Alice sends the state via the same depolarizing channel she uses in the previous analysis (and that she has the same noise on her qubits), we can link the parameter ν with the depolarizing parameter p_{dep} of the channel and to the QBER Q: $\nu = 2p - p^2 = 2Q$. Therefore we get

$$r_{(N-1)\times\mathrm{QKD},\infty} = \frac{1 - h\left(\frac{1}{2} + \frac{1}{2}\sqrt{2(1-2Q)^2 - 1}\right) - h(Q)}{N-1}. \tag{A93}$$

Note that the factor $1/(N-1)$ comes from the fact that the total number of rounds while running $N-1$ DIQKD protocols is $(N-1)n$, where n is the number of rounds for one DIQKD protocol.

The comparison of the key rates of DICKA, Eq. (A92), and $(N-1) \times$ DIQKD, Eq. (A93), for different values of N, are plotted in Fig. 1. The results show that for low noise it is advantageous to use the DICKA protocol. In this comparison we assume that the cost of generation of a GHZ state is the same as the cost to generate one EPR pair. However, in implementations where the GHZ state is created out of EPR pairs that will not be the case. Therefore the cost of creation of these states must be taken into account in the analysis of the particular implementations. Note, also, that in this section we have modeled the implementation for depolarizing channels, however, the security analysis is general and can be adapted for any particular implementation.

[1] A. K. Ekert, Quantum Cryptography Based on Bell's Theorem, Phys. Rev. Lett. **67**, 661 (1991).

[2] C. H. Bennett and G. Brassard, Quantum cryptography: Public key distribution and coin tossing, Theoretical Aspects of Quantum Cryptography–celebrating 30 years of {BB84}, Theoretical Computer Science **560**, 7 (2014).

[3] D. Mayers and A. Yao, Quantum cryptography with imperfect apparatus, In *Proceedings of the 39th Annual Symposium on Foundations of Computer Science*, FOCS '98 (IEEE Computer Society, Washington, DC, 1998), pp. 503–509.

[4] J. Barrett, L. Hardy, and A. Kent, No Signaling and Quantum Key Distribution, Phys. Rev. Lett. **95**, 010503 (2005).

[5] A. Acín, N. Brunner, N. Gisin, S. Massar, S. Pironio, and V. Scarani, Device-Independent Security of Quantum Cryptography Against Collective Attacks, Phys. Rev. Lett. **98**, 230501 (2007).

[6] S. Pironio, A. Acín, N. Brunner, N. Gisin, S. Massar, and V. Scarani, Device-independent quantum key distribution secure against collective attacks, New J. Phys. **11**, 045021 (2009).

[7] L. Masanes, S. Pironio, and A. Acín, Secure device-independent quantum key distribution with causally independent measurement devices, Nat. Commun. **2**, 238 (2011).

[8] C. A. Miller and Y. Shi, Robust protocols for securely expanding randomness and distributing keys using untrusted quantum devices, J. ACM **63**, 33 (2016).

[9] C. A. Miller and Y. Shi, Universal security for randomness expansion from the spot-checking protocol, SIAM J. Comput. **46**, 1304 (2017).

[10] U. Vazirani and T. Vidick, Fully Device-Independent Quantum Key Distribution, Phys. Rev. Lett. **113**, 140501 (2014).

[11] R. Arnon-Friedman, R. Renner, and T. Vidick, Simple and tight device-independent security proofs, arXiv:1607.1797.

[12] J. Ribeiro, L. Phuc Thinh, J. Kaniewski, J. Helsen, and S. Wehner, Device-independence for two-party cryptography and position verification, arXiv:1606.8750.

[13] J. Kaniewski and S. Wehner, Device-independent two-party cryptography secure against sequential attacks, New J. Phys. **18**, 055004 (2016).

[14] J. Silman, A. Chailloux, N. Aharon, I. Kerenidis, S. Pironio, and S. Massar, Fully Distrustful Quantum Bit Commitment and Coin Flipping, Phys. Rev. Lett. **106**, 220501 (2011).

[15] A. Kent, Quantum tagging for tags containing secret classical data, Phys. Rev. A **84**, 022335 (2011).

[16] M. Tomamichel, S. Fehr, J. Kaniewski, and S. Wehner, A monogamy-of-entanglement game with applications to device-independent quantum cryptography, New J. Phys. **15**, 103002 (2013).

[17] M. Burmester and Y. Desmedt, *A Secure and Efficient Conference Key Distribution System* (Springer, Berlin/Heidelberg, 1995).

[18] K. Chen and H.-K. Lo, Conference key agreement and quantum sharing of classical secrets with noisy GHZ states, in *Proceedings of the International Symposium on Information Theory* (IEEE, Piscataway, 2005), pp. 1607–1611.

[19] M. Epping, H. Kampermann, C. Macchiavello, and D. Bruss, Multi-partite entanglement can speed up quantum key distribution in networks, New J. Phys. **19**, 093012 (2017).

[20] N. D. Mermin, Extreme Quantum Entanglement in a Superposition of Macroscopically Distinct States, Phys. Rev. Lett. **65**, 1838 (1990).

[21] M. Ardehali, Bell inequalities with a magnitude of violation that grows exponentially with the number of particles, Phys. Rev. A **46**, 5375 (1992).

[22] A. V. Belinskiĭ and D. N. Klyshko, Interference of light and Bell's theorem, Phys. Usp. **36**, 653 (1993).

[23] R. Gallego, Ll. Masanes, G. De La Torre, C. Dhara, L. Aolita, and A. Acín, Full randomness from arbitrarily deterministic events, Nat. Commun. **4**, 2654 (2013).

[24] G. de la Torre, M. J. Hoban, C. Dhara, G. Prettico, and A. Acín, Maximally Nonlocal Theories Cannot be Maximally Random, Phys. Rev. Lett. **114**, 160502 (2015).

[25] C. A. Miller and Y. Shi, Optimal robust quantum self-testing by binary nonlocal XOR games, arXiv:1207.1819.

[26] J. Kaniewski, Self-testing of binary observables based on commutation, Phys. Rev. A **95**, 062323 (2017).

[27] J. F. Clauser, M. A. Horne, A. Shimony, and R. A. Holt, Proposed Experiment to Test Local Hidden-Variable Theories, Phys. Rev. Lett. **23**, 880 (1969).

[28] R. Renner, Security of Quantum Key Distribution, Ph.D thesis, ETH Zurich, 2005.

[29] C. Portmann and R. Renner, Cryptographic security of quantum key distribution, arXiv:1409.3525.

[30] M. Tomamichel and A. Leverrier, A largely self-contained and complete security proof for quantum key distribution, Quantum **1**, 14 (2017).

[31] J. Barrett, R. Colbeck, and A. Kent, Memory Attacks on Device-Independent Quantum Cryptography, Phys. Rev. Lett. **110**, 010503 (2013).

[32] F. Dupuis, O. Fawzi, and R. Renner, Entropy accumulation, arXiv:1607.1796.

[33] M. Tomamichel, *Quantum Information Processing with Finite Resources–Mathematical Foundations*, Vol. 5, Springer Briefs in Mathematical Physics (Springer International Publishing, Berlin, 2016).

[34] R. F. Werner and M. M. Wolf, Bell's inequalities for states with positive partial transpose, Phys. Rev. A **61**, 062102 (2000).
[35] K. Nagata, M. Koashi, and N. Imoto, Configuration of Separability and Tests for Multipartite Entanglement in Bell-Type Experiments, Phys. Rev. Lett. **89**, 260401 (2002).
[36] J. Ribeiro, G. Murta, and S. Wehner, Fully general device-independence for two-party cryptography and position verification, arXiv:1609.8487.
[37] L. Masanes, Asymptotic Violation of Bell Inequalities and Distillability, Phys. Rev. Lett. **97**, 050503 (2006).
[38] A. Acín, Distillability, Bell Inequalities, and Multiparticle Bound Entanglement, Phys. Rev. Lett. **88**, 027901 (2001).
[39] S. Lee, J. Lee, and J. Kim, Any multipartite entangled state violating the Mermin-Klyshko inequality can be distilled for almost all bipartite splits, Phys. Rev. A **79**, 032309 (2009).
[40] R. Cleve, P. Hoyer, B. Toner, and J. Watrous, Consequences and limits of nonlocal strategies, in *Proceedings of the 19th IEEE Annual Conference on Computational Complexity 2004* (IEEE, Piscataway, 2004), pp. 236–249.
[41] G. Brassard and L. Salvail, *Secret-Key Reconciliation by Public Discussion* (Springer, Berlin/Heidelberg, 1994).
[42] R. Renner and S. Wolf, *Simple and Tight Bounds for Information Reconciliation and Privacy Amplification* (Springer, Berlin/Heidelberg, 2005).
[43] R. Renner and R. König, Universally composable privacy amplification against quantum adversaries, in *Second Theory of Cryptography Conference, TCC 2005*, Vol. 3378, Lecture Notes in Computer Science (Springer, Berlin, 2005).
[44] M. Tomamichel, C. Schaffner, A. Smith, and R. Renner, Leftover hashing against quantum side information, IEEE Trans. Inf. Theory **57**, 5524 (2011).
[45] M. Tomamichel, R. Colbeck, and R. Renner, A fully quantum asymptotic equipartition property, IEEE Trans. Inf. Theory **55**, 5840 (2009).