



**Efficient entanglement distillation for quantum channels with polarization mode dispersion**Liangzhong Ruan <sup>1</sup>, Brian T. Kirby,<sup>2</sup> Michael Brodsky,<sup>2</sup> and Moe Z. Win <sup>1,\*</sup><sup>1</sup>*Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA*<sup>2</sup>*U.S. Army Research Laboratory, Adelphi, Maryland 20783, USA*

(Received 4 October 2019; accepted 15 June 2020; published 19 March 2021)

Quantum entanglement shared by remote network nodes serves as a valuable resource for promising applications in distributed computing, cryptography, and sensing. However, distributing high-quality entanglement via fiber-optic routes could be challenging due to the various decoherence mechanisms in fibers. In particular, one of the primary polarization decoherence mechanisms in optical fibers is polarization mode dispersion (PMD), which is the distortion of optical pulses by randomly varying birefringences. To mitigate the effect of decoherence in entangled particles, quantum entanglement distillation (QED) algorithms have been proposed. One particular class, the recurrence QED algorithms, stands out because it has relatively relaxed requirements both on the size of the quantum circuits involved and on the initial quality of entanglement between particles. However, because the number of required particles grows exponentially with the number of distillation rounds, an efficient recurrence algorithm needs to converge quickly. We present a recurrence QED algorithm designed for photonic qubit pairs affected by PMD-degraded channels. Our proposed algorithm achieves the optimal fidelity as well as the optimal success probability (conditioned on the fact that optimal fidelity is achieved) in every round of distillation. The attainment of the maximal fidelity improves the convergence speed of fidelity with respect to the number of distillation rounds from linear to quadratic and, hence, significantly reduces the number of rounds. Combined with the fact that the optimal success probability is achieved, the proposed algorithm provides an efficient method to distribute entangled states with a high fidelity via optical fibers.

DOI: [10.1103/PhysRevA.103.032425](https://doi.org/10.1103/PhysRevA.103.032425)**I. INTRODUCTION**

Entanglement shared among quantum network nodes is the source of quantum advantage [1–5] for many applications, including teleportation [6–8], dense coding [9–11], quantum key distribution [12–14], and quantum information relay [15–17]. In quantum networks with more than two nodes, entanglement can also be employed to reduce the queuing delay of quantum data [18] or achieve quantum broadcasting [19]. For the task of distributing entanglement in quantum networks, the fiber-optic infrastructure is a natural candidate. In this context, polarization-entangled photon pairs [20] are particularly useful because of the ease with which light polarization can be manipulated using standard instrumentation [21] and the numerous sources of polarization-entangled photons suitable for use with standard fibers [22]. For polarization-entangled photons, the major decoherence mechanism is birefringence [23–25]. The accumulation of randomly varying birefringence in fibers leads to a phenomenon known as polarization mode dispersion (PMD) [26].

To mitigate the effect of decoherence mechanisms on entangled qubit pairs, quantum entanglement distillation (QED) algorithms [27–30] have been proposed to generate qubit pairs in the target entangled state using local operations and classical communication (LOCC). Since high-quality entanglement is the keystone in many important applications of

quantum computation and quantum information, QED has become an essential building block for the development of quantum networks [31,32].

In the literature, three types of QED algorithms have been proposed, namely, asymptotic [33–35], code-based [36–38], and recurrence [39–41] algorithms. Among the three types of algorithms, the recurrence ones require local operations on just one or two qubits and are robust against severe decoherence. The recurrence algorithms operate on two qubit pairs each time, improving the quality of entanglement in one pair at the expense of the other pair, which is then discarded. The algorithms keep repeating this operation to progressively increase the fidelity of the kept qubit pairs with respect to (w.r.t.) the target entangled state. These algorithms can mitigate the effect of strong decoherence by performing multiple distillation rounds. For instance, the recurrence algorithm proposed in [27] can distill partially decoherent qubit pairs into maximally entangled qubit pairs as long as the initial fidelity of the contaminated qubit pairs w.r.t. the target state is greater than 0.5. To summarize, recurrence QED algorithms are preferable in terms of both implementability and robustness. Proof-of-principle experimental demonstrations of these algorithms [42,43] single out their importance in the near-term development of quantum networks.

Despite their advantages, recurrence QED algorithms do have a drawback in terms of efficiency. The efficiency of QED algorithms is measured in terms of *yield*, which is defined as the ratio between the number of highly entangled output qubit pairs and the number of input qubit pairs impaired by

\*Corresponding author: moewin@mit.edu

decoherence effects. Since at least half of the entangled qubit pairs are discarded in each round of distillation, the efficiency of the recurrence algorithms decreases exponentially with the number of rounds. To reduce the required number of distillation rounds, the LOCC adopted in the algorithms need to be designed so that the fidelity of the kept qubit pairs quickly approaches 1 w.r.t. the number of distillation rounds. To achieve this objective, the quantum privacy amplification (QPA) algorithm was proposed in [39] and was shown numerically to require a smaller number of distillation rounds than the algorithm in [27] for qubit pairs impaired by a quantum depolarizing channel. However, the performance of the QPA algorithm was not characterized analytically. In fact, a set of initial states was found in [40] for which the QPA algorithm was less efficient than the algorithm in [27]. In [40], the design of distillation operations was formulated as an optimization problem, which was inherently nonconvex, and consequently, the optimal solution was not found. In [30], an algorithm is designed to numerically upper bound the output fidelity and successful probability of single-round distillation, but the achievability of these bounds remains unknown. Therefore, the issue of improving the efficiency of recurrence QED algorithms remains an interesting challenge.

In this work, we develop an efficient recurrence QED algorithm for entangled photons impaired by the PMD effect. We envision that a key enabler for designing efficient recurrence QED algorithms is to make them adaptive to the key parameters of PMD. Intuitively, compared to generic algorithms, QED algorithms that adapt to channel-specific decoherence effects will better mitigate such effects and hence distill more efficiently. In fact, it has been observed that knowing the channel benefits the performance of quantum error recovery [44], and efficient channel-adaptive quantum error correction schemes [45,46] have been designed. In the context of QED, adaptive recurrence QED algorithm has been designed for channels with two Kraus operators to improve the convergence speed of fidelity w.r.t. the number of distillation rounds [47]. This work optimizes the distillation operations to most efficiently mitigate the effect of PMD while achieving a high success probability.

The rest of this work is organized as follows. Section II analyzes the PMD effect on photon pairs and formulates the optimization problems for recurrence QED algorithms. Section III characterizes the optimal values of these problems, i.e., the maximal output fidelity and the highest success probability, and then designs a recurrence QED algorithm that achieves the characterized optimal values. Section IV provides several numerical tests, showing that by achieving the optimal fidelity and success probability in each round of distillation, the proposed algorithm provides an efficient method to distribute entangled photons with a high fidelity through quantum channels impaired by fiber birefringence. Finally, Sec. V gives the conclusion.

*Notation.*  $a$ ,  $\mathbf{a}$ , and  $\mathbf{A}$  represent scalar, vector, and matrices, respectively.  $\angle$  and  $(\cdot)^*$  denote the phase and conjugate of a complex number, respectively.  $(\cdot)^\dagger$ ,  $\text{rank}\{\cdot\}$ ,  $\det\{\cdot\}$ , and  $\text{tr}\{\cdot\}$  denote the Hermitian transpose, rank, determinant, and trace of a matrix, respectively.  $\text{tr}_{i,j}\{\cdot\}$  denotes the partial trace w.r.t. the  $i$ th and  $j$ th qubits in the system.  $\propto$  denotes the proportional relationship.  $\mathbf{I}_n$  denotes the  $n \times n$  identity matrix, and  $i$  is the unit imaginary number.

## II. SYSTEM MODEL AND PROBLEM FORMULATION

This section presents the system model and then formulates the optimization problems for recurrence QED algorithms.

### A. Effect of PMD on entangled photon pairs

Consider the quantum network illustrated in Fig. 1(a), in which a photon source is connected to two network nodes, i.e., Alice and Bob, via PMD-degraded optical fibers. In the literature, the PMD effect is often modeled using the first-order approximation [23,24], which characterizes the PMD effect as splitting one incident pulse into two orthogonally polarized components delayed relative to each other. As illustrated in Fig. 1(b), the polarization states of these two components are known as the principal states of polarization (PSP) basis  $\{|s_i\rangle, |s'_i\rangle, i \in \{A, B\}\}$ , and the delay between the two components is called the differential group delay (DGD)  $\tau_A, \tau_B$ . Since typical time constants characterizing the decorrelation of PMD in buried optical fibers are as long as days and sometimes months [48], PMD evolution can be considered adiabatic in the context of quantum communication protocols. Therefore, it is reasonable to assume that the parameters of the PMD effect, particularly the PSP basis  $\{|s_A\rangle, |s'_A\rangle, |s_B\rangle, |s'_B\rangle\}$ , can be measured by the network nodes.

Due to the effect of PMD, the density matrix  $\mathfrak{E}$  of the photon pair after passing through fibers is given by (1). This density matrix is written in the ordered basis  $\{|s_A s_B\rangle, |s'_A s_B\rangle, |s_A s'_B\rangle, |s'_A s'_B\rangle\}$ . Please refer to Appendix A for the detailed derivation and the definition of the parameters in (1), i.e.,  $\eta_1, \eta_2, \alpha$ , and the function  $R(\cdot, \cdot)$ . Denote the element in the  $p$ th row and  $q$ th column of  $\mathfrak{E}$  as  $\xi_{pq}$ .

As illustrated in Fig. 1(b) and (A5), with generic PSP, the PMD effect in the two arms leads to four possible coincident arrival times for the two photons, i.e., slow-slow ( $|s_A s_B\rangle$ ), fast-slow ( $|s'_A s_B\rangle$ ), slow-fast ( $|s_A s'_B\rangle$ ), and fast-fast ( $|s'_A s'_B\rangle$ ). This results in a relatively complicated density matrix. As illustrated in Fig. 1(c), to simplify the density matrix, one could align the PSP basis with the photon polarization basis, so that there are only two possible coincident arrival times, i.e., slow-slow and fast-fast. The physical realization of this operation requires a measurement of the PSP for a given fiber and the ability

$$\mathfrak{E} = \frac{1}{2} \begin{bmatrix} |\eta_1|^2 & -\eta_1 \eta_2 e^{-i\alpha} R^*(\tau_A, 0) & \eta_1 \eta_2^* R^*(0, \tau_B) & \eta_1^2 e^{-i\alpha} R^*(\tau_A, \tau_B) \\ -\eta_1^* \eta_2^* e^{i\alpha} R(\tau_A, 0) & |\eta_2|^2 & -(\eta_2^*)^2 e^{i\alpha} R(-\tau_A, \tau_B) & -\eta_1 \eta_2^* R^*(0, \tau_B) \\ \eta_1^* \eta_2 R(0, \tau_B) & -(\eta_2)^2 e^{-i\alpha} R(-\tau_A, \tau_B) & |\eta_2|^2 & \eta_1 \eta_2 e^{-i\alpha} R^*(\tau_A, 0) \\ (\eta_1^*)^2 e^{i\alpha} R(\tau_A, \tau_B) & -\eta_1^* \eta_2 R(0, \tau_B) & \eta_1^* \eta_2^* e^{i\alpha} R(\tau_A, 0) & |\eta_1|^2 \end{bmatrix} \quad (1)$$

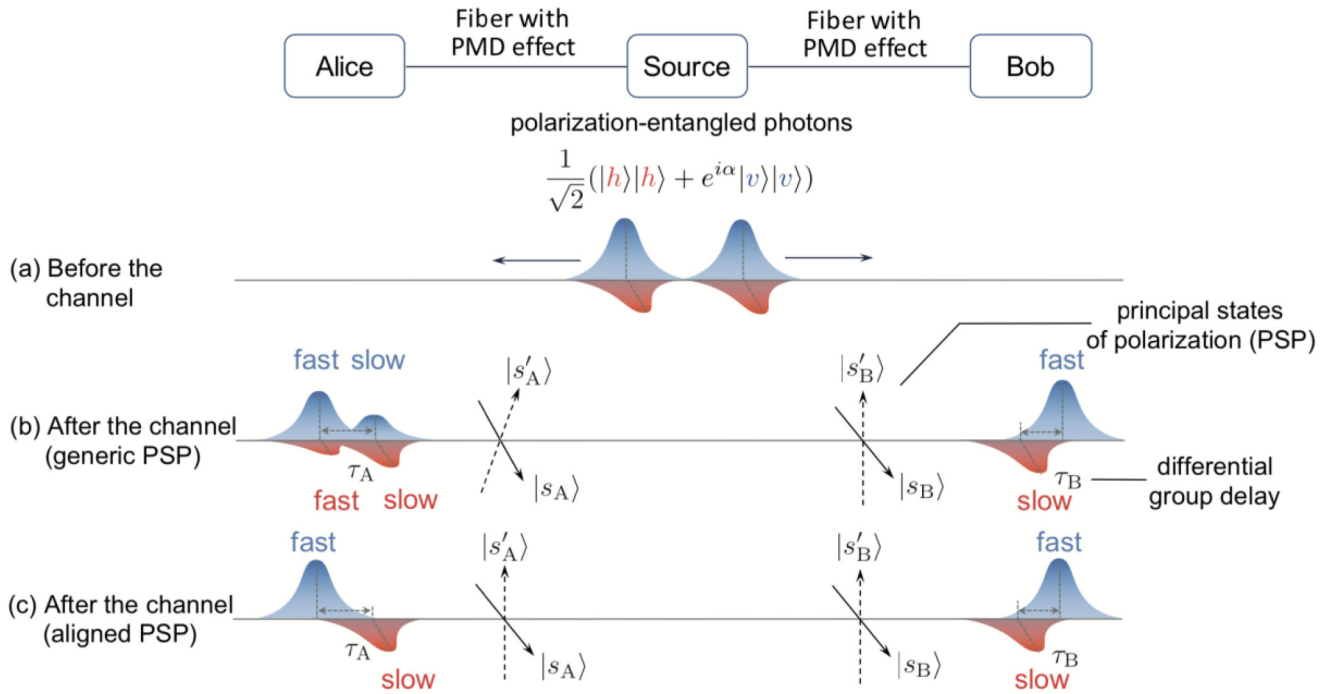


FIG. 1. System model. The overall effect of PMD resembles that of pure birefringence in the sense that it causes an incident pulse to split into two orthogonally polarized components delayed relative to each other [26]. The polarization states of these two components are known as the PSP and the delay between them is called the DGD. Appendix B shows that even with generic PSP, a maximally entangled polarization state prepared by the source can be viewed as if the polarization basis of one of the photons is already aligned with the PSP basis of the channel. Hence, in this figure, the polarization basis of photon B is always aligned with the PSP basis of the channel.

to perform local rotation on the photons before they pass through the fiber. As Appendix B shows, local rotation on one of the photons is sufficient to achieve the alignment of the PSP basis with the photon polarization basis. Existing studies suggest that realignment of these states would be rare, as the PSP in installed optical fibers can remain unchanged for as long as months [48]. In fact, the operation of aligning PSP has also been adopted in the algorithm design for PMD compensation [24] to exploit the advantage of the decoherence-free subspace [23].

When the PSP basis is aligned with the polarization basis, we get  $\eta_1 = 1$  and  $\eta_2 = 0$ . Hence, the density matrix, (1), is simplified to a matrix with four nonzero elements, which are given by

$$\xi_{11} = \xi_{44} = \frac{1}{2}, \quad \xi_{41} = \xi_{14}^* = \frac{1}{2}e^{i\alpha}R(\tau_A, \tau_B),$$

which can be rewritten as

$$\begin{aligned} \mathcal{E} = & \frac{1}{2}(|s_A s_B\rangle\langle s_A s_B| + e^{-i\alpha}R^*(\tau_A, \tau_B)|s_A s_B\rangle\langle s'_A s'_B| \\ & + e^{i\alpha}R(\tau_A, \tau_B)|s'_A s'_B\rangle\langle s_A s_B| + |s'_A s'_B\rangle\langle s'_A s'_B|). \end{aligned} \quad (2)$$

From (2), it can be seen that when the PSP and polarization basis are aligned, the PMD effect is equivalent to a composition of phase-shift and phase-damping channels.

### B. Problem formulation

The network nodes Alice and Bob adopt a recurrence QED algorithm to mitigate the effect of PMD. They operate separately on every two qubit pairs, trying to improve the quality of entanglement in one pair at the expense of the

other pair. This distillation operation  $\mathbb{D}$  can be formulated as follows. Denote the density matrix of a kept qubit pair after the  $k$ th round of distillation as  $\mathcal{E}_k$ , with  $\mathcal{E}_0 = \mathcal{E}$ . Then before the  $k$ th round of distillation, the joint density matrix of two qubit pairs is given by

$$\mathcal{E}_{k-1}^J = \mathcal{E}_{k-1} \otimes \mathcal{E}_{k-1}.$$

Without loss of generality, assume that the network nodes try to keep the first qubit pair. Then the density matrix of the first qubit pair after the distillation operation is given by the partial trace over the third and fourth qubits normalized by the overall trace of the density matrix, i.e.,

$$\mathcal{E}_k = \frac{\text{tr}_{3,4}\{\mathbb{D}\{\mathcal{E}_{k-1}^J\}\}}{\text{tr}\{\mathbb{D}\{\mathcal{E}_{k-1}^J\}\}}, \quad (3)$$

where the distillation operation  $\mathbb{D}$  must be in the class of LOCC, and the probability of successfully keeping the first qubit pair is given by

$$P_k = \text{tr}\{\mathbb{D}\{\mathcal{E}_{k-1}^J\}\}. \quad (4)$$

The fidelity of the kept qubit pairs after the  $k$ th round of distillation w.r.t. the target state is

$$F_k = \langle \phi^+ | \mathcal{E}_k | \phi^+ \rangle, \quad (5)$$

where  $|\phi^+\rangle = \frac{1}{\sqrt{2}}(|h_A h_B\rangle + |v_A v_B\rangle)$ . For notation convenience, define the mapping between the input density matrix  $\mathcal{E}_{k-1}$  and the fidelity of the kept qubit pair  $F_k$  as  $F_{\mathbb{D}}$ , i.e.,

$$F_k = F_{\mathbb{D}}(\mathcal{E}_{k-1}),$$

and denote the mapping between the input density matrix  $\mathfrak{E}_{k-1}$  and the success probability  $P_k$  as  $P_{\mathbb{D}}$ , i.e.,

$$P_k = P_{\mathbb{D}}(\mathfrak{E}_{k-1}).$$

Note that both mappings depend on the distillation operation  $\mathbb{D}$ .

The objective of recurrence QED algorithms is to generate qubit pairs with a sufficiently high fidelity, i.e.,

$$F_K \geq 1 - \epsilon, \quad (6)$$

for some natural number  $K$  and small  $\epsilon > 0$ . With this recurrence QED algorithm, the yield of the algorithm after  $K$  rounds of distillation is given by

$$Y_K = \prod_{k=1}^K \frac{P_k}{2}. \quad (7)$$

It can be seen from (7) that the yield of the algorithm drops by at least half with one more round of distillation. Hence, to improve the yield of the QED algorithm, a primary task is to minimize the required number of distillation rounds, i.e., maximize  $F_k$  at each round. Meanwhile, the success probability  $P_k$  also affects  $Y_K$ . Hence, a secondary task is to maximize  $P_k$  conditional on  $F_k$  being maximized. The problems of fulfilling these two tasks are formulated as follows.

In a certain round of distillation, given the input density matrix  $\mathfrak{E}$ , we maximize the fidelity of the kept qubit pair  $F_{\mathbb{D}}(\mathfrak{E})$  w.r.t. the distillation operation  $\mathbb{D}$ . This problem can be formulated as

$$\mathcal{P}_F: \max_{\mathbb{D} \in \mathcal{D}} F_{\mathbb{D}}(\mathfrak{E}),$$

where  $\mathcal{D}$  is the set of all possible LOCC operations. Denote the optimal fidelity as  $\hat{F}(\mathfrak{E})$ . We maximize the success probability of the distillation operation  $P_{\mathbb{D}}(\mathfrak{E})$  w.r.t. the distillation operation  $\mathbb{D}$  conditioned on the fact that optimal fidelity is achieved. This problem can be formulated as

$$\mathcal{P}_P: \max_{\mathbb{D} \in \mathcal{D}_F} P_{\mathbb{D}}(\mathfrak{E}),$$

where  $\mathcal{D}_F = \{\mathbb{D} : F_{\mathbb{D}}(\mathfrak{E}) = \hat{F}(\mathfrak{E})\}$ .

### III. EFFICIENT QED FOR PMD CHANNELS

This section first characterizes the optimal value of problems  $\mathcal{P}_F$  and  $\mathcal{P}_P$  and then gives an algorithm which achieves the optimal performance in every round of distillation. For conciseness, in the following, both  $|h_A\rangle$  and  $|h_B\rangle$  are denoted as  $|0\rangle$ , and both  $|v_A\rangle$  and  $|v_B\rangle$  are denoted as  $|1\rangle$ . The network node index can be omitted without causing confusion because only local operations are involved in the distillation process.

#### A. Characterization of performance upper bounds

This subsection considers a set of density matrices that includes those given in (2) and characterizes the corresponding optimal performance of problems  $\mathcal{P}_F$  and  $\mathcal{P}_P$ . Specifically, consider the following set of density matrices

$$\mathcal{S} = \{\mathfrak{E} : \mathfrak{E} \text{ is in the form of (8)}\},$$

where

$$\begin{aligned} \mathfrak{E} = & \frac{1}{2}(|ab\rangle\langle ab| + e^{-i\alpha} R^* |ab\rangle\langle a'b'| \\ & + e^{i\alpha} R |a'b'\rangle\langle ab| + |a'b'\rangle\langle a'b'|), \end{aligned} \quad (8)$$

in which

$$\begin{aligned} \langle x|x'\rangle &= 0, \quad x \in \{a, b\}, \\ \alpha &\in [0, 2\pi), \quad \text{and} \\ |R| &\in [0, 1]. \end{aligned}$$

By performing matrix spectral decomposition, the density matrix  $\mathfrak{E}$  can be rewritten as

$$\mathfrak{E} = F|\phi_1\rangle\langle\phi_1| + (1-F)|\phi_2\rangle\langle\phi_2|, \quad (9)$$

where

$$\begin{aligned} F &= \frac{1}{2}(1 + |R|), \\ |\phi_1\rangle &= \frac{1}{\sqrt{2}}(|ab\rangle + e^{i\theta}|a'b'\rangle), \\ |\phi_2\rangle &= \frac{1}{\sqrt{2}}(|ab\rangle - e^{i\theta}|a'b'\rangle), \\ \theta &= \alpha + \angle R. \end{aligned}$$

The following theorem characterizes the optimal fidelity that can be achieved when input density matrix  $\mathfrak{E} \in \mathcal{S}$ .

*Theorem 1. Optimal fidelity.* Let  $\mathfrak{E} \in \mathcal{S}$ . Then the optimal value of  $\mathcal{P}_F$  is given by

$$\hat{F}(\mathfrak{E}) = \frac{F^2}{F^2 + (1-F)^2}. \quad (10)$$

*Proof.* The two network nodes perform the local unitary operations

$$\begin{aligned} U_A &= \frac{|0\rangle + |1\rangle}{\sqrt{2}}\langle a| + \frac{|0\rangle - |1\rangle}{\sqrt{2}}\langle a'|, \\ U_B &= \frac{|0\rangle + |1\rangle}{\sqrt{2}}\langle b| + e^{-i\theta} \frac{|0\rangle - |1\rangle}{\sqrt{2}}\langle b'|, \end{aligned} \quad (11)$$

on a pair of qubits with density matrix  $\mathfrak{E}$ . The updated density matrix is given by

$$\begin{aligned} \check{\mathfrak{E}} &= (U_A \otimes U_B) \mathfrak{E} (U_A \otimes U_B)^\dagger \\ &= F|\phi^+\rangle\langle\phi^+| + (1-F)|\psi^+\rangle\langle\psi^+|, \end{aligned} \quad (12)$$

where

$$|\phi^+\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle), \quad |\psi^+\rangle = \frac{1}{\sqrt{2}}(|01\rangle + |10\rangle).$$

The density matrix in (12) has the structure of the density matrix in Eq. (6) of Ref. [47], with  $\alpha = \beta = \gamma = \delta = \frac{1}{\sqrt{2}}$ . Therefore, one can apply Theorem 2 of Ref. [47] and get

$$\hat{F}(\check{\mathfrak{E}}) = \frac{F^2}{F^2 + (1-F)^2}.$$

Moreover, since unitary operations are reversible, the same optimal fidelity can be achieved starting with either  $\check{\mathfrak{E}}$  or  $\mathfrak{E}$ , i.e.,  $\hat{F}(\check{\mathfrak{E}}) = \hat{F}(\mathfrak{E})$ . This completes the proof.  $\square$

The next theorem characterizes the upper bound of the success probability conditioned on the fact that optimal fidelity is achieved.

*Theorem 2. Optimal probability of success.* Let  $\mathfrak{E} \in \mathcal{S}$  with  $|R| > 0$ . Then the optimal value of  $\mathcal{P}_p$  is given by

$$\hat{P}(\mathfrak{E}) = F^2 + (1 - F)^2. \quad (13)$$

*Proof.* We first prove that the proposed success probability is an upper bound, i.e.,

$$\hat{P}(\mathfrak{E}) \leq F^2 + (1 - F)^2. \quad (14)$$

The statement will be proved by contradiction. Suppose the theorem does not hold, i.e., for some  $\mathfrak{E} \in \mathcal{S}$  with  $|R| > 0$  there exists a distillation operation  $\mathbb{D}$  such that

$$F_{\mathbb{D}}(\mathfrak{E}) = \frac{F^2}{F^2 + (1 - F)^2}, \quad (15)$$

$$P_{\mathbb{D}}(\mathfrak{E}) > F^2 + (1 - F)^2. \quad (16)$$

From (9), the spectral decomposition of the joint density matrix of two qubit pairs is given by

$$\mathfrak{E}^J = F^2 |\phi_1 \phi_1\rangle \langle \phi_1 \phi_1| + F(1 - F) |\phi_1 \phi_2\rangle \langle \phi_1 \phi_2| \\ + (1 - F)F |\phi_2 \phi_1\rangle \langle \phi_2 \phi_1| + (1 - F)^2 |\phi_2 \phi_2\rangle \langle \phi_2 \phi_2|.$$

Define

$$V_{nm} = \text{tr}_{3,4} \left\{ \mathbb{D} \left\{ |\phi_n \phi_m\rangle \langle \phi_n \phi_m| \right\} \right\}, \\ f_{nm} = \langle \phi^+ | V_{nm} | \phi^+ \rangle, \quad p_{nm} = \text{tr} \{ V_{nm} \},$$

where  $n, m \in \{1, 2\}$ . As along as  $\mathbb{D}$  is a valid quantum operation,  $V_{nm}$  must be a positive semidefinite matrix with trace no greater than 1. Therefore,

$$0 \leq f_{nm} \leq p_{nm} \leq 1. \quad (17)$$

It is straightforward that

$$F_{\mathbb{D}}(\mathfrak{E}) = \frac{F^2 f_{11} + F(1 - F)(f_{12} + f_{21}) + (1 - F)^2 f_{22}}{F^2 p_{11} + F(1 - F)(p_{12} + p_{21}) + (1 - F)^2 p_{22}}, \quad (18)$$

$$P_{\mathbb{D}}(\mathfrak{E}) = F^2 p_{11} + F(1 - F)(p_{12} + p_{21}) + (1 - F)^2 p_{22}. \quad (19)$$

Combining (16) and (19), and noting that  $p_{nm} \leq 1$ , it can be derived that

$$p_{12} + p_{21} > 0. \quad (20)$$

Denote

$$S(F) = F^2 f_{11} + F(1 - F)(f_{12} + f_{21}) + (1 - F)^2 f_{22},$$

$$N(F) = F^2(p_{11} - f_{11}) + F(1 - F)(p_{12} + p_{21} - f_{12} - f_{21}) \\ + (1 - F)^2(p_{22} - f_{22}).$$

Then from (15) and (18)

$$F_{\mathbb{D}}(\mathfrak{E}) = \frac{S(F)}{S(F) + N(F)} = \frac{F^2}{F^2 + (1 - F)^2} \\ \Rightarrow \frac{N(F)}{S(F)} = \frac{(1 - F)^2}{F^2}. \quad (21)$$

Note that  $F > \frac{1}{2}$  as  $|R| > 0$ . Hence, one can construct another density matrix  $\tilde{\mathfrak{E}}$  satisfying (9), with a different  $\tilde{F} \in (\frac{1}{2}, F)$ . By repeating the analysis above, it can be derived that

$$F_{\mathbb{D}}(\tilde{\mathfrak{E}}) = \frac{S(\tilde{F})}{S(\tilde{F}) + N(\tilde{F})}. \quad (22)$$

From (17) and (20), either  $p_{12} + p_{21} = f_{12} + f_{21} > 0$ , or  $p_{12} + p_{21} > f_{12} + f_{21} \geq 0$ . If  $p_{12} + p_{21} = f_{12} + f_{21} > 0$ , then

$$S(\tilde{F}) = \frac{\tilde{F}^2}{F^2} \left( F^2 f_{11} + \frac{F^2}{\tilde{F}} (1 - \tilde{F})(f_{12} + f_{21}) + \frac{F^2}{\tilde{F}^2} (1 - \tilde{F})^2 f_{22} \right) \\ > \frac{\tilde{F}^2}{F^2} \left( F^2 f_{11} + F(1 - F)(f_{12} + f_{21}) + (1 - F)^2 f_{22} \right) \\ = \frac{\tilde{F}^2}{F^2} S(F), \quad (23)$$

$$N(\tilde{F}) = \frac{(1 - \tilde{F})^2}{(1 - F)^2} \left( \frac{(1 - F)^2}{(1 - \tilde{F})^2} \tilde{F}^2 (p_{11} - f_{11}) \right. \\ \left. + \tilde{F} \frac{(1 - F)^2}{(1 - \tilde{F})} (p_{12} + p_{21} - f_{12} - f_{21}) \right. \\ \left. + (1 - F)^2 (p_{22} - f_{22}) \right) \\ \leq \frac{(1 - \tilde{F})^2}{(1 - F)^2} \left( F^2 (p_{11} - f_{11}) \right. \\ \left. + F(1 - F)(p_{12} + p_{21} - f_{12} - f_{21}) \right. \\ \left. + (1 - F)^2 (p_{22} - f_{22}) \right) \\ = \frac{(1 - \tilde{F})^2}{(1 - F)^2} N(F). \quad (24)$$

Substituting (21), (23), and (24) into (22), one can get

$$F_{\mathbb{D}}(\tilde{\mathfrak{E}}) > \frac{\tilde{F}^2}{\tilde{F}^2 + (1 - \tilde{F})^2},$$

which leads to

$$\hat{F}(\tilde{\mathfrak{E}}) \geq F_{\mathbb{D}}(\tilde{\mathfrak{E}}) > \frac{\tilde{F}^2}{\tilde{F}^2 + (1 - \tilde{F})^2}. \quad (25)$$

However, (25) contradicts with (10).

Otherwise, if  $p_{12} + p_{21} > f_{12} + f_{21} \geq 0$ , one can use a similar analysis and get

$$S(\tilde{F}) \geq \frac{\tilde{F}^2}{F^2} S(F) \quad \text{and} \quad N(\tilde{F}) < \frac{(1 - \tilde{F})^2}{(1 - F)^2} N(F),$$

which also lead to a contradiction between (25) and (10). This contradiction shows that the success probability given in (13) is indeed an upper bound.

The achievability of (13) is proved by exhibiting the QED algorithm that achieves the upper bound. Please refer to Sec. III B for details.  $\square$

## B. Algorithm design

The two theorems in the previous subsection characterize the optimal fidelity and the corresponding optimal success probability of distillation operations on two pairs of qubits. In this subsection, guided by the insights obtained from the proofs of Theorem 1 and Theorem 2, a recurrence QED algorithm is designed to achieve the optimal fidelity and the



corresponding optimal success probability in every round of distillation.

*Algorithm: Efficient QED for PMD channel.*

- (a) *Local state preparation:* For each qubit pair, the network nodes transform the density matrix to  $\tilde{\mathcal{E}}$  using the local unitary operators  $U_A$  and  $U_B$  defined in (11).
- (b) *Single-round distillation:* The nodes take two of the kept qubit pairs and perform the following operations.
- (i) Each node locally performs a CNOT operation, i.e.,  $U = |00\rangle\langle 00| + |01\rangle\langle 01| + |10\rangle\langle 11| + |11\rangle\langle 10|$ , on the two qubits at hand.
  - (ii) Each node measures the target bit (i.e., the qubit in the second pair) using the operators  $|0\rangle\langle 0|$ ,  $|1\rangle\langle 1|$  and transmits the measurement result to the other node via classical communication.
  - (iii) If their measurement results do not agree, the nodes discard the source qubit pair (i.e., the first pair). Otherwise, the nodes keep the source qubit pair.

The nodes repeat operations (i)–(iii) on all other kept qubits, two pairs at a time.

- (c) *Stopping criterion:* Network nodes recurrently perform the single-round distillation described in (b), until the fidelity of the kept qubit pairs exceeds the required threshold.  $\square$

In the following, we first characterize the performance of the proposed algorithm in Theorem 3, then provide the insights of the theorem in two remarks.

*Theorem 3. Performance of the proposed algorithm.* In the  $k$ th round of distillation, a qubit pair is kept with fidelity

$$F_k = \frac{F_{k-1}^2}{F_{k-1}^2 + (1 - F_{k-1})^2}, \quad (26)$$

probability

$$P_k = F_{k-1}^2 + (1 - F_{k-1})^2, \quad (27)$$

and density matrix

$$\mathcal{E}_k = F_k |\phi^+\rangle\langle\phi^+| + (1 - F_k) |\psi^+\rangle\langle\psi^+|. \quad (28)$$

*Proof.* From (12), after the first step of the algorithm, the joint density matrix of two qubit pairs is given by

$$\begin{aligned} \tilde{\mathcal{E}}^J &= \mathbf{P} \tilde{\mathcal{E}} \otimes \tilde{\mathcal{E}} \mathbf{P}^\dagger \\ &= F^2 |\varphi^{(1)}\rangle\langle\varphi^{(1)}| + F(1 - F) (|\varphi^{(2)}\rangle\langle\varphi^{(2)}| + |\varphi^{(3)}\rangle\langle\varphi^{(3)}|) \\ &\quad + (1 - F)^2 |\varphi^{(4)}\rangle\langle\varphi^{(4)}|, \end{aligned}$$

where  $\mathbf{P}$  is the permutation operator that switches the second and third qubits, and

$$\begin{aligned} |\varphi^{(1)}\rangle &= \frac{1}{2} |0000\rangle + \frac{1}{2} |0101\rangle + \frac{1}{2} |1010\rangle + \frac{1}{2} |1111\rangle, \\ |\varphi^{(2)}\rangle &= \frac{1}{2} |0001\rangle + \frac{1}{2} |0100\rangle + \frac{1}{2} |1011\rangle + \frac{1}{2} |1110\rangle, \\ |\varphi^{(3)}\rangle &= \frac{1}{2} |0010\rangle + \frac{1}{2} |0111\rangle + \frac{1}{2} |1000\rangle + \frac{1}{2} |1101\rangle, \\ |\varphi^{(4)}\rangle &= \frac{1}{2} |0011\rangle + \frac{1}{2} |0110\rangle + \frac{1}{2} |1001\rangle + \frac{1}{2} |1100\rangle. \end{aligned}$$

In the first round of distillation, after both nodes perform the CNOT operation, the joint density matrix of two qubit pairs becomes

$$\begin{aligned} \tilde{\mathcal{E}}^J &= F^2 |\tilde{\varphi}^{(1)}\rangle\langle\tilde{\varphi}^{(1)}| + F(1 - F) (|\tilde{\varphi}^{(2)}\rangle\langle\tilde{\varphi}^{(2)}| \\ &\quad + |\tilde{\varphi}^{(3)}\rangle\langle\tilde{\varphi}^{(3)}|) + (1 - F)^2 |\tilde{\varphi}^{(4)}\rangle\langle\tilde{\varphi}^{(4)}|, \end{aligned} \quad (29)$$

where

$$\begin{aligned} |\tilde{\varphi}^{(1)}\rangle &= \frac{1}{2} |0000\rangle + \frac{1}{2} |0101\rangle + \frac{1}{2} |1111\rangle + \frac{1}{2} |1010\rangle, \\ |\tilde{\varphi}^{(2)}\rangle &= \frac{1}{2} |0001\rangle + \frac{1}{2} |0100\rangle + \frac{1}{2} |1110\rangle + \frac{1}{2} |1011\rangle, \\ |\tilde{\varphi}^{(3)}\rangle &= \frac{1}{2} |0011\rangle + \frac{1}{2} |0110\rangle + \frac{1}{2} |1100\rangle + \frac{1}{2} |1001\rangle, \\ |\tilde{\varphi}^{(4)}\rangle &= \frac{1}{2} |0010\rangle + \frac{1}{2} |0111\rangle + \frac{1}{2} |1101\rangle + \frac{1}{2} |1000\rangle. \end{aligned}$$

From (29), if both measurement results correspond to  $|0\rangle\langle 0|$ , the (unnormalized) density matrix of the source qubit pair is given by

$$\begin{aligned} \mathcal{E}^{00} &= (\mathbf{I}_2 \otimes \langle 0| \otimes \mathbf{I}_2 \otimes \langle 0|) \tilde{\mathcal{E}}^J (\mathbf{I}_2 \otimes |0\rangle \otimes \mathbf{I}_2 \otimes |0\rangle) \\ &= \frac{1}{2} (F^2 |\phi^+\rangle\langle\phi^+| + (1 - F)^2 |\psi^+\rangle\langle\psi^+|). \end{aligned} \quad (30)$$

Similarly, if both measurement results correspond to  $|1\rangle\langle 1|$ , the (unnormalized) density matrix of the source qubit pair is given by

$$\begin{aligned} \mathcal{E}^{11} &= (\mathbf{I}_2 \otimes \langle 1| \otimes \mathbf{I}_2 \otimes \langle 1|) \tilde{\mathcal{E}}^J (\mathbf{I}_2 \otimes |1\rangle \otimes \mathbf{I}_2 \otimes |1\rangle) \\ &= \frac{1}{2} (F^2 |\phi^+\rangle\langle\phi^+| + (1 - F)^2 |\psi^+\rangle\langle\psi^+|). \end{aligned} \quad (31)$$

From (30), and (31), the probability of preserving the source qubit pair is

$$P = \text{tr}\{\mathcal{E}^{00} + \mathcal{E}^{11}\} = F^2 + (1 - F)^2, \quad (32)$$

the fidelity of the kept qubit pairs is

$$F_1 = \frac{\frac{1}{2}F^2 + \frac{1}{2}F^2}{P} = \frac{F^2}{F^2 + (1 - F)^2}, \quad (33)$$

and the density matrix of the kept qubit pair can be written as

$$\mathcal{E}_1 = \frac{\mathcal{E}^{00} + \mathcal{E}^{11}}{P} = F_1 |\phi^+\rangle\langle\phi^+| + (1 - F_1) |\psi^+\rangle\langle\psi^+|. \quad (34)$$

With (32) and (33), the proof for the first round of distillation is complete. For the following distillation rounds, one can take (34) as input and repeat the analysis in (29)–(33). This completes the proof.  $\square$

*Remark 1. Optimality of the proposed algorithm.* In Theorem 3, (28) shows that the proposed algorithm always keeps the density matrix of qubit pairs in set  $\mathcal{S}$ , which means that the results in Theorem 1 and Theorem 2 apply to every round of distillation. Therefore, by comparing (10) and (13) with (26) and (27), one can see that the proposed algorithm achieves the optimal fidelity and the corresponding optimal success probability in every round of distillation. As verified in Sec. IV, this feature enables the proposed algorithm to achieve a high efficiency.  $\square$

*Remark 2. Convergence speed of fidelity.* In terms of the convergence speed of fidelity w.r.t. the number of distillation rounds, the only existing theoretical result was given in [27], which shows that the relation of the fidelity of kept qubit pairs in consecutive rounds is given by

$$F_k = \frac{F_{k-1}^2 + \frac{1}{9}(1 - F_{k-1})^2}{F_{k-1}^2 + \frac{2}{3}F_{k-1}(1 - F_{k-1}) + \frac{5}{9}(1 - F_{k-1})^2}. \quad (35)$$

In this case, when  $F_0 > \frac{1}{2}$ , it can be obtained that

$$\lim_{k \rightarrow \infty} \frac{1 - F_k}{1 - F_{k-1}} = \frac{2}{3}. \quad (36)$$

For the proposed algorithms, it can be shown from (26) that when  $F_0 > \frac{1}{2}$ ,

$$\lim_{k \rightarrow \infty} \frac{1 - F_k}{1 - F_{k-1}} = 0, \quad \lim_{k \rightarrow \infty} \frac{1 - F_k}{(1 - F_{k-1})^2} = 1. \quad (37)$$

Equation (36) shows that with the algorithm proposed in [27], the fidelity of the qubit pairs converges to 1 linearly at rate  $2/3$ , whereas (37) shows that with the proposed algorithms, the fidelity converges to 1 quadratically. Hence, the convergence speed of our algorithm is quadratic in number of distillation rounds, which is a significant improvement over the linear convergence achieved by the recurrence QED algorithm proposed in [27]. On the other hand, the algorithm proposed in [27] applies to generic channels (with  $F_0 > \frac{1}{2}$ ), whereas the proposed algorithm is tailored for the PMD channel. The issue of improving the convergence speed of recurrence QED algorithms for generic channels remains an interesting open question.  $\square$

#### IV. NUMERICAL RESULTS

We now demonstrate the dependence of the proposed recurrence QED algorithm on the parameters of the PMD channel by numerically calculating the yield and output fidelity for different channel configurations. To perform numerical tests, we specify the optical properties of the entanglement source to determine the form of  $R(\tau_A, \tau_B)$  under the generally considered assumption that the pulsed pump laser and frequency response of the filters are Gaussian [24]. Please see the last paragraph in Appendix A for more details.

We compare the yield of our algorithm with that obtained by an existing recurrence QED algorithm [27]. As an additional benchmark, an upper bound of yield derived from distillable entanglement [49,50] is also calculated and plotted. While the achievability of this bound remains unknown, it is arguably the best-known upper bound on the yield of any QED algorithms [51]. We find that our algorithm has a significant performance advantage in parameter regimes where partial PMD compensation occurs [23,24] and achieves a yield close to the theoretical upper bound despite its simple recurrent distillation operations that involve only two qubit pairs. Additionally, we have performed tests to examine the robustness of the proposed algorithm to basis alignment errors.

In the numerical tests, the target fidelity is set to be 0.99. The number of distillation rounds  $K$  is set to be the first round that achieves the target fidelity, and the yield of the algorithm is calculated according to (7). We assume that the photon bandwidths  $B_A$  and  $B_B$  are equal, and we set  $\tau_A B_A = 1$  while varying the DGD on photon  $B$ , given by  $\tau_B$ , the pump laser bandwidth  $B_p$ , and  $\eta$ , which specifies the alignment between the qubit and PSP basis.

Figures 2 and 3 plot the yield as a function of the ratio of the magnitudes of the DGD in each optical path for two different pulse pump bandwidths. Figure 2 plots the case where the pump bandwidth is given by  $B_p = 0.1/\tau_A$ , which corresponds to a relatively long pump duration compared to the DGD. Alternatively, Fig. 3 plots a case where a pump bandwidth is of the order of the DGD, given by  $B_p = 1/\tau_A$ .

In Fig. 2 we see that both algorithms achieve a yield of unity for a finite region of  $\tau_A/\tau_B$  centered around the

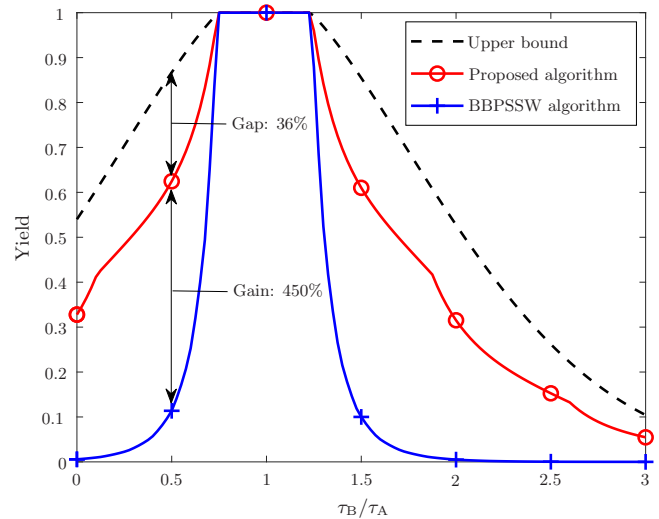


FIG. 2. Comparison of the yield as a function of  $\tau_B/\tau_A$  for the proposed algorithm and the benchmarks, i.e., the upper bound [49] and the recurrence QED algorithm proposed in [27] (referred to as the BBPSSW algorithm here). In this plot,  $B_p = 0.1$ ,  $B_A = B_B = 1$ ,  $\tau_A = 1$ .

decoherence-free subspace at  $\tau_A = \tau_B$  [23,24]. For regions of partial or no compensation, the regions outside of unit yield in Fig. 2 and all of Fig. 3, the proposed algorithm achieves a yield that is significantly higher than the baseline algorithm from [27]. For instance, when  $\tau_B/\tau_A = 0.5$ , the proposed algorithm increases the yield by 450% and 5660% compared to the baseline algorithm, and the yield of the proposed algorithm is 36% and 53% away from the upper bound. Given that the proposed algorithm adopts simple recurrent distillation operations that involve only two qubit pairs, it achieves a desirable balance between efficiency and implementability. We also note that the peak of the yield for both algorithms

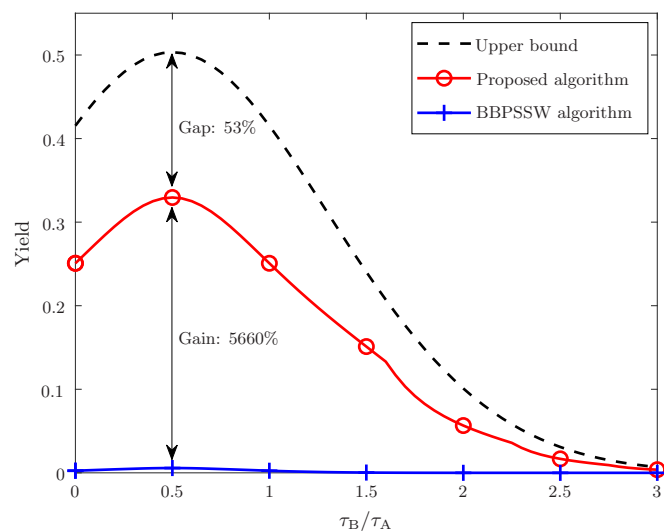


FIG. 3. Comparison of the yield as a function of  $\tau_B/\tau_A$  for the proposed algorithm and the benchmarks, i.e., the upper bound [49] and the BBPSSW algorithm [27]. In this plot,  $B_p = 1$ ,  $B_A = B_B = 1$ ,  $\tau_A = 1$ .

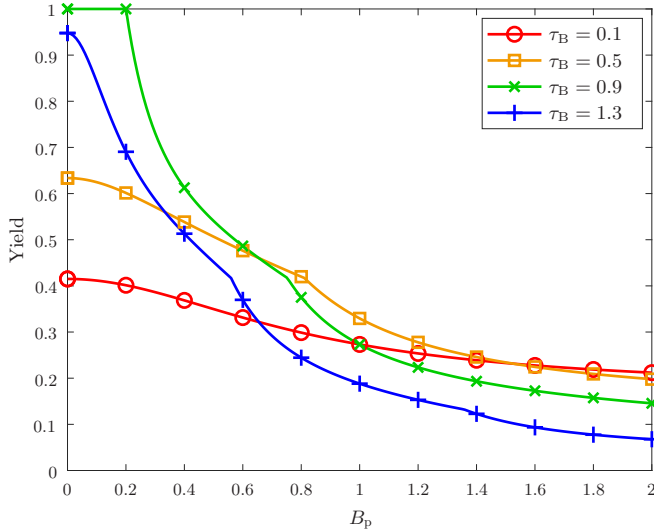


FIG. 4. The efficiency of the proposed algorithm as a function of the bandwidth of the source laser pump. In this figure,  $B_A = B_B = 1$ ,  $\tau_A = 1$ .

in Fig. 3 is shifted away from  $\tau_A = \tau_B$ , as opposed to the peak being centered around this point in Fig. 2. This observation is consistent with those of [24] on PMD compensation, which emphasizes the fact that our algorithm attempts to make use of nonlocal PMD compensation to whatever extent is possible.

To further demonstrate the impact of pump bandwidth on the performance of the proposed algorithm, the yield as a function of  $B_p$  is plotted in Fig. 4 for several values of  $\tau_B$ . From the figure, it can be observed that the yield of the algorithm is a decreasing function of the pump bandwidth  $B_p$ . This is because the larger  $B_p$  is, the more distinguishable are the photon pairs advanced and delayed by PMD. For analogous reasons, we see that when  $B_p$  is large, the yield of the algorithm is likely to decrease when  $\tau_B$  increases. However, when  $B_p$  is small, the yield of the algorithm is highest when the values of  $\tau_A$  and  $\tau_B$  are similar, illustrating the benefits of the decoherence-free subspace created by PMD compensation.

Finally, the performance of the proposed algorithm is evaluated in the presence of basis alignment errors. Until now, perfect alignment between the polarization basis and the PSP basis has been assumed. As mentioned in Sec. II A, such an alignment is not expected to be performed frequently, as the PSP of installed optical fiber has been shown to remain unchanged on the time scale of months [48]. However, any realistic implementation will have to deal with errors in the initial alignment process and the eventual drift of the PSP with time. To help us quantify the effects of implementation error on the performance of the proposed algorithm, we define the misalignment angle between the polarization and the PSP basis as  $\theta$ , where  $\eta_1 = \arcsin(\frac{\theta\pi}{180})$ . In Fig. 5, the output fidelity and the yield of the proposed algorithm are plotted as a function of the misalignment angle  $\theta$  for several values of  $\tau$ , where  $\tau_A = \tau_B = \tau$ . The output fidelities shown in the plot are the maximum achievable fidelity with the proposed algorithm with a required fidelity of 0.99. It can be seen that for all considered values of  $\tau$ , the algorithm can generate qubit

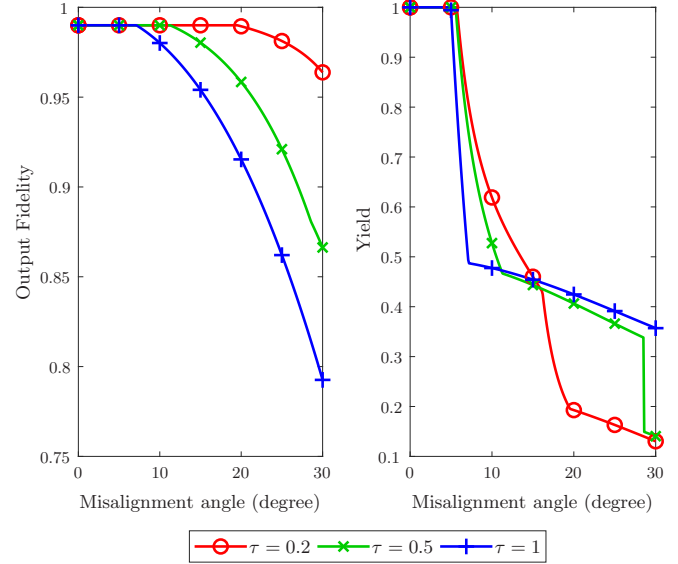


FIG. 5. The output fidelity and the efficiency of the proposed algorithm as a function of the misalignment angle  $\theta$ .  $\eta_1 = \arcsin(\frac{\theta\pi}{180})$ . In this figure,  $B_A = B_B = 1$ ,  $B_p = 0.1$ ,  $\tau_A = \tau_B = \tau$ . The output fidelity is the maximum achievable by the algorithm, up to a fidelity of 0.99.

pairs with the required fidelity when the misalignment angle is no more than  $5^\circ$ . When the misalignment angle  $\theta$  is greater than  $5^\circ$ , the output fidelities are higher for smaller values of  $\tau$ , meaning that the robustness of the algorithm is inversely proportional to the magnitude of the DGD. Finally, it can be observed that the yield of the algorithm drops significantly for misalignment angle  $\theta$  beyond  $5^\circ$ . This means that, even though the algorithm can still obtain photon pairs with high fidelity when  $\theta > 5^\circ$ , it demands a significant increase in resources. This result can be used to bound the precision of local unitary operations needed for an experimental implementation of this algorithm.

Figure 5 also serves as an indication of how the proposed algorithm performs in scenarios with imperfect operations or noise other than PMD. The proposed algorithm will perform well if the effects of operation imperfection or other noise are not significant. Otherwise, both the highest achievable fidelity and the efficiency of the proposed algorithm will drop.

## V. CONCLUSION

This paper presents a recurrence QED algorithm to obtain high-quality entanglement from polarization-entangled photon pairs affected by PMD-degraded channels. For these photon pairs, we have characterized the optimal fidelity that can be achieved by recurrence QED operations as well as the optimal success probability conditioned on the fact that optimal fidelity is achieved. We then proposed a recurrence QED algorithm that achieves both optimal fidelity and optimal success probability in every round of distillation. Analytical results show that the proposed algorithm improves the convergence speed of fidelity w.r.t. the number of distillation rounds from linear to quadratic. Numerical tests show that the proposed algorithm significantly improves the efficiency of



QED in a wide range of operation regions and achieves a yield close to the best-known upper bound for any QED algorithms.

### ACKNOWLEDGMENTS

The authors would like to thank S. Guerrini for his helpful suggestions and careful reading of the manuscript. The fundamental research described in this paper was supported, in part, by the Army Research Office through the MIT Institute for Soldier Nanotechnologies under Contract No. W911NF-13-D-0001.

### APPENDIX A: ANALYSIS OF THE EFFECT OF PMD

The effect of PMD on a polarization-entangled photon pair depends on the way that the photons are generated, in particular, the type of nonlinear media and laser pump. A rigorous treatment dealing with  $\chi^{(3)}$  media and a continuous-wave (CW) pump was given in [23], and that dealing with  $\chi^{(2)}$  media and a pulsed pump was presented in [24]. Here we present an analytical treatment for  $\chi^{(2)}$  media and a pulsed pump. In particular, we consider the limit where the frequency content of the pulse approaches a delta function, effectively becoming a CW beam.

Consider a pair of photons which are entangled in two orthogonal polarizations as well as time. These pairs can be created using parametric down conversion or fiber nonlinearities [52,53] and are notated as

$$|\psi\rangle = |f_{0,0}\rangle \otimes \frac{1}{\sqrt{2}}(|h_A\rangle|h_B\rangle + e^{i\alpha}|v_A\rangle|v_B\rangle), \quad (\text{A1})$$

where  $h_i$  and  $v_i$  are orthogonal polarization basis states of photons A and B. The time-related component  $|f_{0,0}\rangle$  in the state representation (A1) is a special case of

$$|f_{\zeta_A, \zeta_B}\rangle = \iint dt_A dt_B f\left(t_A + \frac{\zeta_A}{2}, t_B + \frac{\zeta_B}{2}\right) |t_A, t_B\rangle, \quad (\text{A2})$$

where  $\zeta_i/2$  represents the shift of the arrival time at node  $i \in \{A, B\}$ . The function  $f(t_A, t_B)$  is normalized so that  $|f(t_A, t_B)|^2$  represents the probability that the two photons overlap in time, and  $\iint dt_A dt_B |f(t_A, t_B)|^2 = 1$ . Since the entanglement is generated via  $\chi^{(2)}$  media, we have

$$f(t_A, t_B) \propto \int dt h_A^*(t - t_A) h_B^*(t - t_B) e_p(t), \quad (\text{A3})$$

where  $h_i(t)$  represents the inverse Fourier transform of the frequency filter  $H_i(\omega)$  at node  $i \in \{A, B\}$  and  $e_p(t)$  is the envelope of the pump signal.

The two types of laser pumps, pulsed and CW, are characterized by the envelope of the pump signal  $e_p(t)$  and its Fourier transform  $E_p(\omega)$ , which describes the frequency content of the input pulse. Experimentally, pulsed pump lasers are convenient because they allow experiments to be broken into discrete detection time bins, resulting in wider bandwidth signal and idler photons, which enables multiple channels. For CW lasers,  $|E_p(\omega)|^2$  approaches a  $\delta$  function, which is a constant in the time domain. In this case,  $f(t_A, t_B)$  becomes a function of only the time difference, hence simplifying the analysis.

The PMD advances or delays photon arrival times, with the maximum and minimum alterations occurring for photons with polarizations equal to the PSP of the fiber [23]. Therefore, it is convenient to write the initial state in terms of the

PSP basis  $\{|s_i\rangle, |s'_i\rangle\}$ ,  $i \in \{A, B\}$ . In this basis the initial state becomes

$$|\psi\rangle = |f_{0,0}\rangle \otimes \left[ \frac{\eta_1}{\sqrt{2}}(|s_A\rangle|s_B\rangle + e^{i\alpha_1}|s'_A\rangle|s'_B\rangle) + \frac{\eta_2}{\sqrt{2}}(|s_A\rangle|s'_B\rangle - e^{i\alpha_2}|s'_A\rangle|s_B\rangle) \right], \quad (\text{A4})$$

where

$$\eta_1 = \langle s_A|h_A\rangle\langle s_B|h_B\rangle + e^{i\alpha}\langle s_A|v_A\rangle\langle s_B|v_B\rangle, \\ \eta_2 = \langle s_A|h_A\rangle\langle s'_B|h_B\rangle + e^{i\alpha}\langle s_A|v_A\rangle\langle s'_B|v_B\rangle,$$

and  $\alpha_i$  satisfies the relation  $\eta_i = |\eta_i|e^{i(\alpha-\alpha_i)/2}$ . Time delays resulting from PMD in the fibers can now be described as

$$|\psi_{\text{PMD}}\rangle = \frac{\eta_1}{\sqrt{2}}|f_{-\tau_A, -\tau_B}\rangle \otimes |s_A s_B\rangle \\ - \frac{\eta_2 e^{i\alpha_2}}{\sqrt{2}}|f_{+\tau_A, -\tau_B}\rangle \otimes |s'_A s_B\rangle \\ + \frac{\eta_2}{\sqrt{2}}|f_{-\tau_A, +\tau_B}\rangle \otimes |s_A s'_B\rangle \\ + \frac{\eta_1 e^{i\alpha_1}}{\sqrt{2}}|f_{+\tau_A, +\tau_B}\rangle \otimes |s'_A s'_B\rangle. \quad (\text{A5})$$

We assume that the coincidence time window of the two photon detectors is much larger than the DGD  $\tau_A, \tau_B$ , so that the photon pair can be detected correctly. To account for the fact that the photodetection process is not sensitive to the photon's time of arrival, the time modes of the two photons are to be traced out. Hence, the polarization state of the two photons can be characterized by a density matrix for two qubits. When written in the ordered basis  $\{|s_A s_B\rangle, |s'_A s_B\rangle, |s_A s'_B\rangle, |s'_A s'_B\rangle\}$ , the density matrix resulting from integration w.r.t. time is given by (1), in which

$$R(\tau_A, \tau_B) = \iint dt_A dt_B f(t_A + \tau_A, t_B + \tau_B) f^*(t_A, t_B), \quad (\text{A6})$$

with the property that  $R(0, 0) = 1$ .

The approach above can also be applied to scenarios involving  $\chi^{(3)}$  media, which changes (A3) and in turn (A6). Since these changes have a minor impact on the analytical results as well as the numerical findings in this paper, we omit the analysis for  $\chi^{(3)}$  here.

In the numerical study, the frequency content of a pulsed pump laser and the frequency response of filters are assumed to be Gaussian. Under this assumption, the form of  $R(\tau_A, \tau_B)$  is given by [25]

$$R(\tau_A, \tau_B) = \kappa \iint d\omega_A d\omega_B |H_A(\omega_A)|^2 |H_B(\omega_B)|^2 \\ \times |E_p(\omega_A + \omega_B)|^2 e^{i(\tau_A \omega_A + \tau_B \omega_B)},$$

where  $E_p(\omega) \propto e^{-\omega/4B_p^2}$  and  $H_i(\omega) \propto e^{-(\omega \pm \Delta\Omega)^2/4B_i^2}$ ,  $i \in \{A, B\}$ , with the  $B_i$  terms representing the root mean square bandwidth of each filter. The central frequency of the pump is set to 0 and Alice and Bob's filters are each offset from it by

$\pm \Delta\Omega$ . The integral results in

$$R(\tau_A, \tau_B) = e^{-\frac{B_A^2 B_B^2 (\tau_A - \tau_B)^2 + B_A^2 B_P^2 \tau_A^2 + B_B^2 B_P^2 \tau_B^2}{2(B_A^2 + B_B^2 + B_P^2)}} e^{-i\Delta\Omega(\tau_A - \tau_B)}.$$

## APPENDIX B: LOCAL ROTATION ON ONE PHOTON IS SUFFICIENT FOR ALIGNMENT

We first prove a lemma and then show that, as a special case of the lemma, local rotation on one of the photons can achieve the alignment of the PSP basis with the photon polarization basis.

*Lemma 1 Basis of maximally entangled states.* Let  $|\phi\rangle$  be a maximally entangled state of two qubits. Then for all qubit basis  $\{|s\rangle, |s'\rangle\}$ , there exists some basis of a qubit  $\{|\tilde{s}\rangle, |\tilde{s}'\rangle\}$  such that

$$|\phi\rangle = \frac{1}{\sqrt{2}}(|\tilde{s}s\rangle + |\tilde{s}'s'\rangle). \quad (\text{B1})$$

*Proof.* Express  $|\phi\rangle$  in the basis of  $\{|s\rangle, |s'\rangle\}$ , i.e.,

$$\begin{aligned} |\phi\rangle &= \alpha_{00}|ss\rangle + \alpha_{01}|ss'\rangle + \alpha_{10}|s's\rangle + \alpha_{11}|s's'\rangle \\ &= (\alpha_{00}|s\rangle + \alpha_{10}|s'\rangle) \otimes |s\rangle + (\alpha_{01}|s\rangle + \alpha_{11}|s'\rangle) \otimes |s'\rangle. \end{aligned} \quad (\text{B2})$$

Denote

$$\mathbf{A} = \begin{bmatrix} \alpha_{00} & \alpha_{01} \\ \alpha_{10} & \alpha_{11} \end{bmatrix},$$

and perform singular value decomposition on  $\mathbf{A}$ , i.e.,

$$\mathbf{A} = \mathbf{U}\mathbf{D}\mathbf{V},$$

where  $\mathbf{U}$ ,  $\mathbf{V}$  are unitary matrices and  $\mathbf{D}$  is a diagonal matrix. Since  $|\phi\rangle$  is a maximally entangled state of two qubits, all the singular values of  $\mathbf{A}$  must be  $\frac{1}{\sqrt{2}}$ . Hence,  $\mathbf{D} = \frac{1}{\sqrt{2}}\mathbf{I}_2$ , and  $\mathbf{A}$  can be rewritten as

$$\mathbf{A} = \frac{1}{\sqrt{2}}\mathbf{U}\mathbf{V} = \frac{1}{\sqrt{2}}\tilde{\mathbf{U}}. \quad (\text{B3})$$

Since  $\mathbf{U}$ ,  $\mathbf{V}$  are unitary matrices, so is  $\tilde{\mathbf{U}}$ . Define

$$\begin{aligned} |\tilde{s}\rangle &= \sqrt{2}(\alpha_{00}|s\rangle + \alpha_{10}|s'\rangle) \\ |\tilde{s}'\rangle &= \sqrt{2}(\alpha_{01}|s\rangle + \alpha_{11}|s'\rangle). \end{aligned} \quad (\text{B4})$$

Then from (B3), since  $\tilde{\mathbf{U}}$  is unitary,  $\{|\tilde{s}\rangle, |\tilde{s}'\rangle\}$  is also a basis of a qubit. Substituting (B4) into (B2), one can obtain (B1). This completes the proof.  $\square$

*Remark 3. Comparison with Schmidt decomposition.* In Lemma 1, the decomposition of the maximally entangled state, i.e., (B1), takes the form of Schmidt decomposition. However, Lemma 1 is not a special case of the Schmidt decomposition theorem. This is because the Schmidt decomposition theorem shows that there exists some basis  $\{|s\rangle, |s'\rangle\}$  and  $\{|\tilde{s}\rangle, |\tilde{s}'\rangle\}$  such that (B1) holds, while Lemma 1 shows that for all qubit bases  $\{|s\rangle, |s'\rangle\}$ , there exists  $\{|\tilde{s}\rangle, |\tilde{s}'\rangle\}$  such that (B1) holds. The ‘‘for all’’ requirement makes a stronger statement that enables us to save photon basis rotation at one node.  $\square$

The photon source generates photon pairs whose polarization state is maximally entangled, i.e.,

$$|\phi\rangle = \frac{1}{\sqrt{2}}(|h_A\rangle|h_B\rangle + e^{i\alpha}|v_A\rangle|v_B\rangle).$$

From Lemma 1, there exists some basis  $\{|\tilde{s}_A\rangle, |\tilde{s}'_A\rangle\}$  such that  $|\phi\rangle$  can be rewritten as

$$|\phi\rangle = \frac{1}{\sqrt{2}}(|\tilde{s}_A\rangle|s_B\rangle + |\tilde{s}'_A\rangle|s'_B\rangle). \quad (\text{B5})$$

From (B5), the polarization state prepared by the source can be viewed as a state in which the polarization basis of photon B is already aligned with the PSP basis of the channel. Hence, rotating photon A to align  $\{|\tilde{s}_A\rangle, |\tilde{s}'_A\rangle\}$  with the PSP basis  $\{|s_A\rangle, |s'_A\rangle\}$  is sufficient to reduce the possible coincident arrival times of the photon pair to two.

- 
- [1] C. H. Bennett and D. P. DiVincenzo, Quantum information and computation, *Nature* **404**, 247 (2000).
- [2] J. P. Dowling and G. J. Milburn, Quantum technology: The second quantum revolution, *Philos. Trans. R. Soc. Lond. A* **361**, 1655 (2003).
- [3] H. J. Kimble, The quantum Internet, *Nature* **453**, 1023 (2008).
- [4] S. Wehner, D. Elkouss, and R. Hanson, Quantum Internet: A vision for the road ahead, *Science* **362**, eaam9288 (2018).
- [5] J. Preskill, Quantum computing in the NISQ era and beyond, *Quantum* **2**, 79 (2018).
- [6] C. H. Bennett, G. Brassard, C. Crépeau, R. Jozsa, A. Peres, and W. K. Wootters, Teleporting an Unknown Quantum State via Dual Classical and Einstein-Podolsky-Rosen Channels, *Phys. Rev. Lett.* **70**, 1895 (1993).
- [7] M. A. Nielsen, E. Knill, and R. Laflamme, Complete quantum teleportation using nuclear magnetic resonance, *Nature* **396**, 52 (1998).
- [8] D. Gottesman and I. L. Chuang, Demonstrating the viability of universal quantum computation using teleportation and single-qubit operations, *Nature* **402**, 390 (1999).
- [9] C. H. Bennett and S. J. Wiesner, Communication via One- and Two-Particle Operators on Einstein-Podolsky-Rosen States, *Phys. Rev. Lett.* **69**, 2881 (1992).
- [10] V. Giovannetti, S. Lloyd, L. Maccone, and P. W. Shor, Entanglement Assisted Capacity of the Broadband Lossy Channel, *Phys. Rev. Lett.* **91**, 047901 (2003).
- [11] J. T. Barreiro, T.-C. Wei, and P. G. Kwiat, Beating the channel capacity limit for linear photonic superdense coding, *Nat. Phys.* **4**, 282 (2008).
- [12] A. K. Ekert, Quantum Cryptography Based on Bell’s Theorem, *Phys. Rev. Lett.* **67**, 661 (1991).
- [13] M. Koashi and J. Preskill, Secure Quantum Key Distribution with an Uncharacterized Source, *Phys. Rev. Lett.* **90**, 057902 (2003).
- [14] M. Epping, H. Kampermann, C. Macchiavello, and D. Bruß, Multi-partite entanglement can speed up quantum key distribution in networks, *New J. Phys.* **19**, 093012 (2017).
- [15] B. T. Kirby, S. Santra, V. S. Malinovsky, and M. Brodsky, Entanglement swapping of two arbitrarily degraded entangled states, *Phys. Rev. A* **94**, 012336 (2016).

- [16] M. Pant, H. Krovi, D. Towsley, L. Tassiulas, L. Jiang, P. Basu, D. Englund, and S. Guha, Routing entanglement in the quantum Internet, *npj, Quantum Inf.* **5**, 25 (2019).
- [17] W. Dai, T. Peng, and M. Z. Win, Optimal remote entanglement distribution, *IEEE J. Sel. Areas Commun.* **38**, 540 (2020).
- [18] W. Dai, T. Peng, and M. Z. Win, Quantum queuing delay, *IEEE J. Sel. Areas Commun.* **38**, 605 (2020).
- [19] W. Dai, T. Peng, and M. Z. Win, Remote state preparation for multiple parties, in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Process* (Brighton, UK, 2019), pp. 7983–7987.
- [20] M. Barbieri, F. De Martini, G. Di Nepi, P. Mataloni, G. M. D’Ariano, and C. Macchiavello, Detection of Entanglement with Polarized Photons: Experimental Realization of an Entanglement Witness, *Phys. Rev. Lett.* **91**, 227901 (2003).
- [21] A. Poppe, A. Fedrizzi, R. Ursin, H. Böhm, T. Lörünser, O. Maurhardt, M. Peev, M. Suda, C. Kurtsiefer, H. Weinfurter *et al.*, Practical quantum key distribution with polarization entangled photons, *Opt. Express* **12**, 3865 (2004).
- [22] S. X. Wang and G. S. Kanter, Robust multiwavelength all-fiber source of polarization-entangled photons with built-in analyzer alignment signal, *IEEE J. Select. Top. Quantum Electron.* **15**, 1733 (2009).
- [23] C. Antonelli, M. Shtaif, and M. Brodsky, Sudden Death of Entanglement Induced by Polarization Mode Dispersion, *Phys. Rev. Lett.* **106**, 080404 (2011).
- [24] M. Shtaif, C. Antonelli, and M. Brodsky, Nonlocal compensation of polarization mode dispersion in the transmission of polarization entangled photons, *Opt. Express* **19**, 1728 (2011).
- [25] M. Brodsky, E. C. George, C. Antonelli, and M. Shtaif, Loss of polarization entanglement in a fiber-optic system with polarization mode dispersion in one optical path, *Opt. Lett.* **36**, 43 (2011).
- [26] J. Gordon and H. Kogelnik, PMD fundamentals: Polarization mode dispersion in optical fibers, *Proc. Natl. Acad. Sci. USA* **97**, 4541 (2000).
- [27] C. H. Bennett, G. Brassard, S. Popescu, B. Schumacher, J. A. Smolin, and W. K. Wootters, Purification of Noisy Entanglement and Faithful Teleportation via Noisy Channels, *Phys. Rev. Lett.* **76**, 722 (1996).
- [28] C. H. Bennett, D. P. DiVincenzo, J. A. Smolin, and W. K. Wootters, Mixed-state entanglement and quantum error correction, *Phys. Rev. A* **54**, 3824 (1996).
- [29] A. E. Ulanov, I. A. Fedorov, A. A. Pushkina, Y. V. Kurochkin, T. C. Ralph, and A. I. Lvovsky, Undoing the effect of loss on quantum entanglement, *Nat. Photon.* **9**, 764 (2015).
- [30] F. Rozpedek, T. Schiet, L. P. Thinh, D. Elkouss, A. C. Doherty, and S. Wehner, Optimizing practical entanglement distillation, *Phys. Rev. A* **97**, 062333 (2018).
- [31] N. H. Nickerson, J. F. Fitzsimons, and S. C. Benjamin, Freely Scalable Quantum Technologies Using Cells of 5-to-50 Qubits with Very Lossy and Noisy Photonic Links, *Phys. Rev. X* **4**, 041041 (2014).
- [32] P. Kómár, E. M. Kessler, M. Bishof, L. Jiang, A. S. S. Rensen, J. Ye, and M. D. Lukin, A quantum network of clocks, *Nat. Phys.* **10**, 582 (2014).
- [33] J. Dehaene, M. Van den Nest, B. De Moor, and F. Verstraete, Local permutations of products of Bell states and entanglement distillation, *Phys. Rev. A* **67**, 022310 (2003).
- [34] K. G. H. Vollbrecht and F. Verstraete, Interpolation of recurrence and hashing entanglement distillation protocols, *Phys. Rev. A* **71**, 062325 (2005).
- [35] E. Hostens, J. Dehaene, and B. De Moor, Asymptotic adaptive bipartite entanglement-distillation protocol, *Phys. Rev. A* **73**, 062337 (2006).
- [36] R. Matsumoto, Conversion of a general quantum stabilizer code to an entanglement distillation protocol, *J. Phys. A: Math. Gen.* **36**, 8113 (2003).
- [37] A. Ambainis and D. Gottesman, The minimum distance problem for two-way entanglement purification, *IEEE Trans. Inf. Theory* **52**, 748 (2006).
- [38] S. Watanabe, R. Matsumoto, and T. Uyem, Improvement of stabilizer-based entanglement distillation protocols by encoding operators, *J. Phys. A: Math. Gen.* **39**, 4273 (2006).
- [39] D. Deutsch, A. Ekert, R. Jozsa, C. Macchiavello, S. Popescu, and A. Sanpera, Quantum Privacy Amplification and the Security of Quantum Cryptography Over Noisy Channels, *Phys. Rev. Lett.* **77**, 2818 (1996).
- [40] T. Opatrný and G. Kurizki, Optimization approach to entanglement distillation, *Phys. Rev. A* **60**, 167 (1999).
- [41] D. Mundarain and M. Orszag, Entanglement preservation by continuous distillation, *Phys. Rev. A* **79**, 052333 (2009).
- [42] D. Abdelkhalek, M. Syllwasschy, N. J. Cerf, J. Fiurášek, and R. Schnabel, Efficient entanglement distillation without quantum memory, *Nat. Commun.* **7**, 11720 (2016).
- [43] N. Kalb, A. A. Reiserer, P. C. Humphreys, J. J. W. Bakermans, S. J. Kamerling, N. H. Nickerson, S. C. Benjamin, D. J. Twitchen, M. Markham, and R. Hanson, Entanglement distillation between solid-state quantum network nodes, *Science* **356**, 928 (2017).
- [44] A. S. Fletcher, P. W. Shor, and M. Z. Win, Optimum quantum error recovery using semidefinite programming, *Phys. Rev. A* **75**, 012338 (2007).
- [45] A. S. Fletcher, P. W. Shor, and M. Z. Win, Structured near-optimal channel-adapted quantum error correction, *Phys. Rev. A* **77**, 012320 (2008).
- [46] A. S. Fletcher, P. W. Shor, and M. Z. Win, Channel-adapted quantum error correction for the amplitude damping channel, *IEEE Trans. Inf. Theory* **54**, 5705 (2008).
- [47] L. Ruan, W. Dai, and M. Z. Win, Adaptive recurrence quantum entanglement distillation for two-Kraus-operator channels, *Phys. Rev. A* **97**, 052332 (2018).
- [48] M. Brodsky, N. J. Frigo, M. Boroditsky, and M. Tur, Polarization mode dispersion of installed fibers, *J. Lightwave Technol.* **24**, 4584 (2006).
- [49] E. M. Rains, Bound on distillable entanglement, *Phys. Rev. A* **60**, 179 (1999).
- [50] E. M. Rains, A semidefinite program for distillable entanglement, *IEEE Trans. Inf. Theory* **47**, 2921 (2001).
- [51] X. Wang and R. Duan, Nonadditivity of Rains’ bound for distillable entanglement, *Phys. Rev. A* **95**, 062322 (2017).
- [52] H. Takesue and K. Inoue, Generation of polarization-entangled photon pairs and violation of Bell’s inequality using spontaneous four-wave mixing in a fiber loop, *Phys. Rev. A* **70**, 031802(R) (2004).
- [53] D. C. Burnham and D. L. Weinberg, Observation of Simultaneity in Parametric Production of Optical Photon Pairs, *Phys. Rev. Lett.* **25**, 84 (1970).