Optimal resource allocation for flexible-grid entanglement distribution networks

Jude Alnas,1,2 Muneer Alshowkan,2 Nageswara S. V. Rao,3 Nicholas A. Peters,2 and Joseph M. Lukens2,*

1Department of Electrical and Computer Engineering,
University of Alabama, 245 7th Avenue, Tuscaloosa, Alabama 35406, USA
2Quantum Information Science Section, Oak Ridge National Laboratory, Oak Ridge, Tennessee 37831, USA
3Advanced Computing Methods Section, Oak Ridge National Laboratory, Oak Ridge, Tennessee 37831, USA

(Dated: April 15, 2022)

We use a genetic algorithm (GA) as a design aid for determining the optimal provisioning of entangled photon spectrum in flexible-grid quantum networks with arbitrary numbers of channels and users. After introducing a general model for entanglement distribution based on frequency-polarization hyperentangled biphotons, we derive upper bounds on fidelity and entangled bit rate for networks comprising one-to-one user connections. Simple conditions based on user detector quality and link efficiencies are found that determine whether entanglement is possible. We successfully apply a GA to find optimal resource allocations in four different representative network scenarios and validate features of our model experimentally in a quantum local area network in deployed fiber. Our results show promise for the rapid design of large-scale entanglement distribution networks.

I. INTRODUCTION

The phenomenon of quantum entanglement has resulted in a plethora of novel quantum technologies. Examples in quantum information science (QIS) include quantum-secured communications [1–3], blind or distributed quantum computing [4, 5], and the emerging quantum internet [6, 7]. Entanglement also promises new techniques beyond computing and communications, such as improved interferometry for radio astronomy [8, 9] and entanglement-enhanced clock synchronization [10, 11] for sensing and navigation [12]. Realizing these technologies will require robust quantum networks, the infrastructure with which quantum information can be transmitted and received. Such a network must also be capable of reliably distributing entangled qubits to network users on demand. Early examples of entanglement distribution for quantum communications can be found in quantum key distribution (QKD) experiments [3, 13–15] that, however, focused on only two directly connected users.

The development of spontaneous parametric down-conversion (SPDC) techniques has facilitated the generation of telecommunication wavelength (∼1550 nm) photons entangled in both the polarization and frequency degrees of freedom [16–20]. Using frequency entanglement, optical techniques from classical networks such as wavelength division multiplexing (WDM) have been leveraged to perform entanglement distribution, forming rudimentary quantum local area networks (QLANs). Much like conventional routing protocols such as the Transmission Control Protocol/Internet Protocol (TCP/IP), the working principles of these QLANs can be described in layers of abstraction [21–24] as in Fig. 1. Under our proposed definitions [22], the physical layer consists of a hyper-entangled light source realized through a pump laser and a nonlinear optical crystal [25] or a waveguide capable of SPDC. This spectrum then enters a wavelength demultiplexer (demux). The link layer is the next abstraction layer and is implemented entirely within the demux, which is configured to partition a hyperentangled spectrum into pairs of frequency-entangled bands. In the subsequent allocation process, users are selectively entangled by multiplexing energy-correlated channel pairs onto their respective physical ports. Users receiving correlated channels share entanglement via the polarization degree of freedom, thus creating a logical network of entanglement that constitutes the transport layer.

Early instances of this approach used fixed-grid technology such as dense wavelength division multiplexing (DWDM) filters to implement the link layer [21, 26–30]. In such a setup, the hyperentangled spectrum is partitioned into a fixed number of fixed-width frequency channels. Likewise, the routing of these channels is essentially fixed, modifiable only by physically disconnecting and reconnecting the DWDM filters or by using supplementary spatial optical switches [31, 32]. Increasing network capacity requires extra frequency channels, which in turn require additional DWDM and switching components. For example, in [29], a fully connected N-user network would require 4N − 1 additional filters to support one more user. Because these channels are fixed in bandwidth, one also faces the potential problem of inefficient spectrum utilization, whereby users receive either more or less bandwidth than required for their desired application.

Flexible wavelength-selective switches (WSSs) based on liquid crystal on silicon technology [33] avoid many of the drawbacks of fixed-grid networks. A WSS provides active wavelength multiplexing and routing in a single device while also allowing for optimal bandwidth utilization. In the context of entanglement distribution, such a device can be used to partition a spectrum of entangled light into an arbitrary number of frequency channels of arbitrary widths (within device limitations), obviating the need for numerous passive components. Channels are then allocated to a physical path connected to a user

* lukensjm@ornl.gov
or a second WSS. Routing is done electronically within the device, allowing rapid reconfiguration of entanglement networks without discrete switching components. As a result, flex-grid entanglement distribution networks are more compact, efficient, and adaptable in comparison to their fixed-grid counterparts. This technology has already found use in recent demonstrations of entanglement distribution in QLANs [22, 34–36].

The dynamic partitioning and allocation facilitated by WSSs also allow for entanglement network optimization. For example, allocations can be adjusted to compensate for losses from low-quality detectors [34] or long-distance transmission [35]. However, problems arise in finding the proper frequency allocations for optimal entanglement distribution. A trial-and-error approach was used in [34] to find allocations of multiple fixed-width channels to equalize coincidence counts in a fully connected four-user network. In [35], an undescribed algorithm and model was used to find the proper channel widths in a fixed allocation to equalize coincidence counts in a five-user network where one user suffered losses due to long-distance transmission. While trial-and-error or exhaustive search methods may be viable for small networks, they are intractable at large scales. Hence, there is a need for tools with which entanglement resource allocations can be easily found to meet global network objectives.

In this paper, we propose a model of an entanglement network for use in tackling the entangled flux allocation (EFA) problem. In Sec. II, we introduce the many-to-many entanglement network model in terms of biphoton fluxes, specialize to the case of one-to-one networks, and derive upper bounds and conditions on entanglement fidelity and entangled bit rate (EBR). In Sec. III, we introduce the EFA optimization problem and demonstrate allocation optimization through the use of a genetic algorithm (GA) [37] for a number of illustrative scenarios. We then follow in Sec. IV with an experimental test of our basic link model in a deployed QLAN on the Oak Ridge National Laboratory campus. Finally, implications and future avenues for research are discussed in Sec. V.

II. ENTANGLEMENT NETWORK MODEL

A. Fundamental model

A polarization entanglement network of \( N \in \mathbb{N}^+ = \{1, 2, \ldots\} \) users and \( L \in \mathbb{N}^+ \) entanglement links can be represented by an undirected graph of \( N \) nodes and \( L \) edges as in the transport layer of Fig. 1. Each edge represents shared entanglement over link \( l \), with density matrix \( \sigma_l \). In our model, link state \( \sigma_l \) is calculated from quantities measurable in a polarized photon counting experiment between users \( m_l \) and \( n_l \), namely, the rate of accidental coincidences \( A_l \) and the rate of correlated coincidences \( C_l \). In this experiment, it is assumed that each user has two thresholding photon detectors of identical system efficiency \( \eta \) and dark count rate \( d \) configured to count photons of orthogonal polarizations in some basis (e.g., rectilinear, diagonal, or circular). Note that \( \eta \) includes all losses due to optical components and \( d \) includes all ambient light apart from the allocated biphoton channels. For example, user \( m_l \) on link \( l \) has two detectors of efficiency \( \eta_{m_l} \) and dark count rate \( d_{m_l} \) (variation between users is allowed, i.e., \( \eta_{m_l} \neq \eta_l \) and \( d_{m_l} \neq d_{n_l} \)). With these definitions, the twofold accidental coincidence rate \( A_l \), summed over all four detector pair combinations, can be written as [38, 39]

\[
A_l = 4\tau \left( \frac{\eta_{m_l}}{2} \sum_{\nu \in L_{m_l}} \bar{\mu}_\nu + d_{m_l} \right) \left( \frac{\eta_{n_l}}{2} \sum_{\nu \in L_{n_l}} \bar{\mu}_\nu + d_{n_l} \right).
\]  

The correlated coincidence rate—i.e., from detection events corresponding to photons from the same entangled pair—is then

\[
C_l = \eta_{m_l} \eta_{n_l} \bar{\mu}_l,
\]  

equal to the total coincidences that would be obtained in the absence of background noise and multi-pair emission. In Eqs. (1) and (2), \( \tau \) is the coincidence window (assumed wide enough to integrate over the entire correlated coincidence peak), and \( L_{m_l} \) is the set of links connected to user \( m_l \). Flux \( \bar{\mu}_l \) is the total biphoton flux over all channels allocated to users \( m_l \) and \( n_l \) on link \( l \), defined at the source (i.e., prior to system loss).

The two-qubit polarization state of the link, \( \sigma_l \), in the coincidence basis for the pair of users, is then expressed as the following density matrix:

\[
\sigma_l = \lambda_l \sum_{k \in \mathcal{B}_l} \frac{\mu_k}{\bar{\mu}_l} U_k \rho_k U_k^\dagger + (1 - \lambda_l) \frac{I_4}{4},
\]  

where \( \lambda_l \) is the visibility given by

\[
\lambda_l = \frac{C_l}{A_l + C_l}.
\]  

In this model, the spectrum of entangled light is partitioned into \( 2K \) entangled frequency bands paired into \( K \) frequency channels (illustrated in Fig. 1); thus, \( K \geq L \) is a prerequisite for allocating flux to all desired links. An entangled photon source emits photon pairs in channel \( k \in \{1, 2, \ldots, K\} \) at a flux of \( \mu_k \), described by the density matrix \( \rho_k \); unitary matrix \( U_k \) encapsulates any propagation effects (e.g., birefringence) experienced by the photons. It is important to emphasize that \( \rho_k \) models the state of exactly two photons (originating from the same pump photon); nonidealities from multi-pair effects and background counts appear through Eqs. (1) and (2). In constructing the complete (noisy) state \( \sigma_l \), the \( \rho_k \) are weighted by the ratio of channel flux \( \mu_k \) to total link flux \( \bar{\mu}_l = \sum_{k \in \mathcal{B}_l} \mu_k \), where \( \mathcal{B}_l \subseteq \{1, 2, \ldots, K\} \) is the set of channels allocated to link \( l \). We assume all noise in the network is uniform background modeled by the maximally mixed two-qubit state \( I_4/4 \) where \( I_4 \) is the \( 4 \times 4 \)
FIG. 1. Example entanglement distribution network of $N = 3$ users and $L = 3$ links. Each user is equipped with a polarization analyzer (red inset). In the physical layer, a hyperentangled spectrum is generated using a continuous-wave (CW) 780 nm pump laser and a periodically poled lithium niobate (PPLN) waveguide. The process of spectrum partitioning and allocation occurs within the wavelength-selective demultiplexing device (demux) and constitutes the link layer. Here, the hyperentangled spectrum is partitioned into $K = 3$ pairs of frequency-correlated bands and allocated to users, forming the logical network of entanglement links that constitutes the transport layer. FPC: fiber polarization controller. QWP: quarter-wave plate. HWP: half-wave plate. PBS: polarizing beam-splitter. SPD: single-photon detector.

Identity matrix. When correlated coincidences dominate, $\lambda_l \to 1$ and $\sigma_l$ approaches a convex sum of entangled states. Conversely, when accidentals dominate, $\lambda_l \to 0$ and $\sigma_l$ approaches the maximally mixed state.

The metrics chosen to quantify network performance are the link state fidelities $F_l$ with respect to some target state $|\Psi_l\rangle$ and the entangled bit rates (EBRs) $R_l$, defined as:

$$F_l = \langle \Psi_l | \sigma_l | \Psi_l \rangle,$$

$$R_l = (A_l + C_l) \log_2(||\sigma_{TA}^T||_1).$$

where the logarithmic term is the logarithmic negativity [Eq. (2) in [40]], $\sigma_{TA}^T$ is the partial transpose of the link state with respect to one photon of the entangled pair, and $|| \cdot ||_1$ is the trace norm. In Eq. (6), $R_l \in \mathbb{R}_{\geq 0}$ is an upper bound on the rate of distillable entanglement received by the users of link $l$ [40]. Intuitively, $F_l$ assesses the quality of two coincident photon detections, whereas $R_l$ incorporates both quality and quantity (rate) into a single number.

Before proceeding further, it is useful to note the regimes under which the model expressed in Eqs. (1–3) remains valid. First, the probability of a single detector click within the coincidence window $\tau$ must be much less than unity: e.g., for user $m_l$, $\tau (\eta m_l \bar{\mu}_l + d_m) \ll 1$ (and analogously for all other users). This condition ensures that the rate of accidental coincidences follows the standard “product-of-singles” formula [38, 39]. Second, the singles detection rates must remain constant regardless of polarization analyzer settings, implying that the traced out marginal quantum state for each user possesses an equal distribution of $H$ and $V$ components; this holds for all Bell and Werner states, for example. Finally, the correlated coincidence rate in Eq. (2) assumes photon pairs that are independent—i.e., distinguishable—of each other, a situation occurring whenever the pump laser in the SPDC process has much lower bandwidth than the narrowest frequency channels, which has been the case in all flex-grid entanglement experiments so far [22, 34, 35]. Importantly, however, the general approach we introduce below for optimal bandwidth allocation is not limited to the current expressions and assumptions, but can be adapted to any physical model as long as it connects fixed user parameters (e.g., $\eta$ and $d$) and allocation decisions (e.g., $B_l$) to the quantities of interest (e.g., $F_l$ and $R_l$).
B. Simplified one-link model

In order to highlight the important features of our method with minimal distractions, we apply three additional assumptions to the model described in Sec. II A for the network scenarios simulated below.

Assumption 1 Users are entangled with only one other.
\[ |\mathcal{L}_m| = |\mathcal{L}_n| = 1 \forall \mathcal{L} \in \{1, 2, ..., L\}. \]

Assumption 2 Channel states are identical.
\[ \rho_k = |\psi\rangle\langle\psi| \forall k \in \{1, 2, ..., K\}. \]

Assumption 3 Channel distortion effects are fully compensated.
\[ U_k = I_4 \forall k \in \{1, 2, ..., K\}. \]

Assumption 1 restricts the logical network topology to one-to-one links, making it sufficient to analyze a single link in isolation to understand the basic behavior. Accordingly, in the initial analysis, we suppress the \( l \) subscripts and treat a single link of flux \( \mu \) consisting of user 1 and user 2, characterized by efficiencies \( \eta_1, \eta_2 \) and dark count rates \( d_1, d_2 \). Although simpler than the fully connected paradigm, whereby each user is entangled with all others in a single physical configuration [21, 30, 34, 35], the one-to-one restriction is more naturally suited to flexgrid capabilities in our view. For, as argued in [22], the fact a WSS can be reconfigured on demand to realize entanglement across all frequency channels of interest and perfect compensation of any birefringence effects in the optical channel. Depending on the nature of the source and transmission medium, these assumptions may be difficult to attain, although they certainly represent the desired situation in practice. For our purposes, they allow us to focus on fundamental effects due to probabilistic photon emission without complications from technical nonidealities, which can be incorporated as needed to reflect an actual experiment.

For states of the form of Eq. (8), fidelity with respect to \( |\psi\rangle \) is simply \( F = (1 + 3\lambda)/4 \). By Eqs. (2,4,7), fidelity can then be expressed as a function of link flux \( \bar{\mu} \)

\[
F(\bar{\mu}) = \frac{1}{4} \left[ 1 + \frac{3\bar{\mu}}{4\tau} \left( \frac{\bar{d}_1}{\eta_1} + \frac{\bar{d}_2}{\eta_2} + \frac{1}{4\tau} \right) \bar{\mu} + \bar{d}_1\bar{d}_2 \right].
\]

A similar expression for EBR can be obtained. For states of the form of Eq. (8), the logarithmic negativity becomes a simple function of \( F \) and thus a function of \( \bar{\mu} \) [40]:

\[
R(\bar{\mu}) = 4\tau \left[ \frac{\eta_1\eta_2}{4} \frac{\bar{\mu}^2}{d_1} + \left( \frac{\eta_1}{2} d_2 + \frac{\eta_2}{2} d_1 + \frac{\eta_1\eta_2}{4\tau} \right) \bar{\mu} + d_1 d_2 \right] \log_2 2F(\bar{\mu}).
\]

C. Dimensionless parametrization

Equations (9) and (10) can be further simplified through a dimensionless parametrization. We define the dimensionless flux \( x \) and noise parameter \( y_n \) as

\[
x := \tau\bar{\mu} \quad ; \quad y_n := \frac{\tau\bar{d}_n}{\eta_n}. \tag{11}
\]

The physical interpretation of \( x \) is the mean number of biphotons produced in a coincidence window \( \tau \), equivalent to a pair production probability per pulse—where the “pulse” is defined in an effective sense for this CW-pumped case. The noise parameter \( y_n \) can be interpreted as a ratio of probabilities. The numerator is the probability of observing a noise event within a coincidence window \( \tau \); the denominator is the probability of detecting a desired photon, given that it was produced within a single window.

After direct substitution of \( x \) and \( y \), Eq. (9) becomes

\[
\mathcal{F}(x) := \frac{1}{4} \left[ 1 + \frac{3x}{4} \left( \frac{y_1}{x_1 + 2y_1 + 2y_2 + 1} + \frac{y_2}{x_2 + 2y_1 + 2y_2 + 1} \right) \times \log_2 2\mathcal{F}(x). \tag{12}
\]

We then define the dimensionless EBR \( \mathcal{R} \) as

\[
\mathcal{R}(x) := \frac{x}{\eta_1\eta_2} \mathcal{R}(\frac{x}{7}) = \frac{\tau}{\eta_1\eta_2} \mathcal{R}(\frac{x}{7}) \times \log_2 2\mathcal{R}(x). \tag{13}
\]

That same window.

Significantly, this parametrization reveals that the dark count rate and efficiency impact the functional form of fidelity and EBR only via their ratio, so that each user can be characterized by a single number \( y_n \), that quantifies the quality of their transmission path and receiver.
Moreover, although $F$ and $R$ are related to each other, they depend on flux $x$ in markedly different ways. Indeed, the interplay between fidelity and EBR, noted in previous quantum networking demonstrations [22], proved one of the key motivations for the current investigation and presents the quantum network engineer with non-trivial tradeoffs in evaluating competing allocations.

D. Model maxima

Fidelity in Eq. (12) is an algebraic function of $x$ so that closed-form expressions for the location and value of the optimum can be found. Setting $\partial F/\partial x = 0$, the maximal fidelity is found to occur at $x_F = 2\sqrt{y_1 y_2}$ with a corresponding maximum of

$$F_{\text{max}}(y_1, y_2) = F(x_F; y_1, y_2) = \frac{1}{4} \left[ 1 + \frac{3}{4\sqrt{y_1 y_2} + 2(y_1 + y_2) + 1} \right].$$

(14)

Figure 2(a) gives the surface plots of $x_F$ and Eq. (14) as well as curves of Eq. (12) for select values of $y_1$ and $y_2$. The surface plots show that as $y_1$ or $y_2$ increases $F_{\text{max}}$ decreases and the location of $F_{\text{max}}$ shifts towards higher $x$. Eventually, $F_{\text{max}}$ approaches 0.5, below which the log-negativity vanishes and the state becomes separable [41].

Unlike fidelity, the expression for EBR in Eq. (13) main due to channel and detector properties and the fundamental fitness function selected for optimization. Intuitively, the optimal flux for fidelity $x_F$ balances the effects of noise from dark counts (which increases with smaller $x$) and multi-pair emission (which grows with increasing $x$). On the other hand, the larger EBR optimum ($x_R > x_F$ whenever $R_{\text{max}} > 0$) results from balancing the tradeoff between state quality (maximized at $x_F$) and detection rate (maximized as $x \to \infty$); this feature of EBR lends credence to our arguments for it as an application-agnostic metric for quantum networking [22, 36]. Nevertheless, as we discuss later, fidelity is likely to remain significant in its own right in near-term quantum networks.

Finally, as an aside, the conclusions found in this section follow logically from previous models such as [42], yet to our knowledge no explicit derivation has been provided in the literature. Accordingly, these findings offer useful design considerations more generally for any SPDC entanglement source, regardless of application. When all experimentally controllable nonidealities are removed (mode mismatch, synchronization, etc.), these limits remain due to channel and detector properties and the fundamental physics of probabilistic pair production.

III. NETWORK OPTIMIZATION

A. Problem formulation & complexity

Using the model given in Sec. II, we now introduce the entangled flux allocation (EFA) problem, which can be formulated as a nonlinear integer optimization problem. Taking as given $K$ energy-correlated pairs of frequency bands, the objective of the EFA problem is to find an allocation of the $K$ channels that optimizes a desired entanglement network performance metric. We can represent an allocation with the vector $\alpha \in \{0, 1, 2, \ldots, L\}^K$ where $\alpha_k = l$ if the $k$th channel pair is assigned to link $l$. If $\alpha_k = 0$, then channel $k$ is allocated to the “reserve link,” a virtual link that holds the flux of any unallocated channels—i.e., channels which are sent to no network users. Assuming each of the $K$ channels is unique, there are $(L + 1)^K$ unique channel allocations.

The complexity of EFA can be determined through comparison with the multiple subset sum problem (MSSP). In the formal definition of MSSP, the objective is to partition a set of $M$ items, $N$ items, $N$ weights, and $M$ bins each with positive integer capacity $c$, “filling” the bins as much as possible without exceeding capacity $c$ [43]. If the integer constraints on the weights and capacity are lifted, the $N$ items, $N$ weights, and $M$ bins of capacity $c$ of MSSP are respectively analogous to the $K$ frequency channels,
FIG. 2. (a) Surface plots in noise parameter space of maximum fidelity $F_{\text{max}}$ (left) and corresponding flux $x_F$ (middle), as well as full plots of $F(x)$ for select noise parameter pairings (right). The dashed gray line marks $F = 0.5$, below which entanglement is lost for the Werner state in Eq. (8). (b) Surface plots in noise parameter space of maximum dimensionless EBR $R_{\text{max}}$ (left) and corresponding flux $x_R$ (middle), along with $R(x)$ for particular noise parameters (right). The boundary curve [Eq. (15)] is drawn in red.

$K$ channel fluxes, and $L$ entanglement links with target flux $\phi_l$ of our simplified EFA model. Whereas all $M$ bins in MSSP have the same capacity $c$, each link may have a distinct $\phi_l$, the exact value of which may be calculated from Eq. (12) or Eq. (13). Furthermore, whereas MSSP forbids bin loads from exceeding capacity $c$, link fluxes are allowed to exceed $\phi_l$ in EFA. Because EFA is obtained by relaxing some constraints of the MSSP, it is a super-problem of MSSP. Further, the formal MSSP is known to be strongly NP-hard [43]. Hence, the complexity of the simplified EFA is at least strongly NP-hard.

As another connection to previous work, we note here that the EFA problem bears some resemblance to the routing and spectrum allocation (RSA) problem in classical flex-grid networks. The objective of RSA is to efficiently utilize a spectrum of frequencies to satisfy traffic demands between source and destination nodes of a physical network. Low-bandwidth demands can be assigned an appropriately smaller portion of the spectrum, effectively increasing network capacity. The RSA problem is generally NP-hard [44, 45] and is typically solved via some form of integer programming [44, 46, 47] or meta-heuristic algorithms [48–51].

In RSA, a route through the physical network must be found in coordination with the assignment of frequency channels. Applications often enforce the continuity and contiguity constraints [52]. The first ensures that frequency channels assigned at the source are not used by nodes en route to the destination. The second ensures that multiple spectral slices allocated to a demand are adjacent in the frequency domain. In contrast, such constraints can be relaxed in the EFA problem considered here. The assumption of a star physical topology whereby each node possesses a direct fiber connection to the source (cf. Fig. 1) ensures that any allocation automatically satisfies the continuity requirement. The reliance on frequency correlations resulting from broadband SPDC means that distributing discontinuous spectra to each user poses no fundamental problem in EFA and, in fact, has been considered experimentally [34].

However, in more complex entanglement distribution networks (e.g., with multiple SPDC sources, nested WSSs, or multi-hop links), the explicit consideration of a continuity constraint will be required in EFA as well. Moreover, spectral contiguity—even if not intrinsically necessary for EFA—will likely prove practically important in order to minimize the effects of wavelength-dependent fiber birefringence; the greater the separation between spectral slices $k$, the less likely their polarization rotations $U_k$ can be compensated in tandem by a single
B. Genetic algorithm

To solve the NP-hard EFA problem, we turn to evolutionary algorithms (EA), which use metaheuristic search strategies modeled after biological processes to find optimal solutions to difficult problems. Since these strategies are independent of the problem itself, they can be applied regardless of problem size, lineararity, or availability of gradients. Well-known examples of EAs include ant colony optimization (ACO) [53, 54], particle swarm optimization (PSO) [55], and GAs [37]. EAs have already found applications in tackling problems in quantum communications [56–58].

We apply the GA as implemented in MATLAB R2021a [59] to tackle the EFA problem. A GA is a stochastic optimization algorithm that maximizes (or minimizes) a so-called fitness (or objective) function by mimicking biological evolution. It begins with a randomly generated population of genes, all of which represent inputs to a user-defined fitness function. The genes that maximize the fitness survive to the next iteration unaltered. A new population is created by randomly crossing over (i.e., combining) or mutating existing genes. These processes are illustrated in Fig. 3(a). A GA is desirable because of its flexibility and adaptability to continuous, integer, or mixed-integer problems of any size. Although discrete versions of ACO [60, 61] and PSO [62–64] exist, they are more amenable to binary problems as opposed to integer problems such as EFA. In our optimizations, we use a population of 200 genes and a crossover fraction of 0.8. The optimization terminates when the maximum fitness achieved does not change for 100 consecutive iterations.

We apply this GA in four different optimization scenarios motivated by realistic network conditions to test its tractability for the EFA problem. Each scenario uses the following fitness function:

\[
 f = \sum_{l=1}^{L} \beta_l \left( R_l; F_l, F_{\text{min}} \right) \tag{16a}
\]

where

\[
 \beta_l \left( R_l; F_l, F_{\text{min}} \right) = \begin{cases} 
 \frac{R_l}{R_{l, \text{max}}}, & F_l \geq F_{\text{min}} \\
 -1, & \text{otherwise} 
\end{cases} \tag{16b}
\]

and \( R_l/R_{l, \text{max}} \) is the link \( l \) EBR normalized to the link’s theoretical maximum EBR. Equation (16) is designed to maximize each link’s EBR while maintaining a fidelity of at least \( F_{\text{min}} \) without enforcing an explicit constraint that would add significant computational complications. For benchmarking purposes, we define \( f_\infty \) as the best possible fitness given constraints and access to infinite resources (e.g., unlimited total flux and channels \( K \)).

Table I provides the values \( L, K, F_{\text{min}}, \) and \( f_\infty \) for each of the four scenarios examined. Figure 3(b) shows the five-link, one-to-one entanglement network used in Scenarios 1–3. Scenario 4 uses a larger, 12-link one-to-one network. For each scenario, optimizations are run with an increasing number of channels \( K \). As GAs are stochastic, five independent optimizations are performed for each \( K \), and the result with the largest fitness is kept. In addition to flux allocation, we include the total biphoton flux \( \mu_{\text{tot}} \) as an optimization variable. Changing this is analogous to changing the pump power of the central provider’s entangled light source. Although our algorithm is designed to handle variable channel fluxes, in the scenarios here we make the simplification that the total flux is uniformly distributed across the \( K \) available channels: \( \mu_k = \mu_{\text{tot}}/K \). Note that this uniform distribution results in redundant allocations. Under such conditions, it does not matter which channels are allocated to a link, only how many channels are allocated. Thus the number of unique allocations reduces from \( (L+1)^K \) to \( \binom{K+L}{L} \).

![Fig. 3](image_url)

**TABLE I.** Parameters for the four optimization scenarios. Given the minimum fidelity threshold \( F_{\text{min}} \), the best achievable fitness is calculated and reported as \( f_\infty \).

| Scenario | L | K | \( F_{\text{min}} \) | \( f_\infty \) |
|----------|---|---|-------------------|-----------|
| 1        | 5 | 5 | [5, 10, 20, 40]   | 0         |
| 2        | 5 | 5 | [5, 10, 20, 40]   | 0.7       |
| 3        | 5 | 5 | [5, 10, 20, 40]   | 0.9       |
| 4        | 12| 12| [12, 24, 48, 96]  | 0.7       |
show that this goal is closely met for any link achieving maximum normalized EBR. Figure 4 similar flux values facilitates this performance. Indeed, the fact that the EBRs for all links are maximized by the solution has fitness constraints $\mathcal{F}_{\text{max}}$ for all five network links without any fidelity.

Eventually, users on near-term quantum networks will likely be expected to request a threshold state quality defined by the particular application, so that optimization results

Using the model in Sec. II, three networks were designed for use in four optimization scenarios. Table II gives the noise parameters and the maximum achievable fidelities for each link for each scenario. As context for these values, for $\tau = 1$ ns we estimate that the superconducting nanowire detector (SNSPD) used by Alice in the experiment of Sec. IV below corresponds to $\gamma_A \approx 1.3 \times 10^{-7}$ ($\gamma_A \approx 8.3 \times 10^{-6}$) with extra channel losses neglected (included); for Bob, who utilizes an InGaAs avalanche photodiode (APD), $\gamma_B \approx 3.5 \times 10^{-5}$ ($\gamma_B \approx 1.7 \times 10^{-2}$) without (with) additional channel losses. In order to explore the impact of heterogeneous nodes on the EFA problem, in Table II we have intentionally selected a much broader range of noise parameters than these experimental examples, leading to links with a wide spread in maximum fidelity $\mathcal{F}_{\text{max}}$. For Scenarios 1, 2, and 4, all links lie in the interval $\mathcal{F}_{\text{max}} \in [0.72, 1]$; for Scenario 3, $\mathcal{F}_{\text{max}} \in [0.93, 1]$. In every case, $\mathcal{F}_{l, \text{max}} > 0.5$ so that $\mathcal{R} > 0$ is possible for each pair of examined nodes. Final optimization results are given in Figs. 4–7. For each scenario, three types of plots are provided: (a) fidelity and normalized EBR achieved via GA optimization, plotted as points on the curves of $\mathcal{F}_{l}(x_l)$ and $\mathcal{R}_{l}(x_l)/\mathcal{R}_{l, \text{max}}$; (b) fitness $f$ for each $K$ as a bar graph with the maximum achievable fitness $f_{\infty}$ marked as a dashed line; and (c) number of channels allocated to each link for each $K$ plotted as a stacked bar graph.

The optimization goal of Scenario 1 is to maximize the EBR of all five network links without any fidelity constraints [$\mathcal{F}_{\text{min}} = 0$ in Eq. (16)]. The best achievable solution has fitness $f_{\infty} = 5$, which corresponds to every link achieving maximum normalized EBR. Figure 4 shows that this goal is closely met for any $K$. For $K < 40$, GA finds allocations with $f = 4.97$, a 0.68% deviation from $f_{\infty} = 5$. This deviation decreases to 0.1% with $K = 40$ where an allocation with $f = 4.995$ is found. The fact that the EBRs for all links are maximized by similar flux values facilitates this performance. Indeed, for $K \in \{5, 10, 20\}$, the optimal allocations divide the channels evenly between all links, and so the fitness obtained in all cases remains approximately the same. It is only at $K = 40$ where the increased spectral granularity leads to a nonuniform provisioning as optimal, with AB and CD receiving 8 channels each and EF, GH, and IJ receiving 7.

In Scenario 1, we concentrated on optimizing EBR only, a promising candidate in our view for a universal metric of an entangled link’s quality. Nonetheless, utilizing this entanglement in practice would demand implementation of entanglement distillation [65–68], an extremely demanding protocol that—although critical for quantum networking in the long-term—is beyond the capabilities of many existing quantum network testbeds. Consequently, users on near-term quantum networks will likely be expected to request a threshold state quality defined by the particular application, so that optimization results...
The bound in Eq. (14). This network was optimized using numerical difficulties with constraints and increased re-
crease is a computational artifact reflecting the growing numerical difficulties with constraints and increased resources.

More stringent fidelity constraints are enforced in Scenario 3. Equation (16) is used with $F_{\text{min}} = 0.9$, and the link noise parameters are chosen such that the maximum possible link fidelities are 1, 0.99, 0.97, 0.95, and 0.93 for AB, CD, EF, GH, and IJ, respectively. The best achievable solution has $F_l = 0.9$ for all links and has a corresponding fitness of $f = 0.911$. Figure 6 shows that all links satisfy fidelity constraints for all $K$. For higher $K$, all links approach the ideal solution. The best fitness achieved is $f = 0.864$ (5.14% deviation).

Scenario 4 tests the capability of the GA when the network size is increased significantly. A 12-link network was designed such that $F_{\text{max}} \geq 0.72$ for all links [according to the bound in Eq. (14)]. This network was optimized using $\mathcal{F}_{\text{min}}$ specified should be used to determine a bandwidth allocation. Thus, in Scenario 2, the same network as in Scenario 1 was optimized, but now with a fidelity threshold. In order to ensure that all links have the potential to exceed the threshold, we must have $\mathcal{F}_{\text{min}} \leq \min_l (\mathcal{F}_{\text{max}})$. For Scenario 2, $\mathcal{F}_{\text{min}}$ must therefore be less than 0.72, and so we select $\mathcal{F}_{\text{min}} = 0.7$ for numerical optimization. Figure 2 shows that in the low-flux regime, EBR increases as fidelity decreases. Hence, the best achievable EBR given a fidelity threshold $\mathcal{F}_{\text{min}}$ will have $F_l = \mathcal{F}_{\text{min}}$ for all links. For Scenario 2, this corresponds to an ideal fitness value of $f = 3.39$. The introduction of a fidelity penalty makes the problem more difficult to solve in comparison to Scenario 1. The best fitness achieved was $f = 3.21$ (5.37% deviation) at $K = 20$. Interestingly, a slightly lower fitness of $f = 3.20$ (5.56% deviation) was achieved with $K = 40$; since fitness should never decrease with $K$ (as allocations at smaller $K$ are a subset of those possible with larger $K$), this slight decrease is a computational artifact reflecting the growing numerical difficulties with constraints and increased resources.

FIG. 6. Scenario 3 optimization results. A five-link entanglement network is optimized to maximize the EBR of all five links while maintaining $F_l \geq 0.9$. (a) GA-achieved fidelity and EBR along model curves [Eqs. (12,13)]. (b) Best fitness achieved for each $K$, compared to the best possible given fidelity constraints and sufficient resources of $f = 0.91$. (c) Distribution of frequency channels to links for each $K$.

FIG. 7. Scenario 4 optimization results. A 12-link entanglement network is optimized to maximize the EBR of all 12 links while maintaining $F_l \geq 0.7$ for all 12 links. (a) GA-achieved fidelity and EBR along model curves [Eqs. (12,13)]. (b) Best fitness achieved for each $K$, compared to the best possible given fidelity constraints and sufficient resources of $f = 7.9$. (c) Distribution of frequency channels to links for each $K$.

Since $\mathcal{F}_{\text{min}} = 0.7$ in Eq. (16). Once again, the best achievable solution here has $F_l = 0.7$ for all links, corresponding to a maximum fitness of $f = 7.9$. Figure 7 shows that all links satisfy the fidelity constraints for $K \geq 48$. The best achieved fitness improved with available resources, ranging from $f = 3.98$ (49.7% deviation) for $K = 12$ to $f = 6.51$ (17.8% deviation) for $K = 96$.

Beyond limited computational resources, a source of difficulty in these scenarios arises whenever there exists a wide imbalance of noise parameters across the nodes, which challenges the source’s ability to simultaneously satisfy the full range of optimal link demands at low channel numbers, i.e. $K \sim \mathcal{O}(L)$. For example, for
$K = 5$ in Scenarios 1–3, all links were allocated one channel and thus received the same amount of flux. Ideally, to maximize the fitness function, $\mu_{\text{tot}}$ (and thus $x_k = \tau \mu_k$) should be chosen such that the constraints for all links are simultaneously satisfied. Yet, for the $F_{\text{max}}$ settings of Scenarios 2 and 3, the low-noise links like AB require significantly more flux to reach their fidelity-constrained optimal EBR rates than noisy links such as link JI. This noise imbalance means that $x_k$ cannot exceed the maximum flux that satisfies the noisiest link (or else violate the minimum fidelity condition). Mathematically speaking, we must have $x_k \leq \min \{\phi_l\}$ where $\phi_l$ is the link flux that optimizes the contribution of link $l$ to the fitness: $\phi_l = \arg\max_{x \in \mathbb{R}^+} f_l(x)$, where $\mathcal{S} = \{x \in \mathbb{R}^+: F_l(x) \geq F_{\text{min}}\}$. Once $K$ increases, more channels can be allocated to links like AB and CD to compensate for the limited $x_k$. This is reflected in the (c) panels of Figs. 5 and 6, which show that the low-noise links receive proportionally more channels as $K$ increases.

These same considerations lead to interesting results in Scenario 4. Figure 7 shows that for $K \in \{12, 24\}$, higher fitness is achieved by allocating no channels to link WX, resulting in an undefined fidelity, $R_{WX} = 0$, and leaving more flux for lower-noise links. Because Scenario 4 has 12 links contributing to the fitness function as opposed to five in the previous three scenarios, the negative effect of limiting $x_k \leq \min \{\phi_l\}$ for 11 of the 12 links outweighs the benefit of satisfying the fidelity constraint for the noisiest link, WX. With $K \in \{48, 96\}$, sufficient granularity in flux allocation is achieved, and it becomes possible to satisfy the fidelity threshold on all links simultaneously.

IV. EXPERIMENTAL TEST

The entanglement distribution model we have formulated here relies on the physical situation under test; we point the reader to [71] for details on this model for Bayesian inference.

Fidelity and EBR results for pump powers from 2.5 mW to 40 mW (the maximum we can achieve with current equipment) are plotted in Fig. 8. The experimental mean fidelity falls in the interval $[0.87, 0.93]$ for all cases, while the EBR increases linearly with power, showing no signs of the turning point expected in theory with sufficient flux (cf. Fig. 2). An accurate estimate of the experimental pair generation flux $x$ requires the pair-production efficiency $\eta_{\text{pair}}$, defined in our case as the ratio of biphoton pairs produced to the measured input pump flux. This quantity is difficult to obtain due to unknowns in quantum conversion efficiency and the inability to measure facet and waveguide losses independently; our definition $\eta_{\text{pair}}$ will thus be lower than the intrinsic quantum efficiency, since we consider a larger pump power than what actually enters the waveguide. We can employ a fitting procedure using established coincidence formulas to yield the appropriate values for $\eta_A$, $\eta_B$, and $\eta_{\text{pair}}$. Combining measured singles and coincidence rates, our procedure yields $\eta_A = (1.2 \pm 0.2) \times 10^{-2}$, $\eta_B = (2.1 \pm 0.4) \times 10^{-4}$, and $\eta_{\text{pair}} = (2.17 \pm 0.06) \times 10^{-10}$.

Combined with independent measurements of the dark count rates on each detector $d_A = 100$ s$^{-1}$ and $d_B = 3500$ s$^{-1}$ and the experimental coincidence window $\tau = 1 \text{ ns}$, we are able to compare measurements directly against the curves of Eqs. (9) and (10) predicted by our model, which are likewise included in Fig. 8 (using the mean values of the estimates of $\eta_A$, $\eta_B$, and $\eta_{\text{pair}}$, which correspond to $y_A = 8.33 \times 10^{-6}$ and $y_B = 1.67 \times 10^{-2}$).

Our fidelities are indeed lower than the theoretical limits for pump powers from 2.5 to 40 mW, over which our model predicts $\mathcal{F} > 0.95$. This remaining gap can be attributed to technical limitations from the manual alignment procedures, such as imperfect birefringence compensation and temporal drift in the polarization state through the deployed fiber—nonidealities which are intentionally omitted from the EFA model that focuses on the fundamental contributions of accidental coincidences. The above considerations notwithstanding, our experimental findings appear consistent with the entanglement distribution model used in our network design simulations. The linear EBR curve suggests that the current QLAN is operating far below peak performance (theory predicts $R_{\text{max}} = 1.58 \times 10^3$ ebits/s for the system noise parameters), and the limited range of fluxes available at present prevents us from demonstrating high-EBR solutions found from the GA procedure. Moving forward, then, it would be valuable to pursue either significantly higher pump powers or alternative hyperentangled biphoton sources based on type-0 phase matching [18, 32], for which the quantum efficiency should increase by more than 100-fold compared to our type-II source. The model developed and analyzed in this paper thus offers optimism for the near-term realization of flex-grid quantum networks with vastly higher entanglement distribution rates than shown so far; all that is required is an increase
FIG. 8. Experimental tests on a deployed QLAN. (a) Fidelity and (b) dimensioned EBR $R_{AB}$ versus measured input pump power. Black points show experimental results, including ±1σ Bayesian-inferred error bars; blue curves are the model predictions.

in photon flux.

V. CONCLUSION

In this paper, we have derived upper bounds on fidelity and EBR for one-to-one links in an optical entanglement-distribution network. These bounds are dependent on the noise parameters of the network users in the form of a dimensionless quantity $y_n$ defined as the ratio of background count probability to system efficiency. Entanglement between users is only possible under certain conditions, namely that their noise parameters $y_1$ and $y_2$ satisfy Eq. (15). Using this model, example networks were proposed to test flux allocation optimization using a GA. The number of available frequency channels and the distribution of biphoton flux across those channels have a large influence on the ease of finding a bandwidth allocation that attains near-optimal performance for all links. Specifically, noisy links tend to bias the algorithm towards small link fluxes, whereas low-noise links favor higher fluxes. This counteracting pull is absent in networks with links of similar noise parameter pairings. Accordingly, and perhaps unsurprisingly, networks with highly heterogeneous nodes—as evidenced by wide variability in the noise parameters $y_n$—present greater challenges for optimization in a flex-grid quantum network. Irrespective of these challenges, however, by outlining fundamental bounds our model allows us to quantify the closeness between an observed network state and the theoretical ideal, for any bandwidth allocation. This capability should prove invaluable for future network management, enabling a clear distinction between technical and fundamental limitations in reaching a desired configuration.

The model proposed here also has avenues for refinement. The equations and conditions of our dimensionless model are all symmetric to the $y_1 = y_2$ line. This may hint at a model parametrization that combines $y_1$ and $y_2$ into a single parameter. Such a parameter would be a good quantitative metric for qualifying a link, as opposed to the relative descriptions of links as “low” or “high” noise. Finally, although the current experimental results are consistent with the proposed model, they also indicate operation far below the maximum EBRs anticipated as possible on our QLAN due to limitations on the total amount of biphoton flux we can generate. Consequently, future experiments should attempt to achieve much higher pump powers (or more efficient SPDC) in order to measure EBR beyond the linear, low-flux regime. Good agreement would bode well for subsequent application of GAs in optimizing large-scale entanglement networks.

FUNDING

U.S. Department of Energy, Office of Science, Office of Workforce Development for Teachers and Scientists under the Science Undergraduate Laboratory Internship program; U.S. Department of Energy, Office of Science, Advanced Scientific Computing Research, under the Entanglement Management and Control in Transparent Optical Quantum Networks and Early Career Research programs (Field Work Proposals ERKJ378 and ERKJ353).

ACKNOWLEDGMENTS

This research was performed in part at Oak Ridge National Laboratory, managed by UT-Battelle, LLC, for the U.S. Department of Energy under contract no. DE-AC05-00OR22725.

DISCLOSURES

The authors declare no conflicts of interest.

DATA AVAILABILITY

Data available from the authors on request.
[46] K. Christodouloupolous, I. Tomkos, and E. A. Varvarigos, J. Light. Technol. 29, 1354 (2011).
[47] M. Ruiz, M. Pióro, M. Żotkiewicz, M. Klinkowski, and L. Velasco, Photonic Netw. Commun. 26, 53 (2013).
[48] Y. Wang, J. Zhang, Y. Zhao, J. Wang, and W. Gu, Photonic Netw. Commun. 25, 135 (2013).
[49] M. Klinkowski, M. Ruiz, L. Velasco, D. Careglio, V. Lopez, and J. Comellas, IEEE J. Sel. Areas Commun. 31, 26 (2013).
[50] F. Lezama, G. Castañón, A. M. Sarmiento, and I. B. Martins, Photonic Netw. Commun. 31, 129 (2016).
[51] G. Z. Marković, Photonic Netw. Commun. 34, 356 (2017).
[52] L. Velasco, A. P. Vela, F. Morales, and M. Ruiz, J. Light. Technol. 35, 513 (2017).
[53] M. Dorigo, V. Maniezzo, and A. Colorni, IEEE Trans. Syst. Man Cybern. B 26, 29 (1996).
[54] M. Dorigo, Ph. D. Thesis, Politecnico di Milano (1992).
[55] R. Eberhart and J. Kennedy, in Proc. Int. Sym. Micro Mach. Human Sci. (MHS) (1995) pp. 39–43.
[56] L. Gyongyosi and S. Imre, Quantum Inf. Process. 18, 233 (2019), arXiv:1803.02469.
[57] S. Krastanov, V. V. Albert, and L. Jiang, Quantum 3, 123 (2019), arXiv:1712.09762.
[58] F. Ferreira Da Silva, A. Torres-Knoop, T. Coopmans, D. Maier, and S. Wehner, Quantum Sci. Technol. 6, 035007 (2021), arXiv:2010.16373.
[59] Mathworks, “ga,” (2022), https://www.mathworks.com/help/gads/ga.html.
[60] M. Dorigo, G. D. Caro, and L. M. Gambardella, Artif. Life 5, 137 (1999).
[61] M. Schlüter, J. A. Egea, and J. R. Banga, Comput. Oper. Res. 36, 2217 (2009).
[62] J. Kennedy and R. C. Eberhart, in Proc. IEEE Int. Conf. Syst. Man Cybern., Vol. 5 (1997) pp. 4104–4108.
[63] F. Afshinmanesh, A. Marandi, and M. Shahabadi, IEEE Trans. Antennas Propag. 56, 1845 (2008).
[64] J. Alnas, G. Giddings, and N. Jeong, Appl. Sci. 11, 2559 (2021).
[65] C. H. Bennett, G. Brassard, S. Popescu, B. Schumacher, J. A. Smolin, and W. K. Wootters, Phys. Rev. Lett. 76, 722 (1996).
[66] M. Asano, M. Bechu, M. Tame, Ş. Kaya Özdemir, R. Ikuta, D. Ö. Gümey, T. Yamamoto, L. Yang, M. Wegener, and N. Imoto, Sci. Rep. 5, 18313 (2015).
[67] 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).
[68] D.-Y. Chen, Z. Lin, M. Yang, Q. Yang, X.-P. Zang, and Z.-L. Cao, Phys. Rev. A 102, 22425 (2020).
[69] R. Blume-Kohout, New J. Phys. 12, 043034 (2010).
[70] J. M. Lukens, K. J. H. Law, A. Jasra, and P. Lougovski, New J. Phys. 22, 063038 (2020).
[71] H.-H. Lu, K. V. Myilswamy, R. S. Bennink, S. Seshadri, M. S. Alshaykh, J. Liu, T. J. Kippenberg, D. E. Leaird, A. M. Weiner, and J. M. Lukens, arXiv:2108.04124 (2021).