PAPER • OPEN ACCESS

Entanglement generation secure against general attacks

To cite this article: Alexander Pirker *et al* 2017 *New J. Phys.* **19** 113012

New Journal of Physics
The open access journal at the forefront of physics

Deutsche Physikalische Gesellschaft DPG
IOP Institute of Physics

Published in partnership with: Deutsche Physikalische Gesellschaft and the Institute of Physics

OPEN ACCESS

RECEIVED
15 February 2017

REVISED
14 July 2017

ACCEPTED FOR PUBLICATION
18 July 2017

PUBLISHED
8 November 2017

PAPER

Entanglement generation secure against general attacks

Alexander Pirker[1,3], **Vedran Dunjko**[1,2], **Wolfgang Dür**[1] **and Hans J Briegel**[1]

[1] Institut für Theoretische Physik, Universität Innsbruck, Technikerstr. 21a, A-6020 Innsbruck, Austria
[2] Max-Planck-Institute of Quantum Optics, Hans-Kopfermann-Strasse 1, D-85748 Garching, Germany
[3] Author to whom any correspondence should be addressed.

E-mail: alexander.pirker@student.uibk.ac.at, vedran.dunjko@mpq.mpg.de, wolfgang.duer@uibk.ac.at and hans.briegel@uibk.ac.at

Keywords: quantum cryptography, quantum communication, quantum information

Abstract

We present a security proof for establishing private entanglement by means of recurrence-type entanglement distillation protocols over noisy quantum channels. We consider protocols where the local devices are imperfect, and show that nonetheless a confidential quantum channel can be established, and used to e.g. perform distributed quantum computation in a secure manner. While our results are not fully device independent (which we argue to be unachievable in settings with quantum outputs), our proof holds for arbitrary channel noise and noisy local operations, and even in the case where the eavesdropper learns the noise. Our approach relies on non-trivial properties of distillation protocols which are used in conjunction with de-Finetti and post-selection-type techniques to reduce a general quantum attack in a non-asymptotic scenario to an i.i.d. setting. As a side result, we also provide entanglement distillation protocols for non-i.i.d. input states.

1. Introduction

Entanglement is a key resource in quantum information processing. Entanglement can be used to teleport quantum information [1], to implement remote quantum gates [2], or for distributed quantum computation [3]. It allows one to perform tasks that are not possible by classical means, such as secret key expansion vital for secure classical communication. The latter is achieved through the famous and extensively studied quantum key distribution (QKD) protocols [4–10]. In these works, security was proven in a variety of ever more general scenarios, considering noisy channels, imperfect devices and device-independent (DI) settings, where even the local quantum devices are untrusted [11–13].

In contrast, the perhaps equally important task of establishing private entanglement, and the closely related problem of establishing secure quantum channels, has not been resolved in equal generality. The latter has, historically, received significantly less attention [14], until the very recent increase of interest [15–18] in security under ideal settings. The task of establishing private entanglement has been considered in the context of noisy channels and both perfect [19] operations, and operations with local depolarizing noise [20, 21]. In these works, either initial states that are identical and independently distributed (i.i.d.), or asymptotic scenarios are assumed.

Here, we present a comprehensive treatment for the security of distillation protocols. To make our results broadly applicable, we generalize the security model (i.e. powers of the adversary) over standard settings for protocols with quantum outputs. Furthermore, we remove the need for asymptotic, or i.i.d. assumptions, allow for more general noise models, and formulate and prove security criteria which ensure composability—i.e. the security of the protocols when they are used in arbitrary contexts, e.g. as sub-routines of larger protocols.

More specifically, we consider arbitrary attacks employed by an adversary (Eve, the distributer of noisy or corrupt Bell-pairs) and assume noisy communication channels and noisy local operations—essentially arbitrary noise describing imperfect single- and two-qubit gates. We also extend adversarial powers beyond standard: the noisy apparatus may leak all the information about the noise processes which occurred in a run of the protocol to Eve.

Our scenario, by necessity, falls short from full DI, as security under such weakest assumptions is not attainable for protocols with a quantum output—any device used in any protocol with which a client can interact classically, perhaps to test its performance, but which eventually outputs a quantum system, can always

deviate from honest behavior when the final quantum output is eventually demanded (independent of how elaborate the testing may have been). This raises the questions of how DI assumptions can be relaxed such that security becomes possible also for quantum output protocols, or how standard security models can be further extended.

DI assumptions can be understood as an extreme noisy scenario, where Eve has absolute control over the noise processes. Our model relaxes this: Eve's control is not exact (deterministic), but rather probabilistic, however still perfectly heralded—while Eve may fail in her interventions, she still learns the noise realized. In this sense, generalizing the types of noise the protocol is provably secure under in our model, corresponds to scenarios which are ever closer to DI. Naturally, other generalizations of DI settings which make sense for protocols with quantum outputs may be possible[4].

We proceed by first providing a security analysis for i.i.d. inputs, and then generalize to non-i.i.d. states. This is done by employing de-Finetti and post-selection symmetrization-based techniques. However, since we are interested in security in arbitrary contexts, we must go beyond standard scenarios considered in entanglement distillation works [19–21] and explicitly consider the adversarial quantum systems (containing e.g. purifications of all quantum states) as well. Therefore the symmetrization-based techniques cannot be straightforwardly applied, but need to be adapted. We present and discuss the required additional steps of preprocessing, and provide entanglement distillation protocols that are not restricted to i.i.d. inputs, but are capable of dealing with general inputs. The latter is related to recent results in [22–24].

2. Structure of the paper

The paper is organized as follows. In section 3 we introduce the basic concepts, specify the overall setting and define the confidentiality of entanglement distillation protocols. Next, we summarize our main contribution in section 4. In section 5 we show confidentiality of recurrence-type entanglement distillation protocols by proving confidentiality for i.i.d. inputs in section 5.1 and we extend this results to arbitrary initial states in sections 5.2 and 5.3. Finally we prove confidentiality whenever the noise transcripts leak to Eve in section 5.4. We summarize and discuss our results in section 6.

3. The model and security guarantees

Entanglement distillation is modeled by considering three players, Alice and Bob, who wish to generate a shared Bell pair, and Eve, who provides the initial pairs. Thus, Eve is connected to Alice and to Bob via a (generally noisy) quantum channel which may be completely under her control. Alice and Bob are connected by a classical authenticated, but not confidential, channel. In entanglement distillation protocols Alice and Bob apply local, in general noisy, quantum operations to their pairs. To model this noise, we extend the approach of [20], where a noise register, referred to as the 'lab demon' (L) register L is used to store classical information about the local noise history, is appended to Alice and Bob's pairs. In this work, the L register is a quantum register, attached to Alice and Bob. We represent the noisy maps of the entanglement distillation process as unitaries acting on an enlarged Hilbert space. L thereby coherently applies Pauli operators onto the registers of Alice and Bob. Due to the symmetry of Bell states $|B_{00}\rangle = 1/\sqrt{2}\,(|00\rangle + |11\rangle)$, it suffices to consider the case when the noise is applied on Alice's register only. To model the setting where Eve acquires information about the noise transcript during the execution of the protocol, we assume that L informs Eve which noise operator was applied at each step. The setting is illustrated in figure 1. In the remainder of this paper we elaborate further on the full quantum treatment of L and Eve in terms of purifications, going beyond the setting of [20].

The proposed overall protocol under i.i.d. assumption involves several steps. First, Eve distributes n pairs (the *initial states*), to Alice and Bob who apply local 'twirl' operations (random, correlated local operations). Next, Alice and Bob sacrifice some $m \approx \sqrt{n}$ pairs to check whether the fidelity, given with $F(\rho, \sigma) = \mathrm{tr}\sqrt{\rho^{1/2}\sigma\rho^{1/2}}$ for density operators ρ and σ, of the pairs is sufficient for entanglement distillation, via local σ_x and σ_z measurements. If the fidelity F relative to $|B_{00}\rangle$ is insufficient, they abort. Otherwise they proceed with a recurrence-type entanglement distillation protocol to produce a high fidelity Bell-pair from the remaining initial states, which may also be aborted. Finally, Alice and Bob output their final state. For i.i.d. inputs, the twirl ensures that local σ_z and σ_x correlation measurements can be used to estimate the fidelity of individual pairs. This estimate is crucial for ensuring entanglement distillation via recurrence-type entanglement distillation protocols. Later, we will generalize to non-i.i.d. settings by prepending the protocol with symmetrization (permuting of the pairs) and tracing-out steps.

[4] E.g., we assume very primitive, but trusted, quantum devices, such as a device which can either forward an input quantum system, or measure it in one basis. Already such a simple device invalidates our no-go observation.

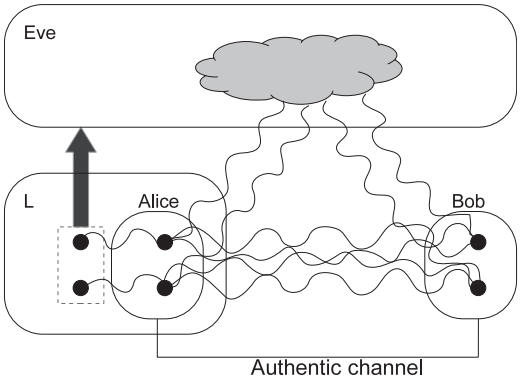

Figure 1. Illustration of the overall setting: Eve provides the initial pairs to Alice and Bob, who run the entanglement distillation protocol. The noisy apparatus may leak the specification of the realized noise map to Eve after every step of the protocol.

To formalize the security requirements, we define the ideal map $\mathcal{F}^{\alpha,l}$, mapping the initial states of Alice and Bob to a single Bell-pair, where α (abstractly) characterizes the noise levels in the channels connecting Eve to Alice and Bob, and also the noise of the local devices, and l indicates that the noise transcripts leak to Eve. The ideal map can intuitively be thought of as a map which simulates a real protocol as follows. In the case of an abort, it replaces the final state with a fixed state $\sigma^{\perp}_{ABE}$. In the non-aborting case, however, it replaces the actual output with a special state $\sigma^{\alpha,\mathcal{P},l}_{ABE}$, which corresponds to the output of a real protocol where the noise transcripts leak to Eve, utilizing distillation protocol $\mathcal{P}$, that was successfully run with asymptotically many high-fidelity i.i.d. initial pairs. This is the best the noisy entanglement distillation protocol $\mathcal{P}$ could ever do. As we show later, $\sigma^{\alpha,\mathcal{P},l}_{ABE}$ is a well-defined state for the entanglement distillation protocols and noise parameters considered here. That is, it depends on the local noise parameters *only*, and not the initial states. Formally, we have for a given real map (that is, the map realized by the execution of a real protocol)

$$(\mathcal{E}^{\alpha,l} \otimes \mathrm{id}_E)(|\psi\rangle\langle\psi|_{ABE}) = p_\rho \sigma_{ABE} \otimes |\mathrm{ok}\rangle\langle\mathrm{ok}|_f + (1 - p_\rho)\sigma^{\perp}_{ABE} \otimes |\mathrm{fail}\rangle\langle\mathrm{fail}|_f \tag{1}$$

a corresponding ideal map

$$(\mathcal{F}^{\alpha,l} \otimes \mathrm{id}_E)(|\psi\rangle\langle\psi|_{ABE}) = p_\rho \sigma^{\alpha,\mathcal{P},l}_{ABE} \otimes |\mathrm{ok}\rangle\langle\mathrm{ok}|_f + (1 - p_\rho)\sigma^{\perp}_{ABE} \otimes |\mathrm{fail}\rangle\langle\mathrm{fail}|_f, \tag{2}$$

where $|\psi\rangle_{ABE}$ is a purification of the initial n-partite ensemble $\rho^{(n)}_{AB}$ provided by Eve, p_ρ is the success probability depending on the initial state $\rho^{(n)}_{AB}$, and $\sigma^{\perp}_{ABE}$ is a fixed state output if the protocol is aborted. Observe that the corresponding success probabilities p_ρ, per definition, are identical for the real and ideal maps $\mathcal{E}^{\alpha,l}$ and $\mathcal{F}^{\alpha,l}$ in (1) and (2) respectively. The two-level flag system f distinguishes the accepting and aborting branches. The state $\sigma^{\alpha,\mathcal{P},l}_{ABE}$ is the asymptotic state of the entanglement distillation protocol $\mathcal{P}$ and is of the form

$$\sigma^{\alpha,\mathcal{P},l}_{ABE} = \left(\sum_{i,j=0}^{1} \omega_{ij}(\alpha, \mathcal{P})|B_{ij}\rangle\langle B_{ij}|_{AB} \otimes |\eta_{ij}\rangle\langle\eta_{ij}|_E\right) \otimes \sigma_E, \tag{3}$$

where $|\eta_{ij}\rangle$ are the leaked noise transcripts of Eve, $|B_{ij}\rangle = (\mathrm{id} \otimes \sigma^j_x \sigma^i_z)|B_{00}\rangle$ the Bell-basis states, and $\omega_{ij}(\alpha, \mathcal{P})$ are probabilities which depend on the noise level of the local devices and the entanglement distillation protocol $\mathcal{P}$. For instance, if the local devices are perfect, then $\omega_{ij} = 1$ if and only if $i = j = 0$, hence AB contains a perfect Bell-pair. Finally, the states $|\eta_{ij}\rangle$ specify the sequences of noise operations, and are orthogonal for different i, j. If the noise transcripts are not leaked to Eve, we denote the ideal protocol by $\mathcal{F}^{\alpha}$. In that case, $|\eta_{ij}\rangle$ in (3) is not accessible to Eve, hence we replace $\sigma^{\alpha,\mathcal{P},l}_{ABE}$ by $\sigma^{\alpha,\mathcal{P}}_{ABE} = (\sum_{i,j}\omega_{ij}(\alpha, \mathcal{P})|B_{ij}\rangle\langle B_{ij}|_{AB}) \otimes \sigma_E$ in (2). Observe that the ideal map $\mathcal{F}^{\alpha,l}$, which mathematically defines the type of process we wish to realize, is a global operation beyond LOCC (local operations and classical communication) which can be decomposed by concatenating the real protocol $\mathcal{E}^{\alpha,l}$ and a replacement map $\mathcal{S}$ (which replaces the final state only if the real protocol succeeds according to the system f in (2)), i.e. $\mathcal{F}^{\alpha,l} = \mathcal{S} \circ \mathcal{E}^{\alpha,l}$.

An entanglement distillation protocol (together with the noise maps), given as a CPTP map $\mathcal{E}^{\alpha,(l)}$, is confidential if it is close to the ideal map:

Definition 1. The protocol $\mathcal{E}^{\alpha,(l)}$ is ε-confidential, if

$$\|(\mathcal{E}^{\alpha,(l)} \otimes \mathrm{id}_E - \mathcal{F}^{\alpha,(l)} \otimes \mathrm{id}_E)(|\psi\rangle\langle\psi|_{ABE})\|_1 \leqslant \varepsilon \tag{4}$$

holds for all initial states $|\psi\rangle_{ABE}$, where $\|\rho\|_1 = \mathrm{tr}\sqrt{\rho\rho^\dagger}$ is the operator 1-norm for a density operators ρ.

The system E above may contain any purification of the initial states Eve provided.

In this work, we use the term security in a generic sense, and the precise meaning depends on the context. For instance, in QKD applications, security means that Alice and Bob establish a perfectly random and secret key which the adversary has negligible information about [5, 6, 9, 35–37]. In recent times, composable security definitions have become commonplace, in which, roughly speaking, security is defined via an ideal process, and security level via the amount by which the process realized by the protocol deviates from the ideal process. In the context of QKD, this distance reduces to the distance on the generated final states of the ideal versus realized protocol. The ideal protocol outputs a completely mixed state on Alice and Bobs system which is in tensor product with Eve. More formally, see also [9], a QKD protocol $\mathcal{Q}$ is said to be ε-secure for initial state ρ_{ABE} if

$$\|\sigma_{S_A S_B CE} - \sigma_{SS} \otimes \sigma_{CE}\|_1 \leqslant \varepsilon \tag{5}$$

holds where $\sigma_{S_A S_B CE} = (\mathcal{Q} \otimes \mathrm{id}_E)(\rho_{ABE})$, S_A and S_B denote the output systems of Alice and Bob (corresponding the generated key), C denotes the classical communication and $\sigma_{SS} = 1/|S| \sum_{s \in S} |s\rangle\langle s| \otimes |s\rangle\langle s|$ for orthogonal states s. The state $\sigma_{SS} \otimes \sigma_{CE}$ corresponds to the output of the ideal protocol.

The confidentiality criterion which we introduce here follows the distance-on-maps approach introduced in the context of QKD like in e.g. [8]. Observe that such an approach is especially tailored to compose different protocols, as the confidentiality definition concerns the distance of the real process with respect to an ideal process. Therefore the real and ideal maps $\mathcal{E}^{\alpha,(l)}$ and $\mathcal{F}^{\alpha,(l)}$ respectively are motivated by abstracting the protocol in terms of processes. It is straightforward to abstract and define the ideal map in terms of input and output relations, reflecting an ideal entanglement distillation process. As we discuss above, the ideal protocol has an ok- and fail-branch. The fail-branch corresponds to the case whenever Alice and Bob abort the procedure, outputting the state $\sigma^{\perp}_{ABE}$. However, if the procedure succeeds then we might think of the ideal map as running the entanglement distillation protocol for infinitely many initial states, ending up in the fixed state $\sigma^{\alpha,\mathcal{P},l}_{ABE}$ of the entanglement distillation protocol $\mathcal{P}$ for noise level α. We observe two important facts regarding that particular state: first, its the best the entanglement distillation protocol $\mathcal{P}$ can do in the presence of noise of level α, and second, as Eve is disentangled from Alice and Bob, this state is useful for applications like quantum teleportation. Hence we refer to this state also as a private state, or equivalently, Alice and Bob share private entanglement. In contrast to (5), the target state $\sigma^{\alpha,\mathcal{P},l}_{ABE}$ in the ok-branch is only in tensor product with respect to Eve if the noise transcripts do not leak to the adversary. In that case a secure quantum channel is feasible in terms of quantum teleportation. Otherwise, that is if the noise transcripts $|\eta_{ij}\rangle$ leak to Eve, she is in a separable state with respect to Alice and Bob, but still enabling for confidential applications. By confidential we mean here that when the final state is used for quantum teleportation no information about the teleported state is leaked, but the final state does not guarantee that Eve cannot change the teleported state. This observation motivates the term confidentiality rather than security.

The classical communication is not correlated to the output of the real protocol, thus it can be ignored, see appendix A for details. The robustness of the protocol[5] is considered in appendix E, which enables us to assume for the subsequent analysis that all basic distillation steps succeed.

4. Main contribution

We summarize the main findings of our paper as follows: recurrence-type entanglement distillation protocols prepended by a symmetrization and a system discarding step enable confidentiality, provided that the noise transcripts do not leak to the adversary for all noise levels α for which distillation would be possible in the i.i.d. case. We also show that this alone implies that the final state in the accepting branch, is close to a tensor product state—Eve is factored out. The results regarding the BBPSSW protocol [28] are analytic whereas for the DEJMPS protocol [19] the results rely on strong numerical evidence. For low noise rates, we achieve better results via the post-selection-based reduction. In that case, no system discarding step is necessary. Finally we find that if an entanglement distillation protocol is confidential when the noise transcripts do not leak, then it also confidential if they do leak to the adversary. In particular, even in the case that Eve picks up information about all the realized noise processes during the protocol, the final output system still enables confidential quantum applications like e.g. quantum teleportation. The paper proceeds as follows. We establish necessary conditions to guarantee confidentiality for recurrence-type entanglement distillation protocols restricted to i.i.d. inputs whenever the noise transcripts are not leaked to Eve. Then, we generalize this to arbitrary initial states via the de-Finetti theorem [25]. Next, we use them to prove the confidentiality criterion (4) for entanglement distillation protocols where the noise transcripts are not leaked. Finally, this will be used to derive the confidentiality bound whenever the noise transcripts are leaked.

[5] The robustness is quantified by the abort probability in the all-honest, but noisy setting.

5. Confidentiality of entanglement distillation protocols

5.1. Entanglement distillation for i.i.d inputs

The basic step of a recurrence-type entanglement distillation protocol is summarized as follows: Alice and Bob share two noisy Bell-pairs, i.e. both have two qubits, each representing a 'half' of a noisy Bell pair, and they first apply local operations to their respective parts of the Bell-pairs; next, they measure one Bell-pair and classically communicate their outcomes. Depending on the entanglement distillation protocol and the outcomes they either keep or discard the unmeasured pair. The basic step is applied to all pairs of the initial states, which comprises one distillation round. This distillation round is iterated where output states of the previous round are used as inputs for the next round. In the limit, a noiseless entanglement distillation protocol outputs a perfect Bell-pair (implying that Eve is factored out).

Here, we allow for any type of noise acting (independently) on the single- and two-qubit gates appearing in the protocol[6]. Using the results of [26], by utilizing random basis changes and adding additional noise, any such general noise can be brought to a standard form: depolarizing noise for imperfect single- and two-qubit CNOT-type operations, see appendix A. Thus, it is sufficient to address noise in such standard form.

For such noise, one can analytically show [27] that for the BBPSSW protocol [28], there exists a unique attracting fixed point of the protocol which only depends on the noise parameters. That is, whenever the fidelity of the initial states is above some minimum fidelity $F_{\min}$, depending on the noise parameters, the protocol converges towards that unique fixed point which we denote by $\sigma_{AB}^{\alpha;\mathrm{B}}$. Observe that $\sigma_{AB}^{\alpha;\mathrm{B}}$ is related to $\sigma_{ABE}^{\alpha,\mathcal{P},l}$ of (3) by letting $\mathcal{P} = \mathrm{B}$ and tracing out Eves system, i.e. $\sigma_{AB}^{\alpha;\mathrm{B}} = \mathrm{tr}_E[\sigma_{ABE}^{\alpha,\mathrm{B},l}]$. In particular, we mean by $\mathcal{P} = \mathrm{B}$ that the BBPSSW protocol is used for entanglement distillation. We find that the output state σ_{AB}^{N}, where $N = \log_2 n$ denotes the number of successfully completed distillation layers, satisfies $\|\sigma_{AB}^{N} - \sigma_{AB}^{\alpha;\mathrm{B}}\|_1 \leqslant \epsilon_{\mathrm{B}}$, where ϵ_{B} is a function of N, and it holds that $\epsilon_{\mathrm{B}} \leqslant F(n) \in O(n^{-b_{\mathrm{B}}(\alpha)})$ and $0 < b_{\mathrm{B}}(\alpha) \leqslant \log_2 3 - 1$.

For the entanglement distillation protocol of Deutsch *et al* [19] (referred to as the DEJMPS protocol) the fixed point analysis is more complicated. In the noiseless case, DEJMPS was proven to have a unique attracting fixed point [29]. For the noisy case, we can only provide extensive numerical evidence that there exists a unique attracting fixed point, depending on the noise parameters only which we denote by $\sigma_{AB}^{\alpha;\mathrm{D}}$, see A.1. Again, observe that $\sigma_{AB}^{\alpha;\mathrm{D}}$ is related to $\sigma_{ABE}^{\alpha,\mathcal{P},l}$ of (3) by setting $\mathcal{P} = \mathrm{D}$ and tracing out Eves system, i.e. $\sigma_{AB}^{\alpha;\mathrm{D}} = \mathrm{tr}_E[\sigma_{ABE}^{\alpha,\mathrm{D},l}]$. We numerically find that for the state σ_{AB}^{N} obtained after successfully completing $N = \log_2 n$ layers of distillation that $\|\sigma_{AB}^{N} - \sigma_{AB}^{\alpha;\mathrm{D}}\|_1 \leqslant \epsilon_{\mathrm{D}}$ where ϵ_{D} is a function of N, and it holds that $\epsilon_{\mathrm{D}} \leqslant F(n) \in O(n^{-b_{\mathrm{D}}(\alpha)})$. $b_{\mathrm{D}}(\alpha)$ is a positive function. We note that a similar analysis, but also with analytic findings for the noiseless DEJMPS protocol was first performed in [29].

We reiterate that we assume for our analysis that all basic distillation steps succeed, since we deal with failures due to the entanglement distillation protocol with a quadratic overhead in terms of initial states, see appendix E.

The final state of the entanglement distillation protocol $\mathcal{P}$ in the ok-branch, σ_{AB}, depends on whether the parameter estimation on $\sqrt{n}$ initial states was accurate or not. The latter occurs with an exponentially small probability in terms of initial states, see the discussion of the robustness of the protocol in appendix E. This in turn implies that the parameter estimation was accurate with probability exponentially close to unity. Therefore the results regarding n i.i.d. initial states as input to the distillation protocol $\mathcal{P}$ above imply that

$$p_\rho \|\sigma_{AB} - \sigma_{AB}^{\alpha;\mathcal{P}}\|_1 \leqslant \epsilon_{\mathcal{P}}(n) + 2p_{\mathrm{PE}} \leqslant \epsilon'_{\mathcal{P}}(n) =: \varepsilon_{\mathcal{P}}(n + \sqrt{n}), \tag{6}$$

where $p_{\mathrm{PE}} \in O(\exp(-\sqrt{n}))$ for all i.i.d. inputs $\rho_{AB}^{\otimes n+\sqrt{n}}$. This equation attains exactly the same form for both protocols with the difference in the labels, so if we substitute $\mathcal{P}$ with B (by writing, for example $\epsilon_{\mathrm{B}}(n)$) we refer to the BBPSSW protocol, where substituting $\mathcal{P}$ with D refers to the DEJMPS protocol. In similar fashion we refer from now by $\epsilon_{\mathcal{P}}(n)$ to $\epsilon'_{\mathcal{P}}(n)$ for the sake of clarity. So to summarize, the distance for $n + \sqrt{n}$ i.i.d. initial states in the ok-branch of the protocol is bounded by $\varepsilon_{\mathcal{P}}(n + \sqrt{n})$.

Since, in the abort case, the outputs of the overall protocol $\mathcal{E}^\alpha$ and the ideal protocol $\mathcal{F}^\alpha$ are identical we obtain that

$$\|(\mathcal{E}^\alpha - \mathcal{F}^\alpha)(\rho_{AB}^{\otimes n})\|_1 = p_\rho \|\sigma_{AB} - \sigma_{AB}^{\alpha;\mathcal{P}}\|_1 \leqslant \varepsilon_{\mathcal{P}}(n), \tag{7}$$

where the probability p_ρ depends on the initial state ρ for both protocols and corresponds to the probability of parameter estimation succeeding and completing $\log_2(n - \sqrt{n})$ distillation layers successfully for initial state ρ. Hence, in both cases, the final distance to the respective fixed points scales polynomial in terms of n.

The functions $b_{\mathrm{B}}(\alpha)$ and $b_{\mathrm{D}}(\alpha)$ of the local noise level α govern the rate of convergence of the real protocol to the ideal protocol in the i.i.d case for entanglement distillation protocols. We numerically found that these functions monotonically increase as the local noise rate α tends to zero appendix A. Thus, increasing the fidelity

[6] We assume that the noise characteristics of the quantum gates are constant throughout the protocol.

of local devices (through e.g. fault tolerance) directly influences the rate of convergence, which in turn governs the confidentiality level.

In contrast to $b_{\mathrm{B}}(\alpha)$, the function $b_{\mathrm{D}}(\alpha)$ is not upper bounded, which implies that for certain noise parameters α the DEJMPS protocol needs to perform fewer distillation rounds than the BBPSSW protocol to achieve the required confidentiality levels. This fast convergence is crucial for the powerful post-selection technique [8] for non i.i.d. initial states, which is not applicable for the BBPSSW protocol.

Now we use the established fixed point properties of entanglement distillation protocols for i.i.d. initial states to show that similar results hold for arbitrary initial states.

5.2. Entanglement distillation for arbitrary inputs

In generalizing the previous results to arbitrary initial states we make use of the de Finetti theorem [25]. The basic de-Finetti results guarantee that the reduced state $\mathrm{tr}_{n-k}(\rho_{AB}^{(n)})$ of a permutation-invariant n-partite state $\rho_{AB}^{(n)}$ is close to an i.i.d state $\int \sigma_{AB}^{\otimes k}\, \mathrm{d}\sigma$, with distance which scales as $O(k/n)$. This enables the following lemma.

Lemma 2. *Let $n,\ k \in \mathbb{N}$ where $k \leqslant n$. Furthermore, let $\mathcal{E}^{s\&t}$ be the real protocol and $\mathcal{F}^{s\&t}$ the ideal protocol including symmetrization and the tracing out of $n - k$ pairs. Moreover, let ρ_{AB} be a bipartite mixed state of n systems shared by Alice and Bob and let $\mathcal{E}$ and $\mathcal{F}$ denote the real and ideal protocol after symmetrization and tracing out $n - k$ pairs. Then*

$$\|\mathcal{E}^{s\&t}(\rho_{AB}) - \mathcal{F}^{s\&t}(\rho_{AB})\|_1 \leqslant \frac{64k}{n} + \max_{\mu_{AB}} \|\mathcal{E}(\mu_{AB}^{\otimes k}) - \mathcal{F}(\mu_{AB}^{\otimes k})\|_1. \tag{8}$$

Proof. Let ρ_{AB} be a mixed state. After Alice and Bob apply a symmetrization they share a permutation invariant state $\tilde{\rho}_{AB}$. Thus we can apply theorem II.7 of [25] and have for $\xi_{AB}^k := \mathrm{tr}_{n-k}[\tilde{\rho}_{AB}]$ the inequality $\|\xi_{AB}^k - \int \mu_{AB}^{\otimes k} \mathrm{d}m(\mu_{AB})\|_1 \leqslant 32k/n$ for some probability measure m on the set of mixed states on AB. Moreover we note that $\mathcal{E}$ and $\mathcal{F}$ are CPTP maps. We define $\tau_k := \int \mu_{AB}^{\otimes k} \mathrm{d}m(\mu_{AB})$. A straightforward computation shows

$$\begin{aligned}
\|\mathcal{E}^{s\&t}(\rho_{AB}) - \mathcal{F}^{s\&t}(\rho_{AB})\|_1 &= \|\mathcal{E}(\xi_{AB}^k) - \mathcal{F}(\xi_{AB}^k)\|_1 \leqslant \|\mathcal{E}(\xi_{AB}^k) - \mathcal{E}(\tau_k)\|_1 + \|\mathcal{E}(\tau_k) - \mathcal{F}(\xi_{AB}^k)\|_1 \\
&\leqslant \|\mathcal{E}(\xi_{AB}^k) - \mathcal{E}(\tau_k)\|_1 + \|\mathcal{E}(\tau_k) - \mathcal{F}(\tau_k)\|_1 + \|\mathcal{F}(\tau_k) - \mathcal{F}(\xi_{AB}^k)\|_1 \\
&\leqslant 2\|\tau_k - \xi_{AB}^k\|_1 + \|\mathcal{E}(\tau_k) - \mathcal{F}(\tau_k)\|_1 \leqslant \frac{64k}{n} + \left\|(\mathcal{E} - \mathcal{F})\left(\int \mu_{AB}^{\otimes k} \mathrm{d}m(\mu_{AB})\right)\right\|_1 \\
&\leqslant \frac{64k}{n} + \max_{\mu_{AB}} \|(\mathcal{E} - \mathcal{F})(\mu_{AB}^{\otimes k})\|_1,
\end{aligned}$$

which completes the proof. □

Therefore the application of the de-Finetti theorem introduces an additive term $\frac{64k}{n}$ when reducing arbitrary initial states to i.i.d. initial states. As the right hand side of (8) is independent of the initial state ρ_{AB}, (8) holds for all initial states ρ_{AB}.

In (8) we have omitted the superscript α characterizing the noise level, and we will use it only if it is specifically needed. Inequality (8) implies that the properties of the fixed point (uniqueness, attractivity, noise-dependence) also hold for arbitrary initial states, if the protocol is prepended by symmetrization and a trace-out step. This enables us to prove the confidentiality criterion of definition 1 for entanglement distillation protocols, where the noise transcripts of L are not leaked, which will, in turn, imply the confidentiality criterion (4) whenever the noise transcripts are leaked.

5.3. Confidentiality of entanglement distillation protocols

The inequality in (7) establishes the local properties of the protocol, and is more-or-less typical for studies of the convergence of entanglement distillation protocols in the i.i.d. case. However, it falls short of the complete characterization captured by the confidentiality criterion (4) in two ways: first, the input states are restricted (i.i.d.); second, it fails to consider the purifying system of Eve[7], vital in cryptographic contexts. While the prior issue is the subject of de-Finetti and post-selection-type reductions, the latter issue can be a problem in general, as small distance of corresponding subsystems does not imply a small distance of the total systems.

However, we can resolve this issue by using the fixed point properties of entanglement distillation protocols. More precisely, we relate the two distances by the following general lemma, proven in the appendix B.1.

[7] Technically, inequality (7) is a statement about the operator norm-induced distance on maps, where expression of (4) is the completely bounded diamond norm, relevant for security statements.

Lemma 3. *Let ρ be an arbitrary mixed state shared by Alice and Bob and let $|\psi\rangle_{ABE}$ be a purification thereof held by Eve. Furthermore, let $\mathcal{P}_1$ correspond to a (distillation-type) real protocol and $\mathcal{P}_2$ correspond to the associated (distillation-type) ideal protocol, i.e.*

$$\mathcal{P}_1(\rho) = p_\rho \sigma_{AB} \otimes |\mathrm{ok}\rangle\langle\mathrm{ok}| + (1 - p_\rho)\sigma^{\perp}_{AB} \otimes |\mathrm{fail}\rangle\langle\mathrm{fail}|,$$
$$\mathcal{P}_2(\rho) = p_\rho \sigma^{\alpha}_{AB} \otimes |\mathrm{ok}\rangle\langle\mathrm{ok}| + (1 - p_\rho)\sigma^{\perp}_{AB} \otimes |\mathrm{fail}\rangle\langle\mathrm{fail}|,$$

where α characterizes the level of the noise, σ^{α}_{AB}, and $\sigma^{\perp}_{AB}$ are two fixed two qubit states. Furthermore, let $\mathcal{P}_1$ and $\mathcal{P}_2$ satisfy the following properties:

(1) The noise transcripts do not leak to Eve.

(2) The protocol $\mathcal{P}_1$ guarantees to converge towards some state σ^{α}_{AB} within the ok-branch of the protocol and $\max_{\mu_{AB}} \|(\mathcal{P}_1 - \mathcal{P}_2)(\mu_{AB})\|_1 \leqslant \varepsilon$.

Then it holds that

$$\|(\mathcal{P}_1 \otimes \mathrm{id}_E - \mathcal{P}_2 \otimes \mathrm{id}_E)(|\psi\rangle\langle\psi|_{ABE})\|_1 \leqslant (34 \cdot 4^8 + 1)\varepsilon. \tag{9}$$

The factor $34 \cdot 4^8 + 1$ arises as an upper bound on the distance of the given states from states in product form based on the notion of non-steerability we introduce (see appendix B.1 for details). In our computations we managed to prove the key lemma in a manner which is proportional to the dimension of the systems, more precisely, the overall size of the corresponding density matrix. It may be the case that the bound of lemma 3 could hold without the dependence on the system size (and indeed, with smaller constants), however this was not necessary for our purposes.

Lemma 3 is vital as it allows us to employ the de-Finetti theorem [25]. Hence, for the protocols $\mathcal{E}^{s\&t}$ and $\mathcal{F}^{s\&t}$, by combining lemma 2 with lemma 3, we obtain the following theorem.

Theorem 4 (de-Finetti-based reduction technique). *Let $\mathcal{E}^{s\&t}$ be the real protocol and $\mathcal{F}^{s\&t}$ the ideal protocol including symmetrization and the tracing out of $n - k$ pairs, taking n input pairs and $k \leqslant n$ and utilizing entanglement distillation protocol $\mathcal{P}$. Then we have*

$$\max_{|\psi\rangle_{ABE}} \|(\mathcal{E}^{s\&t} \otimes \mathrm{id}_E)(|\psi\rangle\langle\psi|) - (\mathcal{F}^{s\&t} \otimes \mathrm{id}_E)(|\psi\rangle\langle\psi|)\|_1 \leqslant (34 \cdot 4^8 + 1)\left(\frac{64k}{n} + \varepsilon_{\mathcal{P}}(k)\right), \tag{10}$$

where $\varepsilon_{\mathcal{P}}(k)$ denotes the maximum distance of the real and ideal protocol without symmetrization and tracing out step using entanglement distillation protocol $\mathcal{P}$ in the ok-branch for k i.i.d. initial states, i.e. equation (7).

Proof. Suppose Eve prepares a purification $|\psi\rangle_{ABE}$ of the state ρ_{AB} shared by Alice and Bob. Recall that the real and ideal protocol including symmetrization and the tracing out of $n - k$ pairs applied to initial state ρ_{AB} read as

$$\mathcal{E}^{s\&t}(\rho_{AB}) = p_\rho \sigma_{AB} \otimes |\mathrm{ok}\rangle\langle\mathrm{ok}| + (1 - p_\rho)\sigma^{\perp}_{AB} \otimes |\mathrm{fail}\rangle\langle\mathrm{fail}|,$$
$$\mathcal{F}^{s\&t}(\rho_{AB}) = p_\rho \sigma^{\alpha,\mathcal{P}}_{AB} \otimes |\mathrm{ok}\rangle\langle\mathrm{ok}| + (1 - p_\rho)\sigma^{\perp}_{AB} \otimes |\mathrm{fail}\rangle\langle\mathrm{fail}|$$

and observe that we have for the initial state ρ_{AB} by lemma 2 that

$$\|(\mathcal{E}^{s\&t} - \mathcal{F}^{s\&t})(\rho_{AB})\|_1 = p_\rho\|\sigma_{AB} - \sigma^{\alpha,\mathcal{P}}_{AB}\|_1 \leqslant \left(\frac{64k}{n} + \max_{\mu_{AB}} \|(\mathcal{E} - \mathcal{F})(\mu^{\otimes k}_{AB})\|_1\right), \tag{11}$$

where $\mathcal{E}$ and $\mathcal{F}$ denote the real and ideal protocol after symmetrization and tracing out $n - k$ pairs. Since the right-hand side of (11) is independent of the initial state ρ_{AB} it holds for all initial states of the protocol. Therefore, the properties of the fixed point (unique, attracting and depending on the noise parameters only) translate from i.i.d. initial states to arbitrary initial states. Hence the protocol guarantees that it converges towards the fixed point of the entanglement distillation protocol.

Additionally, by inserting (7) in (11) we find

$$\|(\mathcal{E}^{s\&t} - \mathcal{F}^{s\&t})(\rho_{AB})\|_1 \leqslant \left(\frac{64k}{n} + \varepsilon_{\mathcal{P}}(k)\right). \tag{12}$$

This implies that the real protocol indeed converges towards the fixed point, and, thus we can apply lemma 3 to the protocols $\mathcal{E}^{s\&t}$ and $\mathcal{F}^{s\&t}$ for the purification $|\psi\rangle_{ABE}$ of ρ_{AB} and we find by using (12) that

$$\|(\mathcal{E}^{s\&t} \otimes \mathrm{id}_E)(|\psi\rangle\langle\psi|) - (\mathcal{F}^{s\&t} \otimes \mathrm{id}_E)(|\psi\rangle\langle\psi|)\|_1 \leqslant (34 \cdot 4^8 + 1)\left(\frac{64k}{n} + \varepsilon_{\mathcal{P}}(k)\right). \tag{13}$$

Taking the maximum in (13) completes the proof. □

Thus, we can reach arbitrary confidentiality levels, however at the cost of wasting some pairs. The scaling of the confidentiality parameter, i.e. the right-hand side of (10), is linear in the number of initial states n, due to the use of the 'basic' de Finetti approach.

If the local noise is low, we can do better in terms of scaling and efficiency, using the post-selection technique [8]. For that purpose, we first establish a result similar to (9) by using the fact that the resulting state of the protocol, including L, is pure, see appendix A. More precisely, we have the following lemma, proven in the appendix B.2.

Lemma 5. *Let $\mathcal{E}$ be the real protocol which guarantees to converge towards a unique and attracting fixed point depending on the noise parameter only and let $\mathcal{F}$ be the ideal protocol. Furthermore let* ρ *be a mixed state (consisting of* n *systems) shared by Alice and Bob. If the extension of $\mathcal{E}$ and $\mathcal{F}$ to the system of* L *satisfies* $\|\mathcal{E}_L(\rho) - \mathcal{F}_L(\rho)\|_1 \leqslant \varepsilon(n)$, *then*

$$\|(\mathcal{E} \otimes \mathrm{id}_{E'})(|\psi\rangle\langle\psi|_{ABE'}) - (\mathcal{F} \otimes \mathrm{id}_{E'})(|\psi\rangle\langle\psi|_{ABE'})\|_1 \leqslant 4\sqrt{\varepsilon(n)}$$

for all purifications $|\psi\rangle_{ABE'}$ of ρ.

This lemma allows us to prove the closeness on any purification from the closeness of the reduced systems, and finally to derive confidentiality from the performance of the ideal protocol via the following theorem.

Theorem 6 (Post-selection-based reduction technique). *Let $\mathcal{E}^s$ be the real protocol and $\mathcal{F}^s$ the ideal protocol preceded by a symmetrization step operating on n input pairs. Furthermore let* $\max_{\mu_{AB}} \|\mathcal{E}(\mu_{AB}^{\otimes n}) - \mathcal{F}(\mu_{AB}^{\otimes n})\|_1 \leqslant \varepsilon_{\mathcal{P}}(n)$, *see (7), where $\mathcal{E}$ and $\mathcal{F}$ denote the sub-protocols after symmetrization (i.e. the protocols without the symmetrization step) and $\mathcal{P}$ the entanglement distillation protocol. Then we have*

$$\max_{|\psi\rangle_{ABE'}} \|(\mathcal{E}^s \otimes \mathrm{id}_{E'})(|\psi\rangle\langle\psi|) - (\mathcal{F}^s \otimes \mathrm{id}_{E'})(|\psi\rangle\langle\psi|)\|_1 \leqslant 4\sqrt{2}\, g_{n,d} \sqrt[4]{\varepsilon_{\mathcal{P}}(n)}\,, \tag{14}$$

where $g_{n,d} = \binom{n+15}{n}$.

Proof. We observe that $\mathcal{E}^s$ and $\mathcal{F}^s$ are permutation invariant maps due to the symmetrizazion step. Thus we can apply the post-selection technique of [8] which implies

$$\begin{aligned}\max_{|\psi\rangle_{ABE'}} &\|(\mathcal{E}^s \otimes \mathrm{id}_{E'})(|\psi\rangle\langle\psi|) - (\mathcal{F}^s \otimes \mathrm{id}_{E'})(|\psi\rangle\langle\psi|)\|_1 \\ &\leqslant g_{n,d}\|(\mathcal{E}^s \otimes \mathrm{id}_{E'})(|\tau\rangle\langle\tau|_{ABE'}) - (\mathcal{F}^s \otimes \mathrm{id}_{E'})(|\tau\rangle\langle\tau|_{ABE'})\|_1\,,\end{aligned} \tag{15}$$

where $|\tau\rangle_{ABE'}$ is a purification of the de-Finetti Hilbert–Schmidt state, hence $\mathrm{tr}_{E'}[|\tau\rangle\langle\tau|_{ABE'}] = \int \mu_{AB}^{\otimes n} \mathrm{d}\eta(\mu) =: \tau'$ where η is the measure induced by the Hilbert–Schmidt metric on $\mathrm{End}(\mathbb{C}^4)$. Furthermore, we note that we have for the extensions of $\mathcal{E}^s$ and $\mathcal{F}^s$ to L, i.e. the maps $\mathcal{E}^s_L$ and $\mathcal{F}^s_L$, that

$$\|\mathcal{E}^s_L(\tau') - \mathcal{F}^s_L(\tau')\|_1 = \left\|(\mathcal{E}^s_L - \mathcal{F}^s_L)\left(\int \mu_{AB}^{\otimes n} \mathrm{d}\eta(\mu)\right)\right\|_1 \leqslant \max_{\mu_{AB}} \|(\mathcal{E}_L - \mathcal{F}_L)(\mu_{AB}^{\otimes n})\|_1. \tag{16}$$

According to appendix A.1.1, which implies that the distance including L scales as the square root of the 1-norm induced distance without L, i.e. Alice and Bob only, we find for (16) by using the assumption $\max_{\mu_{AB}} \|\mathcal{E}(\mu_{AB}^{\otimes n}) - \mathcal{F}(\mu_{AB}^{\otimes n})\|_1 \leqslant \varepsilon_{\mathcal{P}}(n)$ that

$$\|(\mathcal{E}_L - \mathcal{F}_L)(\mu_{AB}^{\otimes n})\|_1 \leqslant 2\sqrt{\|(\mathcal{E} - \mathcal{F})(\mu_{AB}^{\otimes n})\|_1} \leqslant 2\sqrt{\varepsilon_{\mathcal{P}}(n)}\,. \tag{17}$$

As $|\tau\rangle_{ABE'}$ is a purification of τ' we can apply lemma 5 which gives, for (15),

$$\begin{aligned}\max_{|\psi\rangle_{ABE'}} \|(\mathcal{E}^s \otimes \mathrm{id}_{E'})(|\psi\rangle\langle\psi|) - (\mathcal{F}^s \otimes \mathrm{id}_{E'})(|\psi\rangle\langle\psi|)\|_1 &\leqslant 4g_{n,d}\sqrt{\max_{\mu_{AB}} \|(\mathcal{E}_L - \mathcal{F}_L)(\mu_{AB}^{\otimes n})\|_1} \\ &\leqslant 4g_{n,d}\sqrt{2\sqrt{\varepsilon_{\mathcal{P}}(n)}} \\ &= 4\sqrt{2}\, g_{n,d} \sqrt[4]{\varepsilon_{\mathcal{P}}(n)}\,,\end{aligned}$$

which completes the proof. □

Observe that $\varepsilon_{\mathcal{P}}(n)$, which governs the rate of convergence of the overall protocol, relates to the rate of convergence of the entanglement distillation protocol $\mathcal{P}$ via $\varepsilon_{\mathcal{P}}(n) = \epsilon_{\mathcal{P}}(n - \sqrt{n})$, as $\sqrt{n}$ initial states are used for parameter estimation. We remind the reader that the preprocessing steps (symmetrization, tracing out) of the entanglement distillation protocol and the lemmas of this section are non-trivial and crucial for the proof of the de-Finetti-based and post-selection-based reduction technique.

Furthermore we point out that the proof regarding the BBPSSW protocol is analytic and necessarily relies on the de-Finetti-based reduction technique because of its slow convergence rate. The rate of convergence for the BBPSSW protocol can easily be derived, see appendix A for details. For the DEJMPS protocol it turns out that we have polynomial scaling depending on the noise parameter α, i.e. $\max_{\sigma_{AB}} \|(\mathcal{E}-\mathcal{F})(\sigma_{AB}^{\otimes n})\|_1 \leq \varepsilon_{\mathrm{D}}(n) \leqslant O(n^{-b_{\mathrm{D}}(\alpha)})$, see (7).

However, the protocol needs to converge sufficiently quickly, as the post-selection technique incurs a multiplicative increase in the effective distance between real and ideal protocols, which scales as a (15 degree) polynomial in n, see (14). The resulting confidentiality level scales therefore as $O(n^{15-b_{\mathrm{D}}(\alpha)/4})$, which leads to an acceptable noise level that is rather low, e.g. about 10^{-19} for the DEJMPS protocol in the setting of binary pairs[8], see appendix A.1.1. This very low rate is due to the polynomial factor introduced by applying the post-selection technique, i.e. $g_{n,d}$ in (14) with $d = 4$. Observe that these small rates are determined by properties of recurrence-type entanglement distillation protocols, i.e. $b(\alpha)$ for the recurrence-type entanglement distillation protocols studied here, and may be improved by either considering hashing-type protocols [30] or through fault-tolerant constructions. Indeed, the noise threshold for fault-tolerant quantum computation also applies to this case, yielding a tolerable noise level of about 10^{-4}. We reiterate that the post-selection technique is not applicable to the BBPSSW protocol, due to its slow convergence.

5.4. Confidentiality of entanglement distillation protocols when the noise transcripts leak

Finally, we provide confidentiality guarantees for entanglement distillation protocols when the noise transcripts are leaked to Eve. For that purpose, we relate the confidentiality criterion (4) for protocols where the noise transcripts are leaked to the earlier results. More formally, we have the following theorem.

Theorem 7. *Let $\mathcal{E}$ be the real protocol and $\mathcal{F}$ be the ideal protocol satisfying the assumptions of lemma 3. Furthermore, let $\mathcal{E}^l$ denote the real and $\mathcal{F}^l$ the ideal protocol when the noise transcripts leak to Eve. Then*

$$\|(\mathcal{E}\otimes \mathrm{id}_E - \mathcal{F}\otimes \mathrm{id}_E)(|\psi\rangle\langle\psi|)\|_1 \leqslant \varepsilon(n), \text{ implies} \tag{18}$$

$$\|(\mathcal{E}^l\otimes \mathrm{id}_E - \mathcal{F}^l\otimes \mathrm{id}_E)(|\psi\rangle\langle\psi|)\|_1 \leqslant 2\sqrt{\varepsilon(n)}$$

for all purifications $|\psi\rangle_{ABE}$ of initial state ρ_{AB} consisting of n systems.

The proof, see appendix C, uses the unitary equivalence of purifications. Theorem 7 establishes via (18) that if an entanglement distillation protocol is ε-confidential according to definition 1 then the protocol is $2\sqrt{\varepsilon}$-confidential if the noisy apparatus leaks the noise transcripts.

6. Discussion

We have shown that recurrence-type entanglement distillation protocols ensure private entanglement without referring to the asymptotic limit. This holds true even when the local devices are noisy, and when the potential eavesdropper is able to completely monitor the operation of these devices in run-time (i.e., the noisy apparatus leaks information about the realized noise processes). If the noise transcripts are not leaked, Eve is 'factored out'—in tensor product with Alice and Bob, and only classically correlated otherwise. Our protocol can, for instance, be used to realize confidential quantum channels by means of teleportation—the only information that may leak to Eve after teleportation is which noise map was applied to the sent state, but nothing about the state itself (see appendix F for details). More generally, our results imply the confidentiality of the protocols in arbitrary settings (beyond the application to quantum channels), thus opening the way for the confidential realization of various quantum tasks: from establishing quantum channels and quantum networks, to applications such as distributed quantum computation. Aside from cryptographic aspects, the proposed protocol can be used to generate high quality entanglement from non-i.i.d. sources.

Acknowledgments

We acknowledge the support by the Austrian Science Fund (FWF) through the SFB FoQuS F 4012 and project P28000-N27. AP and VD are grateful to Christopher Portmann for useful discussions, comments, and advice concerning technical aspects of this work.

[8] For this simplified analysis we assumed that no parameter estimation is necessary.

Appendix A. Entanglement distillation for i.i.d. inputs

A.1. The DEJMPS protocol

We first provide an overview of the DEJMPS protocol [19] and then extend the description incrementally to our proposed setting (including L and Eve).

The DEJMPS protocol is a recurrence-type entanglement distillation protocol which combines several noisy copies of a mixed state ρ to distill a state arbitrarily close to the maximally entangled state $|B_{00}\rangle$, where $|B_{ij}\rangle = (\mathrm{id} \otimes \sigma_x^j \sigma_z^i)(|00\rangle + |11\rangle)/\sqrt{2}$ for $i \in \{0, 1\}$ and $j \in \{0, 1\}$, provided that the fidelity $F = \langle B_{00}| \rho |B_{00}\rangle$ satisfies $F > 1/2$ for the noiseless case. If the apparatus is noisy, then the minimal required fidelity F needs to satisfy $F > F_{\min}$ (where $F_{\min}$ depends on the noise level of the apparatus) to achieve distillation. For more details on recurrence-type entanglement distillation protocols in general we refer the interested reader to [31]. A basic step of the DEJMPS protocol is as follows:

Protocol 1. Basic step of the DEJMPS protocol

Require: Input state of Alice and Bob: $\rho^{(a_1,b_1)} \otimes \rho^{(a_2,b_2)}$

1: Alice and Bob apply the local basis change $U_x = \mathrm{e}^{-\mathrm{i}\pi/4\sigma_x^{(a_1)}} \otimes \mathrm{e}^{\mathrm{i}\pi/4\sigma_x^{(b_1)}} \otimes \mathrm{e}^{-\mathrm{i}\pi/4\sigma_x^{(a_2)}} \otimes \mathrm{e}^{\mathrm{i}\pi/4\sigma_x^{(b_2)}}$:

$$U_x(\rho^{(a_1,b_1)} \otimes \rho^{(a_2,b_2)})U_x^\dagger.$$

2: Alice and Bob apply a bilateral CNOT (BCNOT):

$$(\mathrm{CNOT}_{a_1\to a_2} \otimes \mathrm{CNOT}_{b_1\to b_2})\rho^{(a_1,b_1)} \otimes \rho^{(a_2,b_2)}(\mathrm{CNOT}_{a_1\to a_2} \otimes \mathrm{CNOT}_{b_1\to b_2})^\dagger.$$

3: Alice and Bob apply a $\sigma_z^{(a_2)} = \sigma_z \otimes \mathrm{id}$ and a $\sigma_z^{(b_2)} = \mathrm{id} \otimes \sigma_z$ measurement

4: Alice and Bob communicate their measurement outcomes, z_a and z_b respectively, over a classical authentic channel

5: **if** $z_a = z_b$ **then**

6: Alice and Bob keep the subsystems a_1 and b_1 of step 2

7: Alice and Bob discard the measured subsystems a_2 and b_2

8: **else**

9: Alice and Bob discard both pairs

10: **end if**

Hence, we can write one basic distillation step of the DEJMPS protocol as the linear map $O_{2-\mathrm{EPP}}(\rho \otimes \rho) = O'_{2-\mathrm{EPP}}(\rho \otimes \rho)O'^\dagger_{2-\mathrm{EPP}}$ where

$$O'_{2-\mathrm{EPP}} = (\mathrm{id}_{a_1,b_1} \otimes P_z^{(a_2)} \otimes P_z^{(b_2)})(\mathrm{CNOT}_{a_1\to a_2} \otimes \mathrm{CNOT}_{b_1\to b_2})U_x$$

modulo a normalization factor and where $P_z = |z\rangle\langle z|$, $z \in \{0, 1\}$ denotes the respective outcome of step 3 of protocol 1.

The basic step is applied to all initial pairs, which comprises one distillation round. This distillation round is iterated where output states of the previous round are used as inputs for the next round. So we summarize the DEJMPS protocol as follows:

Protocol 2. DEJMPS protocol

Require: Input state of Alice and Bob: $\bigotimes_{i=1}^{2^n} \rho^{(a_i,b_i)}$ where $F = \langle B_{00}| \rho^{(a_i,b_i)} |B_{00}\rangle > 1/2$ for all $i \in \{1, .., 2^n\}$

1: **while**Pairs left for distillation **do**

2: Apply protocol 1 to all pairs

3: Use the outputs of the previous step as input for the next distillation round

4: **end while**

We remind the reader that the recurrence relations of the protocol (i.e. update functions of the coefficients of an ensemble) are central for the convergence analysis of the DEJMPS protocol. For Bell-diagonal states, i.e. states of the form

$$\rho = p_{00}|B_{00}\rangle\langle B_{00}| + p_{11}|B_{11}\rangle\langle B_{11}| + p_{01}|B_{01}\rangle\langle B_{01}| + p_{10}|B_{10}\rangle\langle B_{10}|,$$

where $\sum_{ij} p_{ij} = 1$, $p_{ij} \geqslant 0$, a straightforward computation yields the recurrence relations for the DEJMPS protocol to be

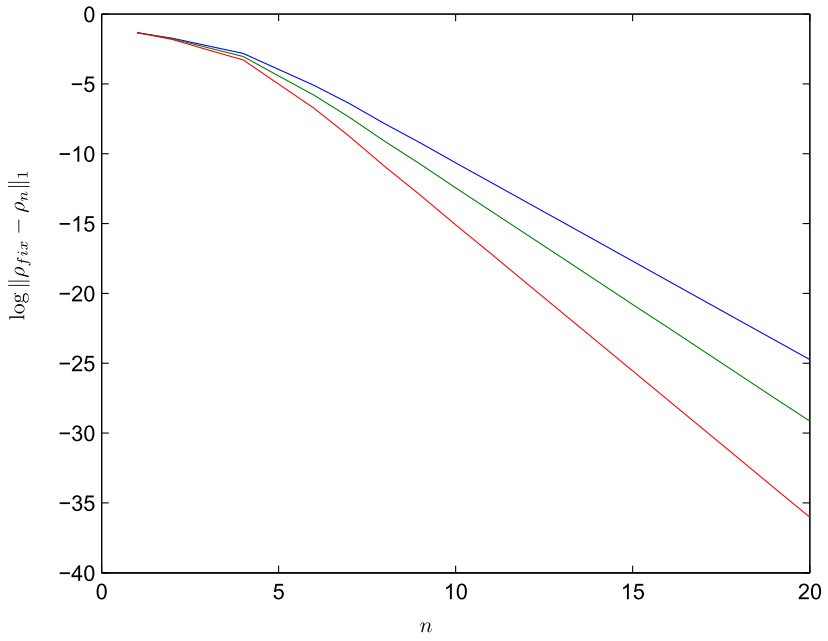

Figure A1. The figure illustrates $\log \|\rho_{\text{fix}} - \rho_n\|_1$ for different noise parameters $f = 0.97$ (blue), $f = 0.98$ (green) and $f = 0.99$ (red). The fixed point ρ_{fix} was evaluated for 500 iterations of the DEJMPS protocol.

$$\tilde{p}_{00} = \frac{p_{00}^2 + p_{11}^2}{N}, \qquad \tilde{p}_{11} = \frac{2p_{01}p_{10}}{N},$$
$$\tilde{p}_{01} = \frac{p_{01}^2 + p_{10}^2}{N}, \qquad \tilde{p}_{10} = \frac{2p_{00}p_{11}}{N}, \tag{A.1}$$

where $N = (p_{00} + p_{11})^2 + (p_{01} + p_{10})^2$, see e.g. [19].

In [29] it has been shown analytically that the recurrence relations (A.1) converge towards a unique and attracting fixed point provided the initial fidelity with $|B_{00}\rangle$, p_{00}, is above $1/2$.

The recurrence relations of the DEJMPS protocol taking independent single qubit white noise, i.e. noise of the form $N\rho = f\rho + (1-f)/4(\rho + \sigma_x\rho\sigma_x + \sigma_y\rho\sigma_y + \sigma_z\rho\sigma_z)$ acting on each qubit of Alice into account, read far more complex. In the presence of noise we have strong numerical evidence that the DEJMPS protocol converges towards a unique and attracting fixed point depending on the noise level f only.

From figure A1 we suggest a linear relationship between $\log \|\rho_{\text{fix}} - \rho_n\|_1$ (where ρ_{fix} and ρ_n denote the fixed point and the state after successfully completing n distillation rounds respectively) and the number of successful distillation rounds n. We immediately observe that the slope only depends on the noise parameter f, i.e. we have that

$$\log \|\rho_{\text{fix}} - \rho_n\|_1 = a(f) - nb(f).$$

Using $\log_2 N = n$, where N denotes the number of input pairs, this implies $\|\rho_{\text{fix}} - \rho_n\|_1 = e^{a(f)}e^{-b(f)\log_2 N} = a'(f)N^{-b'(f)}$, i.e. $\|\rho_{\text{fix}} - \rho_n\|_1$ scales as $F(N) \in O(N^{-b'(f)})$ as mentioned in the main text. Furthermore we numerically find that the function $b'(f)$ monotonically grows for $f \to 1$.

For two qubit correlated noise, we refer the reader to the analysis including L, as the fixed point and the scaling can be recovered from that analysis by tracing out the system of L.

A.1.1. Detailed analysis including L. We outline the remainder of this section as follows: first we derive the recurrence relations of the DEJMPS protocol in the most general setting, taking the noise applied by L into account as well as assuming that Eve receives the leaked noise transcripts of L. We use those recurrence relations in the next subsection to provide analytical results regarding the fixed point of the recurrence relations, where the inputs are binary pairs and L only applies either id or σ_x operators. We close the section with numerical results for general i.i.d. Bell-diagonal pairs and the most general noise maps of L.

The recurrence relations

For i.i.d. input states the state of each system subject to distillation at an intermediate distillation round of the DEJMPS protocol is of the form $|\Psi\rangle_{ABEL} = \sum_{i,j,k,l} P_{ijkl}\,|B_{ij}\rangle_{AB}|kl\rangle_L|ijkl\rangle_E$, where P_{ijkl} are probability amplitudes, if we assume the noise is leaked to Eve after every distillation round. The system AB models the pair of Alice and Bob, L the system of L (where the content of the register corresponds to the effective noise introduced to AB) and

E the system of Eve. L applies the noise processes before a basic protocol step to the systems of Alice. Moreover, L keeps track of the effective noise introduced using its system in a sense we clarify later.

In the following we use the notation

$$\sigma_{0,0} = \mathrm{id}, \quad \sigma_{0,1} = \sigma_x, \quad \sigma_{1,0} = \sigma_z, \quad \sigma_{1,1} = \sigma_y$$

for the four Pauli-operators. Furthermore we denote by superscripts in brackets particle labels and by superscripts without brackets the power of an operator.

L introduces the noise maps $U_{\alpha_1,\beta_1,\alpha_2,\beta_2} = U^{(a_1)}_{\alpha_1,\beta_1} \otimes U^{(a_2)}_{\alpha_2,\beta_2}$ where $U^{(a_k)}_{\alpha,\beta} = \sigma^{(a_k)}_{\alpha,\beta} \otimes ((\sigma^{\alpha}_x) \otimes (\sigma^{\beta}_x))^{(L_k)}$. We observe that applying the noise map $U_{\alpha_1,\beta_1,\alpha_2,\beta_2}$ might flip the contents of the registers L_1 and L_2 depending on the values of α_1, β_1, α_2 and β_2. This enables L to keep track of the noise introduced to a pair.

There are two approaches how L can apply the noise maps $U_{\alpha_1,\beta_1,\alpha_2,\beta_2}$: stochastically in terms of CPTP maps, or coherently in terms of unitaries acting on an enlarged Hilbert space. Here we assume the latter approach, but provide the analysis of the noisy DEJMPS protocol in terms of CPTP maps and purifications.

To show that these are equivalent, first suppose that L owns a register H set to the state $\sum_{\alpha_1,\beta_1,\alpha_2,\beta_2} \sqrt{\tilde{f}_{\alpha_1,\beta_1,\alpha_2,\beta_2}}\, |\alpha_1\beta_1\alpha_2\beta_2\rangle_H$ where $\tilde{f}_{\alpha_1,\beta_1,\alpha_2,\beta_2}$ are the probabilities of applying the respective noise map $U_{\alpha_1,\beta_1,\alpha_2,\beta_2}$. L uses the register H to apply the noise maps $U_{\alpha_1,\beta_1,\alpha_2,\beta_2}$ coherently controlled to the input state $|\Psi\rangle_{ABEL}$. We observe that tracing out H after applying all the noise maps $U_{\alpha_1,\beta_1,\alpha_2,\beta_2}$ in a controlled fashion yields

$$\sum_{\alpha_1,\beta_1,\alpha_2,\beta_2} \tilde{f}_{\alpha_1,\beta_1,\alpha_2,\beta_2} U_{\alpha_1,\beta_1,\alpha_2,\beta_2}(|\Psi\rangle\langle\Psi| \otimes |\Psi\rangle\langle\Psi|) U^{\dagger}_{\alpha_1,\beta_1,\alpha_2,\beta_2}.$$

On the other hand, assume that L applies the noise process in terms of a CPTP map N, i.e.

$$N\rho = \sum_{\alpha_1,\beta_1,\alpha_2,\beta_2} \tilde{f}_{\alpha_1,\beta_1,\alpha_2,\beta_2} U_{\alpha_1,\beta_1,\alpha_2,\beta_2}(|\Psi\rangle\langle\Psi| \otimes |\Psi\rangle\langle\Psi|) U^{\dagger}_{\alpha_1,\beta_1,\alpha_2,\beta_2}.$$

We observe that $N\rho$ will be, in general, a mixed state, thus there exists a purification on a larger Hilbert space. As all purifications are unitarily equivalent, see e.g. [32], we choose the purification

$$|\Phi\rangle = \sum_{\alpha_1,\beta_1,\alpha_2,\beta_2} \sqrt{\tilde{f}_{\alpha_1,\beta_1,\alpha_2,\beta_2}}\, U_{\alpha_1,\beta_1,\alpha_2,\beta_2}|\Psi\rangle \otimes |\Psi\rangle \otimes |\alpha_1\beta_1\alpha_2\beta_2\rangle_H.$$

Hence $\mathrm{tr}_H[|\Phi\rangle\langle\Phi|] = N\rho$. Furthermore, we observe that the pure state $|\Phi\rangle$ can be generated by applying the unitaries $U_{\alpha_1,\beta_1,\alpha_2,\beta_2}$, coherently controlled by the register H, to $|\Psi\rangle \otimes |\Psi\rangle \otimes (\sum_{\alpha_1,\beta_1,\alpha_2,\beta_2} \sqrt{\tilde{f}_{\alpha_1,\beta_1,\alpha_2,\beta_2}}\, |\alpha_1\beta_1\alpha_2\beta_2\rangle_H)$.

This equivalence allows us to assume that L introduces the noise as a CPTP map, applying $U_{\alpha_1,\beta_1,\alpha_2,\beta_2}$ with respective probabilities $f_{\alpha_1,\beta_1,\alpha_2,\beta_2}$ and purifying the state after the basic distillation step is executed by Alice and Bob.

Since the noise of L is applied before the basic distillation step is executed by Alice and Bob, the result of one noisy distillation step reads as

$$\rho' = \sum_{\alpha_1,\beta_1,\alpha_2,\beta_2} \tilde{f}_{\alpha_1,\beta_1,\alpha_2,\beta_2} U_u O'_{2-\mathrm{EPP}}(U^{(a_1)}_{\alpha_1,\beta_1} \otimes U^{(a_2)}_{\alpha_2,\beta_2})(|\Psi\rangle\langle\Psi| \otimes |\Psi\rangle\langle\Psi|)(U^{(a_1)}_{\alpha_1,\beta_1} \otimes U^{(a_2)}_{\alpha_2,\beta_2})^{\dagger} O'^{\dagger}_{2-\mathrm{EPP}} U^{\dagger}_u, \quad \text{(A.2)}$$

which needs finally to be purified.

In order to evaluate (A.2), we proceed as follows:

- Step 1: We first compute

  $$O'_{2-\mathrm{EPP}}(U^{(a_1)}_{\alpha_1,\beta_1} \otimes U^{(a_2)}_{\alpha_2,\beta_2})|\Psi\rangle \otimes |\Psi\rangle,$$

  which corresponds to the state after the noise map $U^{(a_1)}_{\alpha_1,\beta_1} \otimes U^{(a_2)}_{\alpha_2,\beta_2}$ is applied by L and the basic distillation step of the entanglement distillation protocol is executed by Alice and Bob.

- Step 2: We apply the unitary U_u, which acts only on L's systems and whose purpose we clarify later, to the previous equality.

- Step 3: We have to determine the purification held by Eve if the noise is leaked to her. In doing so, we trace out Eve and then provide her with the purification of the resulting state (which corresponds to leaking the noise transcripts to Eve).

Step 1:We observe that applying the noise map $U^{(a_1)}_{\alpha,\beta}$ to $|\Psi\rangle$ yields

$$\begin{aligned} U^{(a_1)}_{\alpha,\beta}|\Psi\rangle &= U^{(a_1)}_{\alpha,\beta}\sum_{i,j,k,l} P_{ijkl}\,|B_{ij}\rangle_{AB}|kl\rangle_L|ijkl\rangle_E \\ &= \sum_{i,j,k,l} P_{ijkl}\,|B_{(i\oplus\alpha)(j\oplus\beta)}\rangle_{AB}|(k\oplus\alpha)(l\oplus\beta)\rangle_L|ijkl\rangle_E \\ &= \sum_{i,j,k,l} P_{(i\oplus\alpha)(j\oplus\beta)(k\oplus\alpha)(l\oplus\beta)}\,|B_{ij}\rangle_{AB}|kl\rangle_L|(i\oplus\alpha)(j\oplus\beta)(k\oplus\alpha)(l\oplus\beta)\rangle_E. \end{aligned} \tag{A.3}$$

This observation suggests the following notational simplifications:

$$P^{\alpha\beta}_{ijkl} = P_{(i\oplus\alpha)(j\oplus\beta)(k\oplus\alpha)(l\oplus\beta)} \quad \text{and} \quad |e^{\alpha\beta}_{ijkl}\rangle_E = |(i\oplus\alpha)(j\oplus\beta)(k\oplus\alpha)(l\oplus\beta)\rangle_E.$$

Using this notation we rewrite (A.3) as $U^{(a_1)}_{\alpha,\beta}|\Psi\rangle = \sum_{i,j,k,l} P^{\alpha\beta}_{ijkl}\,|B_{ij}\rangle_{AB}|kl\rangle_L|e^{\alpha\beta}_{ijkl}\rangle_E$. This is the state of Alice, Bob, L, and Eve after the noise map $U^{(a_1)}_{\alpha,\beta}$ is applied by L to the first pair. In order to compute (A.2) we define

$$\begin{aligned} |\Psi''_{\alpha_1,\beta_1,\alpha_2,\beta_2}\rangle &= (U^{(a_1)}_{\alpha_1,\beta_1}\otimes U^{(a_2)}_{\alpha_2,\beta_2})|\Psi\rangle|\Psi\rangle \\ &= \sum_{i_1,j_1,i_2,j_2}\sum_{k_1,l_1,k_2,l_2} A^{\alpha_1\beta_1}_{i_1j_1k_1l_1}P^{\alpha_2\beta_2}_{i_2j_2k_2l_2}\,|B_{i_1j_1}\rangle_{AB_1}|B_{i_2j_2}\rangle_{AB_2}|k_1l_1\rangle_{L_1}|k_2l_2\rangle_{L_2} \\ &\quad\otimes|e^{\alpha_1\beta_1}_{i_1j_1k_1l_1}\rangle_{E_1}|e^{\alpha_2\beta_2}_{i_2j_2k_2l_2}\rangle_{E_2}, \end{aligned}$$

which corresponds to the state after the noise map $U^{(a_1)}_{\alpha_1,\beta_1}\otimes U^{(a_2)}_{\alpha_2,\beta_2}$ is applied and

$$|\Psi'_{\alpha_1,\beta_1,\alpha_2,\beta_2}\rangle = U_u O_{2-\mathrm{EPP}}|\Psi''_{\alpha_1,\beta_1,\alpha_2,\beta_2}\rangle, \tag{A.4}$$

which is the state after the noise map $U^{(a_1)}_{\alpha_1,\beta_1}\otimes U^{(a_2)}_{\alpha_2,\beta_2}$, one basic distillation step and the update of L's noise register by U_u. Thus we rewrite (A.2) as

$$\rho' = \sum_{\alpha_1,\beta_1,\alpha_2,\beta_2}\tilde{f}_{\alpha_1,\beta_1,\alpha_2,\beta_2}|\Psi'_{\alpha_1,\beta_1,\alpha_2,\beta_2}\rangle\langle\Psi'_{\alpha_1,\beta_1,\alpha_2,\beta_2}|. \tag{A.5}$$

According to (A.4) Alice and Bob apply one basic distillation step of the DEJMPS protocol to the state $|\Psi''_{\alpha_1,\beta_1,\alpha_2,\beta_2}\rangle$. Recall that step 1 of protocol 1 maps $|B_{ij}\rangle$ to $|B_{i(i\oplus j)}\rangle$ and that step 2 maps $|B_{ij}\rangle|B_{i'j'}\rangle$ to $|B_{(i\oplus i')j}\rangle|B_{i'(j\oplus j')}\rangle$. Thus we conclude that after step 1 and 2 of protocol 1 the state of Alice, Bob, L, and Eve is

$$\sum_{i_1,j_1,i_2,j_2}\sum_{k_1,l_1,k_2,l_2} P^{\alpha_1\beta_1}_{i_1j_1k_1l_1}P^{\alpha_2\beta_2}_{i_2j_2k_2l_2}\,|B_{(i_1\oplus i_2)(i_1\oplus j_1)}\rangle_{AB_1}|B_{i_2(i_1\oplus j_1\oplus i_2\oplus j_2)}\rangle_{AB_2}|k_1l_1\rangle_{L_1}|k_2l_2\rangle_{L_2}|e^{\alpha_1\beta_1}_{i_1j_1k_1l_1}\rangle_{E_1}|e^{\alpha_2\beta_2}_{i_2j_2k_2l_2}\rangle_{E_2}. \tag{A.6}$$

Following protocol 1, a σ_z-measurement of the target pair of the BCNOT, i.e. the subsystem AB_2, is applied to (A.6). Next Alice and Bob communicate their respective measurement outcomes over a classic authentic channel. If the measurement outcomes coincide, Alice and Bob keep the source pair, i.e. subsystem AB_1 of step 2, else they discard both subsystems AB_1 and AB_2. We assume that both measurements yield the outcome 1. If both measurement outcomes yield 0, no phase factor $(-1)^{i_2}$ would be required in the expression (A.7). The coinciding measurement outcomes imply $i_1\oplus j_1\oplus i_2\oplus j_2 = 0$. To summarize, the state post-selected on the measurement outcomes 1 of Alice and Bob is

$$\sum_{i_1,j_1,i_2,j_2}\sum_{k_1,l_1,k_2,l_2}(-1)^{i_2}P^{\alpha_1\beta_1}_{i_1j_1k_1l_1}P^{\alpha_2\beta_2}_{i_2(i_1\oplus j_1\oplus i_2)k_2l_2}\,|B_{(i_1\oplus i_2)(i_1\oplus j_1)}\rangle_{AB_1}|k_1l_1\rangle_{L_1}|k_2l_2\rangle_{L_2}|e^{\alpha_1\beta_1}_{i_1j_1k_1l_1}\rangle_{E_1}|e^{\alpha_2\beta_2}_{i_2(i_1\oplus j_1\oplus i_2)k_2l_2}\rangle_{E_2}. \tag{A.7}$$

Step 2: Recall that L stores in its register attached to the pair of Alice and Bob the effective noise introduced. For that purpose we introduce the unitary U_u as well as an ancilla system L_3 set to the state $|00\rangle_{L_3}$. Applying U_u to all three registers of L yields $U_u|00\rangle|i\rangle|j\rangle|i'\rangle|j'\rangle = |u(i,\,j,\,i',\,j')\rangle|i\rangle|j\rangle|i'\rangle|j'\rangle$ where u is the so called flag update function defined in [20]. The function u returns the effective noise introduced on the source pair of step 2 of protocol 1. Applying U_u to (A.7) gives

$$\begin{aligned} |\Psi'_{\alpha_1,\beta_1,\alpha_2,\beta_2}\rangle &= \sum_{i_1,j_1,i_2,j_2}\sum_{k_1,l_1,k_2,l_2}(-1)^{i_2}P^{\alpha_1\beta_1}_{i_1j_1k_1l_1}P^{\alpha_2\beta_2}_{i_2(i_1\oplus j_1\oplus i_2)k_2l_2}\,|B_{(i_1\oplus i_2)(i_1\oplus j_1)}\rangle_{AB_1}|k_1l_1\rangle_{L_1}|k_2l_2\rangle_{L_2}|u(k_1,\,l_1,\,k_2,\,l_2)\rangle_{L_3} \\ &\quad\otimes|e^{\alpha_1\beta_1}_{i_1j_1k_1l_1}\rangle_{E_1}|e^{\alpha_2\beta_2}_{i_2(i_1\oplus j_1\oplus i_2)k_2l_2}\rangle_{E_2}. \end{aligned}$$

We remind the reader that $|\Psi'_{\alpha_1,\beta_1,\alpha_2,\beta_2}\rangle$ is the state after the application of (i) the noise map $U^{(a_1)}_{\alpha_1,\beta_1}\otimes U^{(a_2)}_{\alpha_2,\beta_2}$, (ii) a basic distillation step, and (iii) the update of L's noise register by U_u.

Step 3: Since the noise transcripts—by assumption for this analysis—leak to Eve, we attribute the systems L_1 and L_2 to Eve. In order to treat the most general situation, we assume that Eve holds a purification of $\mathrm{tr}_{L_1,L_2,E_1,E_2}[\rho']$. We determine this purification by computing $\rho'_1 = \mathrm{tr}_{L_1,L_2}[\rho']$ and $\rho'_2 = \mathrm{tr}_{E_1,E_2}[\rho'_1]$ and attribute the purification of ρ'_2 to Eve.

By the linearity of the partial trace we have

$$\rho_1' = \mathrm{tr}_{L_1,L_2}[\rho'] = \sum_{\alpha_1,\beta_1,\alpha_2,\beta_2} \tilde{f}_{\alpha_1,\beta_1,\alpha_2,\beta_2}\, \mathrm{tr}_{L_1,L_2}[|\Psi'_{\alpha_1,\beta_1,\alpha_2,\beta_2}\rangle\langle\Psi'_{\alpha_1,\beta_1,\alpha_2,\beta_2}|].$$

It is useful to define $\rho'_{\alpha_1,\beta_1,\alpha_2,\beta_2} = \mathrm{tr}_{L_1,L_2}[|\Psi'_{\alpha_1,\beta_1,\alpha_2,\beta_2}\rangle\langle\Psi'_{\alpha_1,\beta_1,\alpha_2,\beta_2}|]$ which evaluates to

$$\begin{aligned}\rho'_{\alpha_1,\beta_1,\alpha_2,\beta_2} &= \mathrm{tr}_{L_1,L_2}[|\Psi'_{\alpha_1,\beta_1,\alpha_2,\beta_2}\rangle\langle\Psi'_{\alpha_1,\beta_1,\alpha_2,\beta_2}|]\\ &= \sum (-1)^{i_2\oplus i_2'} P^{\alpha_1\beta_1}_{i_1j_1k_1l_1} P^{\alpha_2\beta_2}_{i_2(i_1\oplus j_1\oplus i_2)k_2l_2} (P^{\alpha_1\beta_1}_{i_1'j_1'k_1l_1} P^{\alpha_2\beta_2}_{i_2'(i_1'\oplus j_1'\oplus i_2')k_2l_2})^* |B_{(i_1\oplus i_2)(i_1\oplus j_1)}\rangle\langle B_{(i_1'\oplus i_2')(i_1'\oplus j_1')}|\\ &\quad\otimes |u(k_1,\,l_1,\,k_2,\,l_2)\rangle\langle u(k_1,\,l_1,\,k_2,\,l_2)| \otimes |e^{\alpha_1\beta_1}_{i_1j_1k_1l_1}\rangle\langle e^{\alpha_1\beta_1}_{i_1'j_1'k_1l_1}| \otimes |e^{\alpha_2\beta_2}_{i_2(i_1\oplus j_1\oplus i_2)k_2l_2}\rangle\langle e^{\alpha_2\beta_2}_{i_2'(i_1'\oplus j_1'\oplus i_2')k_2l_2}|.\end{aligned}$$

In the previous expression we neglected the indices appearing in the sum for simplicity, but it is understood that the sum ranges over all indices except α_1, β_1, α_2 and β_2.

In order to determine the state of Alice, Bob, and L which Eve finally purifies we have to compute $\rho_2' = \mathrm{tr}_{E_1,E_2}[\rho_1']$. Again, the linearity of the partial trace yields

$$\rho_2' = \mathrm{tr}_{E_1,E_2}[\rho_1'] = \sum_{\alpha_1,\beta_1,\alpha_2,\beta_2} \tilde{f}_{\alpha_1,\beta_1,\alpha_2,\beta_2}\, \mathrm{tr}_{E_1,E_2}[\rho'_{\alpha_1,\beta_1,\alpha_2,\beta_2}]. \tag{A.8}$$

We remind the reader that $|e^{\alpha\beta}_{ijkl}\rangle_{E_1} = |(i\oplus\alpha)(j\oplus\beta)(k\oplus\alpha)(l\oplus\beta)\rangle_{E_1}$. Hence, for fixed α_1 and β_1, we have $\mathrm{tr}|e^{\alpha_1\beta_1}_{i_1j_1k_1l_1}\rangle\langle e^{\alpha_1\beta_1}_{i_1'j_1'k_1l_1}| = \delta_{i_1i_1'}\delta_{j_1j_1'}$, which implies that $i_1' = i_1$ and $j_1' = j_1$. Thus, we also have

$$\mathrm{tr}|e^{\alpha_2\beta_2}_{i_2(i_1\oplus j_1\oplus i_2)k_2l_2}\rangle\langle e^{\alpha_2\beta_2}_{i_2'(i_1'\oplus j_1'\oplus i_2')k_2l_2}| = \mathrm{tr}|e^{\alpha_2\beta_2}_{i_2(i_1\oplus j_1\oplus i_2)k_2l_2}\rangle\langle e^{\alpha_2\beta_2}_{i_2'(i_1\oplus j_1\oplus i_2')k_2l_2}| = \delta_{i_2i_2'}.$$

Hence

$$\begin{aligned}&\mathrm{tr}_{E_1,E_2}[\rho'_{\alpha_1,\beta_1,\alpha_2,\beta_2}]\\ &= \sum_{i_1,i_2,j_1}\sum_{k_1,l_1,k_2,l_2} P^{\alpha_1\beta_1}_{i_1j_1k_1l_1} P^{\alpha_2\beta_2}_{i_2(i_1\oplus j_1\oplus i_2)k_2l_2} (P^{\alpha_1\beta_1}_{i_1j_1k_1l_1} P^{\alpha_2\beta_2}_{i_2(i_1\oplus j_1\oplus i_2)k_2l_2})^*\\ &\times |B_{(i_1\oplus i_2)(i_1\oplus j_1)}\rangle\langle B_{(i_1\oplus i_2)(i_1\oplus j_1)}| \otimes |u(k_1,\,l_1,\,k_2,\,l_2)\rangle\langle u(k_1,\,l_1,\,k_2,\,l_2)|\\ &= \sum_{i_1,i_2,j_1}\sum_{k_1,l_1,k_2,l_2} |P^{\alpha_1\beta_1}_{i_1j_1k_1l_1} P^{\alpha_2\beta_2}_{i_2(i_1\oplus j_1\oplus i_2)k_2l_2}|^2\\ &\times |B_{(i_1\oplus i_2)(i_1\oplus j_1)}\rangle\langle B_{(i_1\oplus i_2)(i_1\oplus j_1)}| \otimes |u(k_1,\,l_1,\,k_2,\,l_2)\rangle\langle u(k_1,\,l_1,\,k_2,\,l_2)|.\end{aligned} \tag{A.9}$$

By inserting (A.9) in (A.8) we get

$$\begin{aligned}\rho_2' &= \mathrm{tr}_{E_1,E_2}[\rho_1']\\ &= \sum_{\alpha_1,\beta_1,\alpha_2,\beta_2} \tilde{f}_{\alpha_1,\beta_1,\alpha_2,\beta_2}\, \mathrm{tr}_{E_1,E_2}[\rho'_{\alpha_1,\beta_1,\alpha_2,\beta_2}]\\ &= \sum_{\alpha_1,\beta_1,\alpha_2,\beta_2} \tilde{f}_{\alpha_1,\beta_1,\alpha_2,\beta_2} \sum_{i_1,i_2,j_1}\sum_{k_1,l_1,k_2,l_2} |P^{\alpha_1\beta_1}_{i_1j_1k_1l_1} P^{\alpha_2\beta_2}_{i_2(i_1\oplus j_1\oplus i_2)k_2l_2}|^2\\ &\quad\times |B_{(i_1\oplus i_2)(i_1\oplus j_1)}\rangle\langle B_{(i_1\oplus i_2)(i_1\oplus j_1)}| \otimes |u(k_1,\,l_1,\,k_2,\,l_2)\rangle\langle u(k_1,\,l_1,\,k_2,\,l_2)|\\ &= \sum_{i_1,i_2,j_1} |B_{(i_1\oplus i_2)(i_1\oplus j_1)}\rangle\langle B_{(i_1\oplus i_2)(i_1\oplus j_1)}| \otimes \sum_{\gamma_0,\gamma_1}\left(\sum_{\substack{\alpha_1,\beta_1,\alpha_2,\beta_2,k_1,l_1,k_2,l_2\\ u(k_1,l_1,k_2,l_2)=(\gamma_0,\gamma_1)}} \tilde{f}_{\alpha_1,\beta_1,\alpha_2,\beta_2} |P^{\alpha_1\beta_1}_{i_1j_1k_1l_1} P^{\alpha_2\beta_2}_{i_2(i_1\oplus j_1\oplus i_2)k_2l_2}|^2\right)\\ &\quad\times |\gamma_0\gamma_1\rangle\langle\gamma_0\gamma_1|.\end{aligned}$$

Rearranging the sum over i_1, i_2 and j_1 in the previous equation gives

$$\begin{aligned}&\sum_{\delta_0,\delta_1} |B_{\delta_0\delta_1}\rangle\langle B_{\delta_0\delta_1}| \otimes \sum_{\gamma_0,\gamma_1}\left(\sum_{\substack{i_1,i_2,j_1\\ i_1\oplus i_2=\delta_0, i_1\oplus j_1=\delta_1}} \sum_{\substack{\alpha_1,\beta_1,\alpha_2,\beta_2,k_1,l_1,k_2,l_2\\ u(k_1,l_1,k_2,l_2)=(\gamma_0,\gamma_1)}} \tilde{f}_{\alpha_1,\beta_1,\alpha_2,\beta_2} |P^{\alpha_1\beta_1}_{i_1j_1k_1l_1} P^{\alpha_2\beta_2}_{i_2(i_1\oplus j_1\oplus i_2)k_2l_2}|^2\right)\\ &\quad |\gamma_0\gamma_1\rangle\langle\gamma_0\gamma_1|.\end{aligned} \tag{A.10}$$

Using the definition

$$|\tilde{P}_{\delta_0\delta_1\gamma_0\gamma_1}|^2 = \sum_{\substack{i_1,i_2,j_1\\ i_1\oplus i_2=\delta_0, i_1\oplus j_1=\delta_1}} \sum_{\substack{\alpha_1,\beta_1,\alpha_2,\beta_2,k_1,l_1,k_2,l_2\\ u(k_1,l_1,k_2,l_2)=(\gamma_0,\gamma_1)}} \tilde{f}_{\alpha_1,\beta_1,\alpha_2,\beta_2} |P^{\alpha_1\beta_1}_{i_1j_1k_1l_1} P^{\alpha_2\beta_2}_{i_2(i_1\oplus j_1\oplus i_2)k_2l_2}|^2, \tag{A.11}$$

where δ_0, δ_1, γ_0, $\gamma_1 \in \{0,\,1\}$ and omitting the normalization factor for clarity, (A.10) simplifies to

$$\sum_{\delta_0,\delta_1} |B_{\delta_0\delta_1}\rangle\langle B_{\delta_0\delta_1}| \otimes \sum_{\gamma_0,\gamma_1} |\tilde{P}_{\delta_0\delta_1\gamma_0\gamma_1}|^2 |\gamma_0\gamma_1\rangle\langle\gamma_0\gamma_1|,$$

which is the state of Alice, Bob, and L after one noisy distillation step. Since this final state is purified by Eve with the leaked noise transcripts and all purifications are unitarily equivalent, the state of Alice, Bob, L, and Eve after one noisy distillation step can be written without loss of generality as

$$|\psi^{\mathrm{DEJMPS}}\rangle = \sum_{\delta_0,\delta_1,\gamma_0,\gamma_1} \tilde{P}_{\delta_0,\delta_1,\gamma_0,\gamma_1} |B_{\delta_0,\delta_1}\rangle_{AB} |\gamma_0\gamma_1\rangle_L |\delta_0\delta_1\gamma_0\gamma_1\rangle_E .$$

This also implies that (A.11) are the recurrence relations of the noisy DEJMPS protocol.

Fixed point and convergence—binary pairs

First we study the scaling of the systems of Alice, Bob, and L and extend those results then to the (possibly leaked) noise transcripts of Eve in terms of purifications.

Suppose that the initial i.i.d. pairs of Alice and Bob are mixtures of $|B_{00}\rangle$ and $|B_{01}\rangle$ and that L applies either the identity or a σ_x-operator with respective probabilities $\tilde{f}_0$ and $\tilde{f}_1 = 1 - \tilde{f}_0$ independently to each pair. We remind the reader that Eve purifies the state of Alice, Bob, and L with the leaked noise transcripts, i.e. each individual state taking Eve into account at an intermediate round of the DEJMPS protocol reads as $\sum_{i,j} P_{ij} |B_{0i}\rangle_{AB} \otimes |\eta_j\rangle_L \otimes |\eta_{ij}\rangle_E$. Using $p_{ij} = |P_{ij}|^2$, the recurrence relations (A.11) for the setting we are concerned with here simplify to

$$\tilde{p}_{00} = 1/N(\tilde{f}_0^2(p_{00}^2 + 2p_{00}p_{01}) + \tilde{f}_1^2(p_{11}^2 + 2p_{10}p_{11}) + 2\tilde{f}_0\tilde{f}_1(p_{11}p_{00} + p_{10}p_{00} + p_{11}p_{01})), \tag{A.12}$$

$$\tilde{p}_{01} = 1/N(\tilde{f}_0^2 p_{01}^2 + 2\tilde{f}_0\tilde{f}_1 p_{10}p_{01} + \tilde{f}_1^2 p_{10}^2), \tag{A.13}$$

$$\tilde{p}_{10} = 1/N(\tilde{f}_0^2(p_{10}^2 + 2p_{10}p_{11}) + \tilde{f}_1^2(p_{01}^2 + 2p_{00}p_{01}) + 2\tilde{f}_0\tilde{f}_1(p_{01}p_{10} + p_{00}p_{10} + p_{01}p_{11})), \tag{A.14}$$

$$\tilde{p}_{11} = 1/N(\tilde{f}_0^2 p_{11}^2 + 2\tilde{f}_0\tilde{f}_1 p_{00}p_{11} + \tilde{f}_1^2 p_{00}^2), \tag{A.15}$$

where $N = (\tilde{f}_0^2 + \tilde{f}_1^2)((p_{00} + p_{01})^2 + (p_{10} + p_{11})^2) + 4\tilde{f}_0\tilde{f}_1(p_{00} + p_{01})(p_{10} + p_{11})$. In the following we denote the recurrence relations (A.12)–(A.15) by the vector-valued mapping $\mathbf{f}$, i.e. $\mathbf{p} \xrightarrow{\mathbf{f}} \tilde{\mathbf{p}}$, where $\mathbf{p} = (p_{00}, p_{01}, p_{10}, p_{11})$. A simple computation yields the following fixed points of $\mathbf{f}$:

$$p_{00}^\infty = 1/2 + \sqrt{4\tilde{f}_0 - 3}\big/(4\tilde{f}_0 - 2) \quad p_{01}^\infty = p_{10}^\infty = 0 \quad p_{11}^\infty = 1 - p_{00}^\infty, \tag{A.16}$$

$$p_{00}^\infty = 1/2 - \sqrt{4\tilde{f}_0 - 3}\big/(4\tilde{f}_0 - 2) \quad p_{01}^\infty = p_{10}^\infty = 0 \quad p_{11}^\infty = 1 - p_{00}^\infty, \tag{A.17}$$

$$p_{00}^\infty = p_{11}^\infty = 1/2 \quad p_{01}^\infty = p_{10}^\infty = 0. \tag{A.18}$$

The parameter estimation phase guarantees that the fidelity F with $|B_{00}\rangle$ is sufficiently high for distillation. Hence the fixed point of interested is (A.16), i.e.

$$\mathbf{p}^\infty = (1/2 + \sqrt{4\tilde{f}_0 - 3}\big/(4\tilde{f}_0 - 2), 0, 0, 1/2 - \sqrt{4\tilde{f}_0 - 3}\big/(4\tilde{f}_0 - 2)). \tag{A.19}$$

From (A.19) we observe that in the limit the 'cross-probabilities' p_{01} and p_{10}, vanish, hence L is fully correlated to AB.

It is of central importance, regarding convergence that the fixed point $\mathbf{p}^\infty$ is an attractor, as only this ensures convergence towards that fixed point. Note that $\mathbf{p}^\infty$ is an attractor if and only if the largest eigenvalue $\lambda_{\max}$ of $\mathbf{f}'(\mathbf{p}^\infty)$ satisfies $\lambda_{\max} < 1$. We easily find that $\lambda_{\max} = (\tilde{f}_0\sqrt{4\tilde{f}_0 - 3} - \tilde{f}_0)\big/(2\tilde{f}_0 - 1) < 1$ for $0.78 \leqslant \tilde{f}_0 \leqslant 1$.

The fixed point $\mathbf{p}^\infty$ enables us to determine the rate of convergence. For that purpose, we expand $\mathbf{f}$ in terms of its Taylor series around the fixed point $\mathbf{p}^\infty$, i.e. $\tilde{\mathbf{p}} = \mathbf{f}(\mathbf{p}) \approx \mathbf{f}(\mathbf{p}^\infty) + \mathbf{f}'(\mathbf{p}^\infty)(\mathbf{p} - \mathbf{p}^\infty)$. Hence by defining $\mathbf{e} = \mathbf{p} - \mathbf{p}^\infty$ we find $\tilde{\mathbf{e}} = \mathbf{f}'(\mathbf{p}^\infty)\mathbf{e}$, providing an estimate of the error propagation for one successful distillation round. The state of Alice, Bob, and L after n successful distillation rounds and at the fixpoint read as $\rho_n = \sum_{ij} p_{ij}^{(n)} |B_{0i}\rangle\langle B_{0i}|_{AB} \otimes |\eta_j\rangle\langle\eta_j|_L$ and $\rho_{\mathrm{fix}} = \sum_i p_{ii}^\infty |B_{0i}\rangle\langle B_{0i}|_{AB} \otimes |\eta_i\rangle\langle\eta_i|_L$ respectively, which implies for their distance induced by the 1-norm

$$\epsilon_n = \|\rho_n - \rho_{\mathrm{fix}}\|_1 = \left\| \sum_{i,j} (p_{ij}^{(n)} - p_{ij}^\infty) |B_{0i}\rangle\langle B_{0i}|_{AB} \otimes |\eta_j\rangle\langle\eta_j|_L \right\|_1 = \underbrace{\sum_{i,j} |p_{ij}^{(n)} - p_{ij}^\infty|}_{\|\mathbf{e}_n\|_{1;\mathrm{v}}} \leqslant \|\mathbf{f}'(\mathbf{p}^\infty)^{n-1}\| \|\mathbf{e}_1\|_{1;\mathrm{v}}, \tag{A.20}$$

where $\|\mathbf{x}\|_{1;\mathrm{v}} = \sum_{i=1}^k |x_i|$ denotes the 1-norm of vectors in $\mathbb{C}^k$.

Equation (A.20) only concerns the systems of Alice, Bob, and L. To complete the analysis we recall that Eve purifies ρ_n and ρ_{fix} with the leaked noise transcripts of L. If we take this purifying system, E, into account, i.e. consider $\| |\psi^n\rangle\langle\psi^n|_{ABEL} - |\psi^\alpha\rangle\langle\psi^\alpha|_{ABEL} \|_1$ where $\rho_n = \mathrm{tr}_E[|\psi^n\rangle\langle\psi^n|_{ABEL}]$, $|\psi^\alpha\rangle_{ABEL} = \sum_{i,j} P_{ij}^\infty |B_{0i}\rangle_{AB} \otimes |\eta_j\rangle_L \otimes |\eta_{ij}\rangle_E$ with $|P_{ij}^\infty|^2 = p_{ij}^\infty$ and $\rho_{\mathrm{fix}} = \mathrm{tr}_E[|\psi^\alpha\rangle\langle\psi^\alpha|_{ABEL}]$, we find

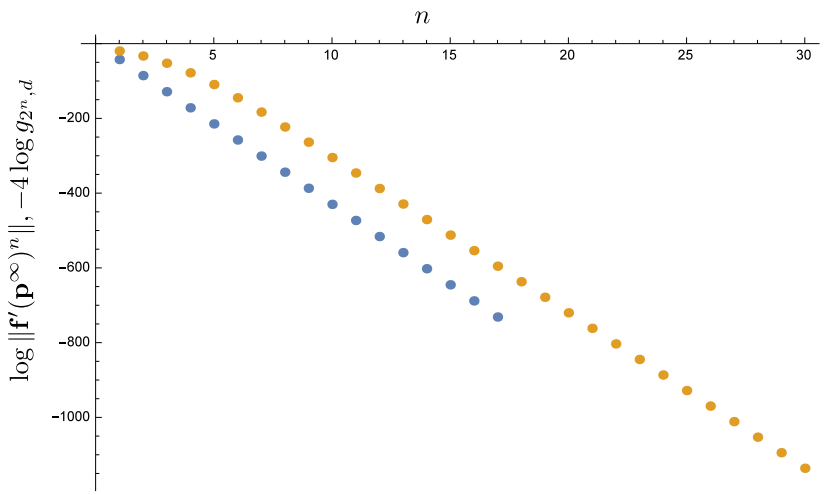

Figure A2. The figure illustrates $\log \|\mathbf{f}'(\mathbf{p}^\infty)^n\|$ (blue) and $-4\log g_{2^n,d}$ (yellow) for the binary pairs and $\tilde{f}_0 = 1 - 10^{-19}$.

$$\| |\psi^n\rangle\langle\psi^n|_{ABEL} - |\psi^\alpha\rangle\langle\psi^\alpha|_{ABEL} \|_1 \leqslant \sqrt{\epsilon_n} \tag{A.21}$$

since purifications scale with a square root.

In order to apply the post-selection-based reduction, we need to relate the previously obtained results for i.i.d. input pairs to general ensembles. As stated in the main text, we exclude the parameter estimation step on $\sqrt{n}$ initial states for simplicity. We remind the reader, as we have stated in the main text, that for all purifications $|\psi\rangle_{ABE'}$ of a n-partite input state ρ_{AB} we have

$$\|(\mathcal{E}\otimes \mathrm{id}_{E'})(|\psi\rangle\langle\psi|_{ABE'}) - (\mathcal{F}\otimes \mathrm{id}_{E'})(|\psi\rangle\langle\psi|_{ABE'})\|_1 \leqslant 4 g_{n,d}\sqrt{\max_{\sigma_{AB}} \|(\mathcal{E}_L - \mathcal{F}_L)(\sigma_{AB}^{\otimes n})\|_1}\,, \tag{A.22}$$

where $g_{n,d} = \binom{n+d^2-1}{n}$. Thus, inserting the previous result for 2^n i.i.d. input states (necessary to achieve n rounds of distillation) in (A.22) yields

$$\|(\mathcal{E}\otimes \mathrm{id}_{E'})(|\psi\rangle\langle\psi|_{ABE'}) - (\mathcal{F}\otimes \mathrm{id}_{E'})(|\psi\rangle\langle\psi|_{ABE'})\|_1 \leqslant 4 g_{2^n,d}\,\epsilon_n^{1/4}.$$

One square root in the expression above arises from inequality (A.21) and the other square root appears from inequality (A.22).

Hence, for confidentiality we necessarily need $g_{2^n,d}\,\epsilon_n^{1/4} \to 0$ for $n \to \infty$. Thus $\epsilon_n^{1/4}$ should decay faster than $g_{2^n,d}$ grows in n. Numerical simulations suggest that, for $\tilde{f}_0 = 1 - 10^{-19}$, this turns out to be true, i.e. the post-selection-based reduction is applicable (see figure A2). As stated in the main text such rates are unlikely to be achievable on the physical level, but they are, at least in principle, possible through fault-tolerant constructions.

Fixed point and convergence—general pairs

In the following we show that the previous established results also hold true for the general i.i.d. setting where L applies all four Pauli operators and each individual pair is arbitrary. We remind the reader that the recurrence relations for states $\sum_{i,j,k,l} P_{ijkl}\,|B_{ij}\rangle_{AB}\otimes|\eta_{kl}\rangle_L\otimes|\eta_{ijkl}\rangle_E$ (i.e. Eve purifies $\rho_n = \sum_{i,j,k,l}|P_{ijkl}|^2|B_{ij}\rangle\langle B_{ij}|_{AB}\otimes|\eta_{kl}\rangle\langle\eta_{kl}|_L$ with the leaked noise transcripts) read (by denoting $|P_{ijkl}|^2 = p_{ijkl}$) as

$$\tilde{p}_{\delta_0\delta_1\gamma_0\gamma_1} = \sum_{\substack{i_1,i_2,j_1\\ i_1\oplus i_2=\delta_0, i_1\oplus j_1=\delta_1}} \sum_{\substack{\alpha_1,\beta_1,\alpha_2,\beta_2,k_1,l_1,k_2,l_2\\ u(k_1,l_1,k_2,l_2)=(\gamma_0,\gamma_1)}} \tilde{f}_{\alpha_1,\beta_1,\alpha_2,\beta_2}\, p_{(i_1\oplus\alpha_1)(j_1\oplus\beta_1)(k_1\oplus\alpha_1)(l_1\oplus\beta_1)}\, p_{(i_2\oplus\alpha_2)(i_1\oplus j_1\oplus i_2\oplus\beta_2)(k_2\oplus\alpha_2)(l_2\oplus\beta_2)}$$

modulo the normalization factor $\sum_{\delta_0\delta_1\gamma_0\gamma_1}\tilde{p}_{\delta_0\delta_1\gamma_0\gamma_1}$.

For simplicity we assume independent single qubit white noise, i.e. $\tilde{f}_{\alpha_1,\beta_1,\alpha_2,\beta_2} = \tilde{f}_{\alpha_1,\beta_1}\tilde{f}_{\alpha_2,\beta_2}$ as well as $\tilde{f}_{\alpha_1,\beta_1} = f$ if $\alpha_1 = \beta_1 = 0$ and $(1-f)/3$ otherwise. Furthermore, we assume that the initial fidelity F with $|B_{00}\rangle$ is sufficiently high for distillation. Numerically iterating the recurrence relations (which we again denote by $\mathbf{p} \xrightarrow{\mathbf{f}} \tilde{\mathbf{p}}$) reveal that, for a sufficiently large number of iterations, the 'cross-probabilities' vanish, i.e. $p^\infty_{ijkl} = 0 \Leftrightarrow i \neq k$ or $j \neq l$. Hence, to obtain a fixed point $\mathbf{p}^\infty = (p^\infty_{ijkl})^1_{i,j,k,l\,=\,0}$ of $\mathbf{f}$, it is reasonable to assume that $p^\infty_{ijkl} = 0 \Leftrightarrow i \neq k$ or $j \neq l$.

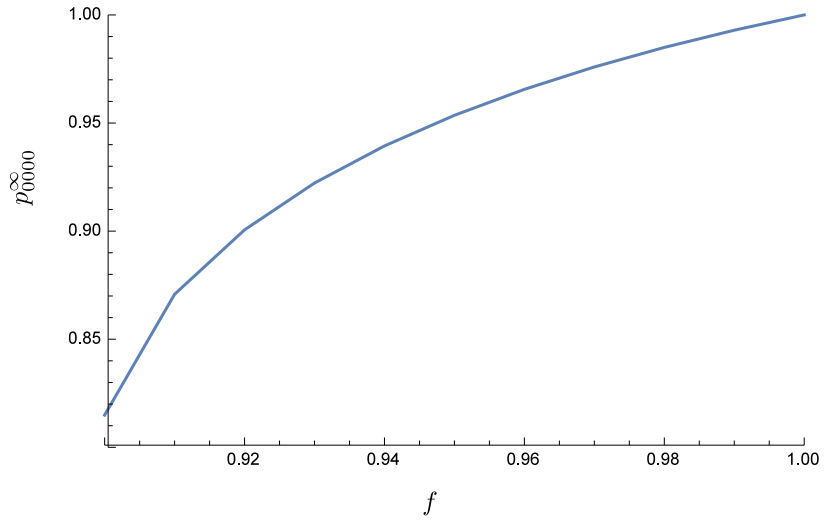

Figure A3. The figure illustrates p^{∞}_{0000} as a function of f. The fidelity with $|B_{00}\rangle$ of the asymptotic state is equal to unity for a perfect apparatus.

Thus the fixed point $\mathbf{p}^{\infty}$ is determined by four equations in four unknowns, namely the equations

$$p_{\delta_0\delta_1\delta_0\delta_1} = \frac{1}{N} \sum_{\substack{i_1,i_2,j_1 \\ i_1\oplus i_2=\delta_0, i_1\oplus j_1=\delta_1}} \sum_{\substack{\alpha_1,\beta_1,\alpha_2,\beta_2 \\ u(i_1,j_1,i_2,i_1\oplus j_1\oplus i_2)=(\delta_0,\delta_1)}} \tilde{f}_{\alpha_1,\beta_1}\tilde{f}_{\alpha_2,\beta_2} p_{(i_1\oplus\alpha_1)(j_1\oplus\beta_1)(i_1\oplus\alpha_1)(j_1\oplus\beta_1)} \cdot p_{(i_2\oplus\alpha_2)(i_1\oplus j_1\oplus i_2\oplus\beta_2)(i_2\oplus\alpha_2)(i_1\oplus j_1\oplus i_2\oplus\beta_2)},$$

where $\delta_0,\ \delta_1 \in \{0,\ 1\}$ and $N = \sum_{\delta_0,\delta_1} p_{\delta_0\delta_1\delta_0\delta_1}$. Figure A3 illustrates the numerical estimate of p^{∞}_{0000} as a function of f.

Similar to the case of binary pairs, we can write the recurrence relations $\mathbf{f}$ in terms of its Taylor series expansion around the fixed point $\mathbf{p}^{\infty}$, i.e. $\tilde{\mathbf{p}} = \mathbf{f}(\mathbf{p}) \approx \mathbf{f}(\mathbf{p}^{\infty}) + \mathbf{f}'(\mathbf{p}^{\infty})(\mathbf{p} - \mathbf{p}^{\infty})$. Hence by defining $\mathbf{e} = \mathbf{p} - \mathbf{p}^{\infty}$ we have $\tilde{\mathbf{e}} = \mathbf{f}'(\mathbf{p}^{\infty})\mathbf{e}$, i.e. as for binary pairs, the error induced by the 1-norm of the state of Alice, Bob, and L after n successful distillation rounds satisfies

$$\|\rho_n - \rho_{\mathrm{fix}}\|_1 = \left\| \sum_{i,j,k,l} (p^{(n)}_{ijkl} - p^{\infty}_{ijkl}) |B_{ij}\rangle\langle B_{ij}|_{AB} \otimes |\eta_{kl}\rangle\langle\eta_{kl}|_L \right\|_1 \leqslant \sum_{i,j,k,l} |p^{(n)}_{ijkl} - p^{\infty}_{ijkl}| \leqslant \|\mathbf{f}'(\mathbf{p}^{\infty})^{n-1}\| \|\mathbf{e}_1\|_{1;\mathrm{v}}. \tag{A.23}$$

Figure A4 suggests a linear relationship between the number of successful distillation rounds n and $\log \|\mathbf{f}'(\mathbf{p}^{\infty})^{n-1}\|$ for each noise level f, i.e. $b(f)n + a(f) = \log \|\mathbf{f}'(\mathbf{p}^{\infty})^{n-1}\|$. As the number N of pairs necessary to achieve n distillation rounds is $N = 2^n$ ($\Leftrightarrow n = \log_2 N$) we have $b(f)\log_2 N + a(f) = \log \|\mathbf{f}'(\mathbf{p}^{\infty})^{n-1}\|$, which is equivalent to

$$\|\mathbf{f}'(\mathbf{p}^{\infty})^{n-1}\| = e^{a(f)} e^{b(f)\log_2 N} = a'(f) N^{b'(f)}.$$

Hence, $\|\mathbf{f}'(\mathbf{p}^{\infty})^{n-1}\|$ scales as $F(N) \in O(N^{b'(f)})$ where $b'(f) < 0$ and $b'(f)$ decays for $f \to 1$.

What is left to show, is that the fixed point $\mathbf{p}^{\infty}$ is an attracting fixed point. For that purpose we numerically compute the largest eigenvalue of $\mathbf{f}'(\mathbf{p}^{\infty})$, see figure A5, and observe that, for noise below 10^{-1}, i.e. $1 - f < 10^{-1}$, the largest eigenvalue $\lambda_{\max}$ of $\mathbf{f}'(\mathbf{p}^{\infty})$ fulfills $\lambda_{\max} < 1$, proving that $\mathbf{p}^{\infty}$ is an attracting fixed point.

This implies that, if the initial fidelity F with $|B_{00}\rangle$ is sufficiently large for distillation, the DEJMPS protocol necessarily converges towards the fixed point $\mathbf{p}^{\infty}$ where the 'cross-probabilities' vanish.

The analysis so far still lacks Eve's system E for the leaked noise transcripts. Suppose $|\psi^n\rangle_{ABEL}$ and $|\psi^f\rangle_{ABEL}$ are purifications of ρ_n and ρ_{fix}, i.e. $\rho_n = \mathrm{tr}_E[|\psi^n\rangle\langle\psi^n|]$ and $\rho_{\mathrm{fix}} = \mathrm{tr}_E[|\psi^f\rangle\langle\psi^f|]$ respectively. This implies $\epsilon_n = \||\psi^n\rangle\langle\psi^n| - |\psi^f\rangle\langle\psi^f|\|_1 \leqslant \sqrt{F(N)}$, i.e. $\epsilon_n \in O(N^{b'(f)/2})$ which we also confirmed with our numeric results.

It is straightforward to extend the analysis above to two-qubit correlated noise introduced by L on the system of Alice and Bob. For that purpose we assume that $\tilde{f}_{\alpha_1,\beta_1,\alpha_2,\beta_2} = \tilde{f} + (1-\tilde{f})/16$ if $\alpha_1 = \beta_1 = \alpha_2 = \beta_2 = 0$ and $(1-\tilde{f})/16$ otherwise. Also in that case we numerically observe that $p^{\infty}_{ijkl} = 0 \Leftrightarrow i \neq k$ or $j \neq l$. Hence it is reasonable to assume that $p^{\infty}_{ijkl} = 0 \Leftrightarrow i \neq k$ or $j \neq l$ in order to obtain a fixed point $\mathbf{p}^{\infty} = (p^{\infty}_{ijkl})^1_{i,j,k,l\ =\ 0}$ of $\mathbf{f}$.

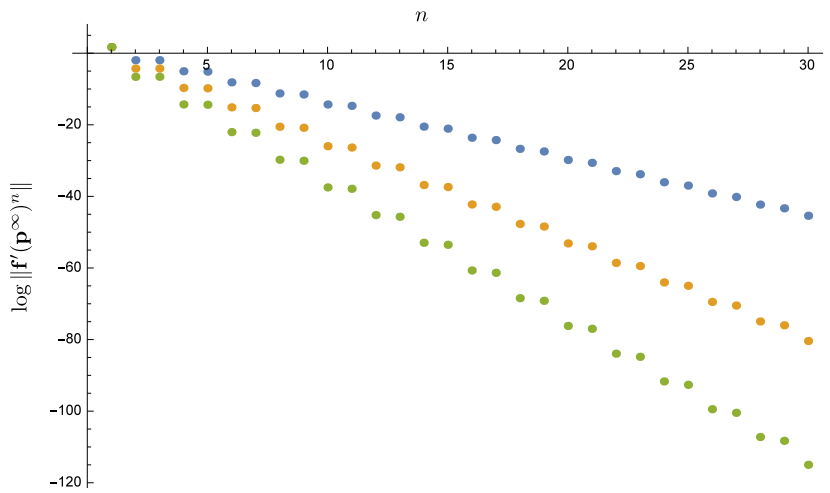

Figure A4. The figure illustrates the value of $\log \|\mathbf{f}'(\mathbf{p}^\infty)^n\|$ as a function of successful distillation rounds for single qubit white noise 10^{-2} (blue), 10^{-3} (yellow) and 10^{-4} (green).

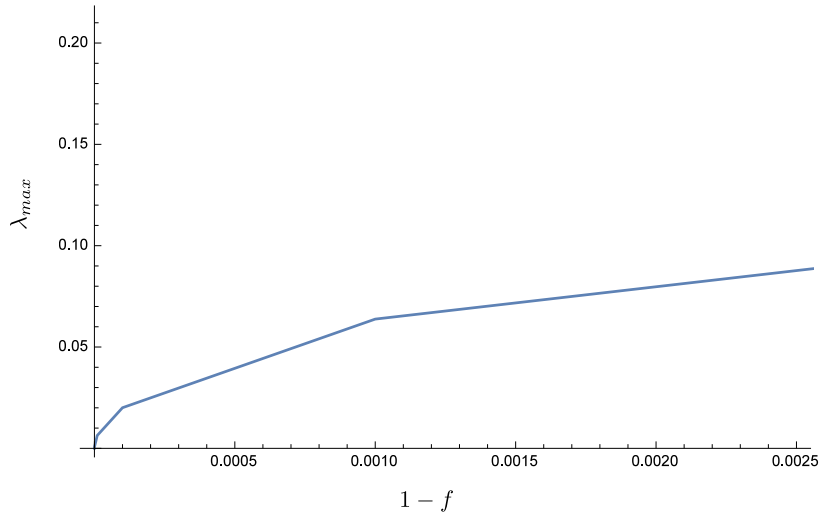

Figure A5. The figure shows the largest eigenvalue of $\mathbf{f}'(\mathbf{p}^\infty)$ (y-axis) for single qubit white noise (x-axis).

The fixed point $\mathbf{p}^\infty$ is determined by four equations in four unknowns, namely the equations

$$p_{\delta_0\delta_1\delta_0\delta_1} = \frac{1}{N} \sum_{\substack{i_1,i_2,j_1 \\ i_1\oplus i_2=\delta_0, i_1\oplus j_1=\delta_1}} \sum_{\substack{\alpha_1,\beta_1,\alpha_2,\beta_2 \\ u(i_1,j_1,i_2,i_1\oplus j_1\oplus i_2)=(\delta_0,\delta_1)}} \tilde{f}_{\alpha_1,\beta_1,\alpha_2,\beta_2} p_{(i_1\oplus\alpha_1)(j_1\oplus\beta_1)(i_1\oplus\alpha_1)(j_1\oplus\beta_1)} \cdot p_{(i_2\oplus\alpha_2)(i_1\oplus j_1\oplus i_2\oplus\beta_2)(i_2\oplus\alpha_2)(i_1\oplus j_1\oplus i_2\oplus\beta_2)},$$

where $\delta_0, \delta_1 \in \{0, 1\}$ and $N = \sum_{\delta_0,\delta_1} p_{\delta_0\delta_1\delta_0\delta_1}$. Figure A6 illustrates the numerical estimate of p^∞_{0000} as a function of $\tilde{f}$.

Furthermore we numerically compute the largest eigenvalue of $\mathbf{f}'(\mathbf{p}^\infty)$ and observe that if $\tilde{f} > 0.8284$, the largest eigenvalue $\lambda_{\max}$ of $\mathbf{f}'(\mathbf{p}^\infty)$ fulfills $\lambda_{\max} < 1$, hence $\mathbf{p}^\infty$ is an attracting fixed point, see figure A7.

Finally, we obtain again a linear relationship between the number of successful distillation rounds n and $\log \|\mathbf{f}'(\mathbf{p}^\infty)^{n-1}\|$ for each noise level $\tilde{f}$, i.e. $b_2(\tilde{f})n + a_2(\tilde{f}) = \log \|\mathbf{f}'(\mathbf{p}^\infty)^{n-1}\|$, see figure A8. This implies, similar to the case of single qubit white noise, that the right-hand side of (A.23) converges polynomial fast towards zero in terms of initial states. The rate of convergence is governed by $\tilde{f}$, i.e. $\|\rho_n - \rho_{\text{fix}}\|_1 \leqslant F_2(N)$ where $F_2(N) \in O(N^{b_2(\tilde{f})})$ and $b_2(\tilde{f}) < 0$ with $b_2(\tilde{f})$ decays for $\tilde{f} \to 1$.

Taking the system of leaking noise transcripts into account,this implies that $\epsilon_n = \| |\psi^n\rangle\langle\psi^n| - |\psi^{\tilde{f}}\rangle\langle\psi^{\tilde{f}}| \|_1 \leqslant \sqrt{F_2(N)}$, i.e. $\epsilon_n \in O(N^{b_2(\tilde{f})/2})$.

To conclude the analysis, we now show that the noise model of two-qubit depolarizing noise is actually sufficient to cover *any* noise process for two-qubit operations. This is the case because for any CNOT-type gate

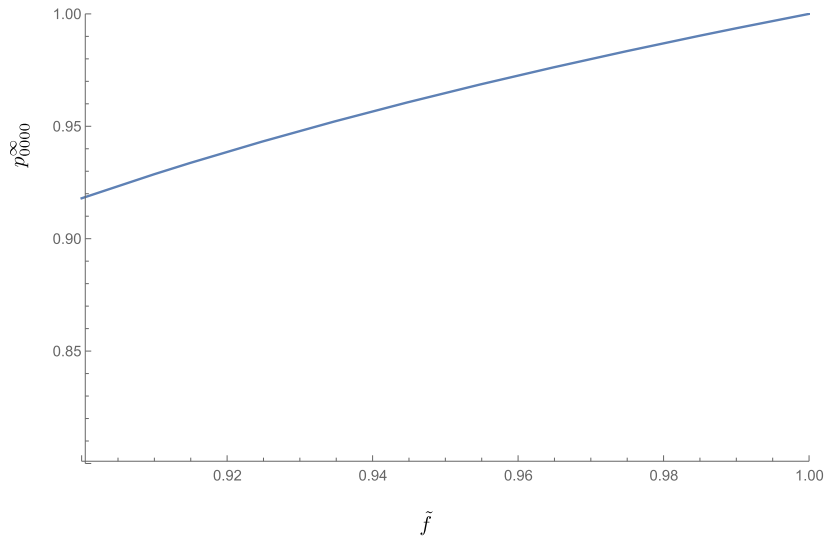

Figure A6. The figure illustrates p^{∞}_{0000} as a function of $\tilde{f}$ for two qubit correlated noise. The fidelity with $|B_{00}\rangle$ of the asymptotic state is equal to unity for a perfect apparatus.

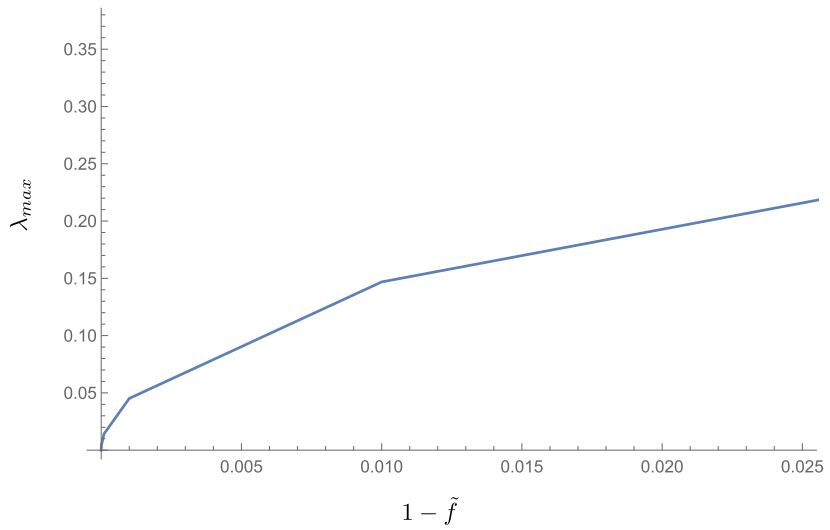

Figure A7. The figure shows the largest eigenvalue of $\mathbf{f}'(\mathbf{p}^{\infty})$ (y-axis) for correlated two qubit noise (x-axis).

(which we need to apply in the case of both recurrence-type entanglement distillation protocols we consider), one can depolarize these gates to a standard form [26]. This is done by randomly applying single-qubit operations before and after the application of the gate, which allows one to reduce any noise characteristics to a specific form with 8 parameters without altering the fidelity of the gate. A further simplification is possible if the noise characteristic of the apparatus is known [26], which could in some cases be achieved through quantum process tomography. In this case, one can add additional (local) noise by randomly choosing to apply the gate, or some other (separable) operation. This allows one to bring any CNOT-type gate (i.e. any two-qubit gate that is equivalent to a CNOT gate up to single qubit unitary operations that are applied before and after the gate) to the standard form

$$\mathcal{E}(\rho) = \tilde{f}\, U\rho U^{\dagger} + \frac{1-\tilde{f}}{16} \sum_{\alpha_1,\beta_1,\alpha_2,\beta_2=0}^{1} \sigma_{\alpha_1,\beta_1}\sigma_{\alpha_2,\beta_2}\rho\sigma_{\alpha_1,\beta_1}\sigma_{\alpha_2,\beta_2}. \tag{A.24}$$

As outlined in [26] this depolarization procedure causes a change in the gate fidelity of the utilized quantum gates. More precisely, if the fidelity of the quantum gate before the depolarization was $F_g = 1 - x$ then the gate fidelity after the depolarization is $F_g' > 1 - 17x$. Thus one reduces the quality of the gate by about an order of magnitude in the worst case by depolarizing to this standard form.

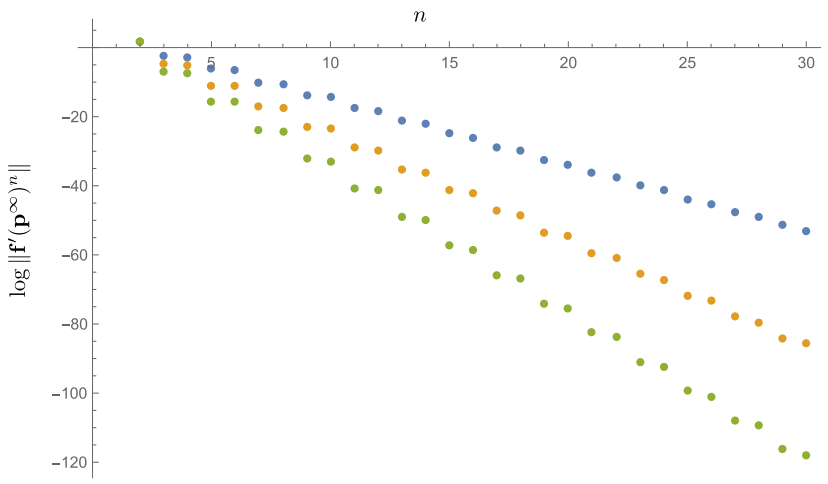

Figure A8. The figure illustrates the value of $\log \|\mathbf{f}'(\mathbf{p}^\infty)^n\|$ as a function of successful distillation rounds for two qubit correlated noise 10^{-2} (blue), 10^{-3} (yellow) and 10^{-4} (green).

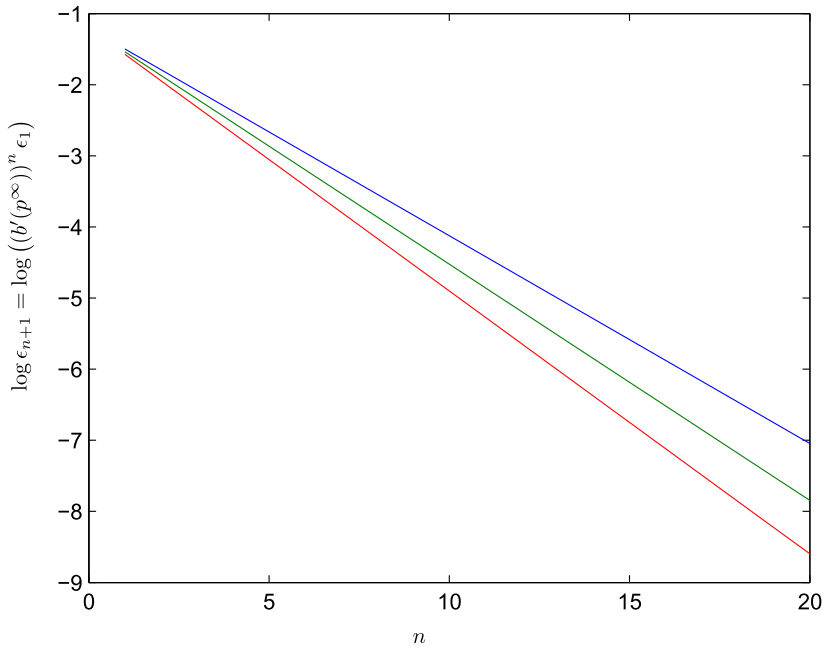

Figure A9. The figure illustrates $\log \epsilon_{n+1}$ for the BBPSSW protocol for different noise parameters $f = 0.97$ (blue), $f = 0.98$ (green) and $f = 0.99$ (red).

We observe that (A.24) can be rewritten as

$$\begin{aligned}\mathcal{E}(\rho) &= \tilde{f}\, U\rho U^\dagger + \frac{1-\tilde{f}}{16} \sum_{\alpha_1,\beta_1,\alpha_2,\beta_2=0}^{1} \sigma_{\alpha_1,\beta_1}\sigma_{\alpha_2,\beta_2}\rho\sigma_{\alpha_1,\beta_1}\sigma_{\alpha_2,\beta_2} \\ &= U\left(\tilde{f}\,\rho + \frac{1-\tilde{f}}{16} \sum_{\alpha_1,\beta_1,\alpha_2,\beta_2=0}^{1} \sigma_{\alpha_1,\beta_1}\sigma_{\alpha_2,\beta_2}\rho\sigma_{\alpha_1,\beta_1}\sigma_{\alpha_2,\beta_2}\right)U^\dagger \\ &= U\left(\sum_{\alpha_1,\beta_1,\alpha_2,\beta_2=0}^{1} \tilde{f}_{\alpha_1,\beta_1,\alpha_2,\beta_2}\sigma_{\alpha_1,\beta_1}\sigma_{\alpha_2,\beta_2}\rho\sigma_{\alpha_1,\beta_1}\sigma_{\alpha_2,\beta_2}\right)U^\dagger, \end{aligned} \tag{A.25}$$

where $\tilde{f}_{0,0,0,0} = \tilde{f} + (1-\tilde{f})/16$ and $\tilde{f}_{\alpha_1,\beta_1,\alpha_2,\beta_2} = (1-\tilde{f})/16$ otherwise. Recall, that one noisy distillation step of the DEJMPS protocol including L is given by (A.2). By introducing $O_{\mathrm{D}} = U_u O'_{2-\mathrm{EPP}}$ we rewrite (A.2) as

$$\rho' = \sum_{\alpha_1,\beta_1,\alpha_2,\beta_2} \tilde{f}_{\alpha_1,\beta_1,\alpha_2,\beta_2} O_{\mathrm{D}}(U^{(a_1)}_{\alpha_1,\beta_1} \otimes U^{(a_2)}_{\alpha_2,\beta_2})(|\Psi\rangle\langle\Psi| \otimes |\Psi\rangle\langle\Psi|)(U^{(a_1)}_{\alpha_1,\beta_1} \otimes U^{(a_2)}_{\alpha_2,\beta_2})^{\dagger} O_{\mathrm{D}}^{\dagger}. \tag{A.26}$$

We observe that the noise maps $U^{(a_1)}_{\alpha_1,\beta_1} \otimes U^{(a_2)}_{\alpha_2,\beta_2}$ in (A.26) act on Alice's part of the systems only. But this is sufficient due to the symmetry of Bell-states—noise on Bobs side can be moved to the other side. Furthermore the additional σ_x-flips introduced on the system(s) of L by the unitaries $U^{(a_1)}_{\alpha_1,\beta_1} \otimes U^{(a_2)}_{\alpha_2,\beta_2}$ are used to keep track of the noise map applied. Because Alice and Bob apply the depolarization procedure as described in [26] and L keeps track of the effective error introduced, we can safely assume that the additional σ_x-flips will be introduced *after* Alice and Bob complete the depolarization procedure, hence it is sufficient to consider two qubit correlated noise introduced at Alice's part of the systems.

A.2. The BBPSSW protocol

The protocol proposed in [28] (also referred to as BBPSSW protocol) is very similar to the DEJMPS protocol. Instead of step 1 of protocol 1 Alice and Bob apply a correlated depolarization procedure (twirl) to their input states which brings them to Werner form.

For the subsequent analysis, suppose that each pair of Alice and Bob is of the form $\rho(p) = p|B_{00}\rangle\langle B_{00}| + (1-p)\frac{1}{4}\mathrm{id}$. We assume that the apparatus applies independent and identical noise of the form $N\rho(p) = f\rho(p) + (1-f)/4(\rho(p) + \sigma_x\rho(p)\sigma_x + \sigma_y\rho(p)\sigma_y + \sigma_z\rho(p)\sigma_z)$ before each distillation step. In similar fashion to the DEJMPS protocol one easily obtains the recurrence relation for the noisy BBPSSW protocol:

$$\tilde{p} = \frac{4p^2f^2 + 2pf}{3p^2f^2 + 3} = b(p).$$

The fixed point p^{∞} of the protocol is obtained by solving the equation $b(p^{\infty}) = p^{\infty}$. A straightforward computation gives the fixed point $p^{\infty} = 2/3 + 1/3\sqrt{4 - 9/f^2 + 6/f}$ (which depends on the noise parameter f). It was shown in [27] that this fixed point is an attractor assuming sufficiently high initial fidelity with $|B_{00}\rangle$ per input pair. Expressing the recurrence relation b in terms of its Taylor series around p^{∞} leads to

$$\tilde{p} = b(p) \approx b(p^{\infty}) + b'(p^{\infty})(p - p^{\infty}). \tag{A.27}$$

Hence, (A.27) provides an approximation of the error in terms of fidelity with $|B_{00}\rangle$ after $n+1$ successful distillation rounds, i.e. $\epsilon_{n+1} = (b'(p^{\infty}))^n\epsilon_1$, see also the plots of figure A9. Moreover, we compute the first derivative of b by

$$b'(p) = \frac{2f(1 + 4fp - f^2p^2)}{3(1 + f^2p^2)^2}.$$

Evaluating b' at p^{∞} yields

$$b'(p^{\infty}) = \frac{9 - 3f}{f(3 + 2(2 + \sqrt{4 - 9/f^2 + 6/f})f)}. \tag{A.28}$$

From this we conclude that, if the apparatus is perfect, i.e. $f = 1$ in (A.28), the error in terms of fidelity with $|B_{00}\rangle$ after $n+1$ successful distillation rounds scales as $\epsilon_{n+1} = (2/3)^n\epsilon_1$.

Using $\log_2 N = n$, where N denotes the number of initial states, we infer for ϵ_{n+1} that

$$\epsilon_{n+1} = \epsilon_1 b'(p^{\infty})^{\log_2 N} = \epsilon_1 (2^{\log_2 b'(p^{\infty})})^{\log_2 N} = \epsilon_1 N^{\log_2 b'(p^{\infty})}.$$

This implies that ϵ_{n+1} scales as $F(N) \in O(N^{\log_2 b'(p^{\infty})})$ and thus $\|\rho_{\mathrm{fix}} - \rho_n\|_1$, where ρ_{fix} and ρ_n denote the fixed point and the state after n successful distillation rounds respectively, scales also as $F(N) \in O(N^{\log_2 b'(p^{\infty})})$ as mentioned in the main text.

For the analysis of two qubit correlated noise we assume that the noisy operations used by the BBPSSW protocol are of the form

$$O_{12}\rho = \tilde{f} O_{12}^{\mathrm{ideal}}\rho + \frac{1-\tilde{f}}{4}\mathrm{tr}_{12}[\rho] \otimes \mathrm{id}_{12}, \tag{A.29}$$

where ρ is a two qubit density operator and O_{12}^{ideal} denotes the ideal two qubit quantum gate. Observe that (A.29) coincides with the standard form of [26]. If the noisy quantum gates are not of the form (A.29) we bring them to that standard form via the same depolarization procedure mentioned in the analysis of the DEJMPS protocol. Hence the following anaylsis is not restricted to this specific noise model, but actually applies to arbitrary noise processes describing noisy two qubit gates.

It has been shown in [27] that the BBPSSW protocol converges for noisy CNOT gates of the form (A.29) to a unique and attracting fixed point if $\tilde{f}$ is sufficiently high. The recurrence relation for the fidelity relative to $|B_{00}\rangle$ obtained in [27] is given by the formula

$$F' = \frac{\tilde{f}^2\left(F^2 + \left(\frac{1-F}{3}\right)^2\right) + \frac{1-\tilde{f}^2}{8}}{\tilde{f}^2\left(F^2 + \frac{2F(1-F)}{3} + 5\left(\frac{1-F}{3}\right)^2\right) + \frac{1-\tilde{f}^2}{2}}. \tag{A.30}$$

Hence one obtains as in [27] the respective fixed points of (A.30) to be

$$F_{\min,\max} = \frac{3 \pm \sqrt{10 - 9/\tilde{f}^2}}{4}.$$

For $F \in (F_{\min}, F_{\max})$ we have that $F' > F$ which shows that $F_{\max}$ is an attracting fixed point. By replacing F' in (A.30) with $\tilde{b}(F)$ we observe similar to (A.27) that the error after $n+1$ successful distillation rounds scales for two qubit correlated noise as $F(N) \in O(N^{\log_2 \tilde{b}'(F_{\max})})$ where N denotes the number of initial states.

Finally we provide a worst case analysis of the BBPSSW protocol. For that purpose assume the following scenario: the noisy apparatus performs with probability f_I the ideal distillation step $\mathcal{E}_I$ and introduces with probability $1 - f_I$ an arbitrary noise map $\mathcal{E}_\perp$. More precisely, we decompose the distillation step taken by Alice and Bob before the measurement of the target system as the CP map

$$\mathcal{E}(\rho) = f_I \mathcal{E}_I(\rho) + (1 - f_I)\mathcal{E}_\perp(\rho),$$

where ρ is a four qubit density operator. Notice that one can always decompose a noisy map in this form, where both maps are completely positive and trace preserving. We remark, however, that the map $\mathcal{E}_I$ denotes the ideal protocol which includes an abort option, i.e. we only keep the first pair if the results of the measurements on the second pair coincide. The map $\mathcal{E}_\perp$ may similarly contain such an abort branch. The noise parameter f_I describes the quality of the overall map[9], i.e. one can think of the process that with probability f_I the desired procedure (including gates and measurements) is performed, while with probability $(1 - f_I)$ something else happens (described by the map $\mathcal{E}_\perp$).

We will now consider the worst case for the map $\mathcal{E}_\perp$ w.r.t. entanglement distillation. The worst case for the BBPSSW protocol is that the apparatus introduces a state orthogonal to $|B_{00}\rangle$ on the source system and the state $|B_{00}\rangle$ on the target system as this will always contribute to the overall success probability of a distillation step of the BBPSSW protocol but lead at the same time to a lower fidelity relative to $|B_{00}\rangle$ after the measurement of the target system compared to the ideal distillation step. One example for such a map is given by $\mathcal{E}_\perp(\rho) = |B_{01}\rangle\langle B_{01}| \otimes |B_{00}\rangle\langle B_{00}|$. Any other map will lead to a larger fidelity after the distillation step followed by depolarization to Werner form. We thus have

$$F' \geqslant \frac{f_I\left(F^2 + \left(\frac{1-F}{3}\right)^2\right)}{f_I\left(F^2 + \frac{2F(1-F)}{3} + 5\left(\frac{1-F}{3}\right)^2\right) + 1 - f_I} \tag{A.31}$$

for the fidelity relative to $|B_{00}\rangle$. This formula can be understood as follows: the ideal protocol is applied with probability f_I, and succeeds with probability f_{suc}, thereby producing a fidelity $\tilde{F}$. The map $\mathcal{E}_\perp$ is applied with probability $(1 - f_I)$, does never abort and does not contribute to the final fidelity (which is clearly the worst case). We thus have $F' \geqslant f_I f_{\text{suc}} \tilde{F}/[f_I f_{\text{suc}} + (1 - f_I)]$.

We now analyze the worst case scenario, i.e. assuming equality in (A.31). Since we know that at each step the actual noise map produces an output density operator with a larger fidelity than the worst-case map, we can conclude that the resulting fidelity of any noise map will be larger than the fixed point which is achieved by the worst-case map. We remark, however, that this does not constitute a full confidentiality proof for arbitrary noise maps, as it is not evident from this analysis that for any fixed noise map a unique fixed point is reached. Assuming equality in (A.31), one can compute that the fixed points of the noisy BBPSSW protocol are in this case given by the solutions of

$$-f_I + (9 - 2f_I)F_\infty - 14f_I F_\infty^2 + 8f_I F_\infty^3 = 0 \tag{A.32}$$

which only depend on the noise parameter f_I. We define $g_{\text{fix}}(x, f_I) = -f_I + (9 - 2f_I)x - 14f_I x^2 + 8f_I x^3$ which implies that (A.32) reads as $g_{\text{fix}}(F_\infty, f_I) = 0$. The question how many solutions of (A.32) are real we easily answer by the discriminant of g_{fix}. We obtain for the discriminant of g_{fix}

$$\Delta(f_I) = -36(648f_I - 873f_I^2 - 212f_I^3 + 436f_I^4). \tag{A.33}$$

Hence if $\Delta(f_I) > 0$ then all three solutions of (A.32) are real. We numerically estimate that $\Delta(f_{I_{\text{crit}}}) = 0$ for $f_{I_{\text{crit}}} \approx 0.9641$, hence for $f_I > f_{I_{\text{crit}}}$ there exist three real solutions of (A.32) because $\Delta(f_I) > 0$ for $f_I > f_{I_{\text{crit}}}$, see figure A10. Thus, for $f_I > f_{I_{\text{crit}}}$, we compute the fixed points of the noisy BBPSSW protocol via solving (A.32). Figure A11 shows the function g_{fix} for different values of f_I. From figures A10 and A11 we infer that we have three possible fixed points for $f_I > f_{I_{\text{crit}}}$. Hence we need to show that the fixed point with the highest fidelity

[9] We remark that a similar analysis can be performed by modeling local operations of Alice and Bob sepearetely in this way.

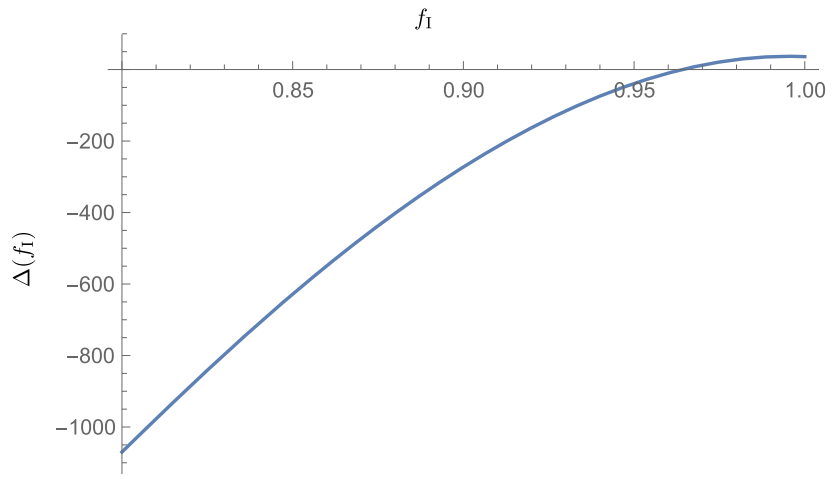

Figure A10. The figure illustrates the discriminant $\Delta(f_I)$ of (A.32). For $f_I > 0.9641$ we have $\Delta(f_I) > 0$.

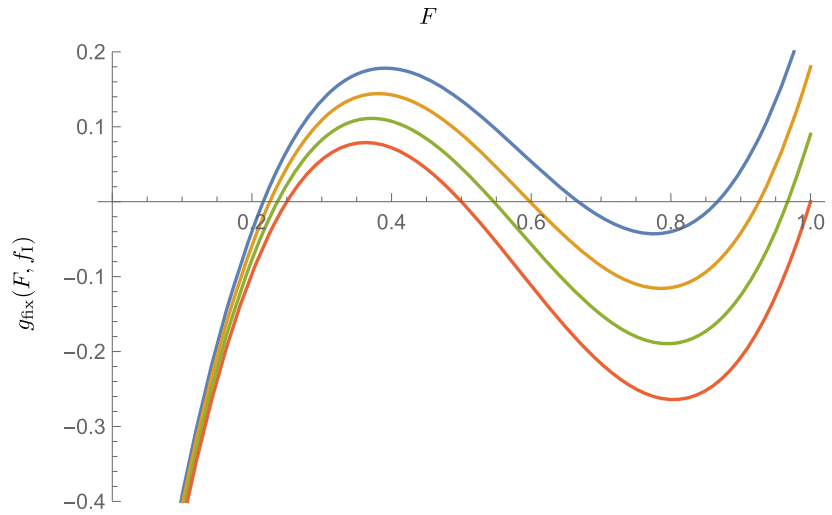

Figure A11. The figure illustrates the function g_{fix} for $f_I = 0.97$ (blue), $f_I = 0.98$ (yellow), $f_I = 0.99$ (green) and $f_I = 1$ (red). The fixed points of the BBPSSW protocol correspond to the zero's of $g_{fix}(F, f_I)$.

relative to $|B_{00}\rangle$ obtained via (A.32) is an attracting fixed point. We solve this issue by showing that $F' > F$ for $F \in (F_{min}, F_{max})$ (where F_{min} denotes the second, and F_{max} the third fixed point in figure A11). From figure A12 we find that $F' - F > 0$ for $f_I > f_{I_{crit}}$, hence $F' > F$ which shows that F_{max} is an attracting fixed point whenever starting with initial fidelity $F > F_{min}$.

Furthermore, by assuming equality in (A.31) and replacing F' with $b_\perp(F)$, we observe similar to (A.27) that the error after $n + 1$ successful distillation rounds scales in this worst case analysis as $F(N) \in O(N^{\log_2 b'_\perp(F_{max})})$ where N denotes the number of initial states.

Appendix B. Confidentiality of entanglement distillation protocols

In this section we provide the proofs of lemmas 3 and 5 of the main text, crucial for the de-Finetti-based and post-selection-based reduction techniques. Both proofs require only one specific property of the real protocol $\mathcal{E}^\alpha$: after passing the parameter estimation phase the entanglement distillation protocol always converges to *one* fixed point, i.e. the fixed point is *unique*, an *attractor* for all the states which pass the parameter estimation and depends on the noise parameters *only*, as this implies that the distance with respect to the 1-norm within the ok-branch of the protocol is bounded and converges towards zero.

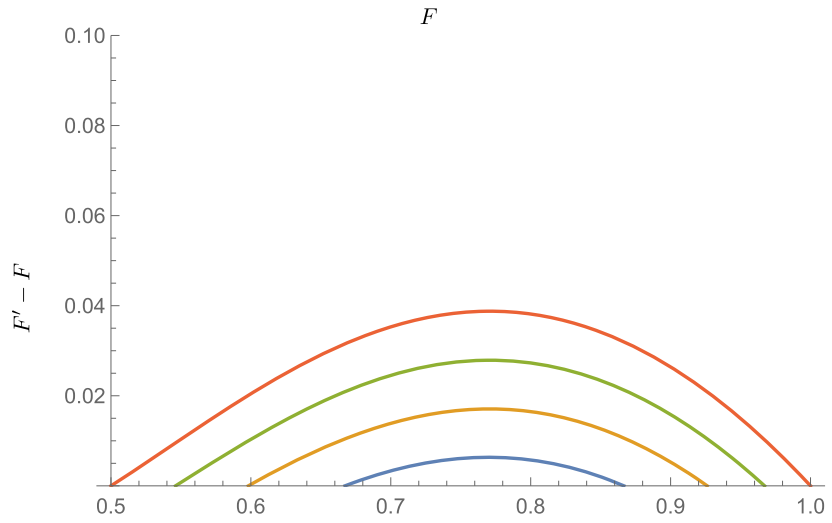

Figure A12. The figure illustrates $F' - F$ for $f_I = 0.97$ (blue), $f_I = 0.98$ (yellow), $f_I = 0.99$ (green) and $f_I = 1$ (red).

B.1. Proof of lemma 3

We first state the following lemma which establishes a connection between measurements on one subsystem of a bipartite state and tensor product states.

Lemma 8 (Steering of local states). *Let ρ_{AB} be a bipartite (in general, mixed) state and let $\rho_A = \mathrm{tr}_B[\rho_{AB}]$ and $\rho_B = \mathrm{tr}_A[\rho_{AB}]$. Furthermore let ρ_B^ϕ be defined as*

$$\rho_B^\phi = \frac{\mathrm{tr}_A[(|\phi\rangle\langle\phi| \otimes I)\rho_{AB}]}{p_A(\phi)},$$

where $|\phi\rangle \in \mathcal{H}_A$ and $p_A(\phi) = \mathrm{tr}(|\phi\rangle\langle\phi|\ \rho_A)$. If $\|\rho_B^\phi - \rho_B\|_1 \leqslant \epsilon$ for all $|\phi\rangle \in \mathcal{H}_A$, then

$$\|\rho_{AB} - \rho_A \otimes \rho_B\|_1 \leqslant 2C\epsilon, \tag{B.1}$$

where C only depends on the dimensions of A and B. In particular, if we fix the number of qubits of A and B to 2 respectively, then we have $C = 4^8$.

Proof. In the following we denote the four Pauli operators by

$$\sigma_0 = \mathrm{id}, \quad \sigma_1 = \sigma_x, \quad \sigma_2 = \sigma_z, \quad \sigma_3 = \sigma_y.$$

First we decompose ρ_{AB} in the Pauli basis, i.e. we have

$$\rho_{AB} = \frac{1}{2^{n+m}} \sum_{\mathbf{i},\mathbf{j}} \alpha_{\mathbf{ij}} \sigma_{\mathbf{i}} \otimes \sigma_{\mathbf{j}}, \tag{B.2}$$

where n and m denote the number of qubits of A and B respectively and we use the notations $\mathbf{i} = (i_1, .., i_n)$ and $\mathbf{j} = (j_1, .., j_m)$ where each i_k and j_k are in $\{0, .., 3\}$ as well as $\sigma_{\mathbf{i}} = \bigotimes_{k=1}^n \sigma_{i_k}$ and $\sigma_{\mathbf{j}} = \bigotimes_{k=1}^m \sigma_{j_k}$. Recall that $\mathrm{tr}(\sigma_0) = 2$ and $\mathrm{tr}(\sigma_1) = \mathrm{tr}(\sigma_2) = \mathrm{tr}(\sigma_3) = 0$. From this one easily computes ρ_A and ρ_B by

$$\rho_A = \mathrm{tr}_B[\rho_{AB}] = \frac{1}{2^{n+m}} \sum_{\mathbf{i},\mathbf{j}} \alpha_{\mathbf{ij}} \sigma_{\mathbf{i}} \mathrm{tr}(\sigma_{\mathbf{j}}) = \frac{1}{2^n} \sum_{\mathbf{i}} \alpha_{\mathbf{i0}} \sigma_{\mathbf{i}}, \tag{B.3}$$

$$\rho_B = \mathrm{tr}_A[\rho_{AB}] = \frac{1}{2^{n+m}} \sum_{\mathbf{i},\mathbf{j}} \alpha_{\mathbf{ij}} \mathrm{tr}(\sigma_{\mathbf{i}}) \sigma_{\mathbf{j}} = \frac{1}{2^m} \sum_{\mathbf{j}} \alpha_{\mathbf{0j}} \sigma_{\mathbf{j}}. \tag{B.4}$$

Using (B.2)–(B.4) we obtain for (B.1)

$$\begin{aligned} \|\rho_{AB} - \rho_A \otimes \rho_B\|_1 &\leqslant \frac{1}{2^{n+m}} \sum_{\mathbf{i},\mathbf{j}} \|(\alpha_{\mathbf{ij}} - \alpha_{\mathbf{i0}}\alpha_{\mathbf{0j}})\sigma_{\mathbf{i}} \otimes \sigma_{\mathbf{j}}\|_1 = \frac{1}{2^{n+m}} \sum_{\mathbf{i},\mathbf{j}} |\alpha_{\mathbf{ij}} - \alpha_{\mathbf{i0}}\alpha_{\mathbf{0j}}| \cdot \|\sigma_{\mathbf{i}} \otimes \sigma_{\mathbf{j}}\|_1 \\ &= \frac{2^{n+m}}{2^{n+m}} \sum_{\mathbf{i},\mathbf{j}} |\alpha_{\mathbf{ij}} - \alpha_{\mathbf{i0}}\alpha_{\mathbf{0j}}| = \sum_{\mathbf{i},\mathbf{j}} |\alpha_{\mathbf{ij}} - \alpha_{\mathbf{i0}}\alpha_{\mathbf{0j}}| = \|\mathbf{a} - \mathbf{a}'\|_{1;\mathbb{C}^{4^{n+m}}}, \end{aligned} \tag{B.5}$$

where $\mathbf{a} = (\alpha_{00}, .., \alpha_{3^n 3^m})$, $\mathbf{a}' = (\alpha_{00}\alpha_{00}, .., \alpha_{3^n 0}\alpha_{03^m})$ and $\|\cdot\|_{1;\mathbb{C}^{4^{n+m}}}$ denotes the 1-norm of vectors in $\mathbb{C}^{4^{n+m}}$. Hence in order to prove (B.1) it is sufficient to prove $\|\mathbf{a} - \mathbf{a}'\|_{1;\mathbb{C}^{4^{n+m}}} \leqslant 2C\epsilon$. By assumption we have for ρ_B^ϕ where $|\phi\rangle \in \mathcal{H}_A$ and $p_A(\phi) = \mathrm{tr}((|\phi\rangle\langle\phi| \otimes I)\rho_{AB})$ that $\|\rho_B^\phi - \rho_B\|_1 \leqslant \epsilon$ for all $|\phi\rangle \in \mathcal{H}_A$. Moreover,

according to theorem 9.1 in [32] we have for all $|\xi\rangle \in \mathcal{H}_B$

$$\frac{1}{2}|p_B(\xi|\phi) - q_B(\xi)| = \frac{1}{2}|\mathrm{tr}(|\xi\rangle\langle\xi| \rho_B^{\phi}) - \mathrm{tr}(|\xi\rangle\langle\xi| \rho_B)| \leqslant \max_{E_m} \frac{1}{2}\sum_m |\mathrm{tr}(E_m\rho_B^{\phi}) - \mathrm{tr}(E_m\rho_B)| = \|\rho_B^{\phi} - \rho_B\|_1 \leqslant \epsilon, \tag{B.6}$$

where $p_B(\xi|\phi)$ denotes the conditional probability of obtaining the outcome ϕ on system A and the outcome ξ on system B and $\{E_m\}$ denotes a POVM on the subsystem of B. Suppose we perform a projective measurement on the systems of A and B denoted by $\{|\psi_k\rangle_{AB}\} = \{|\phi_k\rangle_A \otimes |\xi_k\rangle_B\}$ where $k \in \{1, .., 4^{n+m}\}$ on ρ_{AB} and $\rho_A \otimes \rho_B$. This yields for the respective probabilities $p_{AB}(\psi_k)$ and $q_{AB}(\psi_k)$ of observing outcome k for ρ_{AB} and $\rho_A \otimes \rho_B$

$$\begin{aligned} p_{AB}(\psi_k) &= \mathrm{tr}(|\psi_k\rangle\langle\psi_k| \rho_{AB}) = \mathrm{tr}(|\phi_k\rangle\langle\phi_k|_A \otimes |\xi_k\rangle\langle\xi_k|_B\rho_{AB}) = \mathrm{tr}(|\xi_k\rangle\langle\xi_k|_B\mathrm{tr}_A[(|\phi_k\rangle\langle\phi_k|_A \otimes I)\rho_{AB}]) \\ &= \mathrm{tr}(|\xi_k\rangle\langle\xi_k|_B p_A(\phi_k)\rho_B^{\phi_k}) = p_A(\phi_k)\mathrm{tr}(|\xi_k\rangle\langle\xi_k|_B\rho_B^{\phi_k}) = p_A(\phi_k)p_B(\xi_k|\phi_k), \\ q_{AB}(\psi_k) &= \mathrm{tr}(|\psi_k\rangle\langle\psi_k| \rho_A \otimes \rho_B) = \mathrm{tr}(|\phi_k\rangle\langle\phi_k| \rho_A)\mathrm{tr}(|\xi_k\rangle\langle\xi_k| \rho_B) = q_A(\phi_k)q_B(\xi_k) \end{aligned}$$

where $p_B(\xi_k|\phi_k)$ denotes the conditional probability of obtaining outcome ϕ_k on system A first and obtaining outcome ξ_k on system B. We observe $p_A(\phi_k) = q_A(\phi_k)$. Thus we obtain

$$|p_{AB}(\psi_k) - q_{AB}(\psi_k)| = p_A(\phi_k)|p_B(\xi_k|\phi_k) - q_B(\xi_k)| \leqslant 2\epsilon p_A(\phi_k)$$

using (B.6). In order to compute a bound for (B.5) we use quantum state tomography, see e.g. [33]. For that purpose we perform an informationally complete POVM induced by different separable bases on $\mathcal{H}_A \otimes \mathcal{H}_B$. More precisely, we choose that many POVMs such that we have in total 4^{n+m} different outcomes. We observe for $|\psi_k\rangle_{AB} = |\phi_k\rangle_A \otimes |\xi_k\rangle_B$ that

$$p_{AB}(\psi_k) = \frac{1}{2^{n+m}}\sum_{\mathbf{i},\mathbf{j}} \langle\phi_k| \sigma_{\mathbf{i}}|\phi_k\rangle\langle\xi_k| \sigma_{\mathbf{j}}|\xi_k\rangle\alpha_{\mathbf{ij}} \quad \text{and} \quad q_{AB}(\psi_k) = \frac{1}{2^{n+m}}\sum_{\mathbf{i},\mathbf{j}} \langle\phi_k| \sigma_{\mathbf{i}}|\phi_k\rangle\langle\xi_k| \sigma_{\mathbf{j}}|\xi_k\rangle\alpha_{\mathbf{i0}}\alpha_{\mathbf{0j}}. \tag{B.7}$$

Enumerating (B.7) for $1 \leqslant k \leqslant 4^{n+m}$ yields 4^{n+m} equations for $\mathbf{a}$, i.e.

$$p_{AB}(\psi_1) = \frac{1}{2^{n+m}}\sum_{\mathbf{i},\mathbf{j}} \langle\phi_1| \sigma_{\mathbf{i}}|\phi_1\rangle\langle\xi_1| \sigma_{\mathbf{j}}|\xi_1\rangle\alpha_{\mathbf{ij}}, \tag{B.8}$$

$$\cdots$$

$$p_{AB}(\psi_{4^{n+m}}) = \frac{1}{2^{n+m}}\sum_{\mathbf{i},\mathbf{j}} \langle\phi_{4^{n+m}}| \sigma_{\mathbf{i}}|\phi_{4^{n+m}}\rangle\langle\xi_{4^{n+m}}| \sigma_{\mathbf{j}}|\xi_{4^{n+m}}\rangle\alpha_{\mathbf{ij}} \tag{B.9}$$

as well as 4^{n+m} equations for $\mathbf{a}'$

$$q_{AB}(\psi_1) = \frac{1}{2^{n+m}}\sum_{\mathbf{i},\mathbf{j}} \langle\phi_1| \sigma_{\mathbf{i}}|\phi_1\rangle\langle\xi_1| \sigma_{\mathbf{j}}|\xi_1\rangle\alpha_{\mathbf{i0}}\alpha_{\mathbf{0j}}, \tag{B.10}$$

$$\cdots$$

$$q_{AB}(\psi_{4^{n+m}}) = \frac{1}{2^{n+m}}\sum_{\mathbf{i},\mathbf{j}} \langle\phi_{4^{n+m}}| \sigma_{\mathbf{i}}|\phi_{4^{n+m}}\rangle\langle\xi_{4^{n+m}}| \sigma_{\mathbf{j}}|\xi_{4^{n+m}}\rangle\alpha_{\mathbf{i0}}\alpha_{\mathbf{0j}}. \tag{B.11}$$

We can rewrite the systems of equations (B.8), (B.9) and (B.10), (B.11) using

$$T = \begin{pmatrix} \langle\phi_1| \sigma_0|\phi_1\rangle\langle\xi_1| \sigma_0|\xi_1\rangle & \cdots & \langle\phi_1| \sigma_{3^n}|\phi_1\rangle\langle\xi_1| \sigma_{3^m}|\xi_1\rangle \\ \cdots & \cdots & \cdots \\ \langle\phi_{4^{n+m}}| \sigma_0|\phi_{4^{n+m}}\rangle\langle\xi_{4^{n+m}}| \sigma_0|\xi_{4^{n+m}}\rangle & \cdots & \langle\phi_{4^{n+m}}| \sigma_{3^n}|\phi_{4^{n+m}}\rangle\langle\xi_{4^{n+m}}| \sigma_{3^m}|\xi_{4^{n+m}}\rangle \end{pmatrix}$$

and $\mathbf{p} = (p_{AB}(\psi_1), .., p_{AB}(\psi_{4^{n+m}}))$ and $\mathbf{q} = (q_{AB}(\psi_1), .., q_{AB}(\psi_{4^{n+m}}))$ as

$$\mathbf{p} = \frac{1}{2^{n+m}}T\mathbf{a} \quad \text{and} \quad \mathbf{q} = \frac{1}{2^{n+m}}T\mathbf{a}'$$

respectively. Hence $2^{n+m}(\mathbf{p} - \mathbf{q}) = T(\mathbf{a} - \mathbf{a}')$. Moreover we observe that T is invertible if the POVM is informationally complete, see [33] for details. Thus, inverting T and taking norms on both sides yields

$$\begin{aligned} \|\mathbf{a} - \mathbf{a}'\|_{1;\mathbb{C}^{4^{n+m}}} &\leqslant 2^{n+m}\|T^{-1}\|\|\mathbf{p} - \mathbf{q}\|_{1;\mathbb{C}^{4^{n+m}}} = 2^{n+m}\|T^{-1}\|\sum_k |p_{AB}(\psi_k) - q_{AB}(\psi_k)| \\ &\leqslant 2^{n+m}\|T^{-1}\|\sum_k 2\epsilon p_A(\phi_k) \leqslant 2\|T^{-1}\|4^{n+m}2^{n+m}\epsilon, \end{aligned}$$

which completes the proof for the general case with $C = \|T^{-1}\|4^{n+m}2^{n+m}$.

Before we complete the lemma we need to determine C for the case of $n = m = 2$. We choose $|\phi_{4^3(j_1-1)+4^2(j_2-1)+4(j_3-1)+j_4}\rangle = |\phi'_{j_1}\rangle \otimes |\phi'_{j_2}\rangle \otimes |\phi'_{j_3}\rangle \otimes |\phi'_{j_4}\rangle$ where $j_1, j_2, j_3, j_4 \in \{1, 2, 3, 4\}$ and

$$|\phi'_1\rangle = (|0\rangle + |1\rangle)/\sqrt{2}, \tag{B.12}$$

$$|\phi_2'\rangle = (|0\rangle + i|1\rangle)/\sqrt{2}, \tag{B.13}$$

$$|\phi_3'\rangle = |0\rangle, \tag{B.14}$$

$$|\phi_4'\rangle = (|0\rangle - |1\rangle)/\sqrt{2}, \tag{B.15}$$

which is informationally complete and thus a valid choice. This choice of $|\phi_l'\rangle$ corresponds to a Pauli tomography on a single qubit. We observe that the matrix T is invertible and compute $\|T^{-1}\| = 16$. Thus $C = 4^8$ which completes the proof. □

Roughly speaking lemma 8 states that if all post-selected reduced states of a bipartite state, where each partition consists of two qubits, are η-close then the overall state is $2 \cdot 4^8\eta$ close to a product state.

We gave the lemma in a more general form as it may have utility beyond the scope of this paper. However for our purposes we need a stronger, but more specific result. In the following lemma we will show that we can achieve the same result even if the measurements must succeed above a threshold, which is important in the application of the lemma.

Lemma 9. *In the situation of lemma 8 for $n = m = 2$ it suffice to consider measurements on the subsystem A which have a probability greater than or equal to 1/16.*

More precisely, for every state ρ_{AB} there exists a unitary U acting on system A and a state $\rho'_{AB} = (U \otimes I_B)\rho_{AB}(U \otimes I_B)^\dagger$, such that if the state ρ'_{AB} meets the conditions of lemma 8, i.e. subsystem B is ϵ-non-steerable via measurements on subsystem A for all measurements with probability greater than or equal to 1/16, then

$$\|\rho_{AB} - \rho_A \otimes \rho_B\|_1 \leqslant 2C\epsilon. \tag{B.16}$$

Proof. First we construct the state ρ'_{AB} associated with ρ_{AB} and show that it suffice to consider measurements of probability greater than or equal to $1/16$. Recall the situation of lemma 8. Let ρ_{AB} be a bipartite (in general, mixed) state and let $\rho_A = \mathrm{tr}_B[\rho_{AB}]$ and $\rho_B = \mathrm{tr}_A[\rho_{AB}]$. Furthermore let ρ_B^ϕ be defined as

$$\rho_B^\phi = \frac{\mathrm{tr}_A[(|\phi\rangle\langle\phi| \otimes I)\rho_{AB}]}{p_A(\phi)},$$

where $|\phi\rangle \in \mathcal{H}_A$ and $p_A(\phi) = \mathrm{tr}(|\phi\rangle\langle\phi| \rho_A)$. Then the claim of lemma 8 was: if $\|\rho_B^\phi - \rho_B\|_1 \leqslant \epsilon$ for all $|\phi\rangle \in \mathcal{H}_A$, then

$$\|\rho_{AB} - \rho_A \otimes \rho_B\|_1 \leqslant 2C\epsilon \tag{B.17}$$

where C only depends on the dimensions of A and B. In particular, if we fix the number of qubits of A and B to 2 respectively, then we have $C = 4^8$.

Further recall that the set $|\phi_{4^3(j_1-1)+4^2(j_2-1)+4(j_3-1)+j_4}\rangle = |\phi'_{j_1}\rangle \otimes |\phi'_{j_2}\rangle \otimes |\phi'_{j_4}\rangle \otimes |\phi'_{j_4}\rangle$ where $j_1, j_2, j_3, j_4 \in \{1, 2, 3, 4\}$ of lemma 8, i.e. (B.12)–(B.15), is informationally complete and thus suffice to reconstruct any 4 qubit quantum state where

$$|\phi_1'\rangle = (|0\rangle + |1\rangle)/\sqrt{2}, \tag{B.18}$$

$$|\phi_2'\rangle = (|0\rangle + i|1\rangle)/\sqrt{2}, \tag{B.19}$$

$$|\phi_3'\rangle = |0\rangle, \tag{B.20}$$

$$|\phi_4'\rangle = (|0\rangle - |1\rangle)/\sqrt{2}. \tag{B.21}$$

In order to prove the claim, we use the following observation: the state $\rho_A = \mathrm{tr}_B[\rho_{AB}]$ is a two qubit quantum state, so it can be written as

$$\rho_A = \sum_{j=0}^{3} \lambda_j |\Psi_j\rangle\langle\Psi_j|, \tag{B.22}$$

where the states $|\Psi_j\rangle$ correspond to the (orthogonal) eigenstates of ρ_A for the real non-negative eigenvalues λ_j. Hence there exists at least one $j' \in \{0, 1, 2, 3\}$ such that $\lambda_{j'} \geqslant 1/4$, which corresponds to the maximum of the eigenvalues λ_j. Now we choose a local unitary U such that $U|\Psi_{j'}\rangle = |0\rangle \otimes |0\rangle$. Applying this unitary to (B.22) therefore leads to the state

$$\rho'_A = U\left(\sum_{j=0}^{3} \lambda_j |\Psi_j\rangle\langle\Psi_j|\right)U^\dagger = \lambda_{j'}|00\rangle\langle 00| + \sum_{j\neq j'}^{3} \lambda_j |\varphi_j\rangle\langle\varphi_j|, \tag{B.23}$$

where $|\varphi_j\rangle = U|\psi_j\rangle$. We compute the probability for any projector applied on ρ'_A which is taken from the set (B.18)–(B.21) and of the form $|\phi'\rangle\langle\phi'| = |\phi'_k\rangle\langle\phi'_k| \otimes |\phi'_l\rangle\langle\phi'_l|$ by

$$\begin{aligned}\mathrm{tr}(|\phi'\rangle\langle\phi'|\ \rho'_A) &= \mathrm{tr}(|\phi'_k\rangle\langle\phi'_k|\otimes|\phi'_l\rangle\langle\phi'_l|\ U\rho_A U^\dagger) = \sum_{j=0}^{3}\lambda_j \mathrm{tr}(|\phi'_k\rangle\langle\phi'_k|\otimes|\phi'_l\rangle\langle\phi'_l|\ U|\Psi_j\rangle\langle\Psi_j|\ U^\dagger)\\ &\geqslant \frac{1}{4}\mathrm{tr}(|\phi'_k\rangle\langle\phi'_k|\otimes|\phi'_l\rangle\langle\phi'_l|\ U|\Psi_{j'}\rangle\langle\Psi_{j'}|\ U^\dagger) = \frac{1}{4}\mathrm{tr}(|\phi'_k\rangle\langle\phi'_k|\otimes|\phi'_l\rangle\langle\phi'_l||00\rangle\langle 00|)\\ &= \frac{1}{4}\mathrm{tr}(|\phi'_k\rangle\langle\phi'_k||0\rangle\langle 0|)\mathrm{tr}(|\phi'_l\rangle\langle\phi'_l||0\rangle\langle 0|) \geqslant \frac{1}{4}\frac{1}{2}\frac{1}{2} = \frac{1}{16},\end{aligned} \tag{B.24}$$

where we have used that $\mathrm{tr}(AB) = \mathrm{tr}(BA)$ for matrices A and B and that $\mathrm{tr}(|\phi'_k\rangle\langle\phi'_k||0\rangle\langle 0|) \geqslant 1/2$ for all $k \in \{0, 1, 2, 3\}$.

So we define the state ρ'_{AB} as $\rho'_{AB} = (U\otimes I_B)\rho_{AB}(U\otimes I_B)^\dagger$. Observe that the probabilities of all projectors within the tomographic set (B.18)–(B.21) are greater than or equal to $1/16$ for the state ρ'_A.

Now suppose we perform a measurement from the tomographic set (B.18)–(B.21) on the subsystem A of ρ'_{AB} yielding outcome $|\phi\rangle$. The post-selected state conditioned on $|\phi\rangle$ reads as

$$\rho'^{\phi}_B = \frac{\mathrm{tr}_A[(|\phi\rangle\langle\phi|\otimes I)\rho'_{AB}]}{p_A(\phi)},$$

where $p_A(\phi) \geqslant 1/16$. Furthermore assume as in lemma 8 that $\|\rho'^{\phi}_B - \rho_B\|_1 \leqslant \epsilon$ for all such $|\phi\rangle \in \mathcal{H}_A$. Then lemma 8 implies that

$$\|\rho'_{AB} - \rho'_A\otimes\rho_B\|_1 \leqslant 2C\epsilon. \tag{B.25}$$

The proof completes by observing that ρ'_{AB} and $\rho'_A\otimes\rho_B$ are related by the local unitary U to ρ_{AB} and $\rho_A\otimes\rho_B$ and the unitary equivalence of the trace distance, i.e.

$$\|\rho_{AB} - \rho_A\otimes\rho_B\|_1 = \|(U\otimes I_B)(\rho_{AB} - \rho_A\otimes\rho_B)(U\otimes I_B)^\dagger\|_1 \tag{B.26}$$

$$= \|\rho'_{AB} - \rho'_A\otimes\rho_B\|_1 \leqslant 2C\epsilon. \tag{B.27}$$

□

We observe that, due to the proof of lemma 9, which relies on the informationally complete set (B.18)–(B.21), it suffices to be non-steerable with respect to the measurements within that set for a probability of measurement above or equal to $1/16$. We actually have proven a stronger result, as the actual choice of measurements does not matter, provided the probability of success is above or equal to the threshold $1/16$.

Lemma (Lemma 3 in main text—product form lemma). *Let ρ be an arbitrary mixed state shared by Alice and Bob and let $|\psi\rangle_{ABE}$ be a purification thereof held by Eve. Furthermore, let $\mathcal{P}_1$ correspond to a (distillation-type) real protocol and $\mathcal{P}_2$ correspond to the associated (distillation-type) ideal protocol, i.e.*

$$\mathcal{P}_1(\rho) = p_\rho\sigma_{AB}\otimes|\mathrm{ok}\rangle\langle\mathrm{ok}| + (1-p_\rho)\sigma^{\perp}_{AB}\otimes|\mathrm{fail}\rangle\langle\mathrm{fail}|,$$

$$\mathcal{P}_2(\rho) = p_\rho\sigma^{\alpha}_{AB}\otimes|\mathrm{ok}\rangle\langle\mathrm{ok}| + (1-p_\rho)\sigma^{\perp}_{AB}\otimes|\mathrm{fail}\rangle\langle\mathrm{fail}|,$$

where α characterizes the level of the noise, σ^{α}_{AB}, and $\sigma^{\perp}_{AB}$ are two fixed two qubit states. Furthermore, let $\mathcal{P}_1$ and $\mathcal{P}_2$ satisfy the following properties:

(1) The noise transcripts do not leak to Eve.

(2) The protocol $\mathcal{P}_1$ guarantees to converge towards some state σ^{α}_{AB} within the ok-branch of the protocol and $\max_{\mu_{AB}}\|(\mathcal{P}_1-\mathcal{P}_2)(\mu_{AB})\|_1 \leqslant \varepsilon$.

Then it holds that

$$\|(\mathcal{P}_1\otimes\mathrm{id}_E - \mathcal{P}_2\otimes\mathrm{id}_E)(|\psi\rangle\langle\psi|_{ABE})\|_1 \leqslant (34\cdot 4^8+1)\varepsilon. \tag{B.28}$$

Proof. The proof relies on lemmas 8 and 9. Suppose Eve prepares the pure state $|\psi\rangle_{ABE}$ and let $\mathrm{tr}_E[|\psi\rangle\langle\psi|] = \rho_{AB}$ be the state received by Alice and Bob. Then we have

$$\begin{aligned}(\mathcal{P}_1\otimes\mathrm{id}_E)(|\psi\rangle\langle\psi|) &= p_\rho\sigma_{ABE}\otimes|\mathrm{ok}\rangle\langle\mathrm{ok}| + (1-p_\rho)\sigma^{\perp}_{AB}\otimes\sigma_E\otimes|\mathrm{fail}\rangle\langle\mathrm{fail}|,\\ (\mathcal{P}_2\otimes\mathrm{id}_E)(|\psi\rangle\langle\psi|) &= p_\rho\sigma^{\alpha}_{AB}\otimes\sigma_E\otimes|\mathrm{ok}\rangle\langle\mathrm{ok}| + (1-p_\rho)\sigma^{\perp}_{AB}\otimes\sigma_E\otimes|\mathrm{fail}\rangle\langle\mathrm{fail}|.\end{aligned} \tag{B.29}$$

If we post-select equation (B.29) on the ok-branch we have after normalization

$$\frac{1}{p_\rho}(\mathrm{id}_{ABE}\otimes|\mathrm{ok}\rangle\langle\mathrm{ok}|)(\mathcal{P}_1\otimes\mathrm{id}_E)(|\psi\rangle\langle\psi|) = \sigma_{ABE}\otimes|\mathrm{ok}\rangle\langle\mathrm{ok}|. \tag{B.30}$$

It is obvious from the fact that the protocol is performed by Alice and Bob per definition that any measurement of Eve in the ok-branch can be commuted to the beginning of the protocol $\mathcal{P}_1$ because Eve is not part of the protocol. Hence her measurement only changes the input of the protocol $\mathcal{P}_1$ and thus either cause an abort or not.

We call the final state of Alice and Bob η-Eve-non steerable if for all $\phi \in \mathcal{H}_E$ we have $\|\sigma^{\phi}_{AB} - \sigma_{AB}\|_1 \leqslant \eta$ where $\sigma^{\phi}_{AB} = \mathrm{tr}_E\left[\frac{1}{p_E(\phi)}(\mathrm{id}_{AB} \otimes |\phi\rangle\langle\phi|_E)\sigma_{ABE}\right]$. We sketch the remainder of this proof as follows: we show, that the final state of Alice and Bob is Eve-non steerable in the sense of lemma 8 by making use of the bounded distance of the protocols $\mathcal{P}_1$ and $\mathcal{P}_2$. Furthermore, lemma 9 implies that it suffice to consider measurements of Eve of having probability greater than or equal to $1/16$. Therefore lemmas 8 and 9 completes the proof.

Because the output of Alice and Bob are 2 qubits the purifying system that Eve holds is without loss of generality also a two-qubit system. Hence, according to lemma 9, there exists a state σ'_{ABE}, which is unitarily related to σ_{ABE} via an unitary U on Eve's system only (which is not part of the protocol) and for which it suffice to consider measurements of Eve having probability greater than or equal to $1/16$. Furthermore observe that this local unitary of Eve can not change the success probability of the overall protocol as unitaries are CPTP. In other words, the success probabilities associated with σ_{ABE} and σ'_{ABE} are identical.

More formally, suppose Eve performs a projective measurement on this state σ'_{ABE} (which stems from a purification $|\psi'\rangle_{ABE}$ of ρ'_{AB} which is unitarily related to the purification $|\psi\rangle_{ABE}$ of ρ_{AB} and both having the same success probability, see paragraph above) and observes outcome $|\phi\rangle \in \mathcal{H}_E$ having probability greater than or equal to $1/16$. Then the post-selected state of Alice, Bob, and Eve conditioned on that particular outcome ϕ reads as

$$\begin{aligned}
&\frac{1}{p_E(\phi)}(\mathrm{id}_{AB} \otimes |\phi\rangle\langle\phi|_E)(\sigma'_{ABE} \otimes |\mathrm{ok}\rangle\langle\mathrm{ok}|)\\
&\quad= \frac{1}{p_E(\phi)}(\mathrm{id}_{AB} \otimes |\phi\rangle\langle\phi|_E)\frac{1}{p_\rho}(\mathrm{id}_{ABE} \otimes |\mathrm{ok}\rangle\langle\mathrm{ok}|)(\mathcal{P}_1 \otimes \mathrm{id}_E)(|\psi'\rangle\langle\psi'|_{ABE})\\
&\quad= \frac{1}{p_{\rho^\phi}}(\mathrm{id}_{ABE} \otimes |\mathrm{ok}\rangle\langle\mathrm{ok}|)(\mathcal{P}_1 \otimes \mathrm{id}_E)\underbrace{\left(\frac{\mathrm{id}_{AB} \otimes |\phi\rangle\langle\phi|_E}{p'_E(\phi)}|\psi'\rangle\langle\psi'|_{ABE}\right)}_{=:\rho^\phi_{ABE}}\\
&\quad= \frac{1}{p_{\rho^\phi}}(\mathrm{id}_{ABE} \otimes |\mathrm{ok}\rangle\langle\mathrm{ok}|)(\mathcal{P}_1 \otimes \mathrm{id}_E)(\rho^\phi_{ABE})\\
&\quad= \sigma^\phi_{ABE} \otimes |\mathrm{ok}\rangle\langle\mathrm{ok}|.
\end{aligned}$$

More importantly, we relate the probability of the protocol succeeding for initial state ρ', p_ρ, and the probability of measuring ϕ after the protocol, $p_E(\phi)$, to the probability of the protocol succeeding for the initial state ρ^ϕ_{ABE} (measurement of Eve commuted to the beginning of the protocol), p_{ρ^ϕ}, and the probability of measuring ϕ before the protocol has started, $p'_E(\phi)$, via

$$p_\rho p_E(\phi) = p_{\rho^\phi} p'_E(\phi). \tag{B.31}$$

Observe that (B.31) is equivalent to

$$\frac{p_\rho p_E(\phi)}{p'_E(\phi)} = p_{\rho^\phi}. \tag{B.32}$$

We note that the state σ^ϕ_{ABE} is in the ok-branch of the protocol $\mathcal{P}_1$. The next step is to apply lemma 8 which relates the distances $\|\sigma'_{ABE} - \sigma_{AB} \otimes \sigma'_E\|_1$ and $\|\sigma'_{AB} - \sigma^\phi_{AB}\|_1$. In particular we show that for all measurements of Eve with outcome $|\phi\rangle \in \mathcal{H}_E$ having a probability greater than or equal to $1/16$ we have that $\|\sigma'_{AB} - \sigma^\phi_{AB}\|_1 \leqslant 17\varepsilon/p_\rho$. This then implies using lemma 9 that $\|\sigma_{ABE} - \sigma_{AB} \otimes \sigma_E\|_1 \leqslant 34C\varepsilon/p_\rho$. In detail, using the triangle inequality we compute for the distance between σ'_{AB} and σ^ϕ_{AB}

$$\begin{aligned}
\|\sigma'_{AB} - \sigma^\phi_{AB}\|_1 &\leqslant \|\sigma'_{AB} - \sigma^\alpha_{AB}\|_1 + \|\sigma^\alpha_{AB} - \sigma^\phi_{AB}\|_1 = \frac{1}{p_\rho}\|(\mathcal{P}_1 - \mathcal{P}_2)(\rho'_{AB})\|_1 + \frac{1}{p_{\rho^\phi}}\|(\mathcal{P}_1 - \mathcal{P}_2)(\rho^\phi_{AB})\|_1\\
&\leqslant \left(\frac{1}{p_\rho} + \frac{1}{p_{\rho^\phi}}\right)\max_{\mu_{AB}}\|(\mathcal{P}_1 - \mathcal{P}_2)(\mu_{AB})\|_1.
\end{aligned} \tag{B.33}$$

Now we employ (B.32) in (B.33) which yields

$$\|\sigma'_{AB} - \sigma^{\phi}_{AB}\|_1 \leqslant \left(\frac{1}{p_\rho} + \frac{p'_E(\phi)}{p_\rho p_E(\phi)}\right)\max_{\mu_{AB}}\|(\mathcal{P}_1 - \mathcal{P}_2)(\mu_{AB})\|_1 \leqslant \left(\frac{1}{p_\rho} + \frac{1}{p_\rho p_E(\phi)}\right)\max_{\mu_{AB}}\|(\mathcal{P}_1 - \mathcal{P}_2)(\mu_{AB})\|_1$$
$$= \frac{1}{p_\rho}\left(1 + \frac{1}{p_E(\phi)}\right)\max_{\mu_{AB}}\|(\mathcal{P}_1 - \mathcal{P}_2)(\mu_{AB})\|_1 \leqslant \frac{1}{p_\rho}(1+16)\max_{\mu_{AB}}\|(\mathcal{P}_1 - \mathcal{P}_2)(\mu_{AB})\|_1 \leqslant \frac{17}{p_\rho}\varepsilon \tag{B.34}$$

because $p_E(\phi) \geqslant 1/16$ and $\max_{\mu_{AB}}\|(\mathcal{P}_1 - \mathcal{P}_2)(\mu_{AB})\|_1$ is bounded by ε by assumption. Hence we apply lemma 8 to σ'_{ABE} with $\epsilon = \frac{17}{p_\rho}\varepsilon$ which implies for the distance between σ'_{ABE} and $\sigma_{AB} \otimes \sigma'_E$ that

$$\|\sigma'_{ABE} - \sigma_{AB} \otimes \sigma'_E\|_1 \leqslant \frac{34 \cdot 4^8}{p_\rho}\varepsilon, \tag{B.35}$$

where the factor 4^8 is the constant C of lemma 8 depending on the dimensions of the systems of Alice/Bob and Eve, for which we have $n = m = 2$. Furthermore, this implies via lemma 9 that

$$\|\sigma_{ABE} - \sigma_{AB} \otimes \sigma_E\|_1 \leqslant \frac{34 \cdot 4^8}{p_\rho}\varepsilon \tag{B.36}$$

because σ_{ABE} and $\sigma_{AB} \otimes \sigma_E$ are unitarly related to σ'_{ABE} and $\sigma_{AB} \otimes \sigma'_E$ via the unitary U on Eve's system. Finally, employing (B.36) in (B.28) yields

$$\begin{aligned}\|(\mathcal{P}_1 \otimes \mathrm{id}_E)(|\psi\rangle\langle\psi|) - (\mathcal{P}_2 \otimes \mathrm{id}_E)(|\psi\rangle\langle\psi|)\|_1 &= p_\rho\|\sigma_{ABE} - \sigma^{\alpha}_{AB} \otimes \sigma_E\|_1 \\ &\leqslant p_\rho(\|\sigma_{ABE} - \sigma_{AB} \otimes \sigma_E\|_1 + \|\sigma_{AB} \otimes \sigma_E - \sigma^{\alpha}_{AB} \otimes \sigma_E\|_1) \\ &\leqslant 34 \cdot 4^8\varepsilon + \varepsilon = (34 \cdot 4^8 + 1)\varepsilon.\end{aligned}$$

□

B.2. Proof of lemma 5

Now we turn to the proof of lemma 5 of the main text. For that purpose we remind the reader that the final state after the distillation protocol including the system of L is pure. Thus, the following lemma will turn out to be very useful.

Lemma 10. *Let ρ_{AB} and $\varphi_{AB} = |\varphi\rangle\langle\varphi|_A \otimes \mu_B$ be two mixed states. Furthermore, assume that $\rho_A = \mathrm{tr}_B[\rho_{AB}]$ satisfies $\|\rho_A - |\varphi\rangle\langle\varphi|_A\|_1 \leqslant \varepsilon$ and $\rho_B = \mathrm{tr}_A[\rho_{AB}] = \mu_B$. Then $\|\rho_{AB} - \varphi_{AB}\|_1 \leqslant 4\sqrt{\varepsilon}$.*

Proof. By assumption we have $\|\rho_A - |\varphi\rangle\langle\varphi|_A\|_1 \leqslant \varepsilon$. Moreover, let $|\psi\rangle_{ABR}$ be a purification of ρ_{AB}. According to lemma A.2.7 in [9] there exists a purification $|\varphi\rangle_A \otimes |\xi\rangle_{BR}$ of φ_{AB} such that $\||\psi\rangle_{ABR} - |\varphi\rangle_A \otimes |\xi\rangle_{BR}\|_{\mathrm{vec}} \leqslant \sqrt{\|\rho_A - |\varphi\rangle\langle\varphi|_A\|_1} = \sqrt{\varepsilon}$ where $\||\psi\rangle\|_{\mathrm{vec}} = \sqrt{\langle\psi|\psi\rangle}$ and ${}_{ABR}\langle\psi|\varphi\rangle_A|\xi\rangle_{BR}$ is real and non-negative. Moreover, lemma A.2.3 of [9] gives

$$\||\psi\rangle\langle\psi|_{ABR} - |\varphi\rangle\langle\varphi|_A \otimes |\xi\rangle\langle\xi|_{BR}\|_1 \leqslant 2\||\psi\rangle_{ABR} - |\varphi\rangle_A \otimes |\xi\rangle_{BR}\|_{\mathrm{vec}} \leqslant 2\sqrt{\epsilon}.$$

We define $\xi_B = \mathrm{tr}_R[|\xi\rangle\langle\xi|_{BR}]$. As the 1-norm does not increase under the partial trace we have

$$\|\rho_B - \xi_B\|_1 \leqslant \|\rho_{AB} - |\varphi\rangle\langle\varphi|_A \otimes \xi_B\|_1 \leqslant \||\psi\rangle\langle\psi|_{ABR} - |\varphi\rangle\langle\varphi|_A \otimes |\xi\rangle\langle\xi|_{BR}\|_1 \leqslant 2\sqrt{\epsilon}$$

by construction. Moreover, the assumption $\rho_B = \mu_B$ implies $\|\mu_B - \xi_B\|_1 = \|\rho_B - \xi_B\|_1 \leqslant 2\sqrt{\varepsilon}$. This gives us $\||\varphi\rangle\langle\varphi|_A \otimes \mu_B - |\varphi\rangle\langle\varphi|_A \otimes \xi_B\|_1 = \|\mu_B - \xi_B\|_1 \leqslant 2\sqrt{\varepsilon}$. If we combine these results we obtain

$$\begin{aligned}\|\rho_{AB} - \varphi_{AB}\|_1 = \|\rho_{AB} - |\varphi\rangle\langle\varphi|_A \otimes \mu_B\|_1 &\leqslant \|\rho_{AB} \\ -|\varphi\rangle\langle\varphi|_A \otimes \xi_B\|_1 + \||\varphi\rangle\langle\varphi|_A \otimes \xi_B - |\varphi\rangle\langle\varphi|_A \otimes \mu_B\|_1 &\leqslant 4\sqrt{\varepsilon},\end{aligned}$$

which proves the claim. □

Lemma 10 enables us to prove lemma 5 of the main text.

Lemma (Lemma 5 of the main text). *Let $\mathcal{E}$ be the real protocol which guarantees to converge towards a unique and attracting fixed point depending on the noise parameter only. Let $\mathcal{F}$ be the ideal protocol as defined in the main text. Furthermore let ρ be a mixed state (consisting of n systems) shared by Alice and Bob. If the extension of $\mathcal{E}$ and $\mathcal{F}$ to the system of L satisfies $\|\mathcal{E}_L(\rho) - \mathcal{F}_L(\rho)\|_1 \leqslant \varepsilon(n)$, then*

$$\|(\mathcal{E} \otimes \mathrm{id}_{E'})(|\psi\rangle\langle\psi|_{ABE'}) - (\mathcal{F} \otimes \mathrm{id}_{E'})(|\psi\rangle\langle\psi|_{ABE'})\|_1 \leqslant 4\sqrt{\varepsilon(n)}$$

for all purifications $|\psi\rangle_{ABE'}$ *of* ρ.

Proof. As mentioned in the main text, we introduce a two-level flag system held by Alice which indicates whether they aborted the protocol or not. So we observe

$$\mathcal{E}_L(\rho) = p_\rho \sigma_{ABEL} \otimes |\text{ok}\rangle\langle\text{ok}| + (1-p_\rho)\sigma^{\perp}_{ABEL} \otimes |\text{fail}\rangle\langle\text{fail}|,$$
$$\mathcal{F}_L(\rho) = p_\rho |\psi^f\rangle\langle\psi^f|_{ABEL} \otimes |\text{ok}\rangle\langle\text{ok}| + (1-p_\rho)\sigma^{\perp}_{ABEL} \otimes |\text{fail}\rangle\langle\text{fail}|,$$

where E denotes the system of leaked noise transcripts to Eve. By assumption we have $\|\mathcal{E}_L(\rho) - \mathcal{F}_L(\rho)\|_1 \leqslant \varepsilon(n)$. This is equivalent to $p_\rho\|\sigma_{ABEL} - |\psi_f\rangle\langle\psi_f|_{ABEL}\|_1 \leqslant \varepsilon(n)$ since $\mathcal{E}_L(\rho)$ and $\mathcal{F}_L(\rho)$ are equal on the fail branch. This we can rewrite to $\|\sigma_{ABEL} - |\psi_f\rangle\langle\psi_f|_{ABEL}\|_1 \leqslant \varepsilon(n)/p_\rho$.

Moreover, applying the real and ideal protocol to the purification $|\psi\rangle_{ABE'}$ results in

$$(\mathcal{E}\otimes \text{id}_{E'})(|\psi\rangle\langle\psi|_{ABE'}) = p_\rho \sigma_{ABEE'} \otimes |\text{ok}\rangle\langle\text{ok}| + (1-p_\rho)\sigma^{\perp}_{ABEE'} \otimes |\text{fail}\rangle\langle\text{fail}|,$$
$$(\mathcal{F}\otimes \text{id}_{E'})(|\psi\rangle\langle\psi|_{ABE'}) = p_\rho \sigma^{f}_{ABE} \otimes \rho_{E'} \otimes |\text{ok}\rangle\langle\text{ok}| + (1-p_\rho)\sigma^{\perp}_{ABEE'} \otimes |\text{fail}\rangle\langle\text{fail}|.$$

Again, both expression are equal in the fail branch, thus the 1-norm simplifies to

$$\|(\mathcal{E}\otimes \text{id}_{E'})(|\psi\rangle\langle\psi|_{ABE'}) - (\mathcal{F}\otimes \text{id}_{E'})(|\psi\rangle\langle\psi|_{ABE'})\|_1 = p_\rho\|\sigma_{ABEE'} - \sigma^f_{ABE}\otimes\rho_{E'}\|_1. \tag{B.37}$$

Hence it is sufficient to show $p_\rho\|\sigma_{ABEE'} - \sigma^f_{ABE}\otimes\rho_{E'}\|_1 \leqslant 4\sqrt{\varepsilon(n)}$. We observe that by introducing the system L held by L that

$$p_\rho\|\sigma_{ABEE'} - \sigma^f_{ABE}\otimes\rho_{E'}\|_1 \leqslant p_\rho\|\sigma_{ABELE'} - |\psi^\alpha\rangle\langle\psi^\alpha|_{ABEL}\otimes\rho_{E'}\|_1. \tag{B.38}$$

One easily verifies $\text{tr}_{E'}[\sigma_{ABELE'}] = \sigma_{ABEL}$ and $\text{tr}_{ABEL}[\sigma_{ABELE'}] = \rho_{E'}$ because the system E' is not changed by the protocol $\mathcal{E}$. Moreover, by assumption we have $\|\sigma_{ABEL} - |\psi_f\rangle\langle\psi_f|_{ABEL}\|_1 \leqslant \varepsilon(n)/p_\rho$. Thus we apply lemma 10 to $\rho_{A'B'} := \sigma_{ABELE'}$ and $\varphi_{A'B'} = |\psi_f\rangle\langle\psi_f|_{ABEL}\otimes\rho_{E'}$ where $A' := ABEL$ and $B' := E'$ which implies

$$\|\sigma_{ABELE'} - |\psi_f\rangle\langle\psi_f|_{ABEL}\otimes\rho_{E'}\|_1 \leqslant 4\sqrt{\varepsilon(n)/p_\rho}. \tag{B.39}$$

Employing (B.38) and (B.39) in (B.37) yields

$$\|(\mathcal{E}\otimes \text{id}_{E'})(|\psi\rangle\langle\psi|_{ABE'}) - (\mathcal{F}\otimes \text{id}_{E'})(|\psi\rangle\langle\psi|_{ABE'})\|_1 \leqslant p_\rho 4\sqrt{\varepsilon(n)/p_\rho} = 4\sqrt{p_\rho\varepsilon(n)} \leqslant 4\sqrt{\varepsilon(n)}$$

which completes the proof. □

Appendix C. Confidentiality of entanglement distillation protocols whenever the noise transcripts leak

In this section we show how the confidentiality guarantees regarding an entanglement distillation protocol can be extended to the case whenever the noise transcripts leak to Eve.

We remind the reader that it is not necessary to leak the noise transcripts to Eve after every single distillation round. It is sufficient to copy all noise transcripts at the very end to Eve's register, as L is not accessible and Eve is not part of the protocol being executed by Alice and Bob.

Theorem (Theorem 7 in main text). *Let* $\mathcal{E}$ *be the real protocol and* $\mathcal{F}$ *be the ideal protocol. Furthermore, let* $\mathcal{E}^l$ *be the real and* $\mathcal{F}^l$ *be the ideal protocol when the noise transcripts leak to Eve. Then*

$$\|(\mathcal{E}\otimes \text{id}_E)(|\psi\rangle\langle\psi|_{ABE}) - (\mathcal{F}\otimes \text{id}_E)(|\psi\rangle\langle\psi|_{ABE})\|_1 \leqslant \varepsilon(n)$$

implies that

$$\|(\mathcal{E}^l\otimes \text{id}_E)(|\psi\rangle\langle\psi|_{ABE}) - (\mathcal{F}^l\otimes \text{id}_E)(|\psi\rangle\langle\psi|_{ABE})\|_1 \leqslant 2\sqrt{\varepsilon(n)} \tag{C.1}$$

for all purifications $|\psi\rangle_{ABE}$ *of initial state* ρ_{AB} *consisting of n systems.*

Proof. We observe that

$$(\mathcal{E}\otimes \text{id}_E)(|\psi\rangle\langle\psi|) = p_\rho\sigma_{ABE}\otimes|\text{ok}\rangle\langle\text{ok}| + (1-p_\rho)\sigma^{\perp}_{AB}\otimes\sigma_E\otimes|\text{fail}\rangle\langle\text{fail}|,$$
$$(\mathcal{F}\otimes \text{id}_E)(|\psi\rangle\langle\psi|) = p_\rho\sigma^{\alpha}_{AB}\otimes\sigma_E\otimes|\text{ok}\rangle\langle\text{ok}| + (1-p_\rho)\sigma^{\perp}_{AB}\otimes\sigma_E\otimes|\text{fail}\rangle\langle\text{fail}|.$$

So by assumption we have

$$\|(\mathcal{E}\otimes \text{id}_E)(|\psi\rangle\langle\psi|) - (\mathcal{F}\otimes \text{id}_E)(|\psi\rangle\langle\psi|)\| = p_\rho\|\sigma_{ABE} - \sigma^\alpha_{AB}\otimes\sigma_E\| \leqslant \varepsilon(n),$$

i.e. $\|\sigma_{ABE} - \sigma^\alpha_{AB}\otimes\sigma_E\| \leqslant \varepsilon(n)/p_\rho$.

As outlined in the main text we model L in terms of purifications. Because purifications are unitarly equivalent we choose a particular purification of $\sigma^{\alpha}_{AB}\otimes\sigma_E$. Thus we fix $|\psi_{\mathcal{F}}\rangle_{ABL_1L_2E}=|\psi'\rangle_{ABL_1}\otimes|\psi''\rangle_{L_2E}$ where $|\psi'\rangle_{ABL_1}=\sum_{i,j}\omega_{ij}(\alpha)|B_{ij}\rangle_{AB}|ij\rangle_{L_1}$. The purifying systems L_1 and L_2 we attribute to the Lab Demon. Moreover, according to lemma A.2.7 in [9] there exists a purification $|\psi_{\mathcal{E}}\rangle$ of σ_{ABE} such that $\||\psi_{\mathcal{F}}\rangle_{ABL_1L_2E}-|\psi_{\mathcal{E}}\rangle_{ABL_1L_2E}\|_{\mathrm{vec}}\leqslant\sqrt{\varepsilon(n)/p_\rho}$ where $\||\psi\rangle\|_{\mathrm{vec}}=\sqrt{\langle\psi|\psi\rangle}$ and ${}_{ABL_1L_2E}\langle\psi_{\mathcal{F}}|\psi_{\mathcal{E}}\rangle_{ABL_1L_2E}$ is real and non-negative. Furthermore, lemma A.2.3 of [9] gives

$$\||\psi_{\mathcal{E}}\rangle\langle\psi_{\mathcal{E}}|_{ABL_1L_2E}-|\psi_{\mathcal{F}}\rangle\langle\psi_{\mathcal{F}}|_{ABL_1L_2E}\|_1\leqslant 2\||\psi_{\mathcal{E}}\rangle_{ABL_1L_2E}-|\psi_{\mathcal{F}}\rangle_{ABL_1L_2E}\|_{\mathrm{vec}}\leqslant 2\sqrt{\varepsilon(n)/p_\rho}\,. \tag{C.2}$$

When the noise transcripts leak to Eve, L effectively copies the noise transcripts $|ij\rangle_{L_1}$ to Eve, resulting in the pure state $|\phi\rangle_{ABL_1L_2EE'}=(\sum_{i,j}|B_{ij}\rangle_{AB}|ij\rangle_{L_1}|ij\rangle_{E'})\otimes|\psi\rangle_{L_2E}$. Hence we can model the leakage of the noise transcripts to Eve by a unitary U_M such that $U_M\,|\psi_{\mathcal{F}}\rangle_{ABL_1L_2E}|0\rangle_{E'}=|\phi\rangle_{ABL_1L_2EE'}$. For the protocol when the noise transcripts leak to Eve we have

$$\begin{aligned}
(\mathcal{E}^l\otimes\mathrm{id}_E)(|\psi\rangle\langle\psi|)&=p_\rho\sigma'_{ABE}\otimes|\mathrm{ok}\rangle\langle\mathrm{ok}|+(1-p_\rho)\sigma^{\perp}_{AB}\otimes\sigma_E\otimes|\mathrm{fail}\rangle\langle\mathrm{fail}|\\
&=p_\rho\,\mathrm{tr}_{L_1L_2}[U_M|\psi_{\mathcal{E}}\rangle\langle\psi_{\mathcal{E}}|\,U_M^\dagger]\otimes|\mathrm{ok}\rangle\langle\mathrm{ok}|+(1-p_\rho)\sigma^{\perp}_{AB}\otimes\sigma_E\otimes|\mathrm{fail}\rangle\langle\mathrm{fail}|\\
(\mathcal{F}^l\otimes\mathrm{id}_E)(|\psi\rangle\langle\psi|)&=p_\rho\sigma'^{\alpha}_{ABE}\otimes|\mathrm{ok}\rangle\langle\mathrm{ok}|+(1-p_\rho)\sigma^{\perp}_{AB}\otimes\sigma_E\otimes|\mathrm{fail}\rangle\langle\mathrm{fail}|\\
&=p_\rho\,\mathrm{tr}_{L_1L_2}[U_M|\psi_{\mathcal{F}}\rangle\langle\psi_{\mathcal{F}}|\,U_M^\dagger]\otimes|\mathrm{ok}\rangle\langle\mathrm{ok}|+(1-p_\rho)\sigma^{\perp}_{AB}\otimes\sigma_E\otimes|\mathrm{fail}\rangle\langle\mathrm{fail}|.
\end{aligned}$$

Because the real and the ideal protocol are equal in the fail-branch we obtain by using (C.2)

$$\begin{aligned}
&\|(\mathcal{E}^l\otimes\mathrm{id}_E)(|\psi\rangle\langle\psi|_{ABE})-(\mathcal{F}^l\otimes\mathrm{id}_E)(|\psi\rangle\langle\psi|_{ABE})\|_1\\
&=p_\rho\|\sigma'_{ABE}\otimes|\mathrm{ok}\rangle\langle\mathrm{ok}|-\sigma'^{\alpha}_{ABE}\otimes|\mathrm{ok}\rangle\langle\mathrm{ok}|\|_1\\
&=p_\rho\|\mathrm{tr}_{L_1L_2}[U_M|\psi_{\mathcal{E}}\rangle\langle\psi_{\mathcal{E}}|\,U_M^\dagger]-\mathrm{tr}_{L_1L_2}[U_M|\psi_{\mathcal{F}}\rangle\langle\psi_{\mathcal{F}}|\,U_M^\dagger]\|_1\\
&\leqslant p_\rho\|U_M|\psi_{\mathcal{E}}\rangle\langle\psi_{\mathcal{E}}|\,U_M^\dagger-U_M|\psi_{\mathcal{F}}\rangle\langle\psi_{\mathcal{F}}|\,U_M^\dagger\|_1\\
&=p_\rho\||\psi_{\mathcal{E}}\rangle\langle\psi_{\mathcal{E}}|-|\psi_{\mathcal{F}}\rangle\langle\psi_{\mathcal{F}}|\|_1\leqslant 2\sqrt{\varepsilon(n)p_\rho}\leqslant 2\sqrt{\varepsilon(n)}\,,
\end{aligned}$$

which proves (C.1). □

Thus the confidentiality of a protocol where the noise transcripts leak to Eve is bounded by the confidentiality of the same protocol when they do not.

Appendix D. Quantum one-time padding after the real protocol

In this section we show that a final secret twirl applied to the pair of Alice and Bob decouples Eve completely from the remaining state. Keep in mind that for this Alice and Bob require two classical bits unknown to Eve.

Recall that the state of Alice, Bob, Eve, and L after n distillation rounds is pure and of the form $|\psi\rangle=\sum_{i,j,k,l}P_{ijkl}\,|B_{ij}\rangle_{AB}|\eta_{kl}\rangle_L|\eta_{ijkl}\rangle_E$. Tracing over L yields the mixed state

$$\rho_{ABE}=\sum_{i_1,i_2,j_1,j_2}\sum_{k,l}P_{i_1j_1kl}P^*_{i_2j_2kl}|B_{i_1j_1}\rangle\langle B_{i_2j_2}|\otimes|\eta_{i_1j_1kl}\rangle\langle\eta_{i_2j_2kl}|. \tag{D.1}$$

Suppose Alice and Bob apply a secret twirl $\mathcal{T}$ to (D.1), i.e. they apply stochastically the family of operators $\{\mathrm{id},K_1,K_2,K_1K_2\}$ where $K_1=\sigma_x\otimes\sigma_x$ and $K_2=\sigma_z\otimes\sigma_z$. These are two stabilizers of the Bell state, i.e.,

$$\begin{aligned}
K_1^{r_1}|B_{i_1j_1}\rangle&=(-1)^{i_1r_1}|B_{i_1j_1}\rangle,\\
K_2^{r_2}|B_{i_1j_1}\rangle&=(-1)^{j_1r_2}|B_{i_1j_1}\rangle.
\end{aligned}$$

Hence, applying the secret twirl $\mathcal{T}$ to (D.1) gives

$$\begin{aligned}
\mathcal{T}\rho_{ABE}&=\sum_{\substack{r_1,r_2\\ i_1,i_2,j_1,j_2,k,l}}\frac{1}{4}P_{i_1j_1kl}P^*_{i_2j_2kl}K_1^{r_1}K_2^{r_2}|B_{i_1j_1}\rangle\langle B_{i_2j_2}|K_1^{r_1}K_2^{r_2}\otimes|\eta_{i_1j_1kl}\rangle\langle\eta_{i_2j_2kl}|\\
&=\sum_{\substack{r_1,r_2\\ i_1,i_2,j_1,j_2,k,l}}(-1)^{i_1r_1}(-1)^{j_1r_2}(-1)^{i_2r_1}(-1)^{j_2r_2}\frac{1}{4}P_{i_1j_1kl}P^*_{i_2j_2kl}|B_{i_1j_1}\rangle\langle B_{i_2j_2}|\otimes|\eta_{i_1j_1kl}\rangle\langle\eta_{i_2j_2kl}|\\
&=\sum_{i_1,i_2,j_1,j_2}|B_{i_1j_1}\rangle\langle B_{i_2j_2}|\otimes\frac{1}{4}\sum_{k,l}P_{i_1j_1kl}P^*_{i_2j_2kl}|\eta_{i_1j_1kl}\rangle\langle\eta_{i_2j_2kl}|\sum_{r_1,r_2}(-1)^{(i_1+i_2)r_1}(-1)^{(j_1+j_2)r_2}\\
&=\sum_{i_1,j_1}|B_{i_1j_1}\rangle\langle B_{i_1j_1}|\otimes\sum_{k,l}|P_{i_1j_1kl}|^2|\eta_{i_1j_1kl}\rangle\langle\eta_{i_1j_1kl}|.
\end{aligned}$$

Note that in the resulting state $\sum_{i_1,j_1}|B_{i_1j_1}\rangle\langle B_{i_1j_1}| \otimes \sum_{k,l}|P_{i_1j_1kl}|^2|\eta_{i_1j_1kl}\rangle\langle\eta_{i_1j_1kl}|$ Eve decouples, i.e. Alice/Bob and Eve have a separable state. The obtained resource state can be used to establish a confidential quantum channel by means of quantum teleportation.

Appendix E. Robustness of recurrence-type entanglement distillation protocol

To complete the security characterization of entanglement distillation protocols we also consider the robustness of an entanglement distillation protocol. To define this term precisely we first need the definition of a honest eavesdropper.

Definition 11. We call an eavesdropper honest, if the states sent by the eavesdropper are of the form $|B_{00}\rangle^{\otimes 2^n}$.

It is obvious that a honest eavesdropper is not entangled with the ensemble delivered to Alice and Bob via the noisy quantum channel. Moreover we formally define the robustness of a protocol by:

Definition 12 (Robustness of a protocol). We call a protocol $\mathcal{E}^{\alpha}$ ε_R-robust, if for a honest eavesdropper the probability of aborting the protocol is at most ε_R.

Now we show that we can tune the robustness of a recurrence-type entanglement distillation protocol to be exponentially small in terms of necessary number of input pairs.

Theorem 13. *Let $M \in \mathbb{N}$ such that Alice and Bob achieve ε-confidentiality by succeeding M rounds of a recurrence-type entanglement distillation protocol. Furthermore assume that Alice and Bob receive n pairs from a honest eavesdropper over the quantum channel $\Phi^{\otimes n}$ (where $\Phi(\rho) = \beta\rho + (1-\beta)/4(\sum_{i,j}\sigma_{i,j}\rho\sigma_{i,j})$) such that, after the parameter estimation step of the proposed protocol, $k - \sqrt{k}$ pairs (where $k - \sqrt{k} = c2^M$ and $c = \xi 2^{M+2}$) are left for entanglement distillation. Then, the robustness ε_R of the protocol is bounded by*

$$\varepsilon_R \leqslant \exp\left(-(3\beta - 4F_{\min}(\alpha) - 1)^2\sqrt{k}/128\right) + M\exp(-\xi).$$

Proof. The basic idea of the proof is to request sufficiently many pairs from Eve such that the probabilities of abort during the protocol to be exponentially small while still having enough pairs left to achieve M rounds of a recurrence-type entanglement distillation protocol. We divide the proof into two parts:

- Part 1: We prove that the probability of aborting the recurrence-type entanglement distillation protocol due to parameter estimation is exponentially small.
- Part 2: We prove the same holds true for aborting the protocol during entanglement distillation.

Part 1: Suppose Eve sends the state $|B_{00}\rangle^{\otimes n}$ through the noisy quantum channel $\Phi^{\otimes n}$ to Alice and Bob. Applying Φ to $|B_{00}\rangle\langle B_{00}|$ yields

$$\rho_{AB} = \Phi(|B_{00}\rangle\langle B_{00}|) = (3\beta+1)/4|B_{00}\rangle\langle B_{00}| + (1-\beta)/4(|B_{10}\rangle\langle B_{10}| + |B_{01}\rangle\langle B_{01}| + |B_{11}\rangle\langle B_{11}|). \quad \text{(E.1)}$$

Thus the state Alice and Bob receive is $\rho_{AB}^{\otimes n}$. According to the preceding protocols proposed in the main text, Alice and Bob apply a symmetrization to $\rho_{AB}^{\otimes n}$, and, depending on the noise level of the apparatus, they might have to trace out $n - k$ pairs or not. For the subsequent analysis we assume that this tracing out step is necessary, i.e. the de-Finetti-based reduction needs to be applied. Hence, Alice and Bob continue by applying a twirl to each remaining pair. Since $\rho_{AB}^{\otimes k}$ is invariant under permutations and ρ_{AB} is Bell-diagonal, the remaing state after twirling is equal to $\rho_{AB}^{\otimes k}$.

Next, they apply to $\sqrt{k}$ of the remaining k pairs the parameter estimation for estimating the fidelity of each pair. Necessary for convergence of all recurrence-type entanglement distillation protocols is that the fidelity F of ρ_{AB} with $|B_{00}\rangle$ satisfies $F > F_{\min}(\alpha)$. Hence this step is crucial in order to guarantee successful distillation.

For that purpose, we measure $\lfloor\sqrt{k}\rfloor$ of k pairs by applying two-qubit measurements. To be more precise, we apply a $\sigma_x \otimes \sigma_x$ to the first and $\sigma_z \otimes \sigma_z$ measurement to the second pair. We refer to this measurements by M_1 and M_2 respectively. We observe that the state $|B_{00}\rangle$ is a common eigenstate of M_1 and M_2 with eigenvalue 1. We define to each pair of pairs a random variable X_i for $i \in \{1, .., \lfloor\sqrt{k}\rfloor/2\}$ with $X_i = 1$ whenever both measurements M_1 and M_2 yield outcome 1 and $X_i = 0$ else.

Furthermore we assume for the expected value $\mathbb{E}(X)$ of the fidelity with $|B_{00}\rangle$ that $\mathbb{E}(X) = F_{\min}(\alpha) + \delta$, where $\delta > 0$ will be fixed below. The protocol will be aborted if the estimate is below $F_{\min}(\alpha) + \delta$.

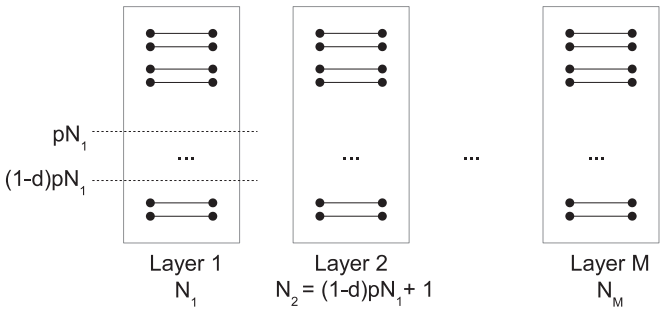

Figure E1. *M* rounds of entanglement distillation.

From (E.1) we observe that, whenever $(3\beta + 1)/4 \leqslant F_{\min}(\alpha)$, the entanglement distillation protocol will not distill any entanglement. This implies for the quantum channel Φ that, if $\beta \leqslant (4F_{\min}(\alpha) - 1)/3$ the parameter estimation step will abort, independent of the input provided by Eve. Thus we assume for the subsequent analysis that $\beta > (4F_{\min}(\alpha) - 1)/3$.

Moreover we define $\eta = \delta/2$. Hence we get by the Hoeffdings inequality [34] for the probability of an error larger than η in our measured estimate $\overline{X}$ for the fidelity the following expression:

$$\mathbb{P}(|\mathbb{E}(X) - \overline{X}| \geqslant \eta) \leqslant \exp(-\eta^2\sqrt{k}/2) =: p_{\text{pe-abort}}.$$

Thus the probability of aborting the protocol due to an error in the parameter estimation is exponentially small in number of necessary input pairs. In order to fix δ we recognize that Alice and Bob abort the protocol whenever $(3\beta + 1)/4 < F_{\min}(\alpha) + \delta$. This is equivalent to $\delta > (3\beta - 4F_{\min}(\alpha) - 1)/4$. Inserting the definition of η yields $\eta > (3\beta - 4F_{\min}(\alpha) - 1)/8$ and thus $p_{\text{pe-abort}} < \exp(-(3\beta - 4F_{\min}(\alpha) - 1)^2\sqrt{k}/128)$.

Part 2: What remains to be shown is that the probability of aborting the protocol in the distillation phase is also exponentially small in the number of input pairs. For that purpose, we assume that the noise level α of the apparatus is such that distillation is feasible. In the following we show that we can force the probability of abort due to entanglement distillation to be exponentially small in terms of requested input pairs.

We assume that Alice and Bob are left with $c2^M$ pairs after parameter estimation. Recall that the Chernoff inequality for a sequence of independent Bernoulli random variables $X_1,\ldots,X_n$ where $\mathbb{P}(X_i = 1) = p$ and $d \in [0, 1]$ reads as

$$\mathbb{P}\left(\sum_i X_i \leqslant (1 - d)pn\right) \leqslant \exp\left(-\frac{d^2}{2}pn\right).$$

Moreover, we observe that a basic distillation step can be modeled by a Bernoulli random variable X_i where $\mathbb{P}(X_i = 1) = p$ is the probability of succeeding (measurement outcomes coincide).

Suppose we perform m rounds of entanglement distillation. Let N_m denote the number of input pairs to the mth round and let $d \in [0, 1]$. Then the Chernoff inequality implies that the probability that less than $(1 - d)pN_m$ basic distillation steps at round m have succeeded is bounded by $\exp\left(-\frac{d^2}{2}pN_m\right)$, i.e.

$$p_{\text{abort},m} = \mathbb{P}\left(\sum_i X_i \leqslant (1 - d)pN_m\right) \leqslant \exp\left(-\frac{d^2}{2}pN_m\right). \tag{E.2}$$

But this also implies that, with probability $1 - p_{\text{abort},m}$, at least $(1 - d)pN_m + 1$ basic distillation steps have succeeded at round m. Thus we may safely assume that $N_{m+1} = (1 - d)pN_m + 1$. The situation is summarized in figure E1. Furthermore we have $N_1 = c2^M$. Eliminating the recurrence relation yields $N_{m+1} = (1 - d)^m p^m c2^M + \sum_{i=0}^{m-1}(1 - d)^i p^i$. This implies for (E.2)

$$p_{\text{abort},m} \leqslant \exp\left(-\frac{d^2}{2}p\left((1 - d)^{m-1}p^{m-1}c2^M + \underbrace{\sum_{i=0}^{m-2}(1 - d)^i p^i}_{>0}\right)\right) \leqslant \exp\left(-\frac{d^2}{2}(1 - d)^{m-1}p^m c2^M\right).$$

Furthermore, we compute the probability of aborting the protocol at distillation round m (assuming that the previous rounds 1, .., $m - 1$ succeeded) by

$$p_{\text{abort at round } m} = p_{\text{abort},m} \underbrace{\prod_{k=1}^{m-1} p_{\text{succeed},k}}_{\leqslant 1} \leqslant p_{\text{abort},m} \leqslant \exp\left(-\frac{d^2}{2}(1-d)^{m-1}p^m c2^M\right). \tag{E.3}$$

The events of aborting the distillation protocol at two different rounds i and j are disjoint. Thus we have for the probability of aborting in any of m rounds $p_{\text{abort in any of } m \text{ rounds}} = \sum_{k=1}^{m} p_{\text{abort at round } k}$. A simple consequence thereof is

$$p_{\text{abort in any of } M \text{ rounds}} = \sum_{k=1}^{M} p_{\text{abort at round } k} \leqslant \sum_{k=1}^{M} \exp\left(-\frac{d^2}{2}(1-d)^{k-1}p^k c2^M\right), \tag{E.4}$$

where we have used (E.3). Inserting $p = 1/2$ and $d = 1/2$ in (E.4) yields

$$\begin{aligned} p_{\text{abort in any of } M \text{ rounds}} &\leqslant \sum_{k=1}^{M} \exp\left(-\frac{1}{8}\frac{1}{2^{2k-1}}c2^M\right) = \sum_{k=1}^{M} \exp\left(-c2^{M-2k-2}\right) \leqslant M \exp\left(-c2^{M-2M-2}\right) \\ &= M \exp\left(-c2^{-(M+2)}\right). \end{aligned} \tag{E.5}$$

By assumption we have $c = 2^{M+2}\xi$ which implies for (E.5)

$$p_{\text{abort in any of } M \text{ rounds}} \leqslant M \exp\left(-\xi 2^{M+2}2^{-(M+2)}\right) = M \exp\left(-\xi\right).$$

Thus, the probability of aborting the protocol satisfies

$$\varepsilon_R \leqslant p_{\text{pe-abort}} + (1 - p_{\text{pe-abort}})p_{\text{abort in any of Mrounds}} \leqslant \exp(-(3\beta - 4F_{\min}(\alpha) - 1)^2 \sqrt{k}/128) + M \exp(-\xi)$$

which completes the proof. □

Appendix F. Establishing a confidential quantum channel

For illustration purposes, we show how confidential quantum channels can be realized using our proposal in conjunction with standard teleportation. By our results, the joint state of Alice, Bob, and Eve after the distillation protocol is ϵ close to the output of the ideal protocol. The latter, since the register of L is not accessible to any of the parties and thus is traced out, yields the state of the form (provided the protocol was not aborted)

$$\rho_{\text{final}} = \sum_{i,j} |\omega_{ij}(\alpha)|^2 |B_{ij}\rangle\langle B_{ij}|_{AB} \otimes |\eta_{ij}\rangle\langle\eta_{ij}|_E. \tag{F.1}$$

The teleportation of any state ρ from Alice to Bob will yield the state

$$\sum_{i,j} |\omega_{i,j}(\alpha)|^2 \sigma_x^j \sigma_z^i \rho \sigma_z^i \sigma_x^j \otimes |\eta_{ij}\rangle\langle\eta_{ij}|_E. \tag{F.2}$$

Thus the only information Eve can obtain is what noise operator was applied on the teleported state, and nothing more—thus, the channel is confidential. Moreover, the probabilities for the different noise processes are not under Eve's control, but depend on the local devices.

References

[1] Bennett C H, Brassard G, Crépeau C, Jozsa R, Peres A and Wootters W K 1993 *Phys. Rev. Lett.* **70** 1895
[2] Bennett C H, DiVincenzo D P, Shor P W, Smolin J A, Terhal B M and Wootters W K 2001 *Phys. Rev. Lett.* **87** 077902
[3] Cirac J I, Ekert A K, Huelga S F and Macchiavello C 1999 *Phys. Rev.* A **59** 4249
[4] Lo H K 2001 A simple proof of the unconditional security of quantum key distribution *J. Phys. A: Math. Gen.* **34** 6957
[5] Gottesman D and Lo H K 2003 Proof of security of quantum key distribution with two-way classical communications *IEEE Trans. Inf. Theor* **49** 457–75
[6] Shor P W and Preskill J 2000 *Phys. Rev. Lett.* **85** 441
[7] Baigneres T 2003 *Quantum Cryptography: On the Security of the BB84 Key-Exchange Protocol* LASEC-STUDENT-2006-001 (https://infoscience.epfl.ch/record/88168)
[8] Christandl M, König R and Renner R 2009 *Phys. Rev. Lett.* **102** 020504
[9] Renner R 2008 Security of quantum key distribution *Int. J. Quantum Inf.* **6** 1–127
[10] Zhao Y B and Yin Z Q 2014 Apply current exponential de Finetti theorem to realistic quantum key distribution *Int. J. Mod. Phys.: Conf. Ser.* **33** 1460370
[11] Acin A, Brunner N, Gisin N, Massar S, Pironio S and Scarani V 2007 *Phys. Rev. Lett.* **98** 230501
[12] Lim C C W, Portmann C, Tomamichel M, Renner R and Gisin N 2013 *Phys. Rev.* X **3** 031006
[13] Vazirani U and Vidick T 2014 *Phys. Rev. Lett.* **113** 140501
[14] Barnum H, Crépeau C, Gottesman D, Smith A and Tapp A 2002 Authentication of quantum messages *The 43rd Annual IEEE Symp. on Foundations of Computer Science, Proc.* pp 449–58
[15] Hayden P, Leung D W and Mayers D 2011 Universal composable security of quantum message authenticationwith key recycling Talk at QCRYPT (Switzerland: Zurich)
[16] Broadbent A and Wainewright E 2016 *Information Theoretic Security: 9th International Conf., ICITS 2016 ((Tacoma, WA, 9–12, August 2016))* ((Cham: Springer International) pp 72–91 Revised Selected Papers 9

[17] Portmann C 2017 Quantum authentication with key recycling *Advances in Cryptology – EUROCRYPT 2017: 36th Annual International Conference on the Theory and Applications of Cryptographic Techniques* April 30 – May 4, 2017((Cham: Springer International) pp 339–68
[18] Garg S, Yuen H and Zhandry M 2016 New security notions and feasibility results for authentication of quantumdata *QCrypt 2016 (Washington, D.C., September 12–16, 2016)*
[19] Deutsch D, Ekert A, Jozsa R, Macchiavello C, Popescu S and Sanpera A 1996 *Phys. Rev. Lett.* **77** 2818
[20] Aschauer H and Briegel H J 2002 *Phys. Rev. Lett.* **88** 047902
[21] Aschauer H and Briegel H J 2002 *Phys. Rev.* A **66** 032302
[22] Brandão F G and Eisert J 2008 Correlated entanglement distillation and the structure of the set of undistillable states *J. Math. Phys.* **49** 042102
[23] Buscemi F and Datta N 2010 Distilling entanglement from arbitrary resources *J. Math. Phys.* **51** 102201
[24] Waeldchen S, Gertis J, Campbell E T and Eisert J 2016 *Phys. Rev. Lett.* **116** 020502
[25] Christandl M, König R, Mitchison G and Renner R 2007 One-and-a-half quantum de Finetti theorems *Commun. Math. Phys.* **273** 473–98
[26] Dür W, Hein M, Cirac J I and Briegel H J 2005 *Phys. Rev.* A **72** 052326
[27] Dür W, Briegel H J, Cirac J I and Zoller P 1999 *Phys. Rev.* A **59** 169–81
[28] Bennett C H, Brassard G, Popescu S, Schumacher B, Smolin J A and Wootters W K 1996 *Phys. Rev. Lett.* **76** 722
[29] Macchiavello C 1998 *Phys. Lett.* A **246** 385–8
[30] Bennett C H, DiVincenzo D P, Smolin J A and Wootters W K 1996 *Phys. Rev.* A **54** 3824
[31] Dür W and Briegel H J 2007 Entanglement purification and quantum error correction *Rep. Prog. Phys.* **70** 1381
[32] Nielsen M A and Chuang I L 2010 *Quantum Computation and Quantum Information* (Cambridge: Cambridge University Press)
[33] Altepeter J B, Jeffrey E R and Kwiat P G 2005 Photonic state tomography *Adv. At. Mol. Opt. Phys.* **52** 105–59
[34] Hoeffding W 1963 Probability inequalities for sums of bounded random variables *J. Am. Stat. Assoc.* **58** 13–30
[35] König R, Renner R, Bariska A and Maurer U 2007 *Phys. Rev. Lett.* **98** 140502
[36] Gottesman D, Lo H K, Lütkenhaus N and Preskill J 2004 *Quantum Inf. Comput.* **4** 325–60
[37] Lo H K and Chau H F 1999 *Science* **283** 2050–6