Efficient entanglement distillation for quantum channels with polarization mode dispersion

Liangzhong Ruan, Brian T. Kirby, Michael Brodsky, and Moe Z. Win

Efficient Entanglement Distillation for Quantum Channels with Polarization Mode Dispersion

Liangzhong Ruan,[1, 2] Moe Z. Win,[2] Brian T. Kirby,[3] and Michael Brodsky[3]

[1] *Tongji University, Shanghai, China*
[2] *Massachusetts Institute of Technology, Cambridge, MA*
[3] *U.S. Army Research Laboratory, Adelphi, MD*
(Dated: June 25, 2020)

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 on both the size of the quantum circuits involved and on the initial quality of entanglement in particles. However, because the number of particles required grows exponentially with the number of rounds of distillation, 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 maximal fidelity as well as the highest success probability (conditioned on the maximal fidelity being achieved) in every round of distillation. The attainment of the maximal fidelity improves the convergence speed of fidelity with respect to the rounds of distillation from linear to quadratic, and hence significantly reduces the number of distillation rounds. Combined with the fact that the highest success probability is achieved, the proposed algorithm provides an efficient method to distribute entangled states with high fidelity via optic fibers.

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 targeted 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 algorithms [39–41]. 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 targeted entangled state. These algorithms can mitigate the effect of strong decoherence by performing multiple rounds of distillations. 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 targeted 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 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 rounds of distillation, 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 rounds of distillation. To achieve this objective, the quantum privacy amplification (QPA) algorithm was proposed in [39], and was shown numerically to require fewer rounds of distillation 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 non-convex, 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 one-round distillation operations, 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 general 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 (QEC) 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 rounds of distillation [47]. This work optimizes the distillation operations to most efficiently mitigate the effect of PMD while achieving high success probability.

Organization: Section II analyzes the PMD effect on photon pairs, then defines 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 value. Section IV provides several numerical tests, which shows that by achieving the optimal fidelity and success probability, the proposed algorithm provides an efficient method to distribute entangled photons with high fidelity through quantum channels impaired by fiber birefringence. Finally, Section V

gives the conclusion.

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

II. SYSTEM MODEL AND PROBLEM FORMULATION

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

A. Effect of PMD on entangled photon pairs

Consider the quantum network illustrated in Fig. 1A, 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. 1B, 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) τ_A, τ_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 $\boldsymbol{\Xi}$ of the photon pair after passing through fibers is given by (1). This density matrix is written in the basis of $|s_A s_B\rangle$, $|s_A s_B'\rangle$, $|s_A' s_B\rangle$, and $|s_A' s_B'\rangle$. Please refer to Appendix A for the detailed derivation and the definition of the parameters in (1), i.e., η_1, η_2, α, and function $R(\cdot, \cdot)$. Denote the element in the p-th row and q-th column of $\boldsymbol{\Xi}$ as ξ_{pq}.

As illustrated in Fig. 1B 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 ($|s_A s_B\rangle$), slow-fast ($|s_A s_B'\rangle$), fast-slow ($|s_A' s_B\rangle$), and fast-fast ($|s_A' s_B'\rangle$). This results in a relatively complicated density matrix. As illustrated in Fig. 1C, to simplify the density matrix, one could align the PSP basis with the photon polarization basis, so that there are only two
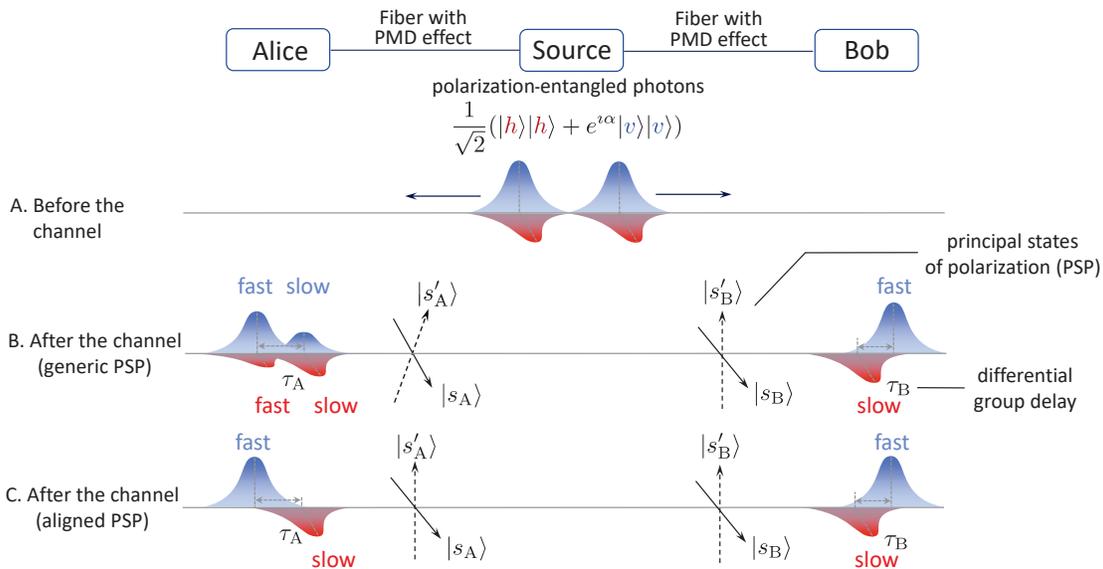
FIG. 1. System Model. The overall effect of PMD resembles that of pure birefringence in the sense that it causes an incident pulse to split into two orthogonally polarized components delayed relative to each other [26]. The polarization states of these two components are known as the PSP and the delay 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.

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

$$\Xi = \frac{1}{2} \begin{bmatrix} |\eta_1|^2 & -\eta_1\eta_2 e^{-\iota\alpha}R^{\dagger}(\tau_A,0) & \eta_1\eta_2^{\dagger}R^{\dagger}(0,\tau_B) & \eta_1^2 e^{-\iota\alpha}R^{\dagger}(\tau_A,\tau_B) \\ -\eta_1^{\dagger}\eta_2^{\dagger}e^{\iota\alpha}R(\tau_A,0) & |\eta_2|^2 & -(\eta_2^*)^2 e^{\iota\alpha}R(\tau_A,-\tau_B) & -\eta_1\eta_2^{\dagger}R^{\dagger}(0,\tau_B) \\ \eta_1^{\dagger}\eta_2 R(0,\tau_B) & -(\eta_2)^2 e^{-\iota\alpha}R^{\dagger}(\tau_A,-\tau_B) & |\eta_2|^2 & \eta_1\eta_2 e^{-\iota\alpha}R^{\dagger}(\tau_A,0) \\ (\eta_1^{\dagger})^2 e^{\iota\alpha}R(\tau_A,\tau_B) & -\eta_1^{\dagger}\eta_2 R(0,\tau_B) & \eta_1^{\dagger}\eta_2^{\dagger}e^{\iota\alpha}R(\tau_A,0) & |\eta_1|^2 \end{bmatrix} \qquad (1)$$

to perform local rotation on the photons before passing 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 realignment of these states would be rare, as the PSP in installed fiber optics 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 decoherence-free subspace (DFS) [23].

elements, which are given by

$$\xi_{11} = \xi_{44} = \frac{1}{2},$$

$$\xi_{41} = \xi_{14}^{\dagger} = \frac{1}{2}e^{\iota\alpha}R(\tau_A,\tau_B),$$

which can be rewritten as

$$\Xi = \frac{1}{2}\big(|s_A s_B\rangle\langle s_A s_B| + e^{-\iota\alpha}R^{\dagger}(\tau_A,\tau_B)|s_A s_B\rangle\langle s_A' s_B'|$$
$$+ e^{\iota\alpha}R(\tau_A,\tau_B)|s_A' s_B'\rangle\langle s_A s_B| + |s_A' s_B'\rangle\langle s_A' s_B'|\big). \quad (2)$$

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 non-zero

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 remove 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 k-th round of distillation as $\boldsymbol{\Xi}_k$, with $\boldsymbol{\Xi}_0 = \boldsymbol{\Xi}$. Then before the k-th round of distillation, the joint density matrix of two kept qubit pairs is given by

$$\boldsymbol{\Xi}_{k-1}^{\mathrm{J}} = \boldsymbol{\Xi}_{k-1} \otimes \boldsymbol{\Xi}_{k-1}.$$

Without loss of generality, assume that the network nodes try to keep the first qubit pair, i.e., the first and second qubits in the system. 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.,

$$\boldsymbol{\Xi}_k = \frac{\mathrm{tr}_{3,4}\left\{\mathbb{D}\left\{\boldsymbol{\Xi}_{k-1}^{\mathrm{J}}\right\}\right\}}{\mathrm{tr}\left\{\mathbb{D}\left\{\boldsymbol{\Xi}_{k-1}^{\mathrm{J}}\right\}\right\}}, \tag{3}$$

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

$$P_k = \mathrm{tr}\left\{\mathbb{D}\left\{\boldsymbol{\Xi}_{k-1}^{\mathrm{J}}\right\}\right\}. \tag{4}$$

Denote the fidelity of the kept qubit pairs after the k-th round of distillation w.r.t. to the targeted state as

$$F_k = \langle \Phi^+ | \boldsymbol{\Xi}_k | \Phi^+ \rangle, \tag{5}$$

where $|\Phi^+\rangle = \frac{1}{\sqrt{2}}(|h_{\mathrm{A}} h_{\mathrm{B}}\rangle + |v_{\mathrm{A}} v_{\mathrm{B}}\rangle)$. For notation convenience, denote the mapping between the input density matrix $\boldsymbol{\Xi}_{k-1}$ and the fidelity of the kept qubit pair F_k as $F_{\mathbb{D}}$, i.e.,

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

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

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

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

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

$$F_K \geq 1 - \epsilon, \tag{6}$$

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

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

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 rounds of distillation, i.e., maximize F_k. 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 $\boldsymbol{\Xi}$, we will maximize the fidelity of the kept qubit pair $F_{\mathbb{D}}(\boldsymbol{\Xi})$ w.r.t. the distillation operation $\mathbb{D}$. This problem can be formulated as

$$\mathscr{P}_F : \quad \max_{\mathbb{D}} F_{\mathbb{D}}(\boldsymbol{\Xi}).$$

Denote the optimal fidelity as $F^*(\boldsymbol{\Xi})$. We will maximize the success probability of the distillation operation $P_{\mathbb{D}}(\boldsymbol{\Xi})$ w.r.t. the distillation operation $\mathbb{D}$ conditional on the optimal fidelity being achieved. This problem can be formulated as:

$$\mathscr{P}_P : \quad \max_{\mathbb{D}} P_{\mathbb{D}}(\boldsymbol{\Xi})$$
$$\mathrm{s.t.} \quad F_{\mathbb{D}}(\boldsymbol{\Xi}) = F^*(\boldsymbol{\Xi}).$$

III. EFFICIENT QED FOR PMD CHANNELS

This section first characterizes the optimal value of problems $\mathscr{P}_F$ and $\mathscr{P}_P$, and then gives an algorithm which achieves the optimal performance in every round of distillation. For conciseness, in the following, both $|h_{\mathrm{A}}\rangle$ and $|h_{\mathrm{B}}\rangle$ are denoted as $|0\rangle$, and both $|v_{\mathrm{A}}\rangle$ and $|v_{\mathrm{B}}\rangle$ 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 the density matrices given in (2), and characterizes the corresponding optimal performance of problems $\mathscr{P}_F$ and $\mathscr{P}_P$. Specifically, the set of density matrices is defined as

$$\mathcal{S} = \{\boldsymbol{\Xi} \text{ that satisfies (8)}\}$$

where

$$\boldsymbol{\Xi} = \frac{1}{2}\big(|ab\rangle\langle ab| + e^{-\imath\alpha} R^{\dagger} |ab\rangle\langle a'b'| \\ + e^{\imath\alpha} R |a'b'\rangle\langle ab| + |a'b'\rangle\langle a'b'|\big), \tag{8}$$

in which

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

The density matrix $\boldsymbol{\Xi}$ can be simplified as follows. By performing matrix spectral decomposition, it can be obtained that

$$\boldsymbol{\Xi} = F|\phi_1\rangle\langle\phi_1| + (1-F)|\phi_2\rangle\langle\phi_2|, \qquad (9)$$

where

$$F = \frac{1}{2}(1 + |R|),$$

$$|\phi_1\rangle = \frac{1}{\sqrt{2}}\big(|ab\rangle + e^{\imath\theta}|a'b'\rangle\big),$$

$$|\phi_2\rangle = \frac{1}{\sqrt{2}}\big(|ab\rangle - e^{\imath\theta}|a'b'\rangle\big),$$

$$\theta = \alpha + \mathrm{Phase}\{R\}.$$

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

Theorem 1 (Optimal fidelity)**:** Let $\boldsymbol{\Xi} \in \mathcal{S}$. Then the optimal value of $\mathscr{P}_F$ is given by

$$F^*(\boldsymbol{\Xi}) = \frac{F^2}{F^2 + (1-F)^2}. \qquad (10)$$

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

$$\boldsymbol{U}_{\mathrm{A}} = \frac{|0\rangle + |1\rangle}{\sqrt{2}}\langle a| + \frac{|0\rangle - |1\rangle}{\sqrt{2}}\langle a'|,$$
$$\boldsymbol{U}_{\mathrm{B}} = \frac{|0\rangle + |1\rangle}{\sqrt{2}}\langle b| + e^{-\imath\theta}\frac{|0\rangle - |1\rangle}{\sqrt{2}}\langle b'|, \qquad (11)$$

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

$$\check{\boldsymbol{\Xi}} = (\boldsymbol{U}_{\mathrm{A}} \otimes \boldsymbol{U}_{\mathrm{B}})\,\boldsymbol{\Xi}\,(\boldsymbol{U}_{\mathrm{A}} \otimes \boldsymbol{U}_{\mathrm{B}})^\dagger$$
$$= F|\Phi^+\rangle\langle\Phi^+| + (1-F)|\Psi^+\rangle\langle\Psi^+|, \qquad (12)$$

where

$$|\Phi^+\rangle = \frac{1}{\sqrt{2}}\big(|00\rangle + |11\rangle\big),$$

$$|\Psi^+\rangle = \frac{1}{\sqrt{2}}\big(|01\rangle + |10\rangle\big).$$

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

$$F^*(\check{\boldsymbol{\Xi}}) = \frac{F^2}{F^2 + (1-F)^2}.$$

Moreover, since unitary operations are reversible, $F^*(\check{\boldsymbol{\Xi}}) = F^*(\boldsymbol{\Xi})$. This completes the proof. $\square$

The next theorem characterizes the upper bound of the success probability conditioned on the optimal fidelity having been achieved.

Theorem 2 (Optimal probability of success)**:** Let $\boldsymbol{\Xi} \in \mathcal{S}$ with $|R| > 0$. Then the optimal value of $\mathscr{P}_P$ is given by

$$P^*(\boldsymbol{\Xi}) = F^2 + (1-F)^2. \qquad (13)$$

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

$$P^*(\boldsymbol{\Xi}) \le F^2 + (1-F)^2. \qquad (14)$$

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

$$F_{\mathbb{D}}(\boldsymbol{\Xi}) = \frac{F^2}{F^2 + (1-F)^2}, \qquad (15)$$

$$P_{\mathbb{D}}(\boldsymbol{\Xi}) > F^2 + (1-F)^2. \qquad (16)$$

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

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

Define

$$\boldsymbol{V}_{nm} = \mathrm{tr}_{3,4}\big\{\mathbb{D}\{|\phi_n\phi_m\rangle\langle\phi_n\phi_m|\}\big\},$$
$$f_{nm} = \langle\Phi^+|\boldsymbol{V}_{nm}|\Phi^+\rangle,$$
$$p_{nm} = \mathrm{tr}\{\boldsymbol{V}_{nm}\},$$

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

$$0 \le f_{nm} \le p_{nm} \le 1. \qquad (17)$$

It is straight forward that

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

$$P_{\mathbb{D}}(\boldsymbol{\Xi}) = F^2 p_{11} + F(1-F)(p_{12} + p_{21}) + (1-F)^2 p_{22}. \qquad (19)$$

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

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

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)

$$F_{\mathbb{D}}(\boldsymbol{\Xi}) = \frac{S(F)}{S(F) + N(F)} = \frac{F^2}{F^2 + (1-F)^2}$$
$$\Rightarrow \frac{N(F)}{S(F)} = \frac{(1-F)^2}{F^2}. \qquad (21)$$

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

$$F_{\mathbb{D}}(\tilde{\boldsymbol{\Xi}}) = \frac{S(\tilde{F})}{S(\tilde{F}) + N(\tilde{F})} = \frac{1}{1 + \frac{N(\tilde{F})}{S(\tilde{F})}}. \qquad (22)$$

From (17) and (20), if $f_{12} + f_{21} = p_{12} + p_{21} > 0$, then

$$
\begin{aligned}
S(\tilde{F}) &= \frac{\tilde{F}^2}{F^2}\Big(F^2 f_{11} + \frac{F^2}{\tilde{F}}(1 - \tilde{F})(f_{12} + f_{21}) \\
&\quad + \frac{F^2}{\tilde{F}^2}(1 - \tilde{F})^2 f_{22}\Big) \\
&> \frac{\tilde{F}^2}{F^2}\Big(F^2 f_{11} + F(1 - F)(f_{12} + f_{21}) + (1 - F)^2 f_{22}\Big) \\
&= \frac{\tilde{F}^2}{F^2}S(F), \qquad (23)
\end{aligned}
$$

$$
\begin{aligned}
N(\tilde{F}) &= \frac{(1 - \tilde{F})^2}{(1 - F)^2}\Big(\frac{(1 - F)^2}{(1 - \tilde{F})^2}\tilde{F}^2(p_{11} - f_{11}) \\
&\quad + \tilde{F}\frac{(1 - F)^2}{(1 - \tilde{F})}(p_{12} + p_{21} - f_{12} - f_{21}) \\
&\quad + (1 - F)^2(p_{22} - f_{22})\Big) \\
&\leq \frac{(1 - \tilde{F})^2}{(1 - F)^2}\Big(F^2(p_{11} - f_{11}) \\
&\quad + F(1 - F)(p_{12} + p_{21} - f_{12} - f_{21}) \\
&\quad + (1 - F)^2(p_{22} - f_{22})\Big) \\
&= \frac{(1 - \tilde{F})^2}{(1 - F)^2}N(F). \qquad (24)
\end{aligned}
$$

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

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

which leads to

$$F^*(\tilde{\boldsymbol{\Xi}}) \geq F_{\mathbb{D}}(\tilde{\boldsymbol{\Xi}}) > \frac{\tilde{F}^2}{\tilde{F}^2 + (1 - \tilde{F})^2}. \qquad (25)$$

However, (25) contradicts with (10).

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

$$S(\tilde{F}) \geq \frac{\tilde{F}^2}{F^2}S(F), \text{ and}$$

$$N(\tilde{F}) < \frac{(1 - \tilde{F})^2}{(1 - F)^2}N(F),$$

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

The achievability of (13) will be proved constructively with the QED algorithm to be proposed. Please refer to Section III B for details. □

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):

- **Local state preparation:** For each qubit pair, the network nodes transform the density matrix to $\tilde{\boldsymbol{\Xi}}$ using local unitary operators $\boldsymbol{U}_A$ and $\boldsymbol{U}_B$ defined in (11).

- **First round distillation:** The nodes take two of the kept qubit pairs and perform the following operations.

  (i) Each node locally performs CNOT operation, i.e., $\boldsymbol{U} = |00\rangle\langle00| + |01\rangle\langle01| + |10\rangle\langle11| + |11\rangle\langle10|$ on the two qubits at hand.

  (ii) Each node measures the target bit (i.e., the qubit in the second pair) using operators $|0\rangle\langle0|$, $|1\rangle\langle1|$, 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.

- **Following rounds:** Network nodes perform the same operations as in the first round, until the fidelity of the kept qubit pairs exceeds the required threshold. □

In the following, we will first characterize the performance of the proposed algorithm in Theorem 3, then explain the implications of this theorem in two remarks.

Theorem 3 (Performance of the proposed algorithm): In the k-th round of distillation, the source qubit pair is kept with fidelity

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

probability

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

and density matrix

$$\boldsymbol{\Xi}^{(k)} = F_k|\Phi^+\rangle\langle\Phi^+| + (1 - F_k)|\Psi^+\rangle\langle\Psi^+|. \qquad (28)$$

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

$$\boldsymbol{\Xi}_{\mathrm{J}} = \boldsymbol{P}\breve{\boldsymbol{\Xi}} \otimes \breve{\boldsymbol{\Xi}}\, \boldsymbol{P}^{\dagger}$$
$$= F^2 |\Omega^{(1)}\rangle\langle\Omega^{(1)}|$$
$$+ F(1-F)\big(|\Omega^{(2)}\rangle\langle\Omega^{(2)}| + |\Omega^{(3)}\rangle\langle\Omega^{(3)}|\big)$$
$$+ (1-F)^2 |\Omega^{(4)}\rangle\langle\Omega^{(4)}|$$

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

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

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

$$\breve{\boldsymbol{\Xi}}_{\mathrm{J}} = F^2 |\breve{\Omega}^{(1)}\rangle\langle\breve{\Omega}^{(1)}| + F(1-F)\big(|\breve{\Omega}^{(2)}\rangle\langle\breve{\Omega}^{(2)}|$$
$$+ |\breve{\Omega}^{(3)}\rangle\langle\breve{\Omega}^{(3)}|\big) + (1-F)^2 |\breve{\Omega}^{(4)}\rangle\langle\breve{\Omega}^{(4)}|, \quad (29)$$

where

$$
\begin{aligned}
|\breve{\Omega}^{(1)}\rangle =\ & \tfrac{1}{2}|0000\rangle + \tfrac{1}{2}|0101\rangle \\
& + \tfrac{1}{2}|1111\rangle + \tfrac{1}{2}|1010\rangle \\
|\breve{\Omega}^{(2)}\rangle =\ & \tfrac{1}{2}|0001\rangle + \tfrac{1}{2}|0100\rangle \\
& + \tfrac{1}{2}|1110\rangle + \tfrac{1}{2}|1011\rangle
\end{aligned}
$$

$$
\begin{aligned}
|\breve{\Omega}^{(3)}\rangle\ =\ & \tfrac{1}{2}|0011\rangle + \tfrac{1}{2}|0110\rangle \\
& + \tfrac{1}{2}|1100\rangle + \tfrac{1}{2}|1001\rangle \\
|\breve{\Omega}^{(4)}\rangle\ =\ & \tfrac{1}{2}|0010\rangle + \tfrac{1}{2}|0111\rangle \\
& + \tfrac{1}{2}|1101\rangle + \tfrac{1}{2}|1000\rangle .
\end{aligned}
$$

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

$$\boldsymbol{\Xi}_{00} = \big(\mathbb{I}_2 \otimes \langle 0| \otimes \mathbb{I}_2 \otimes \langle 0|\big)\, \breve{\boldsymbol{\Xi}}_{\mathrm{J}}\, \big(\mathbb{I}_2 \otimes |0\rangle \otimes \mathbb{I}_2 \otimes |0\rangle\big)$$
$$= \frac{1}{2}\big(F^2 |\Phi^+\rangle\langle\Phi^+| + (1-F)^2 |\Psi^+\rangle\langle\Psi^+|\big). \quad (30)$$

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

$$\boldsymbol{\Xi}_{11} = \big(\mathbb{I}_2 \otimes \langle 1| \otimes \mathbb{I}_2 \otimes \langle 1|\big)\, \breve{\boldsymbol{\Xi}}_{\mathrm{J}}\, \big(\mathbb{I}_2 \otimes |1\rangle \otimes \mathbb{I}_2 \otimes |1\rangle\big)$$
$$= \frac{1}{2}\big(F^2 |\Phi^+\rangle\langle\Phi^+| + (1-F)^2 |\Psi^+\rangle\langle\Psi^+|\big). \quad (31)$$

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

$$P = \mathrm{tr}\{\boldsymbol{\Xi}_{00} + \boldsymbol{\Xi}_{11}\} = F^2 + (1-F)^2, \quad (32)$$

the fidelity of the kept qubit pairs is

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

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

$$\boldsymbol{\Xi}^{(1)} = \frac{\boldsymbol{\Xi}_{00} + \boldsymbol{\Xi}_{11}}{P} = F_1 |\Phi^+\rangle\langle\Phi^+| + (1-F_1)|\Psi^+\rangle\langle\Psi^+|. \quad (34)$$

With (32) and (33), the proof for the first round of distillation is complete. For the following rounds of distillations, one can take (34) as input, and repeat the analysis in (29)–(33). This competes 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), (13) with (26), (27), one can see that the proposed algorithm achieves the optimal fidelity and the corresponding optimal success probability in every round of distillation. As will be verified in Section IV, this feature enables the proposed algorithm to achieve high efficiency. $\square$

Remark 2 (Convergence speed of fidelity)**:** In terms of the convergence speed of fidelity w.r.t. the rounds of distillation, 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

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

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

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

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

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

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 iteration 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. □

IV. NUMERICAL RESULTS

We will now demonstrate the dependence of the proposed recurrence distillation 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 the pulsed pump laser and frequency response of the filters are Gaussian [24]. Please see the last paragraph of 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 of the Rain's bound [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 how robust the proposed algorithm is to basis alignment errors.

In the numerical tests, the targeted fidelity is set to be 0.99. The round of distillation K is set to be the minimum round that achieves the targeted 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 τ_B, the pump laser bandwidth B_p, and η, which specifies the alignment between the qubit and PSP basis.

Figs. 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. Fig. 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 as compared to the DGD. Alternatively, Fig. 3 plots a case where a pump bandwidth is on 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 τ_A/τ_B centered around the DFS 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
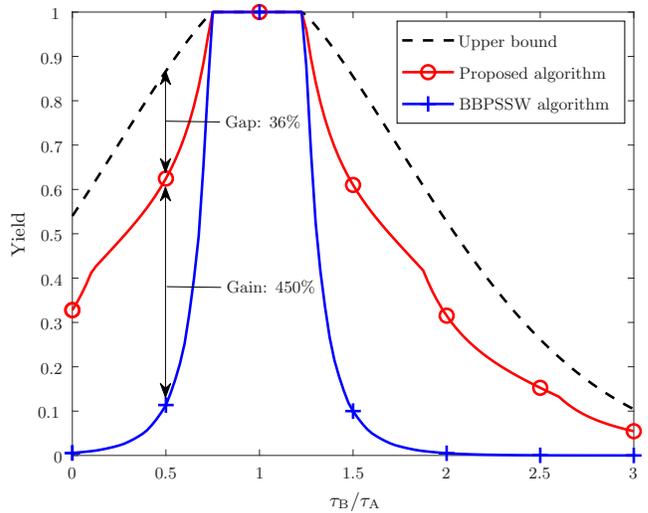
FIG. 2. Comparison of the yield as a function of τ_B/τ_A for the proposed algorithm and the benchmarks, i.e., the upper bound [49] and the recurrence QED algorithm proposed in [27] (referred to as the BBPSSW algorithm here). In this plot, $B_p = 0.1, B_A = B_B = 1$, $\tau_A = 1$.
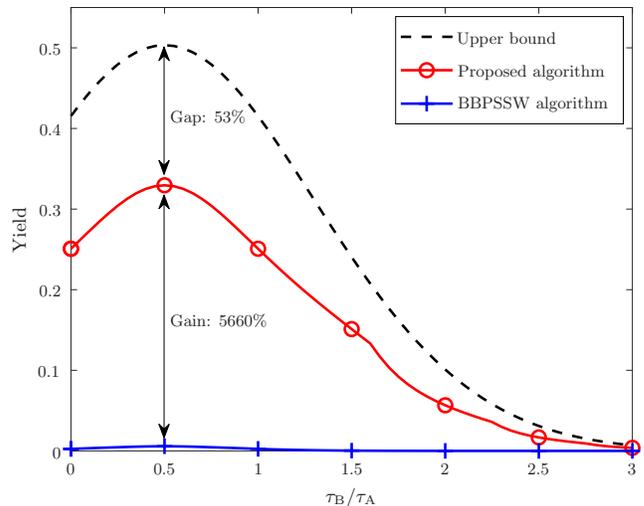
FIG. 3. Comparison of the yield as a function of τ_B/τ_A for the proposed algorithm and the benchmarks, i.e., the upper bound [49] and the BBPSSW algorithm [27]. In this plot, $B_p = 1, B_A = B_B = 1$, $\tau_A = 1$.

significantly higher than the baseline algorithm from [27] and is reasonably close to the best known upper bound. For instance, when $\tau_B/\tau_A = 0.5$, the proposed algorithm increases the yield from 450% to 5660% compared to the baseline algorithm and is 36% to 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 in Fig. 3 is shifted away from $\tau_A = \tau_B$, as opposed to the peak being cen-
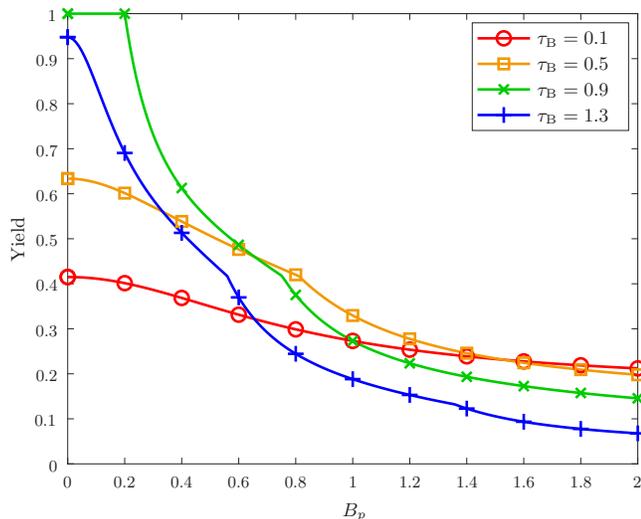
FIG. 4. The efficiency of the proposed algorithm as a function of the bandwidth of the source laser pump. In this figure, $B_A = B_B = 1$, $\tau_A = 1$.

FIG. 5. The output fidelity and the efficiency of the proposed algorithm as a function of the misalignment angle θ. $\eta_1 = \arcsin(\frac{\theta \pi}{180})$. 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.

tered 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 τ_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 τ_B increases. However, when B_p is small, the yield of the algorithm is highest when the values of τ_A, τ_B are similar, illustrating the benefits of the DFS 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 Section II A, such an alignment is not expected to be performed frequently, as the PSP of installed fiber optics has been shown to remain unchanged on the timescale 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 PSP basis as θ, 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 misalignment angle θ for several values of τ, where $\tau_A = \tau_B = \tau$. The output fidelities
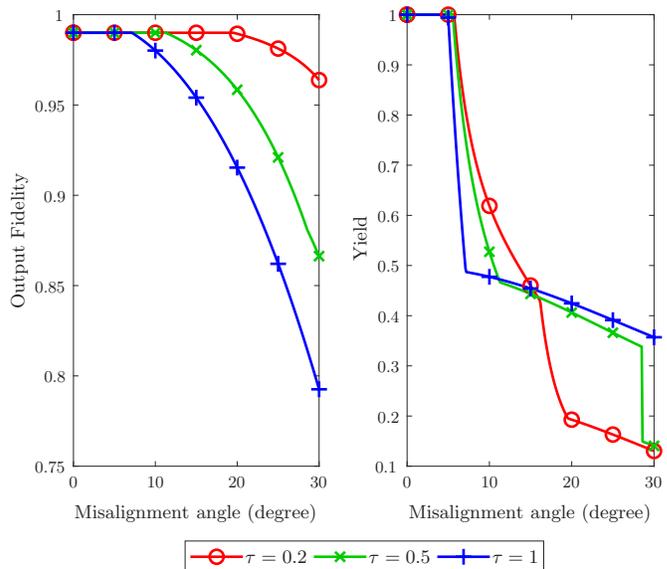
shown in the plot are the maximum achievable fidelity with the proposed algorithm with a required fidelity of 0.99. It can be seen that for all considered values of τ, the algorithm can generate qubit pairs with the required fidelity when the misalignment angle is no more than 5 degrees. When the misalignment angle θ is greater than 5 degrees, the output fidelities are higher for smaller values of τ, 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 when the misalignment angle θ is around 5 degrees. This means that, even though the algorithm can still obtain photon pairs with high fidelity when $\theta > 5°$, 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.

Fig. 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

Recurrence QED algorithms have good implementability and robustness, but improving their efficiency remains an interesting challenge. This work adopts recurrence QED algorithms to obtain high-quality entangle-

ment 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 optimal fidelity being achieved. We then proposed a recurrence QED algorithm that achieves both optimal fidelity and success probability in every round of distillation. Analytical results show that the proposed algorithm improves the convergence speed of fidelity w.r.t. the rounds of distillation 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.

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 the scenario with $\chi^{(2)}$ media and a pulsed pump was analyzed in [24]. Here we present an analytical treatment for $\chi^{(2)}$ media and a pulsed pump, and will also 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

$$|\psi\rangle = |f(t_{\mathrm{A}}, t_{\mathrm{B}})\rangle \otimes \frac{1}{\sqrt{2}}\big(|h_{\mathrm{A}}\rangle|h_{\mathrm{B}}\rangle + e^{\imath\alpha}|v_{\mathrm{A}}\rangle|v_{\mathrm{B}}\rangle\big), \quad (\mathrm{A1})$$

where h_i and v_i are orthogonal polarization basis states of photons A and B. The term $|f(t_{\mathrm{A}}, t_{\mathrm{B}})\rangle$ describes the time component of the state and is given by

$$|f(t_{\mathrm{A}}, t_{\mathrm{B}})\rangle = \int\int dt_{\mathrm{A}}dt_{\mathrm{B}} f(t_{\mathrm{A}}, t_{\mathrm{B}})|t_{\mathrm{A}}, t_{\mathrm{B}}\rangle. \quad (\mathrm{A2})$$

The function $|f(t_{\mathrm{A}}, t_{\mathrm{B}})|^2$ is proportional to the probability that the two photons overlap in time, and $\int dt_{\mathrm{A}}dt_{\mathrm{B}}|f(t_{\mathrm{A}}, t_{\mathrm{B}})|^2 = 1$. Specifically, since the entanglement is generated via a $\chi^{(2)}$ media, this function can

be written as

$$f(t_{\mathrm{A}}, t_{\mathrm{B}}) = \int dt H_{\mathrm{A}}^*(t - t_{\mathrm{A}})H_{\mathrm{B}}^*(t - t_{\mathrm{B}})E_p(t), \quad (\mathrm{A3})$$

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, CW and pulsed, are characterized by the envelope of the pump signal $E_p(t)$ and its Fourier transform $\tilde{E}_p(\omega)$, 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, and can result in wider bandwidth signal and idler photons, which enables multiple channels. For CW lasers, $|\tilde{E}_p(\omega)|^2$ approaches a delta function, which is a constant in the time domain. In this case, $f(t_{\mathrm{A}}, t_{\mathrm{B}})$ becomes a function of only the time difference, removing any absolute reference and hence simplifies analysis.

The effect of PMD is to advance or delay 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

$$
\begin{aligned}
|\psi\rangle = |f(t_{\mathrm{A}}, t_{\mathrm{B}})\rangle \otimes \Big[& \frac{\eta_1}{\sqrt{2}}\big(|s_{\mathrm{A}}\rangle|s_{\mathrm{B}}\rangle + e^{\imath\alpha_1}|s_{\mathrm{A}}'\rangle|s_{\mathrm{B}}'\rangle\big) \\
& + \frac{\eta_2}{\sqrt{2}}\big(|s_{\mathrm{A}}\rangle|s_{\mathrm{B}}'\rangle - e^{\imath\alpha_2}|s_{\mathrm{A}}'\rangle|s_{\mathrm{B}}\rangle\big) \Big],
\end{aligned}
\tag{A4}
$$

where

$$
\begin{aligned}
\eta_1 &= (s_{\mathrm{A}} \cdot h_{\mathrm{A}})(s_{\mathrm{B}} \cdot h_{\mathrm{B}}) + e^{\imath\alpha}(s_{\mathrm{A}} \cdot v_{\mathrm{A}})(s_{\mathrm{B}} \cdot v_{\mathrm{B}}), \\
\eta_2 &= (s_{\mathrm{A}} \cdot h_{\mathrm{A}})(s_{\mathrm{B}}' \cdot h_{\mathrm{B}}) + e^{\imath\alpha}(s_{\mathrm{A}} \cdot v_{\mathrm{A}})(s_{\mathrm{B}}' \cdot v_{\mathrm{B}}),
\end{aligned}
$$

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

$$
\begin{aligned}
|\psi_{\mathrm{PMD}}\rangle = & \frac{\eta_1}{\sqrt{2}}\big|f\big(t_{\mathrm{A}} - \tfrac{\tau_{\mathrm{A}}}{2}, t_{\mathrm{B}} - \tfrac{\tau_{\mathrm{B}}}{2}\big)\big\rangle \otimes |s_{\mathrm{A}}s_{\mathrm{B}}\rangle + \\
& \frac{\eta_2}{\sqrt{2}}\big|f\big(t_{\mathrm{A}} - \tfrac{\tau_{\mathrm{A}}}{2}, t_{\mathrm{B}} + \tfrac{\tau_{\mathrm{B}}}{2}\big)\big\rangle \otimes |s_{\mathrm{A}}s_{\mathrm{B}}'\rangle - \\
& \frac{\eta_2 e^{\imath\alpha_2}}{\sqrt{2}}\big|f\big(t_{\mathrm{A}} + \tfrac{\tau_{\mathrm{A}}}{2}, t_{\mathrm{B}} - \tfrac{\tau_{\mathrm{B}}}{2}\big)\big\rangle \otimes |s_{\mathrm{A}}'s_{\mathrm{B}}\rangle + \\
& \frac{\eta_1 e^{\imath\alpha_1}}{\sqrt{2}}\big|f\big(t_{\mathrm{A}} + \tfrac{\tau_{\mathrm{A}}}{2}, t_{\mathrm{B}} + \tfrac{\tau_{\mathrm{A}}}{2}\big)\big\rangle \otimes |s_{\mathrm{A}}'s_{\mathrm{B}}'\rangle. \quad (\mathrm{A5})
\end{aligned}
$$

We assume that the coincidence time window of the two photon detectors is much larger than the DGD τ_{A}, τ_{B}, so that photon pair can be detected correctly. Then, to account for the fact that the photo detection 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

basis of $|s_A s_B\rangle$, $|s_A s'_B\rangle$, $|s'_A s_B\rangle$, and $|s'_A s'_B\rangle$, the density matrix resulting from integration of time results is given by (1), in which

$$R(\tau_A, \tau_B) = \int \int dt_A dt_B f(t_A + \tau_A, t_B + \tau_B) f^*(t_A, t_B),$$

(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 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 will be assumed to be Gaussian. Under this assumption, the form of $R(\tau_A, \tau_B)$ is given by [25]

$$R(\tau_A, \tau_B) = \kappa \int \int d\omega_A d\omega_B |H_A(\omega_A)|^2 |H_B(\omega_B)|^2$$
$$|\tilde{E}_p(\omega_A + \omega_B)|^2 e^{i(\tau_A \omega_A + \tau_B \omega_B)},$$

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

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

Appendix B: Local Rotation on One Photon is Sufficient for Alignment

We will 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 (The basis of maximally entangled states): $|\phi\rangle$ is 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

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

(B1)

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

$$|\phi\rangle = \alpha_{00}|ss\rangle + \alpha_{01}|ss'\rangle + \alpha_{10}|s's\rangle + \alpha_{11}|s's'\rangle$$
$$= \big(\alpha_{00}|s\rangle + \alpha_{10}|s'\rangle\big) \otimes |s\rangle + \big(\alpha_{01}|s\rangle + \alpha_{11}|s'\rangle\big) \otimes |s'\rangle.$$

(B2)

Denote $\boldsymbol{A} = \begin{bmatrix} \alpha_{00} & \alpha_{01} \\ \alpha_{10} & \alpha_{11} \end{bmatrix}$, and perform singular value decomposition on $\boldsymbol{A}$, i.e.,

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

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

$$\boldsymbol{A} = \frac{1}{\sqrt{2}} \boldsymbol{U}\boldsymbol{V} = \frac{1}{\sqrt{2}} \tilde{\boldsymbol{U}}.$$

(B3)

Since $\boldsymbol{U}$, $\boldsymbol{V}$ are unitary matrices, so is $\tilde{\boldsymbol{U}}$. Denote

$$\big[|\tilde{s}\rangle \quad |\tilde{s}'\rangle\big] = \big[|s\rangle \quad |s'\rangle\big] \tilde{\boldsymbol{U}}.$$

(B4)

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

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 basis $\{|s\rangle, |s'\rangle\}$, 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. ∎

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

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

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

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

(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, Nature **404**, 247 (2000).

[2] J. P. Dowling and G. J. Milburn, Phil. Trans. R. Soc. Lond. A **361**, 1655 (2003).

[3] H. J. Kimble, Nature **453**, 1023 (2008).

[4] S. Wehner, D. Elkouss, and R. Hanson, Science **362**, 1 (2018).

[5] J. Preskill, Quantum **2**, 79 (2018).

[6] C. H. Bennett, G. Brassard, C. Crépeau, R. Jozsa, A. Peres, and W. K. Wootters, Phys. Rev. Lett. **70**, 1895 (1993).

[7] M. A. Nielsen, E. Knill, and R. Laflamme, Nature **396**, 52 (1998).

[8] D. Gottesman and I. L. Chuang, Nature **402**, 390 (1999).

[9] C. H. Bennett and S. J. Wiesner, Phys. Rev. Lett. **69**, 2881 (1992).

[10] V. Giovannetti, S. Lloyd, L. Maccone, and P. W. Shor, Phys. Rev. Lett. **91**, 047901 (2003).

[11] J. T. Barreiro, T.-C. Wei, and P. G. Kwiat, Nature Physics **4**, 282 (2008).

[12] A. K. Ekert, Phys. Rev. Lett. **67**, 661 (1991).

[13] M. Koashi and J. Preskill, Phys. Rev. Lett. **90**, 057902 (2003).

[14] M. Epping, H. Kampermann, C. Macchiavello, and D. Bruß, New Journal of Physics **19**, 093012 (2017).

[15] B. T. Kirby, S. Santra, V. S. Malinovsky, and M. Brodsky, Phys. Rev. A **94**, 012336 (2016).

[16] M. Pant, H. Krovi, D. Towsley, L. Tassiulas, L. Jiang, P. Basu, D. Englund, and S. Guha, npj, Quantum Information **5**, 1 (2019).

[17] W. Dai, T. Peng, and M. Z. Win, IEEE J. Sel. Areas Commun. **38**, 540 (2020).

[18] W. Dai, T. Peng, and M. Z. Win, IEEE J. Sel. Areas Commun. **38**, 605 (2020).

[19] W. Dai, T. Peng, and M. Z. Win, in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Process.* (Brighton, United Kingdom, 2019) pp. 7983–7987.

[20] M. Barbieri, F. De Martini, G. Di Nepi, P. Mataloni, G. M. D'Ariano, and C. Macchiavello, Phys. Rev. Lett. **91**, 227901 (2003).

[21] A. Poppe, A. Fedrizzi, R. Ursin, H. Böhm, T. Lörunser, O. Maurhardt, M. Peev, M. Suda, C. Kurtsiefer, H. Weinfurter, *et al.*, Optics Express **12**, 3865 (2004).

[22] S. X. Wang and G. S. Kanter, IEEE Journal of selected topics in quantum electronics **6**, 1733 (2009).

[23] C. Antonelli, M. Shtaif, and M. Brodsky, Physical review letters **106**, 080404 (2011).

[24] M. Shtaif, C. Antonelli, and M. Brodsky, Optics express **19**, 1728 (2011).

[25] M. Brodsky, E. C. George, C. Antonelli, and M. Shtaif, Optics letters **36**, 43 (2011).

[26] J. Gordon and H. Kogelnik, Proceedings of the National Academy of Sciences **97**, 4541 (2000).

[27] C. H. Bennett, G. Brassard, S. Popescu, B. Schumacher, J. A. Smolin, and W. K. Wootters, Phys. Rev. Lett. **76**, 722 (1996).

[28] C. H. Bennett, D. P. DiVincenzo, J. A. Smolin, and W. K. Wootters, 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, Nature Photonics **9**, 764 (2015).

[30] F. Rozpedek, T. Schiet, L. P. Thinh, D. Elkouss, A. C. Doherty, and S. Wehner, Phys. Rev. A **97**, 062333 (2018).

[31] N. H. Nickerson, J. F. Fitzsimons, and S. C. Benjamin, Phys. Rev. X **4**, 041041 (2014).

[32] P. Kómár, E. M. Kessler, M. Bishof, L. Jiang, A. S. Sørensen, J. Ye, and M. D. Lukin, Nature Physics **10**, 582 (2014).

[33] J. Dehaene, M. V. den Nest, B. D. Moor, and F. Verstraete, Phys. Rev. A **67**, 022310 (2003).

[34] K. G. H. Vollbrecht and F. Verstraete, Phys. Rev. A **71**, 062325 (2005).

[35] E. Hostens, J. Dehaene, and B. D. Moor, Phys. Rev. A **73**, 062337 (2006).

[36] R. Matsumoto, J. Phys. A: Math. Gen. **36**, 8113 (2003).

[37] A. Ambainis and D. Gottesman, IEEE Trans. Inf. Theory **52**, 748 (2006).

[38] S. Watanabe, R. Matsumoto, and T. Uyem, J. Phys. A: Math. Gen. **39**, 4273 (2006).

[39] D. Deutsch, A. Ekert, R. Jozsa, C. Macchiavello, S. Popescu, and A. Sanpera, Phys. Rev. Lett. **77**, 2818 (1996).

[40] T. Opatrný and G. Kurizki, Phys. Rev. A **60**, 167 (1999).

[41] D. Mundarain and M. Orszag, Phys. Rev. A **79**, 052333 (2009).

[42] D. Abdelkhalek, M. Syllwasschy, N. J. Cerf, J. Fiurášek, and R. Schnabel, Nat. Commun. **7**, 1 (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, Science **356**, 928 (2017).

[44] A. S. Fletcher, P. W. Shor, and M. Z. Win, Phys. Rev. A **75**, 012338 (2007).

[45] A. S. Fletcher, P. W. Shor, and M. Z. Win, Phys. Rev. A **77**, 012320 (2008).

[46] A. S. Fletcher, P. W. Shor, and M. Z. Win, IEEE Trans. Inf. Theory **54**, 5705 (2008).

[47] L. Ruan, W. Dai, and M. Z. Win, Phys. Rev. A **97**, 052332 (2018).

[48] M. Brodsky, N. J. Frigo, M. Boroditsky, and M. Tur, Journal of Lightwave Technology **24**, 4584 (2006).

[49] E. M. Rains, Phys. Rev. A **60**, 179 (1999).

[50] E. M. Rains, IEEE Trans. Inf. Theory **47**, 2921 (2001).

[51] X. Wang and R. Duan, Phys. Rev. A **95**, 062322 (2017).

[52] H. Takesue and K. Inoue, Physical Review A **70**, 031802 (2004).

[53] D. C. Burnham and D. L. Weinberg, Physical Review Letters **25**, 84 (1970).