Quantum Science and Technology

PAPER • **OPEN ACCESS**

Advantage distillation for quantum key distribution

To cite this article: Zhenyu Du *et al* 2025 *Quantum Sci. Technol.* **10** 015050

IOP Publishing Quantum Sci. Technol. 10 (2025) 015050 https://doi.org/10.1088/2058-9565/ad9d75

Quantum Science and Technology

OPEN ACCESS

RECEIVED
2 August 2024

REVISED
2 December 2024

ACCEPTED FOR PUBLICATION
11 December 2024

PUBLISHED
20 December 2024

PAPER

Advantage distillation for quantum key distribution

Zhenyu Du, Guoding Liu, Xingjian Zhang and Xiongfeng Ma*

Center for Quantum Information, Institute for Interdisciplinary Information Sciences, Tsinghua University, Beijing 100084, People's Republic of China

* Author to whom any correspondence should be addressed.

E-mail: xma@tsinghua.edu.cn

Keywords: quantum key distribution, advantage distillation, error-correcting code

Abstract

Quantum key distribution promises information-theoretically secure communication, with data post-processing playing a vital role in extracting secure keys from raw data. While hardware advancements have significantly improved practical implementations, optimizing post-processing techniques offers a cost-effective avenue to enhance performance. Advantage distillation, which extends beyond standard information reconciliation and privacy amplification, has proven instrumental in various post-processing methods. However, the optimal post-processing remains an open question. Therefore, it is important to develop a comprehensive framework to encapsulate and enhance these existing methods. In this work, we propose an advantage distillation framework for quantum key distribution, generalizing and unifying existing key distillation protocols. Inspired by entanglement distillation, our framework not only integrates current techniques but also improves upon them. Notably, by employing classical linear codes, we achieve higher key rates, particularly in scenarios where one-time pad encryption is not used for post-processing. Our approach provides insights into existing protocols and offers a systematic way for further enhancements in quantum key distribution.

1. Introduction

Quantum key distribution (QKD) is the most practical field in quantum information science, with seminal protocols like BB84 [1] and 10^{-91} [2] laying its foundation. QKD enables the establishment of secret keys between remote parties, conventionally referred to as Alice and Bob, with the assurance of information-theoretic security rooted in the principles of quantum mechanics [3, 4]. While real-world imperfections in state preparation, transmission, and measurement pose inherent challenges, the development of ingenious techniques like the decoy-state method [5–7] and cutting-edge measurement-device-independent protocols [8–11] has significantly pushed the limits of practical implementations [12]. In parallel, hardware advancements have driven the evolution of QKD systems, exemplified by satellite-based QKD, which harnesses the unique benefits of space-based platforms to overcome terrestrial limitations in establishing secure long-distance communication links [13, 14]. Additionally, the integration of QKD with quantum repeaters and other quantum communication technologies within quantum networks has paved the way for secure communication among multiple parties over extended distances [15–20]. These remarkable strides collectively signify a paradigm shift towards achieving quantum-secured communication on a global scale, promising unprecedented security levels for sensitive information exchange.

The QKD process can be dissected into two distinct phases. In the quantum phase, quantum states are prepared and transmitted through channels, often assumed to be susceptible to adversarial control by Eve. Following potential tampering by Eve, these states are measured by detection devices, yielding raw data. Subsequently, the classical post-processing phase commences, where Alice and Bob execute specific procedures to distill secure key bits from this raw data. This involves data sifting, which is dictated by the chosen QKD protocol, followed by preprocessing, information reconciliation, and privacy amplification. The flowchart of QKD, as illustrated in figure 1, provides a visual representation of this process.

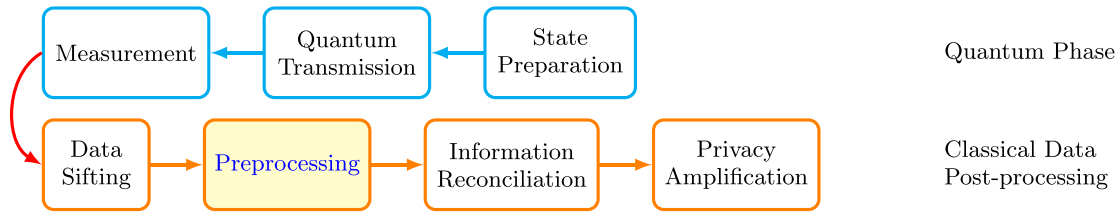

Figure 1. Flowchart for QKD. The process consists of two phases. Top: The quantum phase involves state preparation, transmission through a potentially adversarial channel, and measurement. Bottom: Classical data post-processing encompasses data sifting, preprocessing (the focus of this work), information reconciliation, and privacy amplification. Parameter estimation, utilizing data statistics, is also crucial. For simplicity, procedures like authentication and error verification [21] are omitted.

While hardware advancements undoubtedly contribute to improving QKD performance [12], the development of refined post-processing schemes presents a more economically viable path. Standard post-processing methods typically encompass data sifting, information reconciliation, and privacy amplification. Initially, raw keys are obtained after data sifting, containing both bit and phase errors. The goals of information reconciliation and privacy amplification are to correct bit and phase errors separately, utilizing standard error correction method [22]. Shor and Preskill [4] derive the key formula for the BB84 protocol:

$$1 - h(\delta_b) - h(\delta_p) \tag{1}$$

where δ_b and δ_p represent the bit and phase error rates, respectively, and h is the entropy function. The information-theoretical security of these processes is grounded in entanglement distillation [3], complementary approaches [23], or entropic approaches [24, 25]. Among these, the security proof based on entanglement distillation stands out due to its broader applicability, extending beyond QKD to encompass entanglement distillation and quantum networks. Initially conceived for distilling perfect Einstein–Podolsky–Rosen (EPR) pairs between Alice and Bob, this proof was later ingeniously extended to the prepare-and-measure protocol [4]. In the practical prepared-and-measure protocols, one-time pad (OTP) encryption is usually applied [23, 26]. Nonetheless, the role of OTP in classical communication warrants further investigation.

While data sifting, information reconciliation, and privacy amplification are often standardized, the preprocessing step is where various post-processing schemes strive to enhance QKD performance, thus becoming a focal point for innovation. Preprocessing seeks effective methods to further select keys and alter the bit and phase errors, aiming to improve key rate and error thresholds. A particularly effective method, classical advantage distillation, leverages two-way communication to establish secret keys in scenarios where one-way communication proves insufficient [27]. Gottesman and Lo [28] harnessed two-way communication to enhance the performance of QKD protocols. Their innovative introduction of the 'B step' as a preprocessing operation on raw keys, applied prior to standard information reconciliation and privacy amplification, paved the way for further advancements. Building upon this, Chau's introduction of adaptive privacy amplification [29] further pushed the boundaries, achieving the current state-of-the-art results in tolerable error rates.

In addition to two-way communication techniques, a novel protocol has emerged that involves Alice deliberately adding noise to her raw keys [30]. This strategy, aimed at reducing Eve's knowledge of the keys, is also considered a form of preprocessing. It can be implemented using one-way classical communication, illustrating that innovative one-way schemes can significantly enhance the performance of QKD.

Although various methods have been proposed, achieving optimal post-processing remains an important open question. Given the considerable flexibility inherent in preprocessing techniques, the development of a comprehensive framework that both unifies and enhances these diverse methods becomes imperative. Such a framework would not only deepen our understanding of existing protocols but also serve as a springboard for refining them, ultimately boosting the performance of QKD protocols.

In this paper, we propose an advantage distillation framework for QKD, demonstrating its capacity to seamlessly integrate previous protocols, such as the B step and adding noise, while exhibiting remarkable generality. This integration lays the groundwork for a fundamental understanding that can be harnessed to elevate key rates and error thresholds across a spectrum of preprocessing schemes. To showcase the potency of our framework, we apply it in conjunction with classical linear codes. We rigorously derive key rate formulas for various scenarios, encompassing adding noise and schemes both with and without OTP encryption. Our findings underscore the enhanced efficiency of QKD protocols that forgo OTP encryption

compared to those conventionally employing it. Moreover, we apply our framework to $[n\ 1\ n]$ codes, deriving a tolerable error rate that aligns with the best-known results in the literature [29]. Notably, while our framework seamlessly integrates previous protocols [31], we achieve an improvement in key rate through the systematic utilization of $[n\ 1\ n]$ codes and $[m\ m-1\ 2]$ codes. These insights collectively offer a methodical approach to refining current QKD protocols.

The remainder of this paper is organized as follows. In section 2, we review the basic concepts of error correction and the security proof based on entanglement distillation. In section 3.1, we introduce our QKD advantage distillation framework and apply it to different protocols, including those based on classical linear codes. In section 4, we analyze the role of error-correcting codes in advantage distillation and derive key rate formulas under different scenarios. Our analysis shows that omitting OTP encryption can yield a higher key rate. We also derive the key rate formula for incorporating the adding noise method into our framework. In section 5, we employ our framework to derive the error tolerance threshold that aligns with the best-known result and improves the key rate of previous works. In section 6, we provide the conclusion and outlook.

2. Preliminaries

In this section, we introduce the basic concepts and results of error correction and how they fit into proving the security of QKD protocols under the entanglement distillation picture.

2.1. Classical linear code

We introduce classical linear code [32]. A linear code C encoding k-bit message into an n-bit code space is specified by the $n \times k$ *generator matrix* G whose elements belong to {0,1}. The matrix G maps messages to their encoded equivalent. Thus, the k-bit message x is encoded as Gx. The multiplication and plus operation in linear code is modulo 2. The set of codewords generated by G is also denoted as C, in which

$$C = \left\{ Gx : x \in \{0,1\}^k \right\}. \tag{2}$$

Denote $|x|$ as the Hamming weight of a codeword x, i.e. the number of nonzero entries of x. The Hamming distance between two codewords x, y as $d(x,y) \equiv |x+y|$. The distance d of a linear code C can be defined as

$$d \equiv \min_{x,y \in C, x \neq y} d(x,y) = \min_{x \in C, x \neq 0} |x|. \tag{3}$$

We denoted C as $[n\ k\ d]$ code.

Another important matrix is the $(n-k) \times n$ *parity check matrix* H satisfying $HG = 0$. From equation (2), we have

$$Hx = 0, \forall x \in C. \tag{4}$$

If $Hx \neq 0$, then some errors must cause the encoded message to no longer belong to C.

We introduce some definitions that are used substantially in this work. The *tag* of an n-bit string x is defined as $t = Hx$. An *error pattern* is a bit string e of length n. The tag of an error pattern e is defined as *error syndrome* $s = He$, which has length $n-k$. There are 2^{n-k} different syndrome $\{s^0, s^1, \cdots, s^{2^{n-k}-1}\}$, and for each syndrome s^j, there are 2^k error pattern $\{e_0^j, e_1^j, \cdots, e_{2^k-1}^j\}$ satisfying $He_i^j = s^j$. We will use error syndromes to classify error patterns.

For a linear code C with parity check matrix H and generator matrix G, its *dual code* $C^\perp$ is defined with generator matrix H^T and parity check matrix G^T. A code and its dual code are useful in the security proof based on entanglement distillation, where the parity check matrix H of code C is used to classify bit error patterns and the parity check matrix G^T of code $C^\perp$ is used to classify phase error patterns. For more discussions, see sections 2.3 and 4.

Consider the $[7\ 4\ 3]$ code as an illustrative example, with the parity check matrix H given by:

$$H = \begin{pmatrix} 1 & 0 & 0 & 1 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 1 \end{pmatrix}. \tag{5}$$

The quantum circuit for extracting the tag Hx from a bit string x is provided in figure 2. Assuming the initial state of the data qubits is $|x\rangle$ in the computational basis, the tag can be extracted using CNOT gates from data qubits to ancillary qubits.

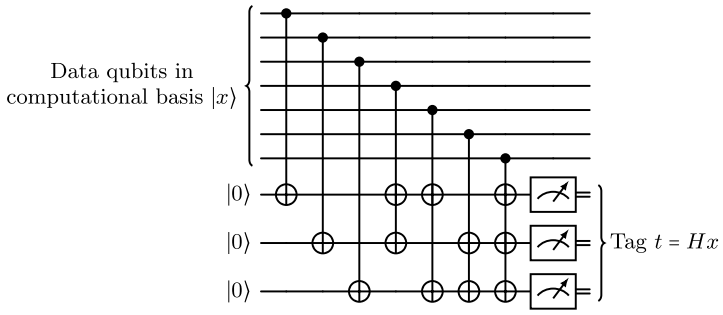

Figure 2. Quantum circuit for extracting the tag $t = Hx$ from data qubits $|x\rangle$. The circuit employs the parity check matrix H of the [7 4 3] code. Ancillary qubits are initialized to $|0\rangle$. The operations on each ancillary qubit correspond to a row of H. Ancillary qubits are measured in the computational basis at the end, yielding the tag $t = Hx$ of the data x.

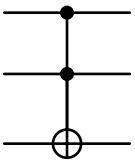

Figure 3. The CCNOT operation, also known as the Toffoli gate, exhibits non-linearity in its treatment of the two control qubits.

Notably, each CNOT gate has one of the data qubits as the control qubit, showcasing a feature of linear codes. This property does not hold for non-linear operations, such as the CCNOT operation depicted in figure 3, which might be utilized in some non-linear codes. While we do not delve into the application of non-linear codes in this work, it is worth mentioning that some of them can be integrated into our framework if they are classically replaceable operations [33]. We will discuss the classically replaceable operation in section 2.3.

2.2. ε-smallest probable set and error correction

We discuss basic concepts about error correction. Suppose Alice and Bob hold two n-bit strings x, y, the error pattern e is defined as $e = x + y$. We show how Alice and Bob can identify the error pattern e. The error pattern e can be seen as a realization of a random variable X distributed on $\{0,1\}^n$, according to a probability distribution p_X. To analyze the behavior of error patterns in the asymptotic limit, we first define the ε-smallest probable set.

Definition 1 (ε**-smallest probable set).** Given $0 \leqslant \varepsilon \leqslant \frac{1}{2}$, a random variable $X \in \mathcal{X}$ which has a finite range $|\mathcal{X}| < \infty$ and probability distribution p_X, the ε-smallest probable set $\mathcal{T}_X^\varepsilon$ of X is defined as the smallest set of values such that a realization of X lies in it with a failure probability no larger than ε,

$$\begin{aligned} \mathcal{T}_X^\varepsilon &= \mathrm{argmin}_{\mathcal{S}} |\mathcal{S}|, \\ s.t.\, &\Pr(X \in \mathcal{S}) \geqslant 1 - \varepsilon. \end{aligned} \tag{6}$$

If $X = (X_1, X_2, \cdots, X_n)$ and $X_1, X_2, \cdots, X_n$ are *i.i.d.*, the cardinality of ε-smallest probable set can be upper bounded by

$$|\mathcal{T}_X^\varepsilon| \leqslant 2^{n(h(X_1)+\varepsilon)} \tag{7}$$

for n sufficiently large [34]. Here, the entropy of random variable $h(X_1)$ is defined as

$$h(X_1) = -\sum_x p(X_1 = x) \log p(X_1 = x). \tag{8}$$

We also define

$$h(x) = -x \log x - (1-x) \log (1-x) \tag{9}$$

to denote the entropy function of a real number x, where $0 \leqslant x \leqslant 1$.

To succeed in correcting the error with a failure probability ε, we only need to guarantee success when the error pattern is in the ε-smallest probable set $\mathcal{T}_X^{\varepsilon}$. To identify the error pattern, we introduce universal hash functions.

Definition 2 (universal hash family [35, 36]). A family of functions $\mathcal{F}$ mapping elements e in a space $\mathcal{T}$ to another space $\mathcal{S}$ is called a universal hash family if the probability of a randomly chosen hash function $f \in \mathcal{F}$ outputting the same hashing result for any two different strings is upper bounded by

$$\forall e_i \neq e_j \in \mathcal{T}, \Pr_{f\in\mathcal{F}}\left[f(e_i) = f(e_j)\right] \leqslant \frac{1}{|\mathcal{S}|}. \tag{10}$$

In our scenario, the elements of $\mathcal{T}$ comprise n-bit error patterns and $|\mathcal{S}| = 2^k$, with k denoting the length of the tag. A crucial step in error pattern identification is selecting universal hashing functions, and random $k \times n$ hashing matrices can be employed for this purpose [26].

To identify the error pattern e, they jointly and randomly select a hash function $f \in \mathcal{F}$. Then, they compute the tag $s_x = f(x)$ and $s_y = f(y)$ independently. Alice then transmits s_x to Bob through an authenticated channel, after which Bob calculates the error syndrome $s = f(e) = f(x \oplus y) = s_x \oplus s_y$. If e is the unique error pattern in $\mathcal{T}$ satisfying $f(e) = s$, then e can be successfully identified. The probability of failure is bounded by

$$\sum_{e'\in\mathcal{T}\setminus\{e\}} \Pr[f(e') = f(e)] \leqslant \frac{|\mathcal{T}|}{|\mathcal{S}|}. \tag{11}$$

Hence, e can be identified with a probability of $1 - \frac{|\mathcal{T}|}{|\mathcal{S}|}$.

From equation (11), to guarantee small failure probability δ on successfully identifying the error pattern in set $\mathcal{T}$, the length of the tag should satisfy

$$k \geqslant \log|\mathcal{T}| + \log\frac{1}{\delta}. \tag{12}$$

Combined with equation (7), we derive the following lemma.

Lemma 1 (One-way error correction [22, 34]). *Suppose Alice holds n-bit string $x \in C$, Bob holds n-bit string $x + e$, e is a realization of a random variable X, and k is the length of the tag. To identify the error pattern e with probability $1 - \varepsilon$, $\frac{k}{n}$ can be arbitrarily close to $\frac{\log|\mathcal{T}_n^{\varepsilon}|}{n}$ for sufficiently large n. If $X = (X_1, X_2, \cdots, X_n)$ and X_is are i.i.d, then for any $\varepsilon > 0$, $\frac{k}{n}$ can be arbitrarily close to $h(X_1)$ for sufficiently large n.*

Proof. From equation (12), to identify e with failure probability $\varepsilon + \delta$, the length of the tag can be chosen as $\frac{k}{n} \geqslant \frac{\log|\mathcal{T}_n^{\varepsilon}| + \log\frac{1}{\delta}}{n}$. δ can be arbitrarily close to 0 and $\frac{k}{n}$ can be arbitrarily close to $\frac{\log|\mathcal{T}_n^{\varepsilon}|}{n}$, provided n is sufficiently large.

If $X = (X_1, X_2, \cdots, X_n)$, X_is are *i.i.d*, then by equation (7), $\frac{k}{n} \leqslant h(X_1) + \varepsilon$ for n sufficiently large. Note that ε can be chosen arbitrarily small. Thus, we draw the conclusion. □

In this error correction scheme, Alice transmits information to Bob, enabling him to identify and correct the error pattern e, ultimately resulting in both parties possessing identical key strings. We designate such a scheme as a one-way error correction scheme.

2.3. Security proof of QKD based on entanglement distillation

We present the security proof of QKD based on entanglement distillation. First, we describe an entanglement distillation protocol that utilizes one-way classical communication [3, 37]. Then, we demonstrate that this protocol, when followed by computational basis measurements, can be reduced to the prepare-and-measure QKD protocol, thereby establishing its security [4, 26].

2.3.1. One-way entanglement distillation

Denote the perfect EPR pair as $|\psi_{00}\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$. Entanglement distillation can be formalized as this: After distributing several noisy EPR pairs through a noisy channel to two parties, Alice and Bob measure some of the EPR pairs to estimate the error, with a $2N$-qubits state ρ^N remained. Based on the error estimation, they perform quantum error correction to transform ρ^N into a state σ^M which can be arbitrarily close to $|\psi_{00}\rangle^{\otimes M}$ in the large N limit. Then, they measure the state σ^M on the Z basis and obtain secure keys.

To perform the error estimation, Alice and Bob can perform X and Z measurements on subsets of EPR pairs and employ the random sampling method [21] to obtain the bit error rate δ_b and phase error rate δ_p of $2N-$qubits state ρ^N, which is defined as:

$$\begin{aligned}\delta_b &= \frac{1}{2}\left(1-\frac{1}{N}\mathrm{tr}\left[\rho^N\sum_{i=1}^{N}\left(\sigma_z^i\otimes\sigma_z^i\right)\right]\right),\\ \delta_p &= \frac{1}{2}\left(1-\frac{1}{N}\mathrm{tr}\left[\rho^N\sum_{i=1}^{N}\left(\sigma_x^i\otimes\sigma_x^i\right)\right]\right),\end{aligned} \tag{13}$$

in which $\sigma_z^i(\sigma_x^i)$ denotes the Pauli operator that perform $\sigma_z(\sigma_x)$ on site i and acts trivially on other sites. Error estimation, which is equivalent to the parameter estimation step in BB84 protocols, can be treated as a separated process preceding the subsequent distillation protocols.

Denote the Bell basis as $|\psi_{ab}\rangle = I\otimes X^aZ^b|\psi_{00}\rangle$, $a,b\in\{0,1\}$. That is,

$$\begin{aligned}|\psi_{00}\rangle &= \frac{1}{2}\left(|00\rangle+|11\rangle\right),\\ |\psi_{10}\rangle &= \frac{1}{2}\left(|01\rangle+|10\rangle\right),\\ |\psi_{01}\rangle &= \frac{1}{2}\left(|00\rangle-|11\rangle\right),\\ |\psi_{11}\rangle &= \frac{1}{2}\left(|01\rangle-|10\rangle\right).\end{aligned} \tag{14}$$

We express the complete state of N noisy EPR pairs in the Bell basis. In appendix A, we show that the off-diagonal terms do not impact the distillation process. Thus, we can restrict our attention to the Bell-diagonal terms. Additionally, the bit and phase errors in each qubit can be made independent and identically distributed (*i.i.d.*) through random permutation [38]. Therefore, without loss of generality, we consider the input state to be a Bell-diagonal state of the form $\rho^{\otimes N}$, where

$$\rho = \mathrm{diag}\left(p_{00},p_{01},p_{10},p_{11}\right) \tag{15}$$

in the Bell basis $|\psi_{00}\rangle,|\psi_{01}\rangle,|\psi_{10}\rangle,|\psi_{11}\rangle$. From equation (13), we have

$$\begin{aligned}\delta_b &= p_{10}+p_{11},\\ \delta_p &= p_{01}+p_{11}.\end{aligned} \tag{16}$$

As ρ is Bell-diagonal, $\rho^{\otimes N}$ is in the state $\bigotimes_{i=1}^{N}\psi_{a_i,b_i}$ with a probability of $\prod_{i=1}^{N}p_{a_i,b_i}$. The bit error pattern is defined as $e_b = a_1a_2\cdots a_N$, and the phase error pattern is defined as $e_p = b_1b_2\cdots b_N$. Alice and Bob apply CNOT gates from $\rho^{\otimes N}$ to the ancillary qubits according to a hashing matrix H and measure the ancillary qubits to identify the bit error pattern. Subsequently, Alice sends her measurement results to Bob. Bob obtains the bit error syndrome $s_b = Me_b$ and deduces the bit error with high probability. According to lemma 1, the length of the tag $\frac{k}{n}$ is $h(\delta_b)$ asymptotically.

In entanglement distillation, bit error correction and phase error correction need to be implemented sequentially. Thus, we need to ensure that the two error correction procedures do not affect each other. As stated in the following lemma, this is guaranteed when we use perfect EPR pairs as ancillary qubits.

Lemma 2 (Bit- and phase-error-correction decoupling [39]). *Using EPR pairs as ancilla qubits, bit error correction does not affect phase errors and vice versa.*

Proof. Consider the effect of bit error correction on the phase error as an example. As shown in figure 4, if Alice and Bob use EPR pairs as ancillary qubits, the equivalent circuit indicates that, in the phase error space, Alice and Bob *simultaneously* apply the XOR operation or do not apply it at all, which keeps the phase error unchanged.

We can also directly prove this statement by considering the effect of the bilateral CNOT operation on EPR pairs:

$$\left(\mathrm{CNOT}_{13}\mathrm{CNOT}_{24}\right)|\psi_{a,b}\rangle\langle\psi_{a,b}|_{12}|\psi_{c,d}\rangle\langle\psi_{c,d}|_{34}\left(\mathrm{CNOT}_{13}\mathrm{CNOT}_{24}\right)^{\dagger} = |\psi_{a,b\oplus d}\rangle\langle\psi_{a,b\oplus d}|_{12}|\psi_{a\oplus c,d}\rangle\langle\psi_{a\oplus c,d}|_{34}. \tag{17}$$

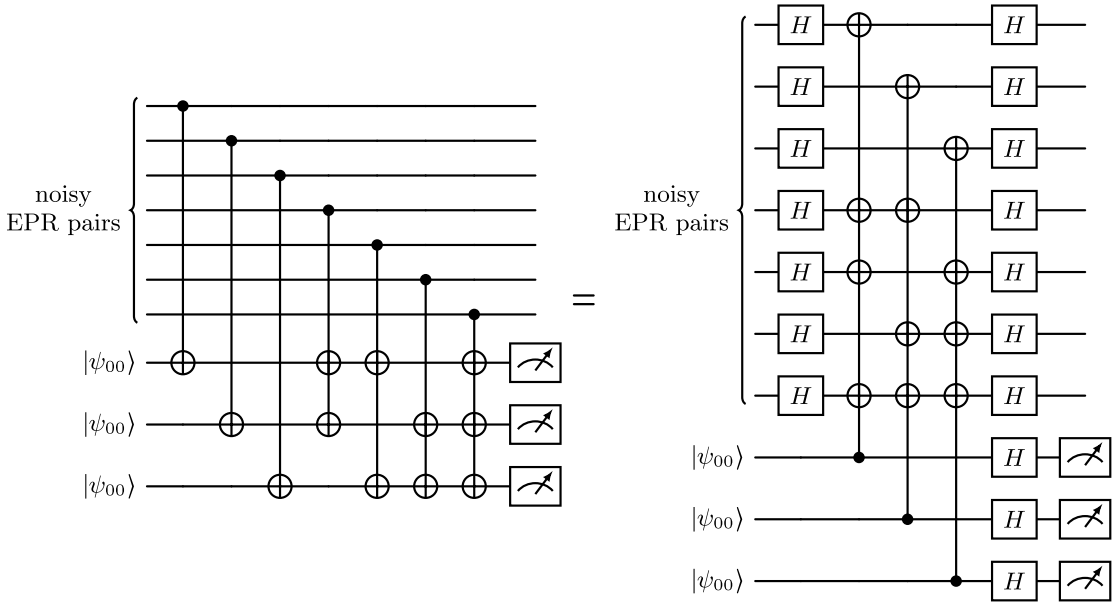

Figure 4. The quantum circuit for hashing and measuring the tag, along with its equivalent circuit. Only one side of Alice and Bob is illustrated. Here, the Hadamard gates transform phase errors into bit errors. Since the ancillary qubits are perfect EPR pairs, Alice and Bob simultaneously apply the XOR operation or not, ensuring that the phase error pattern remains unchanged.

Here, CNOT_{13} (CNOT_{24}) corresponds to the CNOT operation on Alice's (Bob's) side. When the ancillary qubits are initially perfect EPR pairs, meaning $c=0$ and $d=0$, the subsequent CNOT operations only alter c, while keeping $d=0$ unchanged. Thus, b remains unchanged after each CNOT operation, ensuring that the phase error is unaffected.

After identifying the bit error pattern, Bob can apply X gates to correct bit errors in some of the noisy EPR pairs. The X operation does not affect the phase error since $X|\psi_{ab}\rangle = |\psi_{a\oplus 1,b}\rangle$. Thus, bit error correction does not influence the phase error pattern, and vice versa, due to the symmetry between bit and phase error correction. □

In the BB84 protocol, Alice and Bob are aware of both the bit and phase error rates. In the worst-case scenario, these errors are independent. Consequently, correcting the bit errors consumes $nh(\delta_b)$ EPR pairs, while correcting the phase errors consumes $nh(\delta_p)$ EPR pairs. Therefore, a total of $nh(\delta_b)+nh(\delta_p)$ EPR pairs are used to obtain n perfect EPR pairs with high probability. The distillable rate, defined as the net gain of perfect EPR pairs divided by the number of noisy EPR pairs consumed, is given by:

$$r = 1 - h(\delta_b) - h(\delta_p). \tag{18}$$

If Alice and Bob acquire more detailed information about the error terms beyond just the bit error rate δ_b and the phase error rate δ_p, the distillable rate can be further improved. For instance, by performing measurements in the X, Y, and Z bases during parameter estimation-as in the six-state protocol [40]-they can fully characterize all the diagonal terms of the noisy EPR pairs. This additional information increases the tolerable bit error rate up to 12.7%. Notably, this entanglement distillation protocol requires only one-way communication, as it leverages one-way error correction, as described in lemma 1.

2.3.2. Reduction to prepare-and-measure QKD protocol

Suppose Alice and Bob successfully distill M perfect EPR pairs, $|\psi_{00}\rangle^{\otimes M}$. Then, Eve is only left with a state ρ_E such that the total system is in a product state:

$$|\psi_{00}\rangle\langle\psi_{00}|^{\otimes M}\otimes\rho_E. \tag{19}$$

After Alice and Bob measure the perfect EPR pairs in the computational basis, the state becomes:

$$2^{-M}\sum_{k\in\{0,1\}^M}|kk\rangle\langle kk|_{AB}\otimes\rho_E. \tag{20}$$

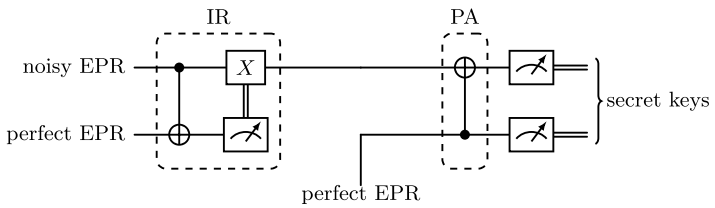

Figure 5. The QKD protocol consists of two essential steps: information reconciliation (IR) and privacy amplification(PA) [4, 26]. Here, we focus on Bob's operations. In the information reconciliation step, Alice and Bob conduct hashing from noisy EPR pairs to perfect EPR pairs to extract the bit error syndrome(illustrated by the CNOT operation from noisy EPR to perfect EPR). Subsequently, Bob corrects his noisy EPR pairs based on the obtained syndrome, employing the controlled-X operation. They proceed to privacy amplification, which uses additional perfect EPR pairs to add randomness to their keys, ensuring the security of the generated keys (illustrated by the CNOT operation from perfect EPR to noisy EPR). Finally, the process concludes with measuring the noisy EPR pairs in the computational basis, obtaining the secret keys.

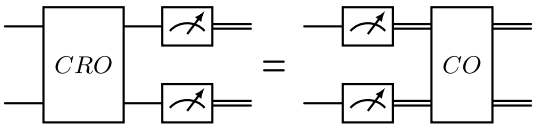

Figure 6. Illustration of CRO. A CRO operation followed by measurement in the computational basis is equivalent to measurement in the computational basis followed by classical operation (CO).

Here, Eve's state ρ_E is independent of the secret keys k, ensuring that the keys are secure. In practice, when Alice and Bob hold states that are very close to perfect EPR pairs, security can still be proven [21, 41, 42]. Consequently, equation (18) provides the key rate for QKD. The maximum tolerable bit error rate for this one-way protocol, which represents the threshold at which Alice and Bob can still obtain a net gain of secure keys, is 11%.

We call the process of bit error correction information reconciliation because it ensures that Alice and Bob ultimately share the same keys. To guarantee the keys are secure, we only need to show that the phase error pattern can be identified without actually calculating it [4]. Consequently, the phase error correction can be replaced with privacy amplification, which introduces additional shared randomness to erase Eve's knowledge about the keys [26]. The QKD process can then be succinctly summarized as information reconciliation followed by privacy amplification, as illustrated in figure 5.

The procedure in figure 5 requires quantum computers. Fortunately, it is equivalent to the prepare-and-measure protocols [4, 26]. This reduction to classical operations can be understood through the framework proposed by [33], which defines the classically replaceable operation.

Definition 3 (Classically Replaceable Operation, CRO [33]). Consider operations right before a measurement. A CRO can be realized by first measuring the input state and then processing the outcomes classically.

The process of replacing quantum operations with classical operations is depicted in figure 6. In the QKD protocol, information reconciliation and privacy amplification are classified as CROs [26, 33]. Consequently, the final measurement in the QKD protocol can be moved to the beginning to obtain raw keys x and y on each side, followed by classical information reconciliation and privacy amplification, as illustrated in figure 7. Using perfect EPR pairs is thus reduced to employing pre-shared secret keys for encrypting communication via OTP. For information reconcillation, Alice uses a pre-shared secret key to encrypt the tag s_x of her raw keys x, $s_x = Hx$. Bob can then use the same pre-shared key to decrypt Alice's tag, obtain the symdrome of bit error pattern $s_b = s_x \oplus s_y$, and correct the bit errors $e_b = x \oplus y$. For the adversary, Eve, the encrypted message appears as a random string, offering no information. For privacy amplification, Alice and Bob use pre-shared secret keys to enhance the secrecy of keys [26]. Ultimately, the QKD protocol is reduced to a prepare-and-measure protocol.

3. Advantage distillation for QKD

As discussed above, Alice and Bob distill perfect EPR pairs through one-way communication and perform computational basis measurements to obtain their secret key. This process ultimately reduces to classical

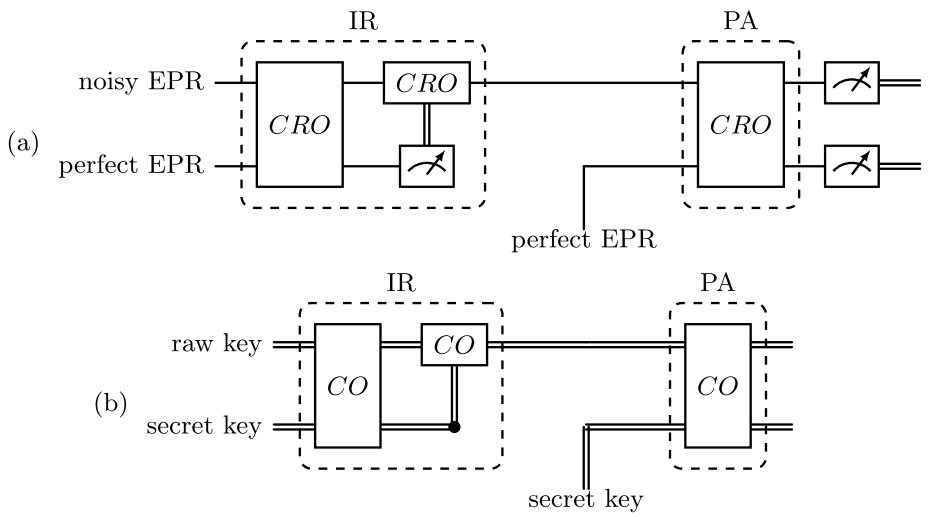

Figure 7. Reducing the QKD protocol to the prepare-and-measure protocol. (a) Information reconciliation (IR) and privacy amplification (PA) are classified as CROs; thus, the measurement can be moved forward to the beginning. (b) After moving the measurement forward, all operations become classical operations (CO). The perfect EPR becomes the secret key shared by Alice and Bob. Finally, the QKD protocol is reduced to a prepare-and-measure protocol.

information reconciliation and privacy amplification using standard one-way hashing techniques. However, not all key distillation schemes rely solely on these two steps. As illustrated in figure 1, additional preprocessing steps that allow both one-way and two-way communication are essential for various schemes to enhance both the tolerable error rate and the key rate.

An illustrative approach is the B step [28, 43], which involves segregating all raw keys into distinct groups. Subsequently, the bit error syndrome is determined in each group through linear hashing and two-way classical communication. Groups exhibiting high bit error rates are discarded, enhancing the overall key rate and error tolerance. This operation can be reduced to a prepare-and-measure protocol [28] and can be seen as a preprocessing step for refining the raw keys before information reconciliation and privacy amplification.

Adding noise is another example where Alice deliberately introduces noise into her raw keys [30, 44]. This intentional noise insertion diminishes Eve's knowledge of the keys, consequently increasing the key rate. Additionally, this method only requires one-way classical communication. While adding noise has not been proven secure through the entanglement distillation method, it has been shown that perfect EPR pairs are not a prerequisite for secure keys. Instead, private states are sufficient to establish secure keys and can be used to demonstrate the security of adding noise [45].

3.1. Advantage distillation framework

Numerous preprocessing techniques can be applied before Alice and Bob perform information reconciliation and privacy amplification. Therefore, developing a general framework to incorporate these techniques is necessary and helpful for further protocol design. Here, we propose the advantage distillation framework, which generalizes the methods above, following the process illustrated in figure 1. The protocols in our framework, developed within the entanglement distillation scenario, are equivalent to prepare-and-measure QKD protocols. This equivalence arises because all operations allowed in our framework are CROs, allowing a reduction similar to that in section 2.3. Consequently, the distillable rate of EPR pairs is equivalent to the key rate of the corresponding QKD protocols.

As depicted in figure 8, our advantage distillation framework comprises three steps. Alice and Bob initially have noisy EPR pairs $\rho^{\otimes N}$ [37]. First, they proceed to a preprocessing operation, which must be a CRO. Next, they discard some EPR pairs. Then, they perform information reconciliation to correct the bit errors. Finally, they proceed to privacy amplification. Ancillary qubits are utilized and measured at each step, and the measurement results can control subsequent operations classically.

We make some remarks on our framework. Firstly, it is possible to conduct multiple preprocessing rounds. As long as a CRO is performed in each round, the preprocessing eventually becomes a CRO. The CRO operation enables moving the measurement forward, effectively reducing the scheme to a prepare-and-measure protocol. Secondly, all the operations can be adaptive based on the measurement

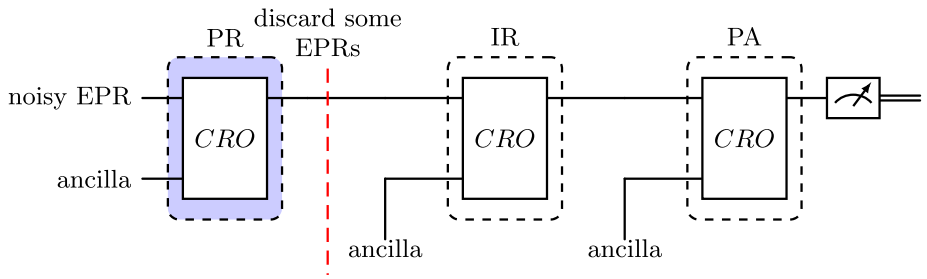

Figure 8. The advantage distillation framework. Initially, Alice and Bob preprocess (PR) the noisy EPR pairs. Subsequently, they discard certain EPR pairs based on the measurement results from preprocessing. They then proceed to information reconciliation (IR) and privacy amplification (PA). Finally, they perform computational basis measurements to obtain the secure key. The type of ancillary qubits is not specified and may vary in different protocols. Ancillary qubits are measured in each step, and all operations can be adaptive based on previous measurement results. We omit the measurement of ancillary qubits and the classical control wires in the figure. The operations must be CRO to enable the reduction to a prepare-and-measure protocol.

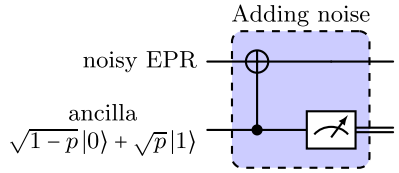

Figure 9. Incorporating adding noise [30, 44] into advantage distillation framework. The ancillary qubit is in the form of $\sqrt{1-p}|0\rangle + \sqrt{p}|1\rangle$ and a CNOT gate is applied from the ancillary qubit to the noisy EPR pair. After measuring the ancillary qubit in the computational basis, this process is equivalent to introducing a bit flip to the noisy EPR pair with a probability of p.

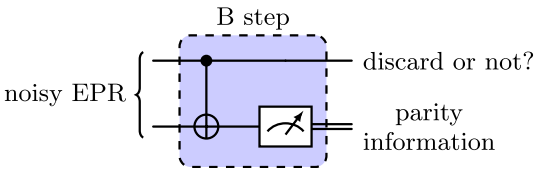

Figure 10. Incorporating the B step [28, 43] into the advantage distillation framework. A CNOT operation is conducted on a pair of noisy EPR pairs. Subsequently, a measurement is performed on one of the two EPR pairs to extract parity information. If the parity differs between the two sides of Alice and Bob, the remaining EPR pair is discarded. Conversely, if the parity is the same, the EPR pair is retained.

outcomes from previous rounds. Thirdly, some EPR pairs may be discarded before entering the information reconciliation process, similar to what is done in the B step. Lastly, the ancillary qubits are not specified, leaving room for devising different protocols. As we will show later, different types of ancillary qubits lead to different key rates, and there are better choices than perfect EPR pairs.

We exemplify our protocol using two previously mentioned examples to illustrate the generality of our approach. Adding noise [30, 44] is achieved by utilizing ancillary qubits in the form of $\sqrt{1-p}|0\rangle + \sqrt{p}|1\rangle$ and applying a CNOT gate from the ancillary qubits to the noisy EPR pairs, as demonstrated in figure 9. After measuring the ancillary qubits in the computational basis, the process is equivalent to adding a bit flip to the noisy EPR pair with a probability of p.

The B step [28, 43] can also be incorporated into our framework, as depicted in figure 10. We present a single round B step for illustration. During this round, a CNOT operation is performed on a pair of noisy EPR pairs, followed by a measurement on one of the two EPR pairs to extract parity information. If the parity differs between Alice and Bob's sides, the remaining EPR pairs are discarded. Conversely, the remaining EPR pair is retained if the parity is the same. As remarked before, conducting the B step in multiple rounds is also a CRO and can be incorporated into our framework.

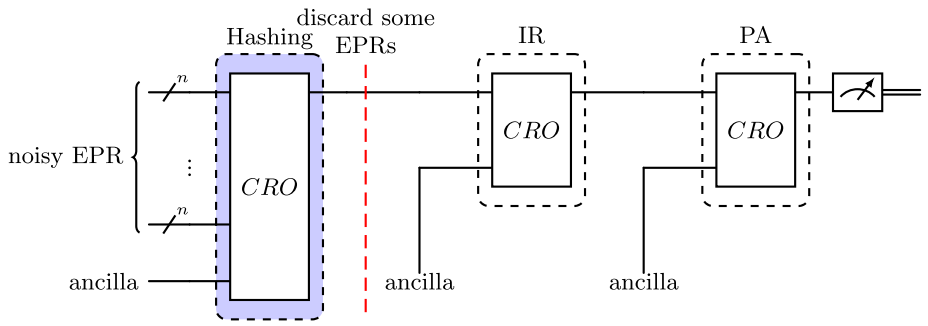

Figure 11. The advantage distillation framework based on classical linear code. Firstly, Alice and Bob divided the noisy EPR pairs into groups, each containing n EPR pairs. They then apply the parity matrix of the classical linear code to each group, extracting the parity information. Then, they engage in two-way classical communication to derive the bit error syndrome (contained in the blue box titled 'Hashing'). Based on the error syndrome, they decide whether to discard this group. Then, they proceed to information reconciliation and privacy amplification to obtain the secure key.

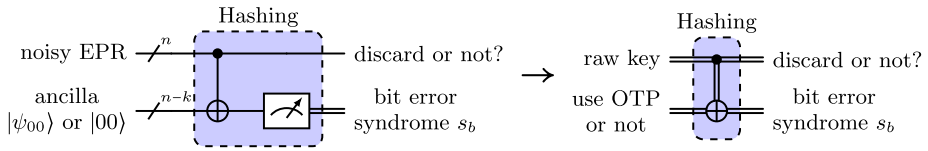

Figure 12. Preprocessing of raw keys using classical linear code. Left: Alice and Bob employ CNOT gates from noisy EPR pairs to ancillary qubits within each group based on the $(n-k)\times n$ parity check matrix H. Following the measurements of ancillary qubits and two-way communication (not depicted), they obtain the bit error syndrome s_b. The decision to discard the group is based on the key rate associated with this bit error syndrome. Right: As the hashing process is a CRO, the measurement can be shifted forward, rendering the entire process classical. The choice of ancillary qubits, either $|\psi_{00}\rangle$ or $|00\rangle$, corresponds to the use of OTP encryption or not, respectively.

3.2. Advantage distillation based on classical linear code

Now, we develop the advantage distillation framework based on classical linear code. Notice that the B step corresponds to the [2 1 2] code, which is not always optimal compared to other codes [46]. Thus, we incorporate the fundamental concept of the B step and extend it to encompass a general error-correcting code, as illustrated in figures 11 and 12 [46].

The EPR pairs are initially divided into groups, each containing n EPR pairs. Let e_b and e_p denote the bit error pattern and phase error pattern within a group. The linear hashing matrix H is applied in each group. Subsequently, Alice and Bob engage in two-way classical communication to derive the bit error syndrome $s_b = He_b$. Based on the error syndrome s_b and the corresponding key rate associated with this syndrome, as derived in theorems 1 and 3, they decide whether to discard the group of EPR pairs. Following this decision, information reconciliation and privacy amplification are applied, and the secure keys are obtained.

The hashing details are illustrated in figure 12. Alice and Bob utilize CNOT gates within each group based on the $(n-k)\times n$ parity check matrix H to operate on the noisy EPR pairs and ancillary qubits. The hashing process is a CRO, allowing the measurement to be shifted forward and rendering the preprocessing classical.

It is worth noting that the ancillary qubits can be selected as either $|\psi_{00}\rangle$ or $|00\rangle$, representing whether Alice and Bob opt to encrypt their messages using OTP during the two-way communication step. As detailed in section 2.3, OTP encryption decouples the bit and phase errors. Surprisingly, we show that the OTP step is unnecessary, and the key rate is *higher* when Alice and Bob do not encrypt their messages, as proved in lemma 4.

In the original B step, hashing is performed between data qubits, as shown in figure 10. Our framework utilizes additional ancillary qubits to carry the hashing value. It turns out that in-place hashing yields the same key rate as using additional ancillary qubits without OTP encryption. We compare these two methods in appendix B.

Our QKD advantage distillation framework based on classical linear code is summarized as follows.

Box 1. Advantage distillation based on classical linear code.

Input:
1. Raw keys obtained after data sifting.
2. BB84 protocol: bit error δ_b, phase error δ_p.
   Six-state protocol: Bell-diagonal terms $p_{00}, p_{01}, p_{10}, p_{11}$.

Advantage distillation:
1. Alice and Bob choose a $[n\ k\ d]$ code, with parity check matrix H and generator matrix G.
2. Alice and Bob divide n-bit raw key into one group.
3. Alice and Bob implement the same hash operation H to raw keys in each group.
4. Alice and Bob compare their results via two-way communication to deduce bit error syndrome s_b^j for each group. They can choose whether to perform OTP encryption.
5. For each bit error syndrome s_b^j, Alice and Bob keep this group if the key rate r^j corresponding to error syndrome s_b^j satisfying $r^j > 0$. Otherwise, they discard the group. The key rate formula for r^j is given in section 4.
6. Alice and Bob proceed to information reconciliation and privacy amplification on the raw keys to obtain secure keys.

4. Key rate formula for advantage distillation

Here, we derive the key rate formula for our QKD advantage distillation framework. Firstly, we introduce some notations. For a multiset S of positive real numbers satisfying $\sum_{x\in S} x = 1$, the entropy of this multiset $h(S)$ is defined as:

$$h(S) = -\sum_{x\in S} x\log x. \tag{21}$$

The entropy of a random variable X with probability mass function $p(X=x)$ is given by $h(X) = h(\{p(X=x)\}_x)$. The conditional entropy of two random variable X, Y is defined as

$$h(Y|X) = \sum_x p(X=x)\, h(Y|X=x) = \sum_x p(X=x)\, h\left(\{p(Y=y|X=x)\}_y\right). \tag{22}$$

Suppose Alice and Bob choose $[n\ k\ d]$ code C with parity check H and generator matrix G. There are totally 2^{n-k} bit error syndromes $\{s_b^j\}_j$. For a specific syndrome, there are 2^k bit error patterns $\{e_{b,i}^j\}_i$ corresponding to this syndrome. Denote the probability of getting bit error syndrome s_b^j as q^j, and the probability of getting bit error pattern $e_{b,i}^j$ as q_i^j. By definition, we have

$$q^j = \sum_{i=0}^{2^k-1} q_i^j. \tag{23}$$

We classify the phase error patterns according to the dual code $C^\perp$ of C. That is, two phase errors $e_{p,i'}^{j'}$ and $e_{p,k'}^{j'}$ corresponding to same syndrome $s_p^{j'}$ if $G^T e_{p,i'}^{j'} = G^T e_{p,k'}^{j'}$. There are 2^k phase error syndrome $\{s_p^{j'}\}_{j'}$ and 2^{n-k} phase error patterns $\{e_{p,i'}^{j'}\}_{i'}$ corresponding to a specific syndrome $s_p^{j'}$. Define $q_i^{jj'}$ as the probability of getting bit error pattern $e_{b,i}^j$ and phase error syndrome $s_p^{j'}$, and $q_{ii'}^{jj'}$ as the probability of getting bit error pattern $e_{b,i}^j$ and phase error pattern $e_{p,i'}^{j'}$. Then, by definition, we have

$$\begin{aligned} q_i^j &= \sum_{j'=0}^{2^k-1} q_i^{jj'}, \\ q_i^{jj'} &= \sum_{i'=0}^{2^{n-k}-1} q_{ii'}^{jj'}. \end{aligned} \tag{24}$$

We list our notations in the following table.

4.1. With OTP encryption

We derive the key rate formula for advantage distillation with OTP encryption in two-way classical communication. Alice and Bob have noisy EPR pairs divided into groups, each containing n EPR pairs. The initial step involves the application of the parity check matrix H to identify the error syndrome s^j. Subsequently, OTP encryption consumes $n-k$ secret keys per group.

After identifying the bit error syndrome s_b^j, the bit error pattern $e_{b,i}^j$ is contained in the set $\{e_{b,i}^j\}_i$. Then, Alice and Bob perform one-way error correction to correct $e_{b,i}^j$. Based on lemma 1, the asymptotic secret key consumption is the entropy of bit error patterns conditional on the bit error syndrome s^j, which can be expressed as

$$I_b^j = h\left(\left\{\frac{q_i^j}{q^j}\right\}_i\right). \tag{25}$$

Upon identification of the bit error pattern, the focus shifts to determining the phase error pattern $e_{p,i'}^{j'}$. As mentioned in section 2.3, to guarantee the keys are secure, Alice and Bob do not need to identify the exact phase error pattern [4] but only need to calculate the length of the tag required to perform phase error correction [26, 47], which gives the key consumption of privacy amplification. According to lemma 1, the length of the tag is given by:

$$\begin{aligned} I_{p,i}^j &= h\left(\left\{\frac{q_{ii'}^{jj'}}{q_i^j}\right\}_{j',i'}\right) \\ &= h\left(\left\{\frac{q_i^{jj'}}{q_i^j}\right\}_{j'}\right) + \sum_{j'}\frac{q_i^{jj'}}{q_i^j} h\left(\left\{\frac{q_{ii'}^{jj'}}{q_i^{jj'}}\right\}_{i'}\right), \end{aligned} \tag{26}$$

where the last equality is derived from the chain rule of the entropy function h. That is, for two random variables X, Y, the entropy of their joint distribution (X, Y) satisfies

$$\begin{aligned} h((X,Y)) &= h(X) + h(Y|X) \\ &= h\left(\{p(X=x)\}_x\right) + \sum_x p(X=x)\, h\left(\{p(Y=y|X=x)\}_y\right). \end{aligned} \tag{27}$$

The cumulative consumption in the one-way error correction of syndrome s_b^j is expressed as:

$$\begin{aligned} R^j &= I_b^j + \sum_i \frac{q_i^j}{q^j} I_{p,i}^j \\ &= h\left(\left\{\frac{q_i^j}{q^j}\right\}_i\right) + \sum_i \frac{q_i^j}{q_j} h\left(\left\{\frac{q_{ii'}^{jj'}}{q_i^j}\right\}_{j',i'}\right) \\ &= h\left(\left\{\frac{q_{i,i'}^{j,j'}}{q^j}\right\}_{i,j',i'}\right). \end{aligned} \tag{28}$$

The first line arises from the fact that the cost of the raw keys R^j, after identifying the syndrome s_b^j, comprises two components: the cost of identifying the bit error pattern, given by I_b^j, and the expected cost of identifying the phase error pattern across different bit error patterns, calculated as $\sum_i \frac{q_i^j}{q^j} I_{p,i}^j$.

The key idea for advantage distillation to increase the key rate and error threshold is as follows: If R^j exceeds n, the corresponding group of keys does not contribute positively to the final secret key and thus will be discarded. Conversely, if R^j is less than n, this group will be retained and produce $n - R^j$ bits of secure keys. Combining with the consumption $n-k$ of OTP encryption for each group, the key rate formula for advantage distillation with OTP encryption is obtained.

Theorem 1 (Key rate formula for advantage distillation with OTP encryption). *Suppose Alice and Bob perform two-way post-processing with one-time pad encryption in the advantage distillation framework using an $[n\ k\ d]$ linear code. The probabilities of getting bit error syndrome s_b^j, bit error pattern $e_{b,i}^j$, phase error syndrome*

Table 1. Notation.

bit error syndrome	s_b^j
bit error pattern	$e_{b,i}^j$
phase error syndrome	$s_p^{j'}$
phase error pattern	$e_{p,i'}^{j'}$
probability for bit error syndrome s_b^j	q^j
probability for bit error pattern $e_{b,i}^j$	q_i^j
probability for bit error pattern $e_{b,i}^j$ and phase error syndrome $s_p^{j'}$	$q_i^{jj'}$
probability for bit error pattern $e_{b,i}^j$ and phase error pattern $e_{p,i'}^{j'}$	$q_{ii'}^{jj'}$

$s_p^{j'}$ and phase error pattern $e_{p,i'}^{j'}$ are given in table 1. Then, the key rate is

$$\begin{aligned} r &= \sum_j q^j r^j - \frac{n-k}{n}, \\ r^j &= \max\left\{1 - \frac{1}{n} h\left(\left\{\frac{q_{i,i'}^{j,j'}}{q^j}\right\}_{i,j',i'}\right), 0\right\}. \end{aligned} \tag{29}$$

We stress that for the six-state protocol, we can calculate the probability of error patterns based on the bit and phase error rates, thereby determining the key rate r. However, for the BB84 protocol, a free parameter exists. Therefore, minimizing the key rate r across all valid free parameters is necessary.

To enhance the key rate formula outlined in theorem 1, we can compress the bit error syndrome during two-way communication with OTP encryption, leveraging lemma 1 once more. This compression reduces the consumption of secret keys for OTP to the entropy of bit error syndromes $h(\{q^j\}_j)$. Such compression enables a more efficient use of OTP, leading to an improved key rate.

Theorem 2 (Key rate formula for advantage distillation with OTP encryption and bit error syndrome hashing). *Suppose Alice and Bob perform two-way post-processing with hashing of the tag and one-time pad encryption in the advantage distillation framework using an $[n\ k\ d]$ linear code. The probabilities of getting bit error syndrome s_b^j, bit error pattern $e_{b,i}^j$, phase error syndrome $s_p^{j'}$ and phase error pattern $e_{p,i'}^{j'}$ are given in table 1. Then, the key rate is*

$$\begin{aligned} r &= \sum_j q^j r^j - \frac{h\left(\{q^j\}_j\right)}{n}, \\ r^j &= \max\left\{1 - \frac{1}{n} h\left(\left\{\frac{q_{i,i'}^{j,j'}}{q^j}\right\}_{i,j',i'}\right), 0\right\}. \end{aligned} \tag{30}$$

In theorems 1 and 2, OTP encryption is employed during information reconciliation, consuming secret keys. However, as shown in the next section (lemma 4), omitting OTP encryption enhances the key rate. Building on this observation, we further enhance the key rate in theorem 1 with the improvements outlined in theorem 3.

We remark that another information reconciliation method, involving directly measuring a portion of the qubits, yields a slight improvement over the results obtained from theorem 2. However, for brevity, we do not present the key rate formula in the main text. A detailed discussion is available in appendix C.

4.2. Without OTP encryption

The following section illustrates how omitting OTP encryption enhances the key rate. The key concept involves understanding the impact on the phase error when Alice and Bob do not encrypt their message while identifying the bit error syndrome. This is equivalent to utilizing $|00\rangle$ as ancillary qubits instead of employing EPR pairs to gather parity information about the noisy EPR pairs. We begin by examining the case where H is the hashing matrix of the $[7\ 4\ 3]$ code:

$$H = \begin{pmatrix} 1 & 0 & 0 & 1 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 1 \end{pmatrix}. \tag{31}$$

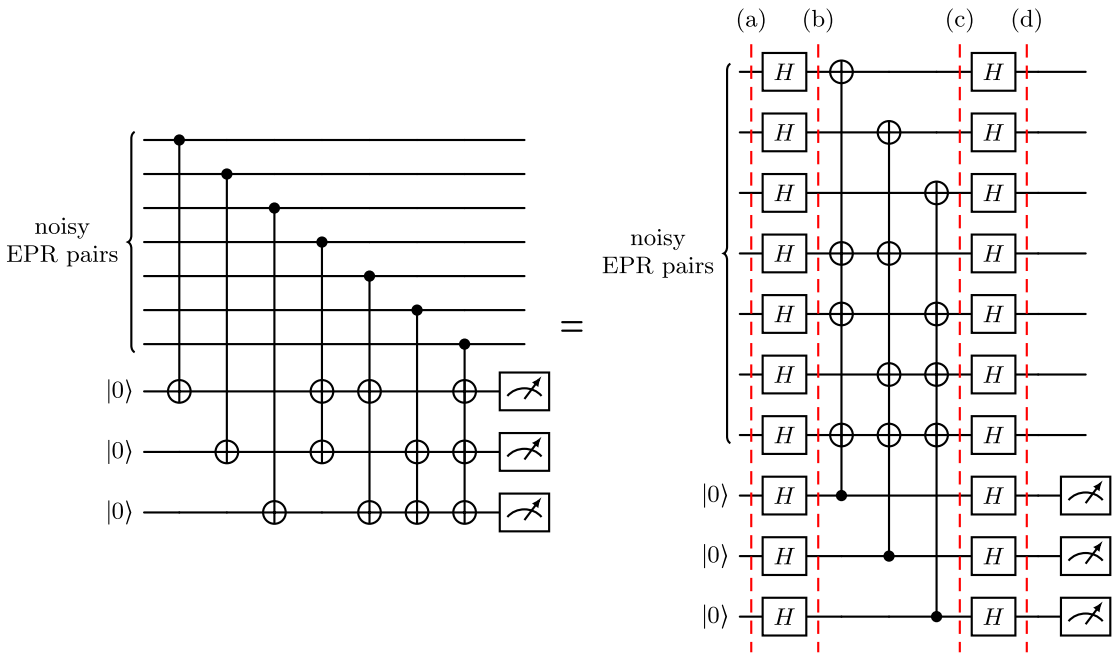

Figure 13. Analysis of the phase error pattern when $|00\rangle$s are employed as ancillary qubits, using the [7 4 3] code as an example. Left: Alice and Bob apply CNOT gates from noisy EPR pairs to ancillary qubits $|00\rangle$. After measuring the ancillary qubits and engaging in two-way communication (not depicted), they determine the bit error syndrome for this group. Right: Equivalence of the quantum circuit. (a) The phase error pattern is e_p. (b) After the Hadamard transform, the phase error pattern e_p transforms into the bit error pattern. (c) The $|00\rangle$ ancillary qubit pairs are the control qubits of CNOT gates to the EPR pairs. As discussed in the main text, $|00\rangle$ can be treated as an equal mixture of $|\psi_{00}\rangle$ and $|\psi_{01}\rangle$. The phase error e_p is converted to $\tilde{e}_p = e_p + H^T a$, where a is uniformly random on $\{0,1\}^{n-k}$. (d) The bit error pattern $\tilde{e}_p$ becomes the phase error pattern.

The process of hashing and measuring the bit error syndrome is depicted in figure 13. Here, $|00\rangle$ is the ancillary qubit. CNOT gates are applied between data qubits and the ancillary qubits, according to the hashing matrix H. Subsequently, the ancillary qubits are measured to extract parity information. This operation does not affect the bit error pattern. We transform the quantum circuit using Hadamard gates to analyze the effect on the phase error pattern. The equivalent circuit is shown in figure 13, where the ancillary qubits become the control qubits of CNOT gates. Each ancillary qubit pair, $|00\rangle$, is a superposition of Bell states: $|00\rangle = \frac{1}{\sqrt{2}}(|\psi_{00}\rangle + |\psi_{01}\rangle)$. Since the non-Bell-diagonal terms do not affect the distillation process (see appendix A), this pair can be treated as passing through a dephasing channel in the Bell basis, resulting in an equal mixture of Bell states $|\psi_{00}\rangle$ and $|\psi_{01}\rangle$. If the ancillary qubit pair is in the state $|\psi_{01}\rangle$, the CNOT gate will alter the phase error pattern. Define a bit string a of length $n-k$ with $a_i = 1$ if the ith ancillary qubits are $|\psi_{01}\rangle$. The bit string a follows a uniform distribution on $\{0,1\}^{n-k}$. This bit string will lead to a change in the phase error by the CNOT gate, as illustrated in figure 13. After measuring the ancilla qubits, the phase error pattern becomes $\tilde{e}_P = e_P + H^T a$. This relationship can be easily extended to an arbitrary linear hashing matrix, leading to the following lemma.

Lemma 3 (Transformation of the phase error pattern without OTP encryption). *Assuming Alice and Bob hold n noisy EPR pairs with the original phase error pattern denoted as e_p. They perform the same linear hash using an $(n-k)\times n$ parity check matrix H between the EPR pairs and ancillary qubits $|00\rangle^{\otimes(n-k)}$. EPR pairs serve as control qubits, while ancillary qubits act as target qubits. If the ancillary qubits are measured, and the results are publicly announced without OTP encryption, the phase error pattern transforms to:*

$$\tilde{e}_P = e_P + H^T a, \tag{32}$$

where a is uniformly distributed on $\{0,1\}^{n-k}$.

Proof. As discussed above, the ancillary qubits can be treated as an equal mixture of $|\psi_{00}\rangle$ and $|\psi_{01}\rangle$. These ancilla qubits can be represented as $\otimes_{i=1}^{n-k}|\psi_{0,a_i}\rangle$, where $a = (a_1, a_2, \ldots, a_{n-k})$ is uniformly distributed over $\{0,1\}^{n-k}$.

Suppose the initial phase error pattern is given by e_p and $H_{ij} = 1$. In this scenario, a bilateral CNOT operation is performed between the ith ancillary qubit pair and the jth noisy EPR pair.

According to equation (17), the bilateral CNOT operation updates the phase error pattern from e_p to $e_p + \hat{v}(j, a_i)$, where $\hat{v}(j, a_i)$ is a vector that has all entries set to zero except for position j, which is set to a_i.

After all the bilateral CNOT operations have been performed, the resulting phase error pattern is updated to $e_p + H^{\mathrm{T}}a$. □

A direct implication of equation (32) is that the key rate without OTP is at least as large as the case with OTP, making the use of OTP encryption for bit error correction unnecessary. This observation is summarized in the following lemma.

Lemma 4 (Unneccessity of OTP encryption). *Assuming Alice and Bob hold n noisy EPR pairs and perform bit error correction using a $(n-k)\times n$ linear hashing matrix H with ancillary qubits $|00\rangle^{\otimes(n-k)}$, the additional consumption of perfect EPR pairs to correct phase error is no more than $(n-k)$ compared to using $|\psi_{00}\rangle^{\otimes(n-k)}$ as ancillary qubits.*

Proof. The key idea is to track the ε-probable set of phase error patterns and show that it expands by at most a factor of 2^{n-k} after hashing. According to lemma 1, this implies that the additional consumption of secret keys for phase error correction in each group will not exceed $n-k$.

Specifically, let $\mathcal{T}$ denote the ε-probable set of phase error patterns before hashing. Consider the set

$$\mathcal{T}' = \left\{ e + H^{\top}a \mid e \in \mathcal{T},\ a \in \{0,1\}^{n-k} \right\}, \tag{33}$$

which satisfies $|\mathcal{T}'| \leqslant 2^{n-k}|\mathcal{T}|$. According to lemma 3, $\mathcal{T}'$ contains all possible phase error patterns after applying the hashing process, given that the original phase error pattern is in $\mathcal{T}$. From the definition of $\mathcal{T}$, it follows that

$$\Pr\left[e_p \in \mathcal{T}'\right] \geqslant 1-\varepsilon. \tag{34}$$

Let $\mathcal{T}''$ denote the ε-probable set of phase error patterns after hashing. Since $\mathcal{T}''$ is the smallest probable set, we have $|\mathcal{T}''| \leqslant |\mathcal{T}'|$. By lemma 1, the additional consumption of secret keys for phase error correction increases by at most

$$\log|\mathcal{T}''| - \log|\mathcal{T}| \leqslant n-k. \tag{35}$$

□

Note that $n-k$ represents the consumption for OTP encryption. Moreover, we can take H as the hashing matrix in information reconcilliation. Thus, we have proved that the additional consumption for phase error correction when performing bit error correction without OTP-whether during preprocessing or information reconciliation-does not surpass the consumption for OTP encryption. As mentioned earlier, phase error correction can be reduced to privacy amplification in QKD, where we do not need to explicitly know the phase error pattern but only need to account for the consumption required for phase error correction. Therefore, one can always perform information reconciliation without OTP and consume additional secret keys during phase error correction. That is, the key rate without using OTP is no less than the key rate with OTP and has the potential to be higher without OTP through careful analysis, especially for two-way preprocessing.

We now analyze the potential savings achievable in the scheme without OTP encryption through a detailed examination of the change in the phase error pattern. Multiplying G^{T} on both sides of equation (32) and utilizing the fact that $HG = G^{\mathrm{T}}H^{\mathrm{T}} = 0$, we obtain:

$$G^{\mathrm{T}}\tilde{e}_p = G^{\mathrm{T}}e_p. \tag{36}$$

This implies that, after two-way post-processing without OTP encryption, the phase error syndrome remains the same, while the probability of obtaining each phase error pattern within the same phase error syndrome becomes equal. Denoting the probability of obtaining bit error pattern $e^{j}_{b,i}$ and phase error pattern $e^{j'}_{p,i'}$ in this scenario as $\tilde{q}^{jj'}_{ii'}$, then:

$$\tilde{q}^{jj'}_{ii'} = \frac{q^{jj'}_{i}}{2^{n-k}}. \tag{37}$$

The consumption of phase error correction becomes

$$\begin{aligned} \tilde{I}^{j}_{p,i} &= h\left(\left\{\frac{\tilde{q}^{jj'}_{ii'}}{q^{j}_{i}}\right\}_{j',i'}\right) \\ &= h\left(\left\{q^{jj'}_{i}\right\}_{j'}\right) + (n-k). \end{aligned} \tag{38}$$

The total consumption of groups with bit error syndrome s^j becomes:

$$\begin{aligned}\tilde{R}^j &= I_b^j + \sum_i \frac{q_i^j}{q_j} \tilde{I}_{p,i}^j \\ &= h\left(\left\{\frac{q_i^j}{q^j}\right\}_i\right) + \sum_i \frac{q_i^j}{q^j} h\left(\left\{\frac{q_i^{jj'}}{q_i^j}\right\}_{j'}\right) + n - k \\ &= h\left(\left\{\frac{q_i^{jj'}}{q^j}\right\}_{j',i}\right) + n - k.\end{aligned} \tag{39}$$

Next, Alice and Bob discard the groups whose consumption exceeds the gain. By combining the above, we derive the following key rate formula.

Theorem 3 (Key rate formula for advantage distillation, without OTP encryption). *Suppose Alice and Bob perform two-way post-processing without one-time pad encryption in the advantage distillation framework using an [n k d] linear code. The probabilities of getting bit error syndrome s^j, bit error pattern e_i^j, phase error syndrome $s_P^{j'}$ and phase error pattern $e_{P,i'}^{j'}$ are given in table 1. Then, the key rate is*

$$\begin{aligned} r &= \sum_j q^j r^j, \\ r^j &= \max\left\{\frac{k}{n} - \frac{1}{n} h\left(\left\{\frac{q_i^{j,j'}}{q^j}\right\}_{i,j'}\right), 0\right\}. \end{aligned} \tag{40}$$

Now we compare equation (39) with previous consumption R^j, given in equation (28), plus the consumption of OTP, $O = n - k$. For the group with bit error syndrome s^j, the savings of EPR pairs in this group is

$$R^j + (n-k) - \tilde{R}^j = \sum_{i,j'} \frac{q_i^{jj'}}{q_i^j} h\left(\left\{\frac{q_{ii'}^{jj'}}{q_i^{jj'}}\right\}_{i'}\right). \tag{41}$$

This demonstrates that advantage distillation without OTP encryption results in a higher key rate than the scheme with OTP encryption in theorem 1 [48]. However, there is no rigorous guarantee omitting OTP encryption outperforms theorem 2.

lemma 4 can be directly applied to enhance the key rate in theorem 2. Specifically, if we compress the syndrome but do not perform OTP encryption during two-way classical communication, the key rate formula will outperform theorem 2. Watanabe *et al* [31] have previously shown this scheme based on the [2 1 2] code, deriving the key rate formula using a different method. We compare these approaches in section 5.2 and demonstrate that theorem 3 can outperform theirs in the high error rate regime. The key rate formula for a general error-correcting code involving error syndrome compression and omitting OTP encryption is left for future work.

4.3. Adding noise combined with classical linear code

Adding noise has been demonstrated to increase the threshold in QKD protocols [30, 44]. In this approach, Alice deliberately introduces noise to her sifted key. While increasing the key rate in the entanglement-based paradigm may seem challenging, a crucial observation is that a private state is sufficient for secret key generation [45]. By adding noise, Alice reduces the knowledge of Eve about the phase error pattern. Remarkably, it has been demonstrated that Alice and Bob do not need to correct all the phase error patterns, leading to savings in consumption during the phase error correction step [44, 49]. This method has proven effective in simulation results, particularly for large error rates. It can extend the tolerable error rate from 11% to 12.4% for the BB84 protocol and from 12.7% to 14.12% for the six-state protocol [30].

We integrate the adding noise method into the advantage distillation framework. The noise introduced in [30, 44] is *i.i.d.* for each noisy EPR pair. However, as the central idea of advantage distillation is identifying the bit error syndrome of each group of EPR pairs and subsequently implementing error correction and privacy amplification, the bit errors of different EPR pairs within a group are correlated. Consequently, introducing *i.i.d.* noise would alter the bit error syndrome of each group, which may be inappropriate for this scenario. To address this, a natural solution is to introduce structured noise to preserve the bit error syndrome for each group. Thus, Alice needs to add noise patterns to the entire group instead of introducing

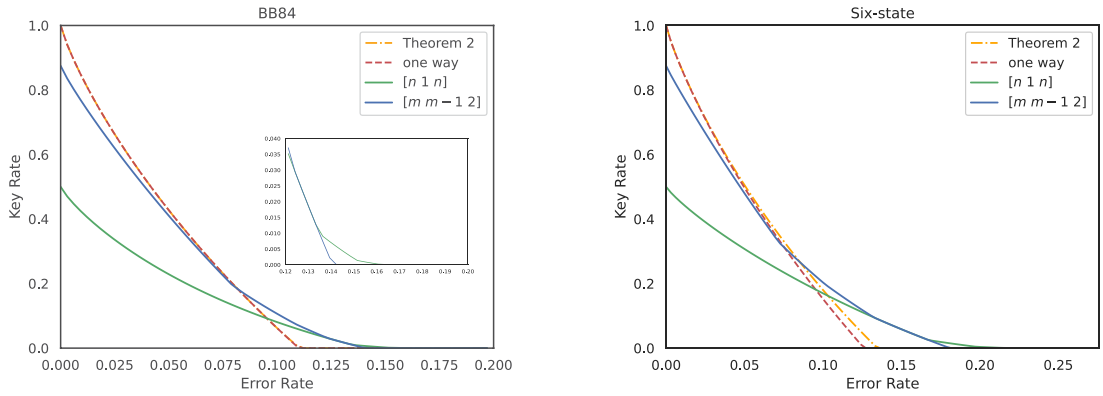

Figure 14. Comparison of the key rate when both OTP and hashing are applied (orange line) with the key rate when OTP is omitted (green line and blue line). The key rate is obtained by considering the best results from $[n\ 1\ n]$ codes and $[m\ m-1\ 2]$ codes, where n and m are less than 9. In the low error rate regime, hashing the syndrome with OTP encryption is more effective. Conversely, the scheme without OTP encryption performs better in the high bit error rate regime. In all regimes, our key rate surpasses the one-way protocol composed of information reconciliation and privacy amplification only.

noise to each EPR pair individually. Appendix E presents the detailed method for adding structured noise and subsequent analysis of the key rate based on private state distillation.

5. Case study

This section applies our advantage distillation framework based on classical linear code to several scenarios. Firstly, we demonstrate that the best-known result for tolerable error rate can be directly derived by combining the $[n\ 1\ n]$ code with theorem 3. Secondly, we enhance the state-of-the-art key rate in the high error rate regime by utilizing $[n\ 1\ n]$ codes and $[m\ m-1\ 2]$ codes.

5.1. Derive the threshold using $[n\ 1\ n]$ code

We show how our framework offers a systematic approach to determining the tolerable error rate of various QKD protocols. Previously, the Shor–Preskill formula [4] establishes an 11% threshold for the BB84 protocol, which is subsequently improved to 12.7% for the six-state protocol by Lo [40]. Gottesman and Lo further enhance the BB84 protocol to tolerate up to 17.9%, and the six-state protocol to 26.4%, by iteratively applying two-way post-processing in both bit error and phase error correction [28]. Chau improves the threshold to 20% for BB84 and 27.6% for the six-state protocol through two-way post-processing coupled with an adaptive privacy amplification method [29].

In our framework, the best known thresholds-20% for BB84 and 27.6% for the six-state protocol, as reported in [29]-can be achieved through a single round of two-way post-processing using an $[n, 1, n]$ code, as detailed in appendix D. While the possibility of further improvements remains open, we conjecture that the $[n\ 1\ n]$ code offers the best threshold within our advantage distillation framework based on classical linear code, owing to its maximum code distance. Potential explorations include trying adaptive coding methods in multiple rounds of preprocessing or combining with the adding noise method. These potential enhancements are left for future research.

5.2. Enhance the key rates using $[n\ 1\ n]$ and $[m\ m-1\ 2]$ codes

In this section, we apply our advantage distillation framework with $[n\ 1\ n]$ and $[m\ m-1\ 2]$ codes, where n and m are less than 9. The protocol steps are outlined in Box 1. The $[n\ 1\ n]$ code is a natural generalization of the B step [28, 43]. Taking the $[n\ 1\ n]$ code as an example, Alice and Bob first divide the raw key bits into groups, each containing n noisy raw key bits. Subsequently, they apply the $[n\ 1\ n]$ code to obtain the bit error syndrome, deciding whether to discard the group based on the key rate of the syndrome provided in theorem 3. Then, information reconciliation and privacy amplification are performed on the remaining raw key bits.

We compare the key rate when both OTP and hashing are applied (theorem 2) with the key rate when OTP is omitted (theorem 3), as illustrated in figure 14. The results indicate that hashing the syndrome with OTP encryption is more effective in the low error rate regime. This is reasonable because, in this regime, the entropy of the error syndrome is low, and only a few groups will be discarded. Thus, the advantage of discarding groups is limited. In contrast, the scheme without OTP performs better in the high error rate regime.

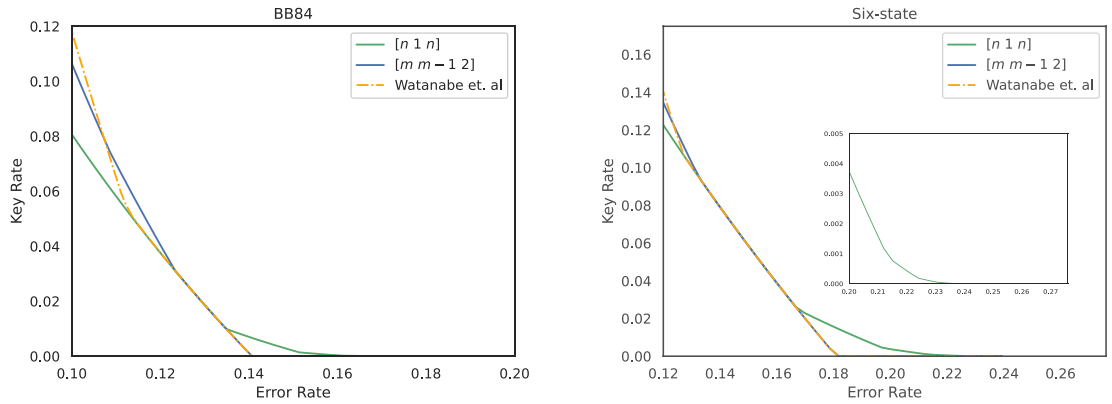

Figure 15. Comparison of the key rate of advantage distillation without OTP (green line and blue line) and the method in [31] (yellow dotted line). The key rate is obtained by considering the best results from $[n\ 1\ n]$ codes and $[m\ m-1\ 2]$ codes, where n and m are less than 9. The method in [31] is [2 1 2] code in advantage distillation framework combined with hashing the syndrome and without using OTP. The results demonstrate that $[m\ m-1\ 2]$ code (specifically, [3 2 2] code) and $[n\ 1\ n]$ codes for $n \geqslant 3$ without hashing can improve the key rate of [31] in the high error rate regime. The optimal codes depicted are given in table 2.

Table 2. Optimal codes for the BB84 protocol (left) and the six state protocol (right) in figure 15.

Error rate	Optimal code
$10.9\% \sim 12.3\%$	[3 2 2]
$13.5\% \sim 15.1\%$	[3 1 3]
$15.1\% \sim 16.1\%$	[4 1 4]
$16.1\% \sim 16.7\%$	[5 1 5]
$16.7\% \sim 17.1\%$	[6 1 6]

Error rate	Optimal code
$12.4\% \sim 13.2\%$	[3 2 2]
$16.8\% \sim 19.7\%$	[3 1 3]
$19.7\% \sim 21.3\%$	[4 1 4]
$21.3\% \sim 22.4\%$	[5 1 5]
$22.4\% \sim 23.1\%$	[6 1 6]

Next, we compare our method with previous work. As discussed in section 4.2, lemma 4 can be directly applied to enhance the key rate in theorem 2, showing that communication without OTP consistently outperforms communication with OTP. Watanabe *et al* [31] combined this approach with a [2 1 2] code, surpassing previous results and achieving state-of-the-art performance in low error rate regimes. Moreover, in high error rate regimes, the method in [25] exhibits currently the highest key rate, as far as we know. Note that the method in [25] is incorporated into the $[n\ 1\ n]$ codes: In the preprocessing step of these schemes, Alice chooses a block length n and uses a private random bit $c \in \{0,1\}$ to encrypt the raw key bits $(x_1, x_2, \ldots, x_n)$ in a group, resulting in $(x_1 \oplus c, x_2 \oplus c, \ldots, x_n \oplus c)$, which are sent to Bob. This is equivalent to sending $(x_1 \oplus c, x_1 \oplus x_2, x_2 \oplus x_3, \ldots, x_{n-1} \oplus x_n)$. Since $x_1 \oplus c$ is completely random to Bob, it carries no useful information, and the relevant message becomes $(x_1 \oplus x_2, x_2 \oplus x_3, \ldots, x_{n-1} \oplus x_n)$. This is exactly the tag s_x Alice sends in the $[n\ 1\ n]$ code without OTP for advantage distillation. Subsequently, Bob calculates the bit error syndrome $s_b = s_x \oplus s_y$ and checks whether $s_b = 0$ to determine whether to retain this group. Thus, the method in [25] corresponds to the $[n\ 1\ n]$ code with a specific syndrome $s_b = (0, 0, \cdots, 0)$.

Our method further enhances the outcomes presented by [31], particularly in the high error rate regime. Our result shows that the $[m\ m-1\ 2]$ code can enhance both methods in certain noise regimes (figure 15). Table 2 presents the optimal codes corresponding to different error rate regimes for the BB84 protocol and the six-state protocol, respectively.

It is worth noting that the method in [31] corresponds to a specific scenario in our framework: the [2 1 2] code with syndrome compression and without OTP encryption. Our results suggest that [3 2 2] code and $[n\ 1\ n]$ codes for $n \geqslant 3$ without hashing can improve the key rate compared to the [2 1 2] code combined with hashing. We anticipate further improvement in the key rate by compressing the syndrome without OTP, which has the potential to surpass the results of [31] in all error rate regimes by leveraging different error-correcting codes. We leave this for future work.

6. Conclusion and outlook

In this study, we have formulated a comprehensive framework for QKD advantage distillation, generalizing previous work and proposing a novel approach based on classical linear code. Our derivations encompass key rate formulas across various scenarios, highlighting that omitting OTP encryption consistently yields a higher key rate than OTP encryption scenarios. The practical application of our framework involves utilizing

$[n\ 1\ n]$ codes to derive the tolerable error rate, which aligns with the current state-of-the-art. Additionally, we employ $[n\ 1\ n]$ and $[m\ m-1\ 2]$ codes to enhance the key rate compared to prior findings systematically.

Moreover, we extend our investigation by integrating the adding noise method into our framework, resulting in an enhanced key rate formula. This collective effort contributes valuable insights into optimizing QKD protocols and advancing the field.

We anticipate implementing our advantage distillation method across a broader range of QKD protocols. Our method can be directly applied to protocols that utilize entanglement distillation-based security proofs, such as measurement-device-independent QKD protocols [8]. Additionally, for practical QKD systems that typically employ coherent states, our method can be integrated with the decoy-state technique to enhance performance in high-noise regimes [43, 50, 51], thereby extending the maximum secure transmission distance of practical QKD systems. Developing the advantage distillation framework with finite-size analysis is crucial for practical implementation and warrants further investigation. Finally, although this study focuses on discrete-variable QKD, there are many opportunities to apply our method to enhance the performance of continuous-variable QKD [52, 53].

There are numerous ways that are promising for further improving our methods. Firstly, while our exploration has focused on some specific codes, there is significant potential to investigate the efficiency of more sophisticated codes. By carefully designing codes, we may enhance threshold values and key rates beyond those achieved by existing protocols. Secondly, opportunities for improvement lie in synergies with additional methods, such as adaptive coding [29]. Thirdly, we envisage enhancing theorem 2 by considering scenarios where OTP encryption is not utilized, refining our understanding of key rates under varied conditions.

Furthermore, while figure 1 encapsulates most post-processing schemes, deviations from this structure are possible, such as altering the order of privacy amplification and information reconciliation [26]. Moreover, investigating the implications of our work on the relation between classical and quantum key distillation via CROs is an intriguing direction for future research [54]. By enhancing the performance of QKD through advantage distillation protocols, we anticipate advancing both the practicality and security of QKD systems, thereby fostering their increased adoption in real-world applications.

Data availability statement

No new data were created or analysed in this study.

Acknowledgments

The authors acknowledge Yizhi Huang and Pei Zeng for their helpful discussions. This work was supported by the National Natural Science Foundation of China Grants No. 12174216 and the Innovation Program for Quantum Science and Technology Grants No. 2021ZD0300804 and 2021ZD0300702.

Appendix A. Entanglement distillation for non-Bell-diagonal states

In the main text, we focus on entanglement distillation and key distillation protocols for Bell-diagonal noisy EPR states. In this section, we extend our analysis to non-Bell-diagonal states and demonstrate that the same protocols remain effective. Specifically, we prove that the entire entanglement distillation process depends solely on the Bell-diagonal components of the quantum state. The effectiveness of key distillation follows directly from our discussion in section 2.

For a two-qubit states ρ, which can be represented in the Bell basis as

$$\rho = \begin{pmatrix} p_{00} & \times & \times & \times \\ \times & p_{01} & \times & \times \\ \times & \times & p_{10} & \times \\ \times & \times & \times & p_{11} \end{pmatrix}, \tag{A1}$$

where the diagonal elements $p_{00}, p_{01}, p_{10}, p_{11}$ represent the probabilities associated with the respective Bell basis states. The off-diagonal elements, denoted by $\times$, capture the coherence between these states. The Bell-diagonal part of ρ, denoted by $\Delta(\rho)$, as the state obtained after dephasing ρ under Bell basis,

$$\Delta(\rho) = \begin{pmatrix} p_{00} & 0 & 0 & 0 \\ 0 & p_{01} & 0 & 0 \\ 0 & 0 & p_{10} & 0 \\ 0 & 0 & 0 & p_{11} \end{pmatrix}. \tag{A2}$$

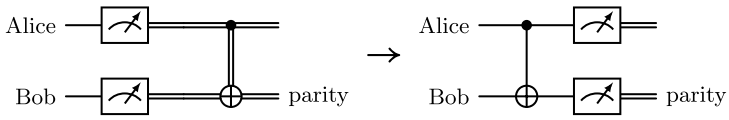

Figure A1. Parity check process. A classical parity check after computational basis measurements is equivalent to first applying a quantum CNOT operation and then measuring the second qubit. The measurement on the first qubit does not affect the distillation process and can be delayed until the entire entanglement distillation protocol succeeds.

Denote the entire state in the entanglement distillation process, which includes both data and ancillary qubits, as σ^N. The proof proceeds in two steps. Firstly, we show that the Bell-diagonal terms of σ^N after applying any unitary operation U within the entanglement distillation protocols in our work depend solely on the Bell-diagonal terms prior to the operation. Formally, this is expressed as

$$\Delta^{\otimes N}\left(U\sigma^N U^\dagger\right) = U\Delta^{\otimes N}\left(\sigma^N\right)U^\dagger. \tag{A3}$$

To establish this, observe that the unitary operations involved are limited to bilateral CNOT, $\mathrm{CNOT}_A \otimes \mathrm{CNOT}_B$, bilateral Hadamard gates $H_A \otimes H_B$, and X_B in error correction, which act as permutations on the Bell states [37]. Thus, the off-diagonal terms do not affect the diagonal terms, thereby proving equation (A3).

Secondly, we show that the parity comparison and post-selection steps in both two-way and one-way distillation only relate to the diagonal terms. This follows from the properties of classically replaceable operations [33]. Specifically, the parity comparison involves Alice and Bob performing Z-basis measurements on the noisy EPR pairs, followed by a classical XOR operation to determine the parity, as illustrated in figure A1.

This operation is equivalent to first performing a CNOT gate followed by a measurement in the computational basis. Notably, only the measurement result of the target qubit affects the subsequent distillation process, allowing the measurement of the source qubit to be delayed. The measurement on the target qubit involves a CNOT operation followed by a computational basis measurement, which is equivalent to measuring the observable Z_1Z_2. This observable is diagonal in the Bell basis. Let Π_0 denote the eigenspace of Z_1Z_2 corresponding to the $+1$ eigenvalue, and Π_1 denote the eigenspace corresponding to the -1 eigenvalue. The probability of obtaining outcome i is given by

$$\mathrm{tr}\left(\sigma^N\Pi_i\right) = \mathrm{tr}\left(\Delta^{\otimes N}\left(\sigma^N\right)\Pi_i\right). \tag{A4}$$

Furthermore, the post-measurement state of the remaining qubits retains the same Bell-diagonal terms:

$$\Delta^{\otimes N}\left(\mathrm{tr}_{AB}\left(\sigma^N\Pi_i\right)\right) = \mathrm{tr}_{AB}\left(\Delta^{\otimes N}\left(\sigma^N\right)\Pi_i\right), \tag{A5}$$

where AB represents the two-qubit system depicted in figure A1 for the parity check.

Combining the above two steps, we demonstrate that non-Bell-diagonal states consistently preserve their Bell-diagonal terms throughout the entire distillation process. The delayed measurement of the source qubit during the parity check does not affect the distillation outcome, as it has already succeeded. Consequently, both entanglement distillation and key distillation protocols remain effective for non-Bell-diagonal states.

Appendix B. Key rate formula for in-place hashing

In the original two-way classical post-processing, hashing is performed without ancilla qubits, constituting an in-place hashing method (see figure 10). Conversely, linear hashing is performed with ancillary qubits in our advantage distillation framework based on classical linear code. It is natural to ask whether this in-place hashing yields a higher key rate. Here, we derive the key rate formula for in-place hashing and demonstrate that both methods produce the same key rate without OTP encryption.

For convenience, we assume the parity check matrix H and generator matrix G are in the following canonical form:

$$\begin{aligned} H &= \begin{bmatrix} A & I_{n-k} \end{bmatrix}, \\ G &= \begin{bmatrix} I_k \\ A \end{bmatrix}, \end{aligned} \tag{B1}$$

where A is a $(n-k)\times k$ matrix. Suppose the bit error pattern is e_b and the phase error pattern is e_p. We write e_b and e_p as the concatenation of two string of length k and $n-k$:

$$\begin{aligned} e_b &= (e_{b,1}, e_{b,2}), \\ e_p &= (e_{p,1}, e_{p,2}). \end{aligned} \tag{B2}$$

Then, the bit error syndrome and phase error syndrome can be expressed as:

$$\begin{aligned} s_b &= Ae_{b,1} + e_{b,2}, \\ s_p &= e_{p,1} + A^T e_{p,2}. \end{aligned} \tag{B3}$$

Upon knowing s_b, there is a one-to-one correspondence between e_b and $e_{b,1}$. Therefore, the consumption of identifying bit error pattern $e_{b,1}$ for the previous k qubits is the same as identifying e_b, which is given in equation (25).

Next, we delve into the consumption of phase error correction. According to equation (32), the phase error pattern of previous k qubits is affected by the phase error pattern of later $n-k$ qubits by

$$\tilde{e}_{p,1} = e_{p,1} + A^{\mathrm{T}} e_{p,2}. \tag{B4}$$

Comparing with equation (B3), we have $\tilde{e}_{p,1} = s_p$. Consequently, the consumption of phase error correction aligns precisely with identifying the phase error syndrome:

$$I^j_{p,i} = h\left(\left\{\frac{q_i^{jj'}}{q_i^j}\right\}_{j'}\right). \tag{B5}$$

Combining the consumption of bit and phase error correction, we derive the consumption for the groups with bit error syndrome s^j:

$$\begin{aligned} R^j &= h\left(\left\{\frac{q_i^j}{q^j}\right\}_i\right) + \sum_i \frac{q_i^j}{q^j} h\left(\left\{\frac{q_i^{jj'}}{q_i^j}\right\}_{j'}\right) \\ &= h\left(\left\{\frac{q_i^{jj'}}{q^j}\right\}_{i,j'}\right). \end{aligned} \tag{B6}$$

Considering the probability of each bit error syndrome, we get the key rate formula for in-place hashing:

$$\begin{aligned} r &= \sum_j q^j r^j, \\ r^j &= \max\left(\frac{k}{n} - \frac{1}{n} h\left(\left\{\frac{q_i^{jj'}}{q^j}\right\}_{i,j'}\right), 0\right). \end{aligned} \tag{B7}$$

This formula is the same as the one presented in theorem 3.

Appendix C. Parity check with OTP

We introduce an alternative parity check method for information reconciliation. First, Alice and Bob hash their raw keys to obtain parities. Alice then compresses this parity information and securely transmits it to Bob using OTP, enabling the derivation of the bit error syndrome. Based on this syndrome, they proceed with further key distillation steps, which may include disclosing specific bit values and discarding parts of the raw keys. For instance, in the B-step scenario, if the error syndrome is 1, Alice and Bob each reveal one bit value while retaining the remaining bits for subsequent privacy amplification.

In a simplified scenario, we assume that either no raw key bits are announced, or the exact bit-error locations are known after announcement. Although partial key announcements and further reconciliation could be considered, this intermediate case is omitted here for simplicity. We derive the key rate formula for this method and demonstrate a marginal improvement over the approach detailed in theorem 2.

When OTP encryption is applied, the phase error in each qubit is independent. The conditional probability of a phase error, based on whether a bit error occurs, is

$$\begin{aligned}\delta_{p0} &= \frac{q_{01}}{q_{01}+q_{00}},\\ \delta_{p1} &= \frac{q_{11}}{q_{10}+q_{11}}.\end{aligned} \tag{C1}$$

The analysis of the error patterns and syndromes aligns with appendix B. Recall from equation (B2) that the bit error pattern is denoted as $e^j_{b,i} = (e_{b,1}, e_{b,2})^j_i$. Alice and Bob correct bit errors by measuring the first k qubits and announcing the results to obtain $e_{b,1}$, from which we can deduce e_b, given the one-to-one correspondence between e_b and $e_{b,1}$ upon knowing s_b. The raw key consumption in this announcement is given by

$$\tilde{I}^j_b = \frac{k}{n}. \tag{C2}$$

Upon knowing the bit error pattern $e^j_{b,i}$, Alice and Bob can determine the phase-error probability in each qubit. They can then correct the phase errors using lemma 1. Denote $l^j_i = |e_{b,2}|$, then the phase error consumption is given by

$$\tilde{I}^j_{p,i} = \frac{n-k-l^j_i}{n} h\left(\delta_{p0}\right) + \frac{l^j_i}{n} h\left(\delta_{p1}\right) \tag{C3}$$

and the key rate for bit error syndrome s^j_b is given by

$$\begin{aligned} r^j &= \max\left(1 - \tilde{I}^j_b - \sum_i \frac{q^j_i}{q^j}\tilde{I}^j_{p,i}, 0\right)\\ &= \max\left(\frac{n-k}{n} - \sum_i \frac{q^j_i}{q^j}\tilde{I}^j_{p,i}, 0\right). \end{aligned} \tag{C4}$$

Combining theorem 2, the key rate is obtained by

$$\begin{aligned} r &= \sum_j q^j r^j - \frac{h\left(\{q^j\}_j\right)}{n},\\ r^j &= \max\left\{1 - \frac{1}{n} h\left(\left\{\frac{q^{j,j'}_{i,i'}}{q^j}\right\}_{i,j',i'}\right), \frac{n-k}{n} - \sum_i \frac{q^j_i}{q^j}\tilde{I}^j_{p,i}, 0\right\}. \end{aligned} \tag{C5}$$

This key rate formula offers a slight improvement over theorem 2 [31, 43], though it does not surpass the results without OTP.

Appendix D. Tolerable error rate for the $[n, 1, n]$ code

In section 5.1, we discuss previous results on the tolerable error rate for BB84 and the six-state protocol. Here, we show how to derive these results using our framework combined with $[n\ 1\ n]$ code. For simplicity, we consider the symmetric case. In BB84, the bit error rate equals the phase error rate, i.e. $\delta_b = \delta_p = q$, and $\rho = \mathrm{diag}(1-2q+q_{11}, q-q_{11}, q-q_{11}, q_{11})$ with the free parameter q_{11}. For the six-state protocol, $\rho = (1-\frac{3q}{2}, \frac{q}{2}, \frac{q}{2}, \frac{q}{2})$. The conditional probability $\delta_{p,0}, \delta_{p,1}$ of a phase error, based on whether a bit error occurs, is given by equation (C1).

According to theorem 3, if a bit error syndrome e^j satisfies $r^j > 0$, then a positive key rate is achieved. For $[n\ 1\ n]$ code, $G = (1, 1, \cdots, 1)^{\mathrm{T}}$. We consider the bit error syndrome $s^0 = (0, 0, \cdots, 0)$ of length $n-1$. The bit error patterns corresponding to this syndrome are in the code space:

$$\begin{aligned} e^0_0 &= (0, 0, \cdots, 0),\\ e^0_1 &= (1, 1, \cdots, 1). \end{aligned} \tag{D1}$$

and the probabilities for each bit error pattern are

$$\begin{aligned} q^0_0 &= (1-q)^n,\\ q^0_1 &= q^n. \end{aligned} \tag{D2}$$

There are only two phase error syndromes; one corresponds to phase error patterns with odd parity, and the other corresponds to patterns with even parity. Given the bit error pattern e_0^0, the probability of getting even-parity phase error pattern is:

$$\begin{aligned}\frac{q_0^{00}}{q_0^0} &= \sum_{r \text{ is even}, 0\leqslant r\leqslant n} \binom{n}{r} \delta_{p0}^r \left(1-\delta_{p0}\right)^{n-r} \\ &= \sum_{r=0}^{n} \frac{1+(-1)^r}{2} \binom{n}{r} \delta_{p0}^r \left(1-\delta_{p0}\right)^{n-r} \\ &= \frac{1+\left(1-2\delta_{p0}\right)^n}{2}.\end{aligned} \tag{D3}$$

The last equality uses the fact that $(1+x)^m = \sum_{i=0}^{m} \binom{m}{i} x^i$. Similarly, we have

$$\frac{q_1^{00}}{q_1^0} = \frac{1+\left(1-2\delta_{p1}\right)^n}{2}. \tag{D4}$$

From theorem 3, we have

$$\begin{aligned}R^0 &= nr^0 \\ &= 1-h\left(\frac{q^n}{(1-q)^n+q^n}\right) - \frac{(1-q)^n}{(1-q)^n+q^n} h\left(\frac{1+\left(1-2\delta_{p0}\right)^n}{2}\right) - \frac{q^n}{(1-q)^n+q^n} h\left(\frac{1+\left(1-2\delta_{p1}\right)^n}{2}\right) \\ &\geqslant 1-h\left(\frac{q^n}{(1-q)^n+q^n}\right) - h\left(\frac{1}{2}+\frac{(1-3q+2q_{11})^n+|q-2q_{11}|^n}{2\left((1-q)^n+q^n\right)}\right),\end{aligned} \tag{D5}$$

where $h(x) = -x\log x - (1-x)\log(1-x)$. The last equation is from the concavity of h and the fact that for the positive secret key, we must have $q < \frac{1}{3}$.

For the BB84 protocol, the lower bound of R^0 can be attained by setting the free parameter q_{11} to zero, which minimizes the right-hand side of equation (D5) from basic calculus. Thus,

$$\begin{aligned}R^0 &\geqslant 1-h\left(\frac{q^n}{(1-q)^n+q^n}\right) - h\left(\frac{1}{2}+\frac{(1-3q)^n+q^n}{2\left((1-q)^n+q^n\right)}\right) \\ &= 1-h\left(\frac{x^n}{1+x^n}\right) - h\left(\frac{1}{2}+\frac{x^n+(1-2x)^n}{2\left(1+x^n\right)}\right),\end{aligned} \tag{D6}$$

where $x = \frac{q}{1-q}$. We have:

$$\begin{aligned}h\left(\frac{x^n}{1+x^n}\right) &= \log\left(1+x^n\right) - \frac{nx^n}{1+x^n}\log x \\ &= \frac{1}{\ln 2}x^n - \frac{nx^n}{1+x^n}\log x + O\left(x^{2n}\right), \\ h\left(\frac{1}{2}+\frac{x^n+(1-2x)^n}{2\left(1+x^n\right)}\right) &= 1-\frac{2}{\ln 2}\left(\frac{x^n+(1-2x)^n}{2\left(1+x^n\right)}\right)^2 + O\left((1-2x)^{3n}\right).\end{aligned} \tag{D7}$$

The first equation comes from direct calculation, and the second expands $h(x)$ at $x = \frac{1}{2}$. Thus,

$$R^0 = \frac{2}{\ln 2}\left(\frac{x^n+(1-2x)^n}{2\left(x^n+1\right)}\right)^2 - \frac{1}{\ln 2}x^n + \frac{nx^n}{1+x^n}\log x + O\left(x^{2n}+(1-2x)^{3n}\right). \tag{D8}$$

When $(1-2x)^2 > x$, the net gain of key R^0 becomes positive for sufficiently large n. This condition corresponds to $x < \frac{1}{4}$ or equivalently $q < \frac{1}{5}$. Therefore, the error rate threshold for the BB84 protocol in this scheme is 20%.

For the six-state protocol, $q_{11} = \frac{q}{2}$. From equation (D5), we have

$$\begin{aligned}R^0 &\geqslant 1-h\left(\frac{q^n}{(1-q)^n+q^n}\right) - h\left(\frac{1}{2}+\frac{(1-2q)^n}{2\left((1-q)^n+q^n\right)}\right) \\ &= 1-h\left(\frac{x^n}{1+x^n}\right) - h\left(\frac{1}{2}+\frac{(1-x)^n}{2\left(1+x^n\right)}\right).\end{aligned} \tag{D9}$$

And we expand h at $\frac{1}{2}$:

$$h\left(\frac{1}{2}+\frac{(1-x)^n}{2\left(1+x^n\right)}\right)=1-\frac{2}{\ln 2}\left(\frac{(1-x)^n}{2\left(1+x^n\right)}\right)^2+O\left((1-x)^{3n}\right). \tag{D10}$$

Similarly as in equation (D8), when $(1-x)^2>x$, $R^0>0$ for sufficiently large n. This condition corresponds to $x<\frac{3-\sqrt{5}}{2}$ or equivalently $q<\frac{3-\sqrt{5}}{5-\sqrt{5}}\approx 27.6\%$. Thus, the six-state protocol has an error rate threshold of 27.6% in this scheme.

Appendix E. Key rate formula for adding noise

Here, we derive the key rate formula for adding noise in the advantage distillation framework. After Alice and Bob identify the bit error syndrome s^j with OTP encryption, we purify the noisy EPR pairs in this group as

$$\sum_{i,j',i'}\sqrt{\frac{q_{ii'}^{jj'}}{q^j}}\mathbb{I}\otimes X_B^{e_i^j}Z_B^{e_{i'}^{j'}}|\psi_{00}^{\otimes n}\rangle_{AB}|e_i^j e_{i'}^{j'}\rangle_E, \tag{E1}$$

where the notation is given in table 1. As mentioned in section 2.3, Alice adds noise under the constraint that will not affect the bit error syndrome. Recall that $s^j=He_i^j$ and $HG=0$. Thus, the bit error pattern after adding noise should be in the form of

$$\tilde{e}_i^j=e_i^j+Gf, f\in\{0,1\}^k. \tag{E2}$$

Alice independently samples each bit of f with a probability p of being 1. The probability of sampling each f is $p_f=p^{|f|}(1-p)^{k-|f|}$. Alice applies a bit flip X_i if $(Gf)_i=1$. After adding noise, the state becomes:

$$\begin{aligned}
&\sum_{j,j',i',f}\sqrt{\frac{q_{ii'}^{jj'}}{q^j}}p_f X_A^{Gf}\otimes X_B^{e_i^j}Z_B^{e_{i'}^{j'}}|\psi_{00}^{\otimes n}\rangle_{AB}|f\rangle_{A'}|e_i^j e_{i'}^{j'}\rangle_E\\
&=\sum_{j,j',i',f}\sqrt{\frac{q_{ii'}^{jj'}}{q^j}}p_f\mathbb{I}\otimes X_B^{e_i^j}Z_B^{e_{i'}^{j'}}X_B^{Gf}|\psi_{00}^{\otimes n}\rangle_{AB}|f\rangle_{A'}|e_i^j e_{i'}^{j'}\rangle_E\\
&=\sum_{j,j',i',f}\sqrt{\frac{q_{ii'}^{jj'}}{q^j}}p_f\mathbb{I}\otimes X_B^{e_i^j+Gf}Z_B^{e_{i'}^{j'}}|\psi_{00}^{\otimes n}\rangle_{AB}Z^{G^T e_{i'}^{j'}}|f\rangle_{A'}|e_i^j e_{i'}^{j'}\rangle_E\\
&=\sum_{j,j',i',f}\sqrt{\frac{q_{ii'}^{jj'}}{q^j}}p_f\mathbb{I}\otimes X_B^{e_i^j+Gf}Z_B^{e_{i'}^{j'}}|\psi_{00}^{\otimes n}\rangle_{AB}Z^{s^{j'}}|f\rangle_{A'}|e_i^j e_{i'}^{j'}\rangle_E.
\end{aligned} \tag{E3}$$

The first equality comes from $X\otimes X|\psi_{00}\rangle=|\psi_{00}\rangle$. The second equality comes from commuting $Z_B^{e_{i'}^{j'}}$ and X_B^{Gf} produce $(-1)^{(e_{i'}^{j'})^T Gf}$, and this is equivalent to apply $Z^{G^T e_{i'}^{j'}}$ on $|f\rangle$. The third equality comes from $G^T e_{i'}^{j'}=s^{j'}$.

The consumption of bit error correction will change due to the change of the distribution on error pattern e_i^j. Denote j_f as the index such that $e_i^{j_f}=e_i^j+Gf$. The probability of getting $\tilde{e}_i^j$ is

$$\tilde{q}_i^j=\sum_f p_f q_i^{j_f}. \tag{E4}$$

The consumption of bit error correction becomes

$$\tilde{I}_e^j=h\left(\left\{\frac{\tilde{q}_i^j}{q^j}\right\}_i\right). \tag{E5}$$

Now, define $|\psi\rangle = \sqrt{(1-p)}|0\rangle + \sqrt{p}|1\rangle$ and $|\psi^{j'}\rangle = Z^{s^{j'}}|\psi\rangle^{\otimes n}$. After bit error correction, the state becomes

$$\begin{aligned}
&\sum_{j,j',i',f} \sqrt{\frac{q_{ii'}^{jj'}}{q^j} p_f} \mathbb{I} \otimes Z_B^{e_{i'}^{j'}} |\psi_{00}^{\otimes n}\rangle_{AB} Z^{s^{j'}} |f\rangle_{A'} |e_i^j + Gf\rangle_{B'} |e_i^j e_{i'}^{j'}\rangle_E \\
&= C_{A',B'} \sum_{j,j',i',f} \sqrt{\frac{q_{ii'}^{jj'}}{q^j} p_f} \mathbb{I} \otimes Z_B^{e_{i'}^{j'}} |\psi_{00}^{\otimes n}\rangle_{AB} Z^{s^{j'}} |f\rangle_{A'} |e_i^j\rangle_{B'} |e_i^j e_{i'}^{j'}\rangle_E \\
&= C_{A',B'} \sum_{j,j',i'} \sqrt{\frac{q_{ii'}^{jj'}}{q^j}} \mathbb{I} \otimes Z_B^{e_{i'}^{j'}} |\psi_{00}^{\otimes n}\rangle_{AB} Z^{s^{j'}} |\psi\rangle_{A'}^{\otimes n} |e_i^j\rangle_{B'} |e_i^j e_{i'}^{j'}\rangle_E \\
&= C_{A',B'} \sum_{j,j',i'} \sqrt{\frac{q_{ii'}^{jj'}}{q^j}} Z_A^{e_{i'}^{j'}} \otimes \mathbb{I} |\psi_{00}^{\otimes n}\rangle_{AB} |\psi^{j'}\rangle_{A'} |e_i^j\rangle_{B'} |e_i^j e_{i'}^{j'}\rangle_E,
\end{aligned} \tag{E6}$$

in which $C_{A'B'}|f\rangle_{A'}|e_i^j\rangle_{B'} = |f\rangle_{A'}|e_i^j + Gf\rangle_{B'}$.

Here comes the key observation of adding noise: If Alice corrects $Z^{e_{i'}^{j'}}$, then the state on subsystem $ABA'B'$ takes the form

$$C_{A'B'} \left[\psi_{00}^{\otimes n} \otimes \rho_{A'B'}\right] C_{A'B'}^{\mathrm{T}}, \tag{E7}$$

which is a private state that can generate secure keys [45]. Alice holds $|\psi^{j'}\rangle$ which contains information of $s^{j'}$. Thus, she can save on the consumption of phase error correction. If this saving surpasses the consumption of the increment in bit error correction, then adding noise can increase the key rate [44].

Now, we consider the consumption of phase error correction. After identifying e_i^j, we must identify phase error pattern $e_{i'}^{j'}$. In fact, with the help of $|\psi^{j'}\rangle$, the saving is $h(\sigma_i^j)$ [44, 49], in which

$$\sigma_i^j = \sum_{j'} \frac{q_i^{jj'}}{q_i^j} \psi^{j'}. \tag{E8}$$

Thus, the phase error consumption is

$$I_{p,i}^j = h\left(\left\{\frac{q_{i,i'}^{j,j'}}{q_i^j}\right\}_{j',i'}\right) - h\left(\sigma_i^j\right). \tag{E9}$$

According to equation (38), if OTP is not performed in the two-way communication step, the consumption is

$$\tilde{I}_{p,i}^j = n - k + h\left(\left\{\frac{q_i^{j,j'}}{q_i^j}\right\}_{j'}\right) - h\left(\sigma_i^j\right). \tag{E10}$$

Thus, the key rate formula for the groups with bit error syndrome s^j is

$$\begin{aligned}
r^j &= \max\left(1 - \frac{1}{n} h\left(\left\{\frac{\tilde{q}_i^j}{q^j}\right\}_i\right) - \frac{1}{n}\sum_i \frac{q_i^j}{q^j}\left[h\left(\left\{\frac{q_{i,i'}^{j,j'}}{q_i^j}\right\}_{j',i'}\right) - h\left(\sigma_i^j\right)\right], 0\right) - \frac{n-k}{n} \quad \text{(with OTP)}, \\
r^j &= \max\left(\frac{k}{n} - \frac{1}{n} h\left(\left\{\frac{\tilde{q}_i^j}{q^j}\right\}_i\right) - \frac{1}{n}\sum_i \frac{q_i^j}{q^j}\left[h\left(\left\{\frac{q_i^{j,j'}}{q_i^j}\right\}_{j'}\right) - h\left(\sigma_i^j\right)\right], 0\right) \quad \text{(without OTP)}.
\end{aligned} \tag{E11}$$

Note that the key rate of the scheme without OTP is always higher than with OTP. The final key rate is

$$r = \sum_j q^j r^j. \tag{E12}$$

ORCID iDs

Zhenyu Du https://orcid.org/0009-0006-5576-0834
Guoding Liu https://orcid.org/0000-0003-3288-1394
Xingjian Zhang https://orcid.org/0000-0003-0677-6996
Xiongfeng Ma https://orcid.org/0000-0002-9441-4006

References

[1] Bennett C H and Brassard G 1984 Quantum Cryptography: Public Key Distribution and Coin Tossing *Proceedings of the Ieee International Conference on Computers, Systems and Signal Processing* (IEEE Press) pp 175–9
[2] Ekert A K 1991 Quantum cryptography based on bell's theorem *Phys. Rev. Lett.* **67** 661
[3] Lo H-K and Chau H F 1999 Unconditional security of quantum key distribution over arbitrarily long distances *Science* **283** 2050
[4] Shor P W and Preskill J 2000 Simple proof of security of the BB84 quantum key distribution protocol *Phys. Rev. Lett.* **85** 441
[5] Hwang W-Y 2003 Quantum key distribution with high loss: toward global secure communication *Phys. Rev. Lett.* **91** 057901
[6] Lo H-K, Ma X and Chen K 2005 Decoy state quantum key distribution *Phys. Rev. Lett.* **94** 230504
[7] Wang X-B 2005 Beating the photon-number-splitting attack in practical quantum cryptography *Phys. Rev. Lett.* **94** 230503
[8] Lo H-K, Curty M and Qi B 2012 Measurement-device-independent quantum key distribution *Phys. Rev. Lett.* **108** 130503
[9] Lucamarini M, Yuan Z L, Dynes J F and Shields A J 2018 Overcoming the rate-distance limit of quantum key distribution without quantum repeaters *Nature* **557** 400–3
[10] Ma X, Zeng P and Zhou H 2018 Phase-matching quantum key distribution *Phys. Rev.* X **8** 031043
[11] Zeng P, Zhou H, Wu W and Ma X 2022 Mode-pairing quantum key distribution *Nat. Commun.* **13** 3903
[12] Xu F, Ma X, Zhang Q, Lo H-K and Pan J-W 2020 Secure quantum key distribution with realistic devices *Rev. Mod. Phys.* **92** 025002
[13] Liao S-K *et al* 2017 Satellite-to-ground quantum key distribution *Nature* **549** 43–47
[14] Liao S-K *et al* 2018 Satellite-relayed intercontinental quantum network *Phys. Rev. Lett.* **120** 030501
[15] Peev M *et al* 2009 The secoqc quantum key distribution network in vienna *New J. Phys.* **11** 075001
[16] Sasaki M *et al* 2011 Field test of quantum key distribution in the tokyo QKD network *Opt. Express* **19** 10387
[17] Razavi M 2012 Multiple-access quantum key distribution networks *IEEE Trans. Commun.* **60** 3071
[18] Chen Y-A *et al* 2021 An integrated space-to-ground quantum communication network over 4,600 kilometres *Nature* **589** 214
[19] Chen T-Y *et al* 2021 Implementation of a 46-node quantum metropolitan area network *npj Quantum Inf.* **7** 134
[20] Zhu H-T *et al* 2024 Field test of mode-pairing quantum key distribution *Optica* **11** 883
[21] Fung C-H F, Ma X and Chau H F 2010 Practical issues in quantum-key-distribution postprocessing *Phys. Rev.* A **81** 012318
[22] Shannon C E 1948 A mathematical theory of communication *Bell Syst. Tech. J.* **27** 379
[23] Koashi M, 2005 Simple security proof of quantum key distribution via uncertainty principle (arXiv:quant-ph/0505108 [quant-ph])
[24] Devetak I and Winter A 2005 Distillation of secret key and entanglement from quantum states *Proc. R. Soc.* A **461** 207–35
[25] Renner R 2008 Security of quantum key distribution *Int. J. Quantum Inf.* **06** 1–127
[26] Huang Y, Zhang X and Ma X 2022 Stream privacy amplification for quantum cryptography *PRX Quantum* **3** 020353
[27] Maurer U 1993 Secret key agreement by public discussion from common information *IEEE Trans. Inf. Theory* **39** 733
[28] Gottesman D and Lo H-K 2003 Proof of security of quantum key distribution with two-way classical communications *IEEE Trans. Inf. Theory* **49** 457
[29] Chau H F 2002 Practical scheme to share a secret key through a quantum channel with a 27.6% bit error rate *Phys. Rev.* A **66** 060302
[30] Kraus B, Gisin N and Renner R 2005 Lower and upper bounds on the secret-key rate for quantum key distribution protocols using one-way classical communication *Phys. Rev. Lett.* **95** 080501
[31] Watanabe S, Matsumoto R, Uyematsu T and Kawano Y 2007 Key rate of quantum key distribution with hashed two-way classical communication *Phys. Rev.* A **76** 032312
[32] Nielsen M A and Chuang I 2002 Quantum computation and quantum information *Am. J. Phys.* **70** 558
[33] Liu G, Zhang X and Ma X 2022 Classically Replaceable Operations *Quantum* **6** 845
[34] Cover T M 1999 *Elements of Information Theory* (Wiley) (https://doi.org/10.1002/047174882X)
[35] Carter J and Wegman M N 1979 universal classes of hash functions *J. Comput. Syst. Sci.* **18** 143
[36] Wegman M N and Carter J 1981 New hash functions and their use in authentication and set equality *J. Comput. Syst. Sci.* **22** 265
[37] Bennett C H, DiVincenzo D P, Smolin J A and Wootters W K 1996 Mixed-state entanglement and quantum error correction *Phys. Rev.* A **54** 3824
[38] Renner R and Cirac J I 2009 de finetti representation theorem for infinite-dimensional quantum systems and applications to quantum cryptography *Phys. Rev. Lett.* **102** 110504
[39] Lo H-K 2003 Method for decoupling error correction from privacy amplification *New J. Phys.* **5** 36
[40] Lo H-K 2001 Proof of unconditional security of six-state quatum key distribution scheme *Quantum Info. Comput.* **1** 81–94
[41] Ben-Or M, Horodecki M, Leung D, Mayers D and Oppenheim J 2005 The Universal Composable Security of Quantum Key Distribution *Theory of Cryptography* (Springe) pp 386–406
[42] Renner R and König R 2005 Universally composable privacy amplification against quantum adversaries *Theory of Cryptography* (Springer) pp 407–25
[43] Ma X, Fung C-H F, Dupuis F, Chen K, Tamaki K and Lo H-K 2006 Decoy-state quantum key distribution with two-way classical postprocessing *Phys. Rev.* A **74** 032330
[44] Renes J M and Smith G 2007 Noisy processing and distillation of private quantum states *Phys. Rev. Lett.* **98** 020502
[45] Horodecki K, Horodecki M, Horodecki P and Oppenheim J 2005 Secure key from bound entanglement *Phys. Rev. Lett.* **94** 160502
[46] Some non-linear operations, such as the CCNOT (figure 3), are also CROs. Thus, there is a possibility that non-linear codes can be incorporated into our framework to further improve the key rate and the allowable bit error threshold.
[47] Gottesman D, Lo H-K, Lutkenhaus N and Preskill J 2004 Security of quantum key distribution with imperfect devices *International Symposium Oninformation Theory, 2004. Isit 2004. Proceedings.* (https://doi.org/10.1109/isit.2004.1365172)
[48] Lemma 4 also applies to the information reconciliation step. However, as the information reconciliation corrects the bit error pattern in all locations, the corresponding hashing is global. We conjecture that there is little room for improving the key rate by refraining from using OTP in this step.

[49] Hausladen P, Jozsa R, Schumacher B, Westmoreland M and Wootters W K 1996 Classical information capacity of a quantum channel *Phys. Rev.* A **54** 1869
[50] Li H-W, Zhang C-M, Jiang M-S and Cai Q-Y 2022 Improving the performance of practical decoy-state quantum key distribution with advantage distillation technology *Commun. Phys.* **5** 53
[51] Huang Y, Du Z and Ma X 2023 Source–replacement model for phase–matching quantum key distribution *Adv. Quantum Technol.* **7** 2300275
[52] Jouguet P, Kunz-Jacques S, Leverrier A, Grangier P and Diamanti E 2013 Experimental demonstration of long-distance continuous-variable quantum key distribution *Nat. Photon.* **7** 378–81
[53] Qi B, Lougovski P, Pooser R, Grice W and Bobrek M 2015 Generating the local oscillator 'locally' in continuous-variable quantum key distribution based on coherent detection *Phys. Rev.* X **5** 041009
[54] Christandl M, Ekert A, Horodecki M, Horodecki P, Oppenheim J and Renner R 2007 Unifying classical and quantum key distillation (arXiv:quant-ph/0608199 [quant-ph])