

Recoil-Induced Errors and Their Correction in Photon-Mediated Entanglement between Atomic Qubits

Jan Apolín^{✉*} and David P. Nadlinger^{✉†}

Department of Physics, University of Oxford, Clarendon Laboratory, Parks Road, Oxford OX1 3PU, United Kingdom



(Received 21 March 2025; revised 24 October 2025; accepted 2 December 2025; published 6 February 2026)

Photonically interconnected matter qubit systems have wide-ranging applications across quantum science and technology, with entanglement between distant qubits serving as a universal resource. While state-of-the-art performance in heralded entanglement generation has thus far been achieved in trapped atomic systems modeled as stationary emitters, the improvements to fidelities and generation rates demanded by large-scale applications require consideration of their motional degrees of freedom. Here, we derive the effects of atomic motion on spontaneous emission coupled into arbitrary optical modes and study the implications for commonly used atom-atom entanglement protocols. We arrive at a coherent physical picture in the form of “kick operators” associated with each instant in the photonic wavepackets, which also suggests a method to mitigate motional errors by disentangling qubit and motion post-hoc. This proposed correction technique removes overheads associated with the thermal motion of atoms and may greatly increase entanglement rates in long-distance quantum network links by allowing single-photon-based protocols to be used in the high-fidelity regime.

DOI: [10.1103/29s3-dzl8](https://doi.org/10.1103/29s3-dzl8)

I. INTRODUCTION

Hybrid quantum systems, combining the complementary strengths of matter-based qubits for stationary storage and processing with photonic qubits as flexible interconnects, are of interest for a broad range of applications, including quantum cryptography and communication tasks [1,2], modular large-scale computing [3–6], tests of fundamental physics [7,8], and distributed optical interferometry [9]. Maximally entangled states between remote qubits serve as a universal resource for such systems; other primitives can be built on top of it, for instance using teleportation protocols [10–13]. Heralded remote entanglement schemes, where success is indicated by the observation of one or more photons emitted from the matter nodes after interference on a beam splitter [14], avoid errors due to link loss and have been demonstrated in a number of platforms, including trapped ions [15], neutral atoms [16], defect centers in diamond [17,18], solid-state quantum dots

[19], and rare-earth defects in crystals [20]. Trapped-ion demonstrations have reached remote entanglement fidelities of 97% [13] and rates of $\sim 200 \text{ s}^{-1}$ [21,22]. While this represents the state of the art in remote entanglement generation to date, both rates and fidelities still fall short of the demands of future large-scale quantum information processing applications [23–26].

Laser-cooled, trapped atoms are well suited as quantum network interfaces: they are indistinguishable, offer long coherence times as quantum memories [27], and support high-fidelity logic gates as part of larger-scale devices [28]. Furthermore, spontaneous emission from short-lived states provides a native optical interface. While optical cavities can additionally be employed to tailor the emission process [16,29–31], here, we focus on free-space configurations. In previous work on remote entanglement generation, trapped atoms have mostly been treated as ideal, stationary dipole emitters, neglecting their quantum-harmonic motion within the confining potential entirely or treating it in a simplified, semiclassical fashion. While phase modulation and recoil effects associated with atomic motion had already been recognized as an inherent source of errors in early proposals [32], other, technical limitations to the fidelity observed in demonstration experiments have overshadowed this issue until recently. As practical realizations are starting to target entangled-state infidelities at or below the percent-level [13,25,26,33,34] while also aiming to maximize the entanglement generation rate—and thus to minimize time spent on laser cooling—a detailed

*Present address: Institute for Quantum Electronics, Department of Physics, ETH Zürich, Otto-Stern-Weg 1, Zürich 8093, Switzerland.

†Contact author: david.nadlinger@physics.ox.ac.uk

Published by the American Physical Society under the terms of the [Creative Commons Attribution 4.0 International](https://creativecommons.org/licenses/by/4.0/) license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

understanding of the role of motion becomes essential. Moreover, even for ground-state-cooled atoms, motional effects would appear to fundamentally limit the achievable fidelity. The recoil energy associated with spontaneously emitted photons tends not to be entirely negligible on the energy scale of the harmonic confinement [35–37]. Furthermore, as the linewidth of the atomic dipole transitions chosen for efficient photon collection is typically larger than the motional frequency, the motional sidebands are not resolved and hence cannot be rejected using spectral filtering; and as photons are generated via spontaneous emission into vacuum following impulsive excitation, techniques to drive single-qubit gates in a motionally insensitive fashion are not applicable [38–40].

In this article, we provide an intuitive framework for describing the effect of atomic motion on photon-mediated entanglement generation. We start by giving a coherent description of spontaneous emission from a trapped emitter into free space (Sec. II). We then consider the effect of collection into a single optical mode (Sec. III) and show that the overall effect of the emission process is described by a coherent superposition of displacement operators, whose rotating-frame phase-space direction depends on the detection time of the emitted photon. We apply this framework to the generation of remote entanglement using a single herald photon (Sec. IV) and to entanglement swapping using two herald photons generated from multiple decay channels (Sec. V) or in multiple time bins (Sec. VI). We discuss the fidelity loss for an initial thermal state of motion, as well as strategies for its correction by disentangling spin and motion *post hoc* following a heralding event.

II. SPIN-MOTION ENTANGLEMENT IN SPONTANEOUS DECAY

Consider a dipole emitter confined in a three-dimensional trapping potential, such as an atomic ion in a radio-frequency trap or an atom trapped in optical tweezers (Fig. 1). For simplicity, we will use “atom” to refer to any dipole emitter throughout the manuscript and “spin” for its discrete internal degrees of freedom. Such a system is modeled by the Hamiltonian

$$H = \underbrace{\sum_i \hbar \omega_i |i\rangle\langle i|}_{\text{internal energy levels}} + \underbrace{\sum_j \hbar \mu_j b_j^\dagger b_j}_{\text{motion in trap}} + \underbrace{\sum_{k\sigma} \hbar c |\mathbf{k}| a_{k\sigma}^\dagger a_{k\sigma}}_{\text{free e.m. field}} - \underbrace{\mathbf{d} \cdot \mathbf{E}(\mathbf{X})}_{\text{dipole interaction}}, \quad (1)$$

where i labels the internal states of the emitter (with energies $\hbar \omega_i$) and j describes the harmonic oscillator modes of the confined motion of the particle (with trap frequencies μ_j and phonon operators satisfying $[b_j, b_{j'}^\dagger] = \delta_{jj'}$)

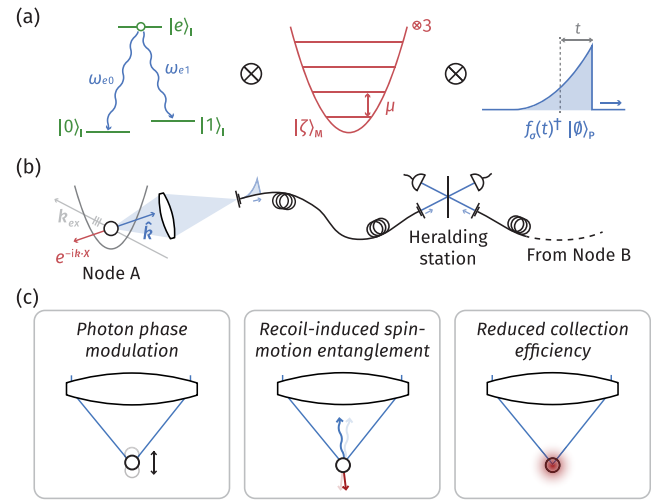


FIG. 1. Spontaneous emission in trapped atoms. (a) We consider atoms decaying from an excited state to one or more stable ground states, involving internal atomic states, center-of-mass motion, and excitations in the photonic multimode vacuum. (b) To generate entanglement between two remote quantum information processing nodes, the trapped atoms are simultaneously excited and the emission is combined at a central heralding station. (c) Atomic motion impacts the remote entanglement generation process through multiple interlinked effects related to recoil during excitation and photon emission; we describe its consequences in a unified framework.

making up the total atom position operator $\mathbf{X}$. We have also expanded the electromagnetic field into plane-wave modes with wavevector $\mathbf{k}$ and polarization σ (satisfying $[a_{k\sigma}, a_{k'\sigma'}^\dagger] = \delta_{kk'} \delta_{\sigma\sigma'}$ in some nominal quantization volume, though we shall shortly take the continuum limit).

Here, we are interested in describing the spontaneous decay dynamics from a single excited state $|e\rangle$ to one or more lower-lying levels $\{|i\rangle\}_i$ assumed to be stable (see Fig. 1). The relevant components of the electric dipole interactions are then

$$\mathbf{d} \cdot \mathbf{E}(\mathbf{X}) = i \sum_{i k \sigma} \mathcal{E}_k g_{i, \hat{\mathbf{k}} \sigma} |i\rangle\langle e| e^{-i \mathbf{k} \cdot \mathbf{X}} a_{k \sigma}^\dagger + \text{h. c.}, \quad (2)$$

where $g_{i, \hat{\mathbf{k}} \sigma} = \langle i | \mathbf{d} | e \rangle \cdot \hat{\mathbf{e}}_{k \sigma}$ is the matrix element of the electric dipole operator projected onto the normalized electric field vector $\hat{\mathbf{e}}_{k \sigma}$ corresponding to the mode $\hat{\mathbf{k}} \sigma$, and $\mathcal{E}_k$ is a field strength prefactor incorporating the mode density (splitting up $\mathbf{k}$ into its magnitude k and unit vector $\hat{\mathbf{k}}$).

The $e^{-i \mathbf{k} \cdot \mathbf{X}}$ term, describing the effect of atomic motion on the coupling to the radiation field, can be interpreted in two equivalent ways: manifestly, as an emitter-position-dependent phase of the resulting field, but because position and momentum are conjugate variables, also as a shift in oscillator momentum due to recoil from the emitted photon. There are two salient parameters for the related dynamics: first, the Lamb-Dicke

parameter $\eta_j = \sqrt{E_{\text{recoil}}/(\hbar\mu_j)} = \sqrt{\hbar k^2/(2M\mu_j)}$ giving the recoil energy for emitter of mass M in oscillator quanta for each motional mode j [41] and second, the ratio μ_j/Γ of motional frequency to emission linewidth ($\Gamma \propto \sum_{i\mathbf{k}\sigma} |\mathcal{E}_k \mathbf{g}_{i,\mathbf{k}\sigma}|^2$) quantifying the amount of oscillator time evolution during the emission process.

In detail, consider the Hamiltonian in the interaction picture regarding the free evolution of internal/motional/field states. Assuming the equilibrium position of the emitter to be in the coordinate origin, we can expand $e^{i\mathbf{k}\cdot\mathbf{X}} = \bigotimes_j D_j(i\eta_j \hat{k}_j)$ in terms of displacement operators $D_j(\beta) = e^{\beta b_j^\dagger - \bar{\beta} b_j}$ and the projection $\hat{k}_j$ of the normalized direction $\hat{\mathbf{k}}$ onto each motional mode. Defining a multi-mode displacement $\mathcal{D}_{i\hat{\mathbf{k}}}(t) := \bigotimes_j D_j(i\eta_j \hat{k}_j)$, we obtain the interaction Hamiltonian in rotating-wave approximation,

$$H_{\text{int}}(t) \approx -i \sum_{i\mathbf{k}\sigma} e^{i(kc - \omega_{ei})t} \mathcal{E}_k \mathbf{g}_{i,\mathbf{k}\sigma} |i\rangle \langle e| a_{\mathbf{k}\sigma}^\dagger \mathcal{D}_{-i\hat{\mathbf{k}}}(t) + \text{h. c.} \quad (3)$$

Let the atom initially be in the excited state $|\psi_0\rangle_{\text{IMP}} = |e\rangle_I |\zeta\rangle_M |\emptyset\rangle_P$, where $|\zeta\rangle_M$ is an arbitrary state of motion (w.l.o.g. assumed to be pure for notational convenience) and $|\emptyset\rangle_P$ is the photonic multimode vacuum. Transitioning to a mode continuum, the evolution under H_{int} is readily solved in the Weisskopf-Wigner approximation [42–45]. At times much longer than Γ^{-1} after the initial excitation, where the excitation has been transferred from the atom to a propagating photon wavepacket, the interaction-picture state is approximately independent of time (see Appendix A). The joint state between spin, motion and photon takes a straightforward form after transforming to time-domain plane-wave modes $\tilde{a}_\sigma^\dagger(t, \hat{\mathbf{k}})$,

$$|\psi\rangle_{\text{IMP}} \approx \sum_i \tilde{\mathcal{E}}_{\omega_{ei}/c} |i\rangle_I \int_0^\infty e^{-\left(\frac{\Gamma}{2} + i\omega_{ei}\right)t} \times \sum_\sigma \int_{S^2} \mathbf{g}_{i,\mathbf{k}\sigma} \mathcal{D}_{-i\hat{\mathbf{k}}}(t) |\zeta\rangle_M \tilde{a}_\sigma^\dagger(t, \hat{\mathbf{k}}) d^2\hat{\mathbf{k}} dt |\emptyset\rangle_P. \quad (4)$$

This is a coherent superposition of the ground states dipole-connected to the excited state $|e\rangle_I$, entangled with photon wavepackets centered on the respective transition frequency, with an exponential envelope given by the excited state lifetime. Each emission time is associated with a recoil momentum kick at that instant, so a photon with time coordinate t along the wavepacket comes with a displacement of the j th mode in the rotating-frame phase-space direction $-ie^{i\mu_j t}$. The displacement amount $\eta_j \hat{k}_j$ depends on the Lamb-Dicke parameter for the given mode and the projection of the emission direction $\hat{\mathbf{k}}$ on the mode direction.

III. COLLECTION OF EMITTED PHOTONS

To access and manipulate the photonic part of the system, a fraction of the dipole radiation emitted into free space is collected using an optical apparatus. The prototypical case is that of a lens objective focusing radiation emitted into a given solid angle quantified by its numerical aperture (NA) $\sin \vartheta_{\text{max}}$ [see Fig. 2(a)]. In the context of remote entanglement generation, single-mode optical fibers or waveguides are typically employed to transmit emitted photons between emitters and enable efficient interference. This has the effect of selecting only a single transversal mode F with time-domain creation operators $\tilde{f}_\sigma(t)^\dagger$ out of all the photonic modes P .

After collection into the single mode F and discarding (tracing out) the other photonic modes, the final state from Eq. (4) becomes

$$\rho_{\text{IMF}} \approx \text{proj} \left(\sum_i |i\rangle_I \int_0^\infty e^{-\left(\frac{\Gamma}{2} + i\omega_{ei}\right)t} \times \sum_\sigma K_{i\sigma}(t) |\zeta\rangle_M \tilde{f}_\sigma^\dagger(t) dt |\emptyset\rangle_F \right) + \rho_{\text{IM}}^{\text{rest}} \otimes |\emptyset\rangle\langle\emptyset|_F, \quad (5)$$

where $\text{proj}(|\psi\rangle)$ denotes a (subnormalized) pure state $|\psi\rangle\langle\psi|$ and

$$K_{i\sigma}(t) := \int_{S^2} w_{i\sigma}(\hat{\mathbf{k}}) \mathcal{D}_{-i\hat{\mathbf{k}}}(t) d^2\hat{\mathbf{k}}, \quad (6)$$

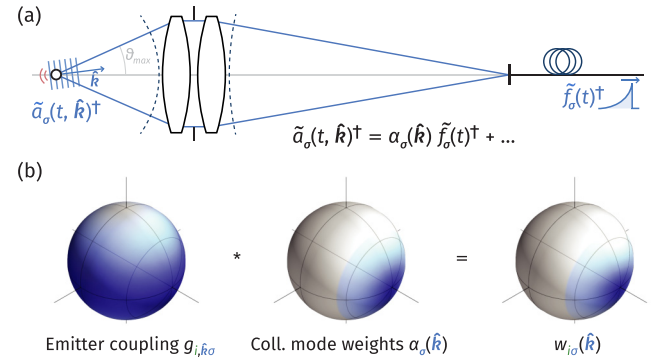


FIG. 2. Collection from a large solid angle into a single mode. (a) Typical optical apparatus for remote entanglement generation, where a lens system (numerical aperture $\sin \vartheta_{\text{max}}$) refocuses emitted photons into a single-mode fiber. In the plane-wave decomposition, each mode $\tilde{a}_\sigma^\dagger(t, \hat{\mathbf{k}})$ is transformed into a superposition of the selected mode $\tilde{f}_\sigma^\dagger(t)$ [with coefficient $\alpha_\sigma(\hat{\mathbf{k}})$, including, e.g., diffraction effects] and some other modes that are subsequently traced out. (b) A photon collected into $\tilde{f}_\sigma$ hence corresponds to a coherent superposition of emission directions $\hat{\mathbf{k}}$, with the weights $w_{i\sigma}(\hat{\mathbf{k}})$ arising as a product of the emission pattern $\mathbf{g}_{i,\mathbf{k}\sigma}$ and the collection mode profile $\alpha_\sigma(\hat{\mathbf{k}})$ (shown for dipole emission from an atom in the center of the sphere, depth of color indicating amplitudes that happen to be real and positive here).

while ρ_{IM}^{rest} elides the exact form of the atomic state where the photon is not collected [46]. In Eq. (6), $w_{i\sigma}(\hat{\mathbf{k}}) \in \mathbb{C}$ gives the contribution of each transversal plane wave mode $\hat{\mathbf{k}}\sigma$ to the single mode selected [47]. It arises as a product of the emission pattern given by $g_{i,\hat{\mathbf{k}}\sigma}$, i.e., the dipole coupling to a particular plane-wave mode, and the contribution $\alpha_\sigma(\hat{\mathbf{k}})$ of the plane wave $a_\sigma^\dagger(\mathbf{k})$ to the extracted mode $f_\sigma^\dagger(k)$ as defined by the transformation describing the action of the imaging system (which we assume to be linear), $a_\sigma^\dagger(\mathbf{k}) = \alpha_\sigma(\hat{\mathbf{k}})f_\sigma^\dagger(k) + \dots$ [see Fig. 2(b) and Appendix B]. Equation (5) is easily generalized from pure initial states of motion $|\zeta\rangle_M$ to arbitrary mixed ρ_M using their spectral decomposition.

With this, we have arrived at the central statement of our theoretical framework: the geometrical aspects of spontaneous emission in the combined system of the internal/motional atomic and photonic degrees of freedom are entirely described by the set of “kick operators” $K_{i\sigma}(t)$ [Eq. (6)] associated with every time offset t within the photonic wavepacket. This description holds for arbitrary initial states of motion, and for arbitrary collection modes, even in the high-NA limit.

$K_{i\sigma}(t)$ provides a unified description of the phase modulation, recoil, and coupling efficiency effects illustrated in Fig. 1(b). One can recognize the first two effects, photon phase modulation and atomic recoil, as two alternative interpretations of the displacement operators in Eq. (6) (which arose as the rotating-frame equivalent of the $e^{-ik\cdot X}$ plane-wave propagators, as discussed in the previous section). The motional dependence of the coupling efficiency is captured by the fact that $K_{i\sigma}(t)$ is not unitary. Only in the small-collection-angle (zero-NA) limit, where the position resolution of the collection process vanishes, is the overall effect approximated by a simple displacement operator (e.g., for collection along the z -axis only, $w_{i\sigma}(\hat{\mathbf{k}}) \propto \delta(\hat{\mathbf{k}} - \hat{\mathbf{z}})$, though with a vanishing prefactor). For nonvanishing collection angles, the nonunitarity of the resulting superposition of displacement operators captures the reduction in collection efficiency from an extended (e.g., thermally smeared-out) wavepacket. In the high-NA case, the projections of the plane-wave components $\hat{\mathbf{k}}$ onto the motional modes will significantly differ across the region where the weights $w_{i\sigma}(\hat{\mathbf{k}})$ are nonzero, such that $K_{i\sigma}(t)$ will create entanglement across the motional modes.

Nevertheless, the time-resolved detection of a single photon in the output mode at a time offset t projects the joint state into one with known kick operators $K_{i\sigma}(t)$ having been applied. If there are multiple decay paths to several ground states $|i\rangle_I$ or a nonexcited state component, this results in spin-motion entanglement. In the following sections, we propose the coherent correction of this and demonstrate that approximately inverting the kick operators with a spin-dependent displacement is enough to recover remote entanglement fidelity given experimentally

relevant parameters. We also make use of the flexibility of $w_{i\sigma}(\hat{\mathbf{k}})$ to express arbitrary collection modes to demonstrate that collection into an appropriate standing wave pattern reduces the motional sensitivity (see Sec. VB).

Note that our discussion has focused specifically on emission into the mode continuum of free space, from which a single mode is then selected afterward (without materially changing the mode structure the atom couples to). To treat emission into the mode of a leaky cavity, a non-Hermitian Schrödinger equation can be employed instead [29,30]. By deriving appropriate kick operators $K_{i\sigma}(t)$ for the propagating output modes, much of the following sections can be generalized to cavity-based systems, but we leave this for future work.

IV. SINGLE-PHOTON-BASED ENTANGLEMENT

Given our understanding of the joint atom-photon state after excitation of a single emitter, we now examine the consequences of the motional degrees of freedom on several protocols for heralded photon-mediated entanglement generation, where emission from two remote atoms is brought together and interferes before being detected. Here, we first analyze a class of protocols based on the detection of a single heralding photon [32,48]. It is well understood that these protocols are sensitive to the optical phase along the optical link connecting the nodes and suffer from an intrinsic rate-fidelity trade-off. Still, such protocols were preferred in several quantum networking demonstrations using solid-state qubits, even if the link phase had to be actively stabilized [19,49–52]. Their reliance on only a single herald photon leads to significantly improved rates compared to two-photon schemes when the overall optical system efficiency is low.

The errors intrinsic to single-photon schemes, including from an uncontrolled but short-term-stable differential link phase, can be perfectly corrected using a simple single-stage entanglement distillation protocol [53,54]. Two Bell pairs are combined into one with theoretically perfect fidelity, which in the high-loss regime can still yield higher rates than the two-photon protocols discussed later [53]. However, the motion of atomic emitters introduces uncorrelated phase errors that lack the structure to be corrected by this distillation scheme. This issue was already described by Cabrillo *et al.* in their original proposal [32], although only for neighboring atoms in free space; the only experimental realization to date, using two co-trapped ions, was indeed limited by motional errors [55]. Here, we show how to describe the motional effects for the technologically relevant case of single-mode collection, and that they can be compensated after the fact.

The canonical realization of the single-photon scheme is based on emitters with two stable internal states, $|0\rangle_I$ and $|1\rangle_I$ [see Fig. 3(a)]. $|0\rangle_I$ does not participate in the decay process, and only a small fraction

$p \in [0, 1]$ is coherently transferred to the short-lived excited state $|e\rangle_I$ (for instance, by preparing a superposition $\sqrt{1-p}|0\rangle_I + \sqrt{p}|1\rangle_I$ and only exciting $|1\rangle_I$ to $|e\rangle_I$). Immediately after excitation (assuming again w.l.o.g. a pure motional state $|\zeta\rangle_M$ for notational clarity), the joint atom-photon state is

$$|\psi\rangle_{IMF} = \sqrt{1-p}|0\rangle_I|\zeta\rangle_M|\emptyset\rangle_F + \sqrt{p}|e\rangle_IL|\zeta\rangle_M|\emptyset\rangle_F, \quad (7)$$

where $L = \mathcal{D}_{i\hat{k}_{ex}}(0)$ is the momentum kick of the excitation laser pulse [56,57] incident on the atom from direction $\hat{k}_{ex}$. After a duration long compared to Γ^{-1} has elapsed, the state reads

$$\begin{aligned} \rho_{IMF} = & \text{proj} \left(\sqrt{1-p}|0\rangle_I|\zeta\rangle_M|\emptyset\rangle_F + \sqrt{p}|1\rangle_I \right. \\ & \times \int_0^\infty e^{-\left(\frac{\Gamma}{2} + i\omega_{el}\right)t} K(t)|\zeta\rangle_M \tilde{f}^\dagger(t) dt |\emptyset\rangle_F \\ & \left. + p|1\rangle_I\langle 1| \dots |\emptyset\rangle\langle\emptyset|_F, \right) \end{aligned} \quad (8)$$

with a combined kick operator $K(t) := K_{1\sigma}(t)L$ including both the laser excitation and decay, and where the dots elide the exact form of the component where the atom was excited but the emitted photon was not collected.

The emission from two nodes A and B is then combined on a central 50:50 beamsplitter [Fig. 3(a)], and the detection of a single photon in either output mode heralds the creation of a remotely entangled state. Concretely, a click for small excitation probabilities p heralds the creation of a maximally entangled state close to $|\Psi^\phi\rangle := (|01\rangle \pm e^{i\phi}|10\rangle)/\sqrt{2}$, with the phase ϕ given by the optical path length difference $\phi_A - \phi_B$ from the nodes A and B to the beamsplitter (shifted by π depending on the clicking detector). For finite p , double excitation leads to an intrinsic error term $\approx p|11\rangle$, leading to a rate-fidelity trade-off even in the absence of motion [32].

To include motional effects, consider two wavepackets of the form (8) and combine the output modes $\tilde{f}_A(t)$ and $\tilde{f}_B(t)$ on a beam splitter, transforming them as $\tilde{f}_A^\dagger(t) \mapsto (\tilde{g}^\dagger(t) + \tilde{h}^\dagger(t))/\sqrt{2}$ and $\tilde{f}_B^\dagger(t) \mapsto (\tilde{g}^\dagger(t) - \tilde{h}^\dagger(t))/\sqrt{2}$. The success condition for the protocol is a single heralding detector click in either $\tilde{g}(t)$ or $\tilde{h}(t)$, but not both, at any time t (relative to the start of the photon wavepacket following initial excitation, retarded by the optical path from atoms to the detectors, which must be matched between the nodes). Our model readily yields analytical solutions for the resulting state in full generality (see Appendix C). For illustrative purposes, we here consider the case where the parameters of both nodes are identical. Assume the state of either node at the beginning of the protocol is $\rho_{IMF} = |0\rangle\langle 0|_I \otimes \rho_M \otimes |\emptyset\rangle\langle\emptyset|_F$ for some general motional state ρ_M , and consider the postherald state

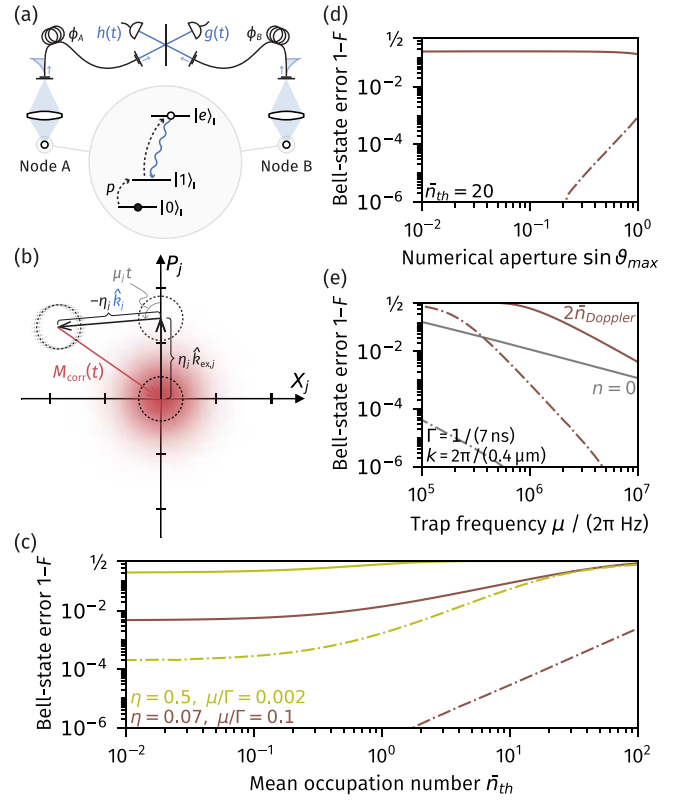


FIG. 3. Single-photon-based remote entanglement. (a) Each node starts in a superposition $\sqrt{1-p}|0\rangle_I + \sqrt{p}|1\rangle_I$ and excites only the $|1\rangle_I$ fraction to the emitting state $|e\rangle_I$. For small p , the detection of a single photon heralds the creation of an approximate maximally entangled state $\propto |01\rangle \pm e^{i(\phi_A - \phi_B)}|10\rangle$. (b) In phase space for the j th motional mode, the $|0\rangle_I$ component of the state is not disturbed, while the $|1\rangle_I$ component is displaced both by the excitation laser pulse and the emitted photon. The displacement amounts are given by the product of Lamb-Dicke parameter η_j and the projection of excitation and emission directions $\hat{k}_{ex}$ and $\hat{k}$ onto the mode direction; the rotating-frame phase-space direction of the emission kick varies with the detection time as $e^{i\mu_j t}$, where μ_j is the motional mode frequency. (c) Heralded Bell-state fidelity for isotropic trap frequency/linewidth μ/Γ and Lamb-Dicke parameters η representative of ion trap (brown) and optical tweezer (olive) experiments. The error from spin-motion entanglement (solid lines) can be considerably reduced by overlapping the motional wavepackets with a spin-dependent force M_{corr} after heralding (dashed-dotted lines). In all panels, collection orthogonal to the excitation pulse with NA 0.6 into an optimally sized Gaussian mode and $p \ll 1$ are assumed. (d) Even outside the zero-NA limit, a spin-dependent displacement suffices to heavily suppress motional errors (shown for $\eta = 0.07$, $\mu/\Gamma = 0.1$). (e) The spin-dependent correction allows high-fidelity operation at twice the Doppler limit $\hbar\Gamma/(2k_B)$ for typical trapped-ion cooling transitions, where motional errors are otherwise significant. Ground-state cooling alone (gray) may not be sufficient.

for an infinite heralding window. Tracing out the motional degree of freedom at the end of the protocol after a click,

we obtain the following compact expression for the fidelity of the produced atom-atom state to the nearest maximally entangled state of the form $|\Psi^\phi\rangle$ [58],

$$F^{\text{eff.} \ll 1} \approx (1-p) \frac{1+C}{2}, \quad C := \frac{\int_0^\infty e^{-\Gamma t} |\text{tr}(K(t) \rho_M)|^2 dt}{\int_0^\infty e^{-\Gamma t} \text{tr}(K(t)^\dagger K(t) \rho_M) dt}, \quad (9)$$

with $0 \leq C \leq 1$, and where the denominator of C gives the collection efficiency. Using Eq. (9), we can easily obtain the fidelity of the generated states for arbitrary motional states and emission/collection geometries through the expectation values of the kick operator. In general, the spatial integral in Eq. (6) can be evaluated numerically (see Appendixes B and E).

In the zero-NA limit, where photons are collected only exactly along one axis (and correspondingly, the success probability is infinitesimally small), $K(t)$ is simply a concatenation of two displacements from laser excitation and recoil, so Eq. (9) is easily evaluated via the characteristic function of the components of ρ_M [59]. Let the excitation laser propagate at angle χ to the collection axis, and, to obtain a compact expression, let the atom be confined in an isotropic potential with resulting uniform motional frequency μ . For an initial thermal state with mean occupation number $\bar{n}_{th}$ in each motional mode, we find

$$C \approx \Gamma \int_0^\infty e^{-\Gamma t} e^{-2\eta^2(1-\cos(\chi)\cos(\mu t))(1+2\bar{n}_{th})} dt. \quad (10)$$

For this special case of zero NA, our results thus coincide with those derived using a different approach already as part of the original proposal by Cabrillo *et al.* for direct interference on a (then vanishingly small) free-space detector [32].

We can understand this as spin-motion entanglement caused by the fact that the kick operator only acts on the emitting $|1\rangle_I$ component, as illustrated in Fig. 3(b). As evident from Eq. (10), an excitation pulse propagating in the collection direction, $\chi = 0$, mitigates this by causing the momentum kicks/phase factors to cancel at least for small t , but this is often impractical as it amounts to pointing the excitation laser into the collection optics. The fidelity curves in Figs. 3(c)–3(e) are given for the moderate case of orthogonal excitation and collection, $\chi = \pi/2$. The impact of this residual entanglement could be reduced by engineering the motional state, e.g., through ground-state cooling or squeezing.

Our analysis suggests a different strategy altogether: after detection of the herald photon at time t , the kick operator $K(t)$ applied to the atom in state $|1\rangle_I$ is known. By undoing it with an appropriate spin-dependent operation in each node, the unwanted spin-motion entanglement can be reduced. An exact inversion might not be possible in the

general high-NA case, as the kick operators are complicated superpositions of multimode displacements. Applying just a spin-displacement to each mode, however, is certainly feasible (spin-dependent forces generated using narrow-linewidth lasers or stimulated Raman transitions are part of the standard trapped-atom toolbox [60]). After receiving the timestamp t of the photon detection from the heralding station, each node can apply

$$U_{\text{corr}}(t) = |0\rangle\langle 0|_I \otimes \mathbb{1}_M + |1\rangle\langle 1|_I \otimes M_{\text{corr}}(t), \quad \text{where} \\ M_{\text{corr}}(t) = \bigotimes_j e^{i\eta_j^2 \hat{k}_{\text{ex},j} k_{\text{avg},j} \sin(\mu_j t)} \\ \times D_j \left(-i\eta_j (\hat{k}_{\text{ex},j} - e^{i\mu_j t} k_{\text{avg},j}) \right), \quad (11)$$

where $k_{\text{avg},j}$ is chosen to compensate the emission kick to maximize the fidelity (so outside the zero-NA limit a compromise “average” displacement). $U_{\text{corr}}(t)$ effectively modifies the kick operator $K(t)$ to $M_{\text{corr}}(t)K(t)$. As shown in Figs. 3(c)–3(e), for parameters typical of trapped-ion experiments such a simple displacement correction can virtually eliminate motional errors, even at Doppler-cooling temperatures.

Thus, our results demonstrate that single-photon-based schemes can feasibly be adapted to achieve low error rates even without ground-state cooling and as such enable a significant increase in entanglement generation rate in the loss-dominated scenarios of interest for long-distance quantum networking.

V. TWO-PHOTON-BASED ENTANGLEMENT (FOLLOWING SINGLE EXCITATION)

We now turn our attention toward remote entanglement protocols based on a single (strong) excitation to $|e\rangle_I$ and subsequent decay to the stable states $|0\rangle_I$ and $|1\rangle_I$ along two possible channels with orthogonal photonic modes. Protocols using time-bin encoding are closely related, but involve two separate excitation events; we shall discuss them in Sec. VI.

In contrast to the single-photon protocol discussed in Sec. IV, two-photon protocols [61,62] are not based on weak excitation and can thus work in the high-efficiency regime without an inherent rate-fidelity trade-off. An (approximately) maximally entangled atom-photon state is generated at each node, and the entanglement is subsequently swapped onto the remote atoms by a Bell-basis measurement on the photons [14]. Typically, atomic structure and collection geometry are chosen such that the decay channels are mapped to orthogonal polarizations in the single-mode fiber linking the nodes [22,63–66], although, e.g., frequency encoding has also been used [15]. The robust nature of two-photon-based protocols, when combined with the well-controlled nature of atomic emitters, has allowed experiments to reach the highest fidelities to

date [13] and, among optical remote entanglement experiments reaching Bell-state fidelities $>70\%$, also the highest rates [21,22].

A key feature aiding the robustness of two-photon protocols is that only the relative phase of the wavepackets corresponding to the two decay channels influences the entangled state phase [61]; as long as the photons remain temporally overlapped on the Bell-state analyser, slow changes to the absolute optical phase between the nodes can be tolerated. Because of this, such protocols are also relatively less sensitive to motional effects when the emission channels are collected along the same path. However, as other error sources in practical realizations are reduced, the effect of atomic motion is no longer negligible [13,55]. Here, we derive analytical expressions for the postherald states and their fidelities in the full picture including motional effects and provide an intuitive explanation of the error mechanisms in the spin-motion-entanglement picture.

Concretely, consider an atom that has just been excited to the state $|e\rangle_I L|\zeta\rangle_M |\emptyset\rangle_P$ and subsequently decays along two possible paths to the stable qubit states $|0\rangle_I$ and $|1\rangle_I$ [see Fig. 4(a)]. To aid clarity of exposition, we shall restrict the model such that the two decay channels give rise to excitations in fully orthogonal collection modes H and V (where the notation hints at the common case of emission from two orthogonal dipoles mapping to orthogonal linear polarizations in a single-mode fiber [22,67]; see Appendix F). The postdecay atom-photon state is then

$$\begin{aligned} \rho_{IMF} = & \text{proj} \left(|0\rangle_I \int_0^\infty e^{-\left(\frac{\Gamma}{2} + i\omega_{e0}\right)t} K_H(t) |\zeta\rangle_M \tilde{f}_H^\dagger(t) dt |\emptyset\rangle_F \right. \\ & + |1\rangle_I \int_0^\infty e^{-\left(\frac{\Gamma}{2} + i\omega_{e1}\right)t} K_V(t) |\zeta\rangle_M \tilde{f}_V^\dagger(t) dt |\emptyset\rangle_F \Big) \\ & + \rho_{IM,\text{dark}} \otimes |\emptyset\rangle\langle\emptyset|_F, \end{aligned} \quad (12)$$

where $\rho_{IM,\text{dark}}$ describes the atomic state for the case where no photon was collected, which is not of interest here due to the later conditioning on a two-photon herald (unless imperfect detectors with dark counts are considered). As in Sec. IV, we have absorbed the momentum kick L from the excitation laser, which now applies to both components of the superposition, into the operators $K_i(t)$.

As illustrated in Fig. 4(a), the emission from two nodes A and B is combined in a photonic Bell-state measurement apparatus, consisting, e.g., of a 50:50 beamsplitter followed by two polarizing beamsplitters to implement polarization-sensitive single-photon detection. Without considering atomic motion, a two-photon coincidence on two detectors of orthogonal polarization projects the input state $(|0H\rangle + |1V\rangle)^{\otimes 2}/2$ into a maximally entangled atom-atom state $(|01\rangle \pm |10\rangle)/\sqrt{2}$ (with the sign depending on the detector combination).

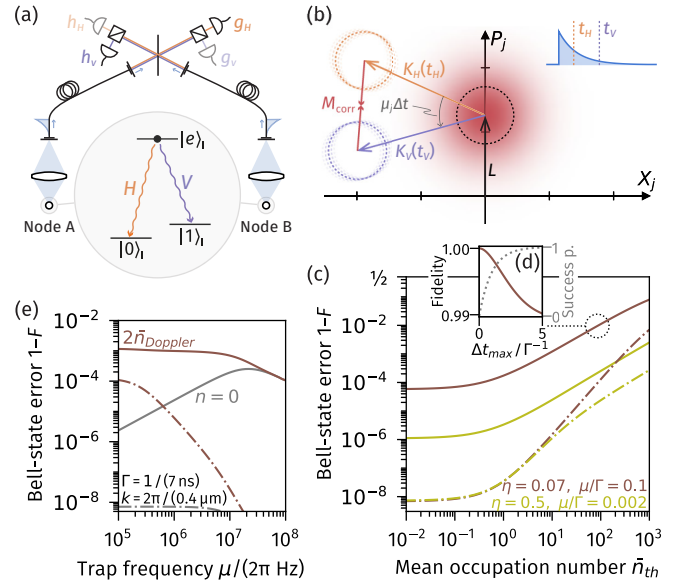


FIG. 4. Two-photon-based remote entanglement. (a) Both parties excite their atom into the emitting state $|e\rangle_I$; the two decay channels to $|0\rangle_I$ and $|1\rangle_I$ are associated with orthogonal photon modes (here labeled H and V). A coincident detection of a photon from both modes in the photonic Bell-state measurement setup heralds the creation of a maximally entangled atom-atom state $(|01\rangle \pm |10\rangle)/\sqrt{2}$. (b) As illustrated in the phase space of the j th motional mode of one of the nodes, a difference $\Delta t := t_V - t_H \neq 0$ in detection times for the photons from either decay channel leads to spin-motion entanglement. (c) Reduction in fidelity assuming a symmetric trap with the same ratio of trap frequency to linewidth μ/Γ and Lamb-Dicke parameter η for all directions, for two sets of parameters characteristic of trapped-ion (brown) and optical tweezer (olive) experiments, given an initial thermal state (excitation pulse orthogonal to the collection direction; collection with numerical aperture 0.6 into an optimally sized Gaussian fiber mode). Solid lines show the fidelity when ignoring atomic motion, dashed-dotted lines the improvement through applying a spin-dependent displacement $M_{\text{corr}}(t_H, t_V)$ to disentangle spin and motion. The inset (d) shows (for $\bar{n}_{th} = 85$) how the error can also be reduced if only heralds with $|\Delta t| < \Delta t_{\text{max}}$ are accepted, but at a significant cost in entanglement generation rate. (e) If the same atomic transition is used for Doppler cooling and atom-photon entanglement, the fidelity is approximately independent of the trap frequency until the resolved-sideband regime is reached (brown, solid). Cooling to the motional ground state (grey) can improve the fidelity in the weakly confined regime. A spin-dependent displacement correction greatly reduces the error in either case (dashed-dotted lines).

To obtain the state in the full picture including atomic motion, consider two copies of the wavepacket from Eq. (12) incident on the Bell-state analyser apparatus (with some arbitrary motional states ρ_{MA} and ρ_{MB} in each node). An ideal beamsplitter network maps $\tilde{f}_{Ai}^\dagger(t) \mapsto (\tilde{g}_i^\dagger(t) + \tilde{h}_i^\dagger(t))/\sqrt{2}$ and $\tilde{f}_{Bi}^\dagger(t) \mapsto (\tilde{g}_i^\dagger(t) - \tilde{h}_i^\dagger(t))/\sqrt{2}$ for either polarization $i \in \{H, V\}$. A herald event then

consists of two-photon coincidence, where after simultaneous excitation, a photon is detected in both one of $\{g_H, h_H\}$ at time t_H and one of $\{g_V, h_V\}$ at time t_V .

The derivation of the full postherald state is given in Appendix D. Here, assume again for compactness that the properties of both nodes are the same, and all herald events up to $t \rightarrow \infty$ are accepted. The fidelity of the resulting state to the closest maximally entangled state is then given by

$$F = \frac{1 + C}{2}, \quad \text{where} \\ C := \frac{\int_0^\infty \int_0^\infty e^{-\Gamma(t_H+t_V)} |\text{tr}(K_V(t_V)^\dagger K_H(t_H) \rho_M)|^2 dt_V dt_H}{\prod_{i=H,V} \int_0^\infty e^{-\Gamma t_i} \text{tr}(K_i(t)^\dagger K_i(t) \rho_M) dt_i}. \quad (13)$$

According to Eq. (13), a difference in $K_V(t_V)$ and $K_H(t_H)$ will generically give rise to a reduction in fidelity. This can occur because of two distinct reasons. First, there might, e.g., be a difference in geometry between the decay channels, such that K_V and K_H are fundamentally different (discussed in Sec. VB). Second, even if the decay channels are perfectly matched, an unaccounted-for difference in detection times still can be an intrinsic source of errors (Sec. VA).

A. Timing-related errors

The detection times t_H and t_V of the two herald photons are both independently distributed, as given by the exponential envelope determined by the excited state lifetime Γ^{-1} . Even if any geometric differences between the kick operators are neglected, $K := K_H \approx K_V$, consider $K(t_H)$ and $K(t_V)$ associated with the two times offsets within the wavepacket (12). In the interaction-picture phase space for the j th motional mode with frequency μ_j [see Fig. 4(b)], the first momentum kick from the excitation laser L is identical for both $|0\rangle$ (H) and $|1\rangle$ (V) components, but the direction of the subsequent emission kick is entangled with the photon wavepacket. A difference $\Delta t := t_V - t_H \neq 0$ in detection times leads to an angle of $\mu_j \Delta t$ between the displacement vectors.

Consequently, detecting photons at different times leaves behind entanglement between internal states and motion, dephasing the entangled state after tracing out the motional degrees of freedom. Figure 4(c) shows numerical results for NA 0.6 collection from two representative configurations for various initial thermal states. To obtain an analytical approximation, consider again the zero-NA limit, where $w_{i\sigma}(\mathbf{k}) \propto \delta(\mathbf{k} - \mathbf{k}_{\text{coll}})$ and the total effect of $K(t)$ is a displacement. Evaluating Eq. (13) for an initial thermal state with mean occupation numbers $\bar{n}_{th,j}$ in each mode j

then gives

$$C = \Gamma \int_0^\infty e^{-\Gamma \Delta t} \prod_j e^{-2\hat{k}_{\text{coll},j}^2 \eta_j^2 (1 - \cos(\mu_j \Delta t))(1 + 2\bar{n}_{th,j})} d\Delta t. \quad (14)$$

Assuming further that the trap potential is isotropic ($\mu_j =: \mu$, $\eta_j =: \eta$ for all j), and $\mu \ll \Gamma$ (such that $\mu_j \Delta t$ remains small in good approximation), this is approximately

$$C \stackrel{\mu \ll \Gamma}{\approx} e^{-2\eta^2 (\mu/\Gamma)^2 (1 + 2\bar{n}_{th})}. \quad (15)$$

In this weakly confined regime, if we consider the motion to be at the Doppler limit temperature $\hbar\Gamma/(2k_B)$ for the transition also used for photon generation,

$$C \stackrel{\mu \ll \Gamma}{\approx} e^{-\hbar k^2 / (M\Gamma)}, \quad \text{Doppler c.} \quad (16)$$

notably independent of the mode frequency. This results from a cancellation between improved cooling for higher motional frequencies, but also a larger angular spread of the recoil displacement directions in rotating-frame phase space. Conversely, for isotropic, strong confinement $\mu \gg \Gamma$, we find

$$C \stackrel{\mu \gg \Gamma}{\approx} e^{-2\eta^2 (1 + 2\bar{n}_{th})} I_0(2\eta^2 (1 + 2\bar{n}_{th})) + O((\mu/\Gamma)^{-2}), \quad (17)$$

with I_n being the modified Bessel function of the first kind. If the mode frequency is increased for given atomic parameters, the thermal occupation number $\bar{n}_{th}$ after cooling will converge to 0 and $C \approx 1 - 2\eta^2 + O(\mu^{-2})$, so the error will reduce as $1 - F \approx \eta^2 = \hbar k^2 / (2M\mu)$. The transition between these regimes is depicted in Fig. 4(e).

Note that for parameters typical of trapped-ion experiments (e.g., Ref. [68]), the dephasing error exceeds 0.5% not too far from the Doppler-cooling limit, so is non-negligible given state-of-the-art fidelities. From Eq. (14), it is clear that the dephasing can be reduced by rejecting events with a detection time difference Δt beyond some cut-off [Fig. 4(d)] or by ground-state-cooling the atoms before excitation. Both these approaches reduce the entangled state generation rate, however. Instead, we propose to correct the resulting spin-motion entanglement through a coherent operation after the herald event takes place. Concretely, in analogy to Eq. (11) for the single-photon protocol, once the photon detection times t_H and t_V are known, both nodes apply

$$U_{\text{corr}}(t_H, t_V) = |0\rangle\langle 0| \otimes M_{\text{corr},H}(t_H) + |1\rangle\langle 1| \otimes M_{\text{corr},V}(t_V), \quad (18)$$

where the spin-dependent corrections are chosen to maximize C in Eq. (13) after $K_H(t_H) \mapsto M_{\text{corr},H}(t_H) K_H(t_H)$,

$K_V(t_V) \mapsto M_{\text{corr},V}(t_V) K_V(t_V)$, that is, aiming to disentangle the internal and motional states. In the zero-NA case, this can be achieved exactly by a displacement to overlap the motional states again [see Fig. 4(b)]. For typical experimental parameters, even such a simple spin-dependent displacement correction still significantly suppresses the error [see Fig. 4(c)]; for instance, for $\eta = 0.07$, $\mu/\Gamma = 0.1$ (typical of prior experiments with $^{88}\text{Sr}^+$ ions [22,68]), the error at $\bar{n}_{th} = 20$ is reduced from 2×10^{-3} to $< 1 \times 10^{-5}$, and at $\bar{n}_{th} = 100$, from 1% to 0.01%.

B. Geometrical errors

In the previous section, we had set $K_H(t) = K_V(t)$ to focus solely on timing-related effects. In general, the position dependence/recoil effects may of course differ between the decay channels. Typically, atom-photon entanglement is generated with two narrowband decay channels of similar photon momentum, and any difference in the kick operators would stem from unequal geometric weights $w_{i\sigma}(\hat{\mathbf{k}})$ in Eq. (6). For the situation common to most experimental demonstrations and illustrated in Fig. 2, where a polarization-insensitive optical system images two decay channels into a single-mode fiber, such effects are small: even for collection of π/σ radiation perpendicular to the magnetic field, where the coupled components correspond to two linear dipoles orthogonal to the collection axis at right angles to each other [22,68], the error from the angular asymmetry of the dipole emission pattern away from its plane of symmetry is very small at realistic parameters (a numerical evaluation of this is given in Appendix F; for π/σ radiation, it is eclipsed by polarization mixing from longitudinal σ dipole components imperfectly suppressed for extended wavepackets).

In this section, we instead illustrate the versatility of our framework using a different, technologically relevant case: that where emission is collected using different optical paths for the two decay channels. This is of particular interest in the context of waveguide photonics integrated into microfabricated ion-trap chips [69,70], where device limitations can make it preferable to collect the two polarization components separately [71,72]. A possible realization is schematically depicted in Fig. 5(a); as long as the different waveguides are stable in optical length relative to each other, the two polarizations then never need to be transmitted in the same waveguide to realize the Bell-basis measurement [72,73].

For moving emitters, however, this introduces a new source of motional errors, which again admits a description in two complementary ways: in the position-dependent-phase picture, motion can now result in a differential phase shift between the components depending on its alignment with the collection directions, while in the phase space recoil picture, the (signed) magnitude of the momentum kick for each of the modes will now differ between the

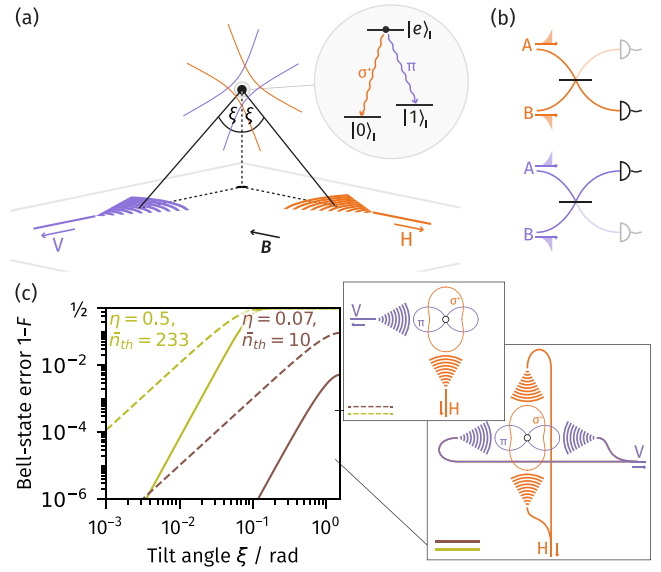


FIG. 5. Two-photon-based entanglement with two modes collected from different directions. (a) An example scenario, where an atom is trapped above a planar device and polarization-encoded photons are collected into integrated waveguide optics using grating couplers. Separating the polarizations into separate waveguides increases efficiency and lessens technical demands, but the angular offset ξ of the modes from a shared axis gives rise to a differential phase sensitivity. $\mathbf{B}$ indicates the atomic quantization direction. (b) The emission from two such zones is measured in the Bell basis, which can be implemented conventionally as in Fig. 4(a), or separated by polarization. (c) Reduction in fidelity for experimental parameters characteristic of trapped-ion (brown) and optical tweezer (olive) experiments, assuming an isotropic trap with Lamb-Dicke parameter η for all directions, as a function of the angle ξ that the collection cones make with the surface normal (thermal state of motion, zero-NA limit). The dashed lines show the Bell-state error when a pair of collection gratings from a single side is used; the solid lines show significantly improved performance from coherently summing the outputs from two symmetrically arranged collection devices.

channels. Either way, this leads to unwanted spin-motion entanglement and hence to errors.

Equation (13) still captures these motional effects with kick operators K_H and K_V now appropriately defined to reflect the different collection directions. Concretely, consider the scenario from Fig. 5(a) with collection along axes each tilted an angle ξ in orthogonal directions from what would be a shared collection axis, and for simplicity, assume an isotropic trap with motional frequency μ and an initial thermal state with occupation number $\bar{n}_{th}$. In the zero-NA approximation, and ignoring the time-dependent effects already described in Sec. V A (i.e., taking $\Gamma \rightarrow \infty$), we find

$$C = e^{-2\gamma}, \quad \text{with } \gamma := \eta^2 (\sin \xi)^2 (1 + 2\bar{n}_{th}). \quad (19)$$

This is illustrated for typical experimental parameters in Fig. 5(c) (dashed curves). Once again, an (approximate) correction of these effects through appropriate spin-dependent operations chosen to disentangle internal and motional states would be viable.

Alternatively, however, as recently also proposed in Ref. [72], a passive improvement is possible as well: by coherently combining the emission from two couplers symmetrically placed around the atom location [see Fig. 5(c)], the emission is effectively collected into a standing wave pattern in the trap plane (though it is still a running wave in the perpendicular direction common to both collection directions). Even in the zero-NA limit, $w(\hat{\mathbf{k}}) \propto (\delta(\hat{\mathbf{k}} - \hat{\mathbf{k}}_{\text{coll}}) + \delta(\hat{\mathbf{k}} + \hat{\mathbf{k}}_{\text{coll}}))/\sqrt{2}$, the collection efficiency now drops as $(1 + e^{-2\gamma})/2$, and the contrast from Eq. (19) instead becomes

$$C = \frac{4e^{-2\gamma}}{(1 + e^{-2\gamma})^2} = (\text{sech } \gamma)^2. \quad (20)$$

For small γ , the Bell-state error $1 - F$ in the one-sided case is $\gamma + O(\gamma^2)$, while the two-sided collection reduces it to $\gamma^2/2 + O(\gamma^4)$.

This reduction in error emphasizes that the correct interpretation of the weights $w_{i\sigma}(\hat{\mathbf{k}})$ for free-space collection into a single mode is not one of an incoherent mixture of different recoil directions, but in fact a coherent superposition. Phase-stable operation of standing waves generated by grating couplers integrated into ion traps was recently demonstrated [74,75], so this could be a valuable technique to reduce errors in intra- or interchip photonic links using integrated optics.

VI. TWO-PHOTON-BASED ENTANGLEMENT USING TIME-BIN ENCODING

By initializing the atoms into an equal superposition of states and subsequently exciting each of the components to the same short-lived emitting state decaying on a closed transition [separated by a π pulse to swap the components, see Fig. 6(a)], the two-photon protocol can be realized using time bins rather than polarization or frequency for the photonic encoding [76]. This can be desirable because of the relative insensitivity of such protocols to common link disturbances such as birefringence or for extension to higher-dimensional quantum systems (qudits with $d > 2$). While the robustness of this encoding scheme allowed fidelities $>90\%$ to be reached in solid-state emitters such as nitrogen vacancy centers [7,17], motion generally complicates the picture.

Broadly, the analysis for two-photon protocols from Sec. V is also applicable here, with the difference that the kick operators corresponding to the qubit states $|0\rangle_I$ and $|1\rangle_I$, which we shall label $K_e(t)$ (early) and $K_l(t)$

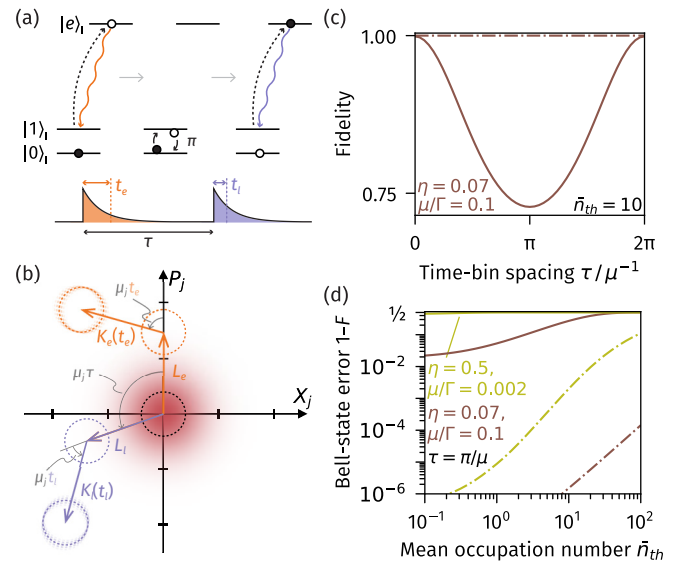


FIG. 6. Time-bin-encoded two-photon-heralded entanglement. (a) To generate entanglement between internal states and photonic time bins, the atom is first prepared in an equal superposition of $|0\rangle_I$ and $|1\rangle_I$, and only the $|1\rangle_I$ is excited to the emitting state $|e\rangle_I$. The states are then swapped using a π pulse, and the process is repeated. A partial Bell-basis measurement (not pictured) on a pair of such maximally entangled states $(|0e\rangle_{IF} + |1l\rangle_{IF})/\sqrt{2}$ in the heralding station yields a maximally entangled atom-atom state. (b) As illustrated in the phase space of the j th motional mode of one of the nodes, there is a differential displacement between the early and late time bins, depending now not only on t_e and t_l , but also on the time-bin spacing τ . (c) The amount of spin-motion entanglement in mode j depends on μ_j ; it is maximal when τ is such that the excitation kicks L_e and L_l happen in opposite directions in the rotating frame, leading to a minimum in fidelity (solid line). A spin-displacement correction (dashed-dotted line) recovers virtually all the loss in fidelity (shown for collection into a Gaussian mode with numerical aperture 0.6, excitation perpendicular to the collection direction, an isotropic trap, and a thermal state of motion). (d) For the worst time-bin spacings, where $\mu\tau$ is an odd multiple of π , the loss in fidelity is considerable even for near-ground-state-cooled atoms (solid lines), but is efficiently suppressed by a spin-dependent postherald displacement (dashed-dotted lines).

(late), now correspond to excitation laser kicks and subsequent emission separated by the time-bin spacing τ . In the rotating-frame phase space for each mode j , the displacements from excitation laser kick and decay for the late time-bin are now rotated by $\mu_j\tau$ [see Fig. 6(b)]. Thus, even if $\Gamma \gg \mu$, such that the spread in decay times is negligible, the momentum kicks from the excitation laser alone can already leave behind spin-motion entanglement. This is similar to the single-photon protocol discussed in Sec. IV, though this effect is periodic in τ , as the late time-bin direction rotates in and out of alignment with that for the early time bin.

Concretely, take the two photon detection times t_e and t_l to be relative to the start of their time bin, and consider an excitation laser pulse along $\hat{\mathbf{k}}_{\text{ex}}$. The kick operators are then

$$\begin{aligned} K_e(t_e) &:= \int_{S^2} w(\hat{\mathbf{k}}) \mathcal{D}_{-i\hat{\mathbf{k}}}(t_e) d^2\hat{\mathbf{k}} \cdot \mathcal{D}_{i\hat{\mathbf{k}}_{\text{ex}}}(0), \\ K_l(t_l) &:= \int_{S^2} w(\hat{\mathbf{k}}) \mathcal{D}_{-i\hat{\mathbf{k}}}(\tau + t_l) d^2\hat{\mathbf{k}} \cdot \mathcal{D}_{i\hat{\mathbf{k}}_{\text{ex}}}(\tau), \end{aligned} \quad (21)$$

and evaluating Eq. (13) in the zero-NA limit, $w(\hat{\mathbf{k}}) \propto \delta(\hat{\mathbf{k}} - \hat{\mathbf{k}}_{\text{coll}})$, for nodes with identical properties and a thermal initial state of motion, we obtain the fidelity as $F = (1 + C)/2$,

$$C = \Gamma^2 \int_0^\infty \int_0^\infty e^{-\Gamma(t_e + t_l)} \prod_j e^{-2\eta_j^2 |\beta_j(t_e, t_l)|^2 (1 + 2\bar{n}_{thj})} dt_l dt_e, \quad (22)$$

where

$$\beta_j(t_e, t_l) := \hat{k}_{\text{ex},j} (1 - e^{i\mu_j \tau}) - \hat{k}_{\text{coll},j} (e^{i\mu_j t_e} - e^{i\mu_j (t_l + \tau)}) \quad (23)$$

quantifies the amount of residual phase space displacement in mode j . If we, for clarity of exposition, assume the confinement to be symmetric and $\hat{\mathbf{k}}_{\text{ex}}$ to be perpendicular to $\hat{\mathbf{k}}_{\text{coll}}$, this further simplifies to

$$C = \frac{\Gamma}{2} \int_{-\infty}^\infty e^{-\Gamma|\Delta t|} e^{-2\eta^2 (2 - \cos(\mu\tau) - \cos(\mu(\Delta t + \tau))) (1 + 2\bar{n}_{th})} d\Delta t, \quad (24)$$

which is similar to the non-time-bin two-photon case Eq. (14), but with τ offsetting the wavepacket-relative arrival time difference $\Delta t := t_l - t_e$. Even if $\Gamma \rightarrow \infty$, such that the spread in emission times is negligible, the contrast is now reduced by $e^{-4\eta^2 (1 - \cos(\mu\tau)) (1 + 2\bar{n}_{th})}$, corresponding to the spin-motion entanglement induced by the excitation and decay kicks at either node [solid lines in Figs. 6(c) and 6(d)]. This can be avoided by choosing the time-bin spacing to be an integer multiple n of the trap period, $\mu\tau = 2\pi n$, but to simultaneously meet this criterion for all motional modes at both nodes at the same time may be a demanding constraint (and for emitters which are part of a register of multiple ions trapped in the same potential, only possible in approximation).

Instead, as discussed in the previous sections, the recoil effect can also be canceled by the application of a spin-dependent force after heralding: if the atom is excited to $|e\rangle_I$ using an approximately plane-wave-like laser beam, the excitation terms in the residual displacement Eq. (23) are known exactly. Depending on μ/Γ , the displacement to correct also depends on the relative timing of t_e and

t_l in the same way as discussed in Sec. V A. Outside the zero-NA limit, this arrival time can only approximately be corrected by separable spin-dependent displacements; nevertheless, this brings the time-bin encoded schemes to the same level of error sensitivity as, e.g., polarization-encoding schemes. As shown in Fig. 6(d), the spin-dependent displacement correction restores the fidelity at $\eta = 0.07$, $\mu/\Gamma = 0.1$, $\bar{n}_{th} = 20$ (fiber-coupled using NA 0.6) from 61% to >99.99% even at the most unfavorable time-bin duration. (This is a feed-forward correction of a known effect, in contrast to a recent proposal for rephasing of unknown energy-level shifts [77].)

In situations where the robustness of time-bin protocols to polarization effects, or their extensibility to higher-dimensional systems, is desirable, our results thus provide a practical guideline toward high-fidelity operation.

VII. CONCLUSION

Our analysis yields collection efficiencies and entangled-state fidelities for a wide variety of protocols and collection geometries. It can be specialized to the case of a vanishing collection angle and thermal states to obtain analytical approximations, but seamlessly generalizes to any arbitrarily structured collection modes (such as imaging into single-mode fibers or multiple combined waveguide couplers). Crucially, it retains its validity in the high-NA limit relevant to technological applications aiming for maximal collection efficiency. We have focused on bipartite entanglement generation between qubits, as this is the most well-explored use case in theory and experiment, but the core kick-operator framework describing the atom-photon wavepackets applies also to entanglement generation between qudits [33] and direct multipartite entanglement generation [73].

For the practical performance of the correction techniques proposed here, motional decoherence and nonideal single-photon detectors (with finite time resolution or dark counts) may become relevant and are readily incorporated into our analysis. In future work, the model could also be augmented to describe system-specific motional dynamics, such as the micromotion of ions in radio-frequency traps (by considering a time-dependent position operator, cf. Ref. [38]). Corrections more accurate than the spin-dependent displacements discussed here appear possible (e.g., using higher-order motional operations such as spin-dependent squeezing, or multimode operations); we leave the extent of possible improvement to future work.

More speculatively, while the current discussion has focused on the detrimental consequences of atomic motion for heralded entanglement generation, the formalism fundamentally describes the heralded application of a spin-motion operator. Thus, the same mechanism might be leveraged to generate complex spin-motion interactions heralded by the detection of a single photon (e.g., in

an interestingly structured light mode) or to mediate interactions between the internal and motional states of multiple emitters.

In summary, we have introduced a flexible framework to describe the role of motion in spontaneous-emission-mediated entanglement generation between trapped atomic qubits. A set of “kick operators” succinctly encapsulates the geometric aspects of the problem, linking together internal, motional, and photonic degrees of freedom. While recoil leads to errors universal to all trapped optical emitters, these errors can be suppressed by reducing the residual spin-motion entanglement after a herald event. We have illustrated simple correction strategies for several entanglement generation protocols; as they only have to be—indeed, can only be—applied after a successful herald, the impact on the entanglement generation rate is minimal. Thus, our results yield improvements in high-rate, high-fidelity remote entanglement generation, whether through relaxation of cooling requirements for high-fidelity operation, through increased flexibility in the use of time-bin encoding, or through enabling the use of single-photon schemes for large rate improvements in link-loss-dominated long-distance settings.

Note: In complementary work, Kikura *et al.* [78] apply the kick operator framework to cavity-assisted photon collection and detail the implications for time-bin photonic entanglement generation between neutral atoms in optical tweezer arrays. During preparation of this manuscript, we also became aware of an experimental demonstration of high-fidelity time-bin-mediated entanglement between trapped ions [33], whose authors give error estimates similar to the zero-NA case presented in Sec. VI in a recent preprint [79].

ACKNOWLEDGMENTS

The authors would like to thank the members of the Oxford Ion Trap Quantum Computing group (in particular R. Srinivas, G. Araneda, P. Drmota, and D. Main) and J. Home for inspiring discussions. We are grateful to A. Martinez, P. Drmota, D. M. Lucas, J. Home, I. Rojko, and G. Giuli for helpful comments on the manuscript. D.P.N. acknowledges Merton College, Oxford for support through a Junior Research Fellowship. J.A. acknowledges support from the Trapped Ion Quantum Information group at ETH Zürich during manuscript revision. This work was supported by the Engineering and Physical Sciences Research Council.

DATA AVAILABILITY

The data that support the findings of this article are not publicly available. The data are available from the authors upon reasonable request.

APPENDIX A: SPONTANEOUS DECAY DERIVATION

In this Appendix, we derive the pure state describing the internal state of the atom, its motion, and the emitted photon after spontaneous decay into free space from a harmonic trap. We start with the interaction picture Hamiltonian (3), looking to solve for the evolution of the initial state $|\psi_0\rangle_{IMP} = |e\rangle_I |\zeta\rangle_M |\emptyset\rangle_P$. We expand the motional state in the Fock basis, with a vector $\vec{n}$ specifying the occupation number for each mode. For instance, for the three modes of a single trapped atom, $|\vec{n}\rangle = |n_1\rangle \otimes |n_2\rangle \otimes |n_3\rangle$, but this remains general in the numbers of modes (e.g., to describe a register of multiple ions trapped in the same potential). Similarly to the standard spontaneous emission derivation in which motion is neglected, we then assume an (interaction picture) Ansatz with an excited contribution and a postdecay contribution with a photon,

$$|\psi(t)\rangle_{IMP} = \sum_{\vec{n}} \alpha_{\vec{n}}(t) |e\rangle_I |\vec{n}\rangle_M |\emptyset\rangle_P + \sum_{\vec{n}} \sum_{i\kappa\sigma} \beta_{i\kappa\sigma\vec{n}}(t) |i\rangle_I |\vec{n}\rangle_M a_{\kappa\sigma}^\dagger |\emptyset\rangle_P. \quad (A1)$$

Plugging this Ansatz into the interaction picture Schrödinger equation, we obtain coupled first-order equations for the excited and decayed amplitudes,

$$\hbar \frac{d\alpha_{\vec{n}}}{dt} = \sum_{i\kappa\sigma} e^{i(\omega_{ei}-k\epsilon)t} \mathcal{E}_{k\mathbf{g}_{i,\hat{k}\sigma}} \sum_{\vec{l}} \beta_{i\kappa\sigma\vec{l}}(t) \langle \vec{n} | \mathcal{D}_{i\hat{k}}(t) | \vec{l} \rangle \quad (A2)$$

$$\hbar \frac{d\beta_{i\kappa\sigma\vec{n}}}{dt} = -e^{i(k\epsilon-\omega_{ei})t} \mathcal{E}_{k\mathbf{g}_{i,\hat{k}\sigma}} \sum_{\vec{l}} \alpha_{\vec{l}}(t) \langle \vec{n} | \mathcal{D}_{-i\hat{k}}(t) | \vec{l} \rangle. \quad (A3)$$

Next, as in the Weisskopf-Wigner treatment [42,80], we transition to continuum modes, make the Markovian approximation for $\alpha_{\vec{n}}(t)$, assume the radiation is closely distributed about the transition frequency ω_{ei} , and extend the integration bounds over negative frequencies k . Then, assuming the Lamb shift is already included in the energy spacings, we recover the exponential decay law for the excited coefficients, $\alpha_{\vec{n}}(t) = e^{-(\Gamma/2)t} \langle \vec{n} | \zeta \rangle$; the motion of the excited contribution is unchanged in the interaction picture. With these solutions in hand, we can solve for the amplitudes related to emission through integration of (A3),

$$\beta_{i\kappa\sigma\vec{n}}(t) = -\frac{\mathcal{E}_{k\mathbf{g}_{i,\hat{k}\sigma}}}{\hbar} \int_0^t e^{i\epsilon'(k\epsilon-\omega_{ei}+i\Gamma/2)t'} \langle \vec{n} | \mathcal{D}_{-i\hat{k}}(t') | \zeta \rangle dt'. \quad (A4)$$

Finally, we insert a resolution of the identity and take the time evolution of the displacement operator outside of the matrix element to find the state after decay,

$$|\psi(t)\rangle_{IMP} = e^{-\Gamma t/2} |e\rangle_I |\zeta\rangle_M |\emptyset\rangle_P - \sum_{i\sigma} \int_{\mathbb{R}^3} \frac{\mathcal{E}_k g_{i,\hat{\mathbf{k}}\sigma}}{\hbar} \langle \vec{l} | \zeta \rangle \frac{\langle \vec{n} | \mathcal{D}_{-\hat{\mathbf{k}}}(0) | \vec{l} \rangle \left(1 - e^{-\Gamma t/2} e^{it(kc - \omega_{ei} - \vec{\mu} \cdot (\vec{l} - \vec{n}))} \right)}{\Gamma/2 + i(\omega_{ei} + \vec{\mu} \cdot (\vec{l} - \vec{n}) - ck)} |i\rangle_I |\vec{n}\rangle_M a_{\vec{k}\sigma}^\dagger |\emptyset\rangle_P d^3\mathbf{k}. \quad (\text{A5})$$

After waiting much longer than the decay lifetime, $t \gg \Gamma^{-1}$, we are left with an interaction-picture state approximately constant in time. From the denominator of the photon amplitude, we can see that the emission spectrum consists of superposed Lorentzians of width Γ centered at $ck = \omega_{ei} + \vec{\mu} \cdot (\vec{l} - \vec{n})$. The relative weights of these peaks depend on the initial motional state.

To illustrate the resulting spectrum, Fig. 7 shows the spectral energy density for both strong and weak confinement in one dimension. In the strong confinement (or resolved-sideband) regime $\mu/\Gamma \gg 1$, transitions between motional states are clearly defined, and conservation of energy leads to peaks in emitted photon frequencies spaced by μ . On the other hand, in the weak confinement regime practical for remote entanglement generation between atoms, the motional changes are less distinguishable. The Lorentzians overlap significantly, and the effect on the spectrum amounts to broadening.

To better interpret the properties of the postemission state (A5) and to obtain further simplifications, we can rewrite it in terms of time-domain modes, $a_\sigma^\dagger(t, \hat{\mathbf{k}}) = \sqrt{c/2\pi} \int_{-\infty}^{\infty} a_\sigma^\dagger(k, \hat{\mathbf{k}}) e^{ikct} dk$. Making repeated use of the narrowband approximation $\Gamma \ll \omega_{ei}$, the state in the long-time limit may be written as

$$|\psi\rangle_{IMP} \approx \sum_i \tilde{\mathcal{E}}_{\omega_{ei}/c} |i\rangle_I \int_0^\infty e^{-(\Gamma/2 + i\omega_{ei})t} \int_{S^2} \sum_\sigma g_{i,\hat{\mathbf{k}}\sigma} \mathcal{D}_{-\hat{\mathbf{k}}}(t) |\zeta\rangle_M a_\sigma^\dagger(t, \hat{\mathbf{k}}) |\emptyset\rangle_P d^2\hat{\mathbf{k}} dt, \quad (\text{A6})$$

with an adjusted field strength prefactor $\tilde{\mathcal{E}}_{\omega_{ei}/c} \propto \mathcal{E}_{\omega_{ei}/c}$ to keep normalization of the state, and the second integral taken over the unit sphere. Note that we are still in the interaction picture where this state is time-independent; t now denotes the time coordinate within the photonic wavepacket.

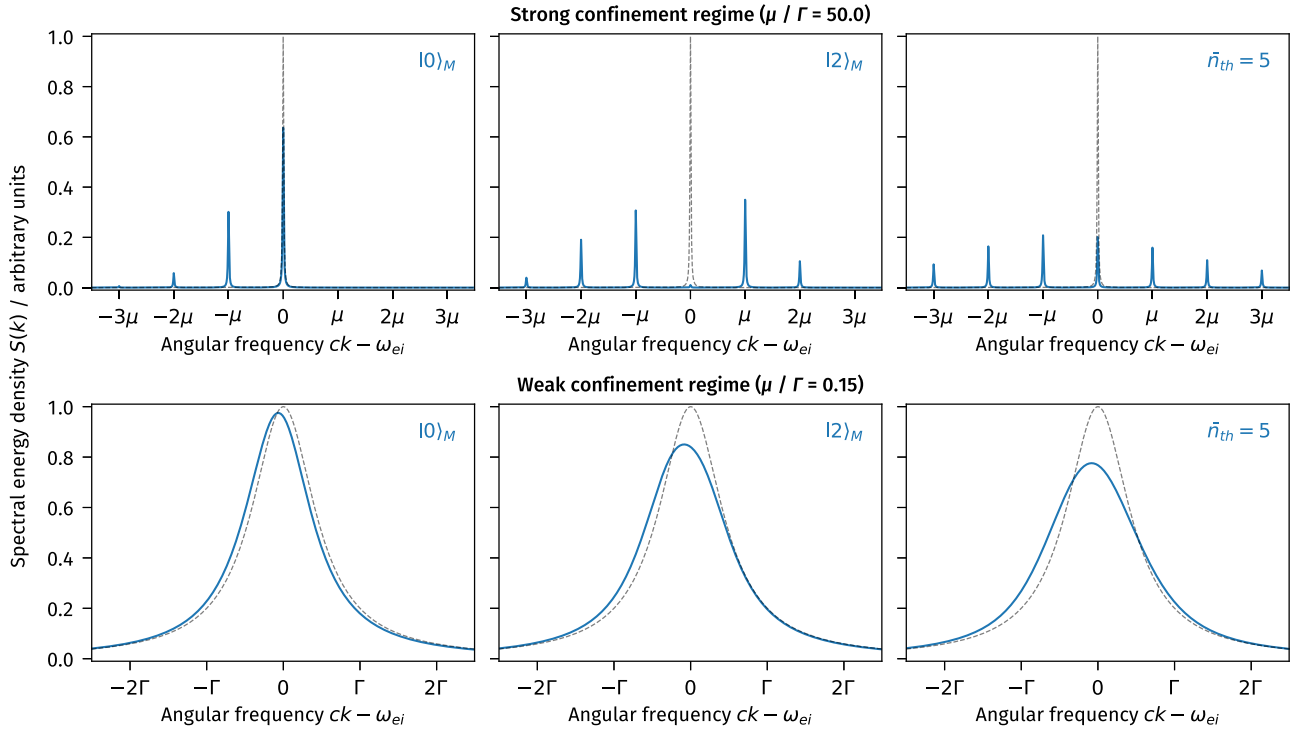


FIG. 7. Spectrum of spontaneous emission from a harmonically trapped atom in one dimension in strong and weak confinement. A large Lamb-Dicke parameter of $\eta = 1.0$ exaggerates the features for illustration purposes. The three columns correspond to different initial motional states (Fock states $|n = 0\rangle$ and $|n = 2\rangle$, and a thermal state with average phonon number $\bar{n}_{th} = 5$). The dashed curve is the Lorentzian lineshape for an infinitely heavy emitter as in the Weisskopf-Wigner treatment. The emission process can either deposit quanta of momentum in the oscillator (strong confinement: peaks with lower angular frequency than the atom transition ω_{ei}), or take them away (peaks with higher frequency than the carrier transition).

APPENDIX B: COUPLING INTO A GAUSSIAN FIBER MODE

To describe the coupling of emitted radiation into the output mode from each node, for instance the approximately Gaussian mode of a single-mode fiber, we are interested in calculating the overlap coefficients $\alpha_\sigma(\hat{\mathbf{k}})$ from Sec. III/ Fig. 2, quantifying how much each plane wave mode contributes to the fiber mode after the collection optics. The desired coefficients form part of the unitary imaging system transformation $a_\sigma^\dagger(\mathbf{k}) = \alpha_\sigma(\hat{\mathbf{k}}) f_\sigma^\dagger(k) + \dots$, where the dots elide modes subsequently traced over, corresponding to emission not coupled into the fiber. To obtain the mode contributions in the high-NA, vector-optics setting, we propagate both modes to a reference sphere at the input of the imaging system [81,82] (around the trap location), where we perform the mode overlap integral.

We approximate the mode of the single-mode fiber propagated back to the aperture plane of the collection lens by a collimated Gaussian beam. Such a mode propagating along $\hat{\mathbf{z}}$ with linear polarization at an angle σ_F is given by

$$\mathbf{e}_{F,\sigma_F}(\rho, \phi) = \sqrt{\frac{2}{\pi w_0^2}} e^{-\left(\frac{\rho}{w_0}\right)^2} (\cos(\sigma_F) \hat{\mathbf{x}} + \sin(\sigma_F) \hat{\mathbf{y}}), \quad (\text{B1})$$

where ρ and ϕ denote the polar coordinates of the beam. This profile is normalized over the x - y plane and is defined up to a phase factor determined by the position along the beam.

Supposing the plane wave vector of the mode $\mathbf{k}_{\sigma_P}$ is given in spherical coordinates, $\mathbf{k} = (k, \vartheta, \varphi)$, the electric field profile at position $\mathbf{r} = (r, \theta, \phi)$ is given by

$$\mathbf{e}_{P,\mathbf{k}_{\sigma_P}}(\mathbf{r}) = e^{i\mathbf{k} \cdot \mathbf{r}} \hat{\mathbf{e}}_{\mathbf{k}_{\sigma_P}}(\vartheta, \varphi) \quad (\text{B2})$$

for a given polarization vector $\hat{\mathbf{e}}_{\mathbf{k}_{\sigma_P}}$.

Consider the lens of focal length f to be positioned along the z axis, with focus at the coordinate origin. It then transforms the incoming Gaussian mode as $\hat{\mathbf{r}} \leftrightarrow \hat{\boldsymbol{\theta}}$, $\rho = f \sin \theta$, and an apodization factor $\sqrt{\cos \theta}$ (Chap. 3 of Ref. [82]). Both field profiles on the reference sphere $r = f$ are then given by

$$\mathbf{e}'_{F,\sigma_F}(f \hat{\mathbf{r}}) = \sqrt{\frac{2}{\pi w_0^2}} \sqrt{\cos \theta} e^{-\left(\frac{f \sin \theta}{w_0}\right)^2} (\cos(\sigma_F - \phi) \hat{\boldsymbol{\theta}} + \sin(\sigma_F - \phi) \hat{\boldsymbol{\phi}}) \quad \text{and} \quad \mathbf{e}_{P,\mathbf{k}_{\sigma_P}}(f \hat{\mathbf{r}}) = e^{ikf \hat{\mathbf{k}} \cdot \hat{\mathbf{r}}} \hat{\mathbf{e}}_{\mathbf{k}_{\sigma_P}}(\vartheta, \varphi). \quad (\text{B3})$$

Now we are ready to calculate the overlap integral over the reference sphere, where the polar angle ranges from 0 to $\vartheta_{\max}$ ($\text{NA} = \sin \vartheta_{\max}$). For given free-space and fiber polarizations, the desired coefficient is given by

$$\alpha(\mathbf{k}) = \int_0^{\vartheta_{\max}} \int_0^{2\pi} (\mathbf{e}'_{F,\sigma_F}(f \hat{\mathbf{r}}) \cdot \mathbf{e}_{P,\mathbf{k}_{\sigma_P}}(f \hat{\mathbf{r}})) \sin(\theta) d\phi d\theta. \quad (\text{B4})$$

Performing the integral over ϕ , we arrive at

$$\alpha(\mathbf{k}) = \sqrt{\frac{8\pi}{w_0^2}} \int_0^{\vartheta_{\max}} \sin(\theta) \sqrt{\cos(\theta)} e^{-\left(\frac{f \sin \theta}{w_0}\right)^2} e^{ifk \cos \vartheta \cos \theta} \begin{pmatrix} \cos^2(\theta/2) \cos \sigma_F J_0(\xi) + \sin^2(\theta/2) \cos(2\varphi - \sigma_F) J_2(\xi) \\ \cos^2(\theta/2) \sin \sigma_F J_0(\xi) + \sin^2(\theta/2) \sin(2\varphi - \sigma_F) J_2(\xi) \\ -i \sin \theta \cos(\varphi - \sigma_F) J_1(\xi) \end{pmatrix} \cdot \hat{\mathbf{e}}_{\mathbf{k}_{\sigma_P}}(\vartheta, \varphi) d\theta, \quad (\text{B5})$$

having defined $\xi := fk \sin \vartheta \sin \theta$ for convenience.

This expression for the mode coupling coefficients contains highly oscillatory integrals from diffraction on the aperture stop,

$$\mathcal{I}_\nu(\vartheta) := \int_0^{\vartheta_{\max}} e^{ifk \cos \vartheta \cos \theta} J_\nu(\xi) S(\theta) d\theta \quad (\text{B6})$$

for $fk \gg 1$, $\nu = 0, 1, 2$, and slowly varying, well-behaved functions $S(\theta)$. By the method of stationary phase (Ref. [83] (Chap. 6) and Ref. [84]), such an integral can be written in the following asymptotic form for $fk \rightarrow \infty$:

$$\mathcal{I}_\nu(\vartheta) \approx \begin{cases} \frac{1}{fk \sin \vartheta} S(\vartheta) e^{i(fk + \frac{\pi}{2}(3\nu-1))} + O((fk)^{-3/2}), & 0 < \vartheta < \vartheta_{\max}, \\ O((fk)^{-3/2}), & \vartheta_{\max} < \vartheta < \frac{\pi}{2}. \end{cases} \quad (\text{B7})$$

If necessary, the $O((fk)^{-3/2})$ terms can be evaluated through integration by parts; they produce an oscillation on top of the slowly varying term from the stationary phase approximation, but such diffraction effects can often be ignored in practice. By plugging (B7) into (B5), we can therefore obtain the coupling coefficients linking atom emission in the plane-wave basis to the output in a single-mode fiber. Finally, the frequency-domain modes can be converted to time-domain modes as in Appendix A.

APPENDIX C: SINGLE-PHOTON SCHEME GENERAL POSTHERALD STATE

In the single-photon scheme, the atoms at both nodes are weakly excited, and the emitted photon wavepackets collected and interfered on a beamsplitter. A detector click then heralds the creation of an entangled state of the two atoms. In this Appendix, we give the expressions for the postherald state of the atoms and motion and the resulting fidelity of the scheme.

The internal structure of the two atoms is assumed to be identical and the only relevant decay route considered is $|e\rangle \rightarrow |1\rangle$. Additionally, we only consider the radiation coupled to one polarization in the single-mode fiber and pure motional states (soon to be generalized) that may differ between the two nodes. In the above scenario, the incoming density matrices from Eq. (8) for each of the nodes $s \in \{A, B\}$ are of the form

$$\begin{aligned} \text{proj} \left(\sqrt{1-p_s} |0\rangle_{I^s} |\zeta_s\rangle_{M^s} |\emptyset\rangle_P + \sqrt{p_s} |1\rangle_{I^s} \int_0^\infty e^{-\left(\frac{\Gamma}{2} + i\omega_{e1}\right)t} K_s(t) |\zeta_s\rangle_{M^s} \tilde{f}_s^\dagger(t) dt |\emptyset\rangle_P \right) \\ + p_s (1 - E_s) |1\rangle_{I^s} \otimes \rho_{M^s, \text{dark}} \otimes |\emptyset\rangle_P, \end{aligned} \quad (\text{C1})$$

where we have written the second term representing the postdecay but uncollected component using the collection efficiency $E_s := \int_0^\infty e^{-\Gamma t} \text{tr}(K_s(t)^\dagger K_s(t) |\zeta_s\rangle_{M^s} \langle \zeta_s|_{M^s}) dt$ such that $\rho_{M^s, \text{dark}}$ is normalized. When we generalize to arbitrary motional states, we will replace $|\zeta_s\rangle_{M^s} \langle \zeta_s|_{M^s}$ with the relevant density operator everywhere. After combining the two (so far separable) systems, we apply the beamsplitter as in Sec. IV, noting that the photonic operators $\tilde{f}_A^\dagger(t)$ and $\tilde{f}_B^\dagger(t)$ commute.

Assuming that at time t after the excitations there was precisely one click at the perfect detector sensitive to mode g and none in the other one, the subnormalized state of both atoms and their motion is given by

$$\begin{aligned} \rho_{\wedge g; I^A M^A I^B M^B}(t) \\ = \frac{e^{-\Gamma t}}{2} \text{proj} \left(\sqrt{1-p_A} |0\rangle_{I^A} |\zeta_A\rangle_{M^A} \otimes \sqrt{p_B} |1\rangle_{I^B} K_B(t) |\zeta_B\rangle_{M^B} + \sqrt{p_A} |1\rangle_{I^A} K_A(t) |\zeta_A\rangle_{M^A} \otimes \sqrt{1-p_B} |0\rangle_{I^B} |\zeta_B\rangle_{M^B} \right) \\ + p_A p_B \frac{e^{-\Gamma t}}{2} |1\rangle_{I^A} |1\rangle_{I^B} \left((1 - E_A) \rho_{M^A, \text{dark}} \otimes K_B(t) |\zeta_B\rangle_{M^B} \langle \zeta_B|_{M^B} K_B(t)^\dagger + K_A(t) |\zeta_A\rangle_{M^A} \langle \zeta_A|_{M^A} K_A(t)^\dagger \otimes (1 - E_B) \rho_{M^B, \text{dark}} \right). \end{aligned} \quad (\text{C2})$$

If instead detector h went off, the only change would be a relative sign change in the superposition. The subnormalized state corresponding to two clicks in the same detector g or h at times t and t' is

$$\begin{aligned} \rho_{\wedge gg; I^A M^A I^B M^B}(t, t') = p_A p_B \frac{e^{-\Gamma(t+t')}}{4} |1\rangle_{I^A} |1\rangle_{I^B} \left(K_A(t) |\zeta_A\rangle_{M^A} \langle \zeta_A|_{M^A} K_A(t)^\dagger \otimes K_B(t') |\zeta_B\rangle_{M^B} \langle \zeta_B|_{M^B} K_B(t')^\dagger \right. \\ \left. + K_A(t) |\zeta_A\rangle_{M^A} \langle \zeta_A|_{M^A} K_A(t')^\dagger \otimes K_B(t') |\zeta_B\rangle_{M^B} \langle \zeta_B|_{M^B} K_B(t)^\dagger \right) + (t \leftrightarrow t'). \end{aligned} \quad (\text{C3})$$

Equations C2 and C3 illustrate sources of error in the single-photon scheme: the first line in Eq. (C2) is related to the desired, maximally entangled atom-atom state, the second line contains terms from double excitation in which only one photon was collected, and Eq. (C3) corresponds to double excitation and detection of two photons in one mode. These equations are straightforwardly generalized to arbitrary initial motional states ρ_{M^A} and ρ_{M^B} by considering their spectral decompositions $\rho_{M^A} = \sum_{m^A} p_{m^A} |\zeta_{m^A}\rangle_{M^A} \langle \zeta_{m^A}|$, $\rho_{M^B} = \sum_{m^B} p_{m^B} |\zeta_{m^B}\rangle_{M^B} \langle \zeta_{m^B}|$. Tracing out the motion (thus erasing any specific

dependence on $\rho_{M^S, \text{dark}}$) and writing out the atom-atom density matrix in the computational basis, we obtain explicitly

$$\rho_{\wedge g; I^A I^B}(t) = \frac{e^{-\Gamma t}}{2} \times \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & (1-p_A)p_B \text{tr}(K_B(t) \rho_{M^B} K_B(t)^\dagger) & \sqrt{p_A p_B (1-p_A)(1-p_B)} \text{tr}(\rho_{M^A} K_A(t)^\dagger) \text{tr}(K_B(t) \rho_{M^B}) & 0 \\ 0 & \sqrt{p_A p_B (1-p_A)(1-p_B)} \text{tr}(K_A(t) \rho_{M^A}) \text{tr}(\rho_{M^B} K_B(t)^\dagger) & p_A(1-p_B) \text{tr}(K_A(t) \rho_{M^A} K_A(t)^\dagger) & 0 \\ 0 & 0 & 0 & p_A p_B \delta \end{pmatrix}, \quad (\text{C4})$$

with $\delta := (1 - E_A) \text{tr}(K_B(t) \rho_{M^B} K_B(t)^\dagger) + (1 - E_B) \text{tr}(K_A(t) \rho_{M^A} K_A(t)^\dagger)$, whereas the two-photon term gives

$$\rho_{\wedge gg; I^A I^B}(t, t') = \frac{e^{-\Gamma(t+t')}}{4} p_A p_B |11\rangle\langle 11|_{I^A I^B} \left(\text{tr}(K_A(t) \rho_{M^A} K_A(t)^\dagger) \text{tr}(K_B(t') \rho_{M^B} K_B(t')^\dagger) \right. \\ \left. + \text{tr}(K_A(t) \rho_{M^A} K_A(t')^\dagger) \text{tr}(K_B(t') \rho_{M^B} K_B(t)^\dagger) \right) + (t \leftrightarrow t'). \quad (\text{C5})$$

In practice, the two-photon term can be neglected entirely, as it would only be significant in the regime of high system (collection and detection) efficiency, where the use of two-photon-herald schemes would be more favorable (as they avoid the rate-fidelity trade-off inherent in single-photon schemes). This is effectively the approximation we make to obtain Eq. (9) in Sec. IV. To illustrate this point in a self-consistent fashion, we calculate the Bell-state fidelity including the two-photon term here and only take the appropriate limit in the end.

To do this, we need to choose a model for the response of the detectors to multiple photons; several choices are plausible. First, consider a detector with infinite time resolution and no dead-time, and consider a finite detection efficiency or other losses to have been pushed into the kick operators instead. In that case, the optimal herald signal would consist of having detected *exactly one* photon in mode g at some time t and none at any other time or the other detector h . Without any restrictions on the detection time t , the subnormalized atom-atom density matrix corresponding to this herald event is simply $\rho_{\wedge g, \text{one}} = \int_0^\infty \rho_{\wedge g}(t) dt$. We can then calculate the normalized state $\rho_{|g, \text{one}} = \rho_{\wedge g, \text{one}} / \text{tr}(\rho_{\wedge g, \text{one}})$ and use it to get an explicit expression for the fidelity to the closest Bell-like state $|\Psi^\phi\rangle$ (for the optimal choice of phase ϕ). We obtain

$$F(\rho_{|g, \text{one}}) = \frac{1}{2} \frac{p_A E_A (1 - p_B) + (1 - p_A) p_B E_B + 2\sqrt{p_A p_B (1 - p_A)(1 - p_B)} \left| \int_0^\infty e^{-\Gamma t} \text{tr}(K_A(t) \rho_{M^A}) \text{tr}(\rho_{M^B} K_B(t)^\dagger) dt \right|}{p_A E_A + p_B E_B - 2p_A p_B E_A E_B}. \quad (\text{C6})$$

The above expressions simplify considerably if we assume that both systems are identical. In that case, the fidelity is given by

$$F(\rho_{|g, \text{one}}) = \frac{1}{2} \frac{1 - p}{1 - pE} \left(1 + \frac{\int_0^\infty e^{-\Gamma t} |\text{tr}(K(t) \rho_M)|^2 dt}{E} \right) \quad \text{with } E := \int_0^\infty e^{-\Gamma t} \text{tr}(K(t)^\dagger K(t) \rho_M) dt, \quad (\text{C7})$$

which yields Eq. (9) in the limit of low collection efficiency E .

A second plausible detection model is that of detectors without any time or photon number resolution. The relevant postherald condition is then that of *at least* one photon in the output mode g and none in h , given by $\rho_{\wedge g, \text{all}} = \int_0^\infty \rho_{\wedge g}(t) dt + \int_0^\infty \int_0^\infty \rho_{\wedge g}(t, t') dt dt'$. In this case, the general fidelity is

$$F(\rho_{|g, \text{all}}) = \frac{1}{2} \frac{p_A E_A (1 - p_B) + (1 - p_A) p_B E_B + 2\sqrt{p_A p_B (1 - p_A)(1 - p_B)} \left| \int_0^\infty e^{-\Gamma t} \text{tr}(K_A(t) \rho_{M^A}) \text{tr}(\rho_{M^B} K_B(t)^\dagger) dt \right|}{p_A E_A + p_B E_B - p_A p_B (E_A E_B - \int_0^\infty \int_0^\infty e^{-\Gamma(t+t')} \text{tr}(K_A(t) \rho_{M^A} K_A(t')^\dagger) \text{tr}(K_B(t') \rho_{M^B} K_B(t)^\dagger) dt dt')}, \quad (\text{C8})$$

and for symmetric system parameters, we obtain

$$F(\rho_{|g, \text{all}}) = \frac{1}{2} \frac{1 - p}{1 - \frac{pE}{2} (1 - \gamma)} \left(1 + \frac{\int_0^\infty e^{-\Gamma t} |\text{tr}(K(t) \rho_M)|^2 dt}{E} \right) \quad \text{with } \gamma := \frac{1}{E^2} \int_0^\infty \int_0^\infty e^{-\Gamma(t+t')} |\text{tr}(K(t) \rho_M K(t')^\dagger)|^2 dt dt'. \quad (\text{C9})$$

The term γ is related to the purity of the emitted wave packets: two identical, pure photons will always be found on the same port of the 50:50 beamsplitter—the Hong-Ou-Mandel effect—, so without number- or time-resolving detectors, can never be distinguished from just a single photon (so the double-excitation probability fully contributes to a $|11\rangle\langle 11|$ error in the heralded state). On the other hand, if the photon state is impure (such as from spin-motion entanglement), there is a chance that the two photons from the double-excitation case will exit the 50:50 beam splitter at opposite ports, which can be rejected as an invalid herald, thereby slightly reducing the $|11\rangle\langle 11|$ component (though the lack of purity will have reduced the contrast numerator).

We stress that both Eqs. (C7) and (C9) reduce to the fidelity given in Eq. (9) in the main text in the limit of low system efficiency ($E \ll 1$), such that the exact detection model is unlikely to be of practical relevance.

APPENDIX D: TWO-PHOTON SCHEME GENERAL POSTHERALD STATE

In this Appendix, we give the general expressions for the atom-atom state resulting from two-photon-heralded protocols. In these protocols, emission from two nodes is combined in a partial Bell-basis measurement setup [see Figs. 4(a) and 5(a)], such that a detection of a pair of photons in two different detectors heralds the creation of an atom-atom entangled state. Where detectors can easily be made sensitive to the photonic qubit state, for instance through polarizing beamsplitters in the case of polarization encoding, this allows a particular Bell state to be identified with probability one half (in the other half, the state is nonentangled), though the model also applies, e.g., to frequency encoding where a state can only be identified with probability one quarter. Here, as in the main text, we treat the case of polarization encoding as the paradigmatic example and label the two modes as H and V .

As our focus is the role of the motional degrees of freedom, we assume that the necessary beamsplitters and detectors are perfect (apart from some amount of loss/limited quantum efficiency, which does not modify the resulting states, only the overall efficiency, and can be pushed into the kick operators); if desired, the kick operators are easily integrated into discussions of Bell-state measurement imperfections from elsewhere in the literature (e.g., Sec. 4.5 of Ref. [85]). For ideal detectors, the postherald state only involves components where a photon was present from each source; in the wavepackets from Eq. (12), the $\rho_{IM,\text{dark}}$ component can be neglected, and we can consider the pure input wavepackets for a pure initial motional state $|\zeta_s\rangle_{Ms}$ (the $s \in \{A, B\}$ superscript labels the two nodes),

$$|0\rangle_{Is} \int_0^\infty e^{-\left(\frac{\Gamma}{2} + i\omega_{e0}^s\right)t} K_{sH}(t) |\zeta_s\rangle_{Ms} \tilde{f}_{sH}^\dagger(t) dt |\emptyset\rangle_F + |1\rangle_{Is} \int_0^\infty e^{-\left(\frac{\Gamma}{2} + i\omega_{e1}^s\right)t} K_{sV}(t) |\zeta_s\rangle_{Ms} \tilde{f}_{sV}^\dagger(t) dt |\emptyset\rangle_F. \quad (\text{D1})$$

Here, as in Sec. V, we have taken the kick operators $K_{si}(t)$ to already include the effect of the excitation laser kick. We also assume the wavepackets from both nodes to be aligned temporally in the heralding station, and the qubit frequencies to be matched, $\omega_{ei}^A = \omega_{ei}^B$ (here, as in the entire document, we work in the rotating frame with respect to the internal state, so this equality causes the time-dependent phase in the wavepacket to factor out into a global phase).

For a perfect Bell-basis measurement setup, the heralded entangled state is the same for both H/V detector combinations with the triggered single-photon detectors on the same side of the central 50:50 beamsplitter and similarly for both where the clicks occur on opposite sides. Considering the same-side cases, the subnormalized state of the internal state and motion conditional on a detection of the photons at times $t_H, t_V \geq 0$ is

$$\rho_{\wedge\text{same}}(t_H, t_V) = e^{-\Gamma(t_H+t_V)} \text{proj} \left(|0\rangle_{IA} K_{AH}(t_H) |\zeta_A\rangle_{MA} \otimes |1\rangle_{IB} K_{BV}(t_V) |\zeta_B\rangle_{MB} + |1\rangle_{IA} K_{AV}(t_V) |\zeta_A\rangle_{MA} \otimes |0\rangle_{IB} K_{BH}(t_H) |\zeta_B\rangle_{MB} \right). \quad (\text{D2})$$

Note that the argument to $\text{proj}(\dots) = |\dots\rangle\langle\dots|$ is itself in general not normalized, as the kick operators contain, e.g., the collection efficiency. The trace of $\rho_{\wedge\text{same}}(t_H, t_V)$ gives the probability density for observing herald photons at the given times, and the normalized state after having detected such a pair of photons is $\rho_{|\text{same}}(t_H, t_V) = \rho_{\wedge\text{same}}(t_H, t_V) / \text{tr}(\rho_{\wedge\text{same}}(t_H, t_V))$. The state for detector clicks on opposite sides, $\rho_{\wedge\text{opp}}(t_H, t_V)$, matches $\rho_{\wedge\text{same}}(t_H, t_V)$ except for a minus sign in the superposition in Eq. (D2); we will discuss the $\rho_{\wedge\text{same}}(t_H, t_V)$ case for brevity.

As in Appendix C, Eq. (D2) is easily generalized to arbitrary initial motional states through their spectral decomposition. After discarding the motional states, the postthermal state in the computational basis becomes

$$\rho_{\wedge \text{same}}(t_H, t_V) = e^{-\Gamma(t_H+t_V)} \times \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & \text{tr}(K_{AH}(t_H) \rho_{MA} K_{AH}(t_H)^\dagger) \text{tr}(K_{BV}(t_V) \rho_{MB} K_{BV}(t_V)^\dagger) & \text{tr}(K_{AH}(t_H) \rho_{MA} K_{AV}(t_V)^\dagger) \text{tr}(K_{BV}(t_V) \rho_{MB} K_{BH}(t_H)^\dagger) & 0 \\ 0 & \text{tr}(K_{AV}(t_V) \rho_{MA} K_{AH}(t_H)^\dagger) \text{tr}(K_{BH}(t_H) \rho_{MB} K_{BV}(t_V)^\dagger) & \text{tr}(K_{AV}(t_V) \rho_{MA} K_{AV}(t_V)^\dagger) \text{tr}(K_{BH}(t_H) \rho_{MB} K_{BH}(t_H)^\dagger) & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}. \quad (\text{D3})$$

Averaging over all possible detection times gives $\rho_{\wedge \text{same,all}} = \int_0^\infty \int_0^\infty \rho_{\wedge \text{same}}(t_H, t_V) dt_V dt_H$ and $\rho_{|\text{same,all}} = \rho_{\wedge \text{same,all}} / \text{tr}(\rho_{\wedge \text{same,all}})$.

The fidelity of the normalized density matrix to the closest maximally entangled state is $F(\rho_{|\text{same,all}}) = \frac{1}{2} (1 + |C|)$, where

$$C = 2 \int_0^\infty \int_0^\infty e^{-\Gamma(t_H+t_V)} \text{tr}(K_{AH}(t_H) \rho_{MA} K_{AV}(t_V)^\dagger) \text{tr}(K_{BV}(t_V) \rho_{MB} K_{BH}(t_H)^\dagger) dt_V dt_H / \left(\int_0^\infty e^{-\Gamma t_H} \text{tr}(K_{AH}(t_H) \rho_{MA} K_{AH}(t_H)^\dagger) dt_H \int_0^\infty e^{-\Gamma t_V} \text{tr}(K_{BV}(t_V) \rho_{MB} K_{BV}(t_V)^\dagger) dt_V + \int_0^\infty e^{-\Gamma t_V} \text{tr}(K_{AV}(t_V) \rho_{MA} K_{AV}(t_V)^\dagger) dt_V \int_0^\infty e^{-\Gamma t_H} \text{tr}(K_{BH}(t_H) \rho_{MB} K_{BH}(t_H)^\dagger) dt_H \right). \quad (\text{D4})$$

If the two systems and initial states are identical, this simplifies to

$$C = \frac{\int_0^\infty \int_0^\infty e^{-\Gamma(t_H+t_V)} |\text{tr}(K_H(t_H) \rho_M K_V(t_V)^\dagger)|^2 dt_V dt_H}{\int_0^\infty e^{-\Gamma t_H} \text{tr}(K_H(t_H) \rho_M K_H(t_H)^\dagger) dt_H \int_0^\infty e^{-\Gamma t_V} \text{tr}(K_V(t_V) \rho_M K_V(t_V)^\dagger) dt_V}, \quad (\text{D5})$$

and C is then real and non-negative, giving Eq. (13) from the main text.

Clearly, the postthermal states and fidelities for constrained detection times can be obtained by adjusting the integral bounds accordingly. In practice, a nominal cutoff for t_H and t_V will certainly be implemented (e.g., a few time-constants after excitation) to proceed to the next generation attempt after a finite amount of time. The accepted range of photon timestamps (or their difference) can be further constrained to reduce the impact of nonidealities such as detector dark counts or motional effects; an example is shown in Fig. 4(d).

If the motional degrees of freedom are not immediately ignored, but an operation is applied by either node to its atom's motional state, this is straightforwardly multiplied from the left to the non-normalized pure state in Eq. (D2), effectively modifying the kick operators. As discussed in Sec. V, by making this motional operation conditional on the internal state and adjusting it in real-time according to the photon detection times t_H and t_V , one can attempt to disentangle spin and motion as well as possible given the constraints on the realizable correction operators. If the single-photon detectors do not have a fine time resolution compared to the relevant time scale (given by the motional frequencies or the mismatch in qubit frequencies, if any), this limits the accuracy with which t_H and t_V can be known, and hence the effectiveness of the applied correction.

Note that Eq. (D1) is not the most general wavepacket possible, as we have taken the two decay channels to correspond to entirely orthogonal output modes. This gives rise to the simple forms of Eqs. (D2) and (D3), where the states remain restricted to the $\{|01\rangle, |10\rangle\}$ subspace. If this is not a good approximation, either because of intrinsic limitations in a particular emission/collection setup, or because some symmetry properties of the system might no longer hold in the presence of motion (such as discussed in Appendix F), the above can be easily adapted by expanding $|0\rangle_{IA} K_{AH}(t_H) \mapsto (|0\rangle_{IA} K_{A0H}(t_H) + |1\rangle_{IA} K_{A1H}(t_H))$, etc. in Eq. (D2).

If necessary, the model for motional effects discussed here can straightforwardly be extended by nonideal optical elements, temporal misalignment, a mismatch in qubit frequencies, detectors with dark counts, or similar imperfections; see Chap. 4.4 of Ref. [85].

APPENDIX E: NUMERICAL EVALUATION OF FIDELITY EXPRESSIONS

In this Appendix, we give additional detail on the numerical calculations used to evaluate the fidelity of the generated Bell states, such as in Figs. 3(c)–3(e), 4(c) and 4(d), 5(c), 6(c) and 6(d), and 8.

For all data presented in the figures, we consider the case where the parameters for both nodes are identical, and in particular, all motional modes are at the same harmonic oscillator frequency μ . Our analysis, including the proposed correction technique, does not rely on this fact; as there are not any cancellation effects of relevance (except for the time-bin spacing τ discussed in Sec. VI), this merely simplifies the presentation. Equations (9), (13), and (22) and following show this simplified symmetric case as well; general expressions for the fidelity can be found in Appendixes C and D.

Our analytic expressions apply to arbitrary initial motional states ρ_M . To obtain concrete numerical results for the fidelity, we consider ρ_M to be a thermal state as parametrized by the average occupation number $\bar{n}_{th}$ in each of the relevant modes. Such states are produced by the use of laser cooling before an entanglement attempt is made. In practical realizations, it may be preferable to (in the absence of immediate heralding success) perform multiple attempts in sequence before interrupting the protocol to recool the trapped atoms. If the repetition rate of the protocol is not close to a motional mode (or a integer fraction thereof), the repeated excitation kicks do not build up a significant coherent excitation, and we find that taking a thermal state with an $\bar{n}$ representative of the average case is a reasonable approximation. Exact results can be obtained by considering ρ_M to be the mixture of initial motional states in the relevant protocol rounds—appropriately weighted by the probability that a given round is the final, successful one—as affected by the cumulative effect of previous rounds of state preparation, excitation, and decay, combined with any background heating or decoherence processes. If significant motional heating is expected to occur across a generation round, the collection efficiency (as given by the $\text{tr}(K^\dagger K \rho_M)$ terms) will also drop as unsuccessful rounds accumulate [21,85]; to obtain an accurate description of the success probabilities, a numerical round-by-round calculation may then be required.

As a concrete example, consider the single-photon case from Sec. IV/ Fig. 3. To calculate the expected protocol fidelity F —apart from the rate-fidelity trade-off inherent in single-photon protocols, as given by the $(1 - p)$ factor—the two integrals in the definition of the contrast term C from Eq. (9) need to be evaluated. Consider the expectation value $\text{tr}(K(t) \rho_M)$. To obtain accurate results outside the zero-NA approximation made to derive Eq. (10), we expand $K(t)$ according to Eq. (6) following the plane wave decomposition of the emission/collection process with weights $w(\hat{\mathbf{k}})$ (eliding the $i\sigma$ indices, as only one transition is relevant here),

$$\text{tr}(K(t) \rho_M) = \int_{S^2} w(\hat{\mathbf{k}}) \text{tr}(\mathcal{D}_{-i\hat{\mathbf{k}}}(t) L \rho_M) d^2\hat{\mathbf{k}} = \int_{S^2} w(\hat{\mathbf{k}}) \prod_j \text{tr}(D_j(-ie^{i\mu t}\eta \hat{\mathbf{k}}_j) D_j(i\eta \hat{\mathbf{k}}_{ex,j}) \rho_{M,j}) d^2\hat{\mathbf{k}}. \quad (\text{E1})$$

Here, we have expanded the excitation laser ($\hat{\mathbf{k}}_{ex}$) and collection ($\hat{\mathbf{k}}$) directions, as well as the initial state $\rho_M = \bigotimes_j \rho_{M,j}$, into the projections onto the harmonic oscillator modes j . Combining the two displacement operators [considering that $D(\alpha)D(\beta) = e^{i\text{Im}(\alpha\bar{\beta})}D(\alpha + \beta)$] and using a standard result about the characteristic function of thermal states, $\text{tr}(D(\alpha) \rho_{th}) = e^{-|\alpha|^2(\bar{n}_{th} + \frac{1}{2})}$ [Eq. (4.4.23) of Ref. [59]], we obtain

$$\text{tr}(K(t) \rho_M) = \int_{S^2} w(\hat{\mathbf{k}}) e^{-i\eta^2 \hat{\mathbf{k}} \cdot \hat{\mathbf{k}}_{ex} \sin(\mu t)} e^{-\eta^2 |\hat{\mathbf{k}}_{ex} - e^{i\mu t} \hat{\mathbf{k}}|^2 (\bar{n}_{th} + \frac{1}{2})} d^2\hat{\mathbf{k}}. \quad (\text{E2})$$

For the weights $w(\hat{\mathbf{k}}) = g_{\hat{\mathbf{k}}} \alpha(\hat{\mathbf{k}})$, we take $g_{\hat{\mathbf{k}}}$ to match the prototypical case of a dipole emitter polarized orthogonal to the collection direction and choose $\alpha(\hat{\mathbf{k}})$ to correspond to a Gaussian fiber mode treated using full vectorial optics [Eqs. (B5) and (B7)]. The Gaussian mode extent w_0 is chosen to optimize the mode matching (coupling efficiency) for a stationary dipole emitter at the given numerical aperture $\sin \vartheta_{\max}$.

We evaluate this expression for a given wavepacket time coordinate t using numerical quadrature and subsequently the full numerator $\int_0^\infty e^{-\Gamma t} |\text{tr}(K(t) \rho_M)|^2 dt$. Calculation of the collection efficiency denominator $\int_0^\infty e^{-\Gamma t} \text{tr}(K(t)^\dagger K(t) \rho_M) dt$ proceeds analogously, where the two kick operators now result in nested spatial integrals over the collection cone,

$$\text{tr}(K(t)^\dagger K(t) \rho_M) = \int_{S^2} \int_{S^2} \overline{w(\hat{\mathbf{k}}')} w(\hat{\mathbf{k}}) e^{2i\eta^2 (\hat{\mathbf{k}}' - \hat{\mathbf{k}}) \cdot \hat{\mathbf{k}}_{ex} \sin(\mu t)} e^{-\eta^2 |\hat{\mathbf{k}}' - \hat{\mathbf{k}}|^2 (\bar{n}_{th} + \frac{1}{2})} d^2\hat{\mathbf{k}} d^2\hat{\mathbf{k}}'. \quad (\text{E3})$$

The time coordinate only enters this expression as in the limit of very large Lamb-Dicke parameters, the excitation laser could induce a large enough displacement for the following harmonic motion to modulate the fiber coupling efficiency across the $1/\Gamma$ duration of the wavepacket.

Expressions such as Eqs. (E2) and (E3) highlight the fact that the appropriate picture of collection into a single output mode is that of a coherent superposition of plane waves, not a classical mixture. For instance, it is the coherent summation of plane-wave contributions across a nonzero solid angle that affords any spatial sensitivity of collection; in the zero-NA

limit [say, $w(\hat{\mathbf{k}}) \propto \delta(\hat{\mathbf{k}} - \hat{\mathbf{z}})$], the collection efficiency $\int_0^\infty e^{-\Gamma t} \text{tr}(K(t)^\dagger K(t) \rho_M) dt$ is constant irrespective of the however large the thermal state may be. Similarly, placing a “photon bucket” detector spanning an equivalent numerical aperture, giving an incoherent mixture over the emission directions, is not equivalent to single-mode collection.

As discussed in the main text, the proposed spin-dependent displacement correction technique, where each node applies the operator from Eq. (11) after a successful herald at time t , can be viewed as modifying the kick operator $K(t) \mapsto M_{\text{corr}}(t)K(t) =: \tilde{K}(t)$. Ideally, $\tilde{K}(t)$ would be proportional to the identity operator, giving $C = 1$ and hence unit fidelity [up to the $(1 - p)$ double-excitation error]. However, with a simple spin-dependent displacement $M_{\text{corr}}(t)$, such a choice is only possible in the zero-NA limit. If we apply a correction corresponding to an “average” mode projection $\mathbf{k}_{\text{avg}}$, the contrast numerator is in general given by

$$\text{tr}(M_{\text{corr}}(t)K(t) \rho_M) = \int_{S^2} w(\hat{\mathbf{k}}) e^{-2i\eta^2(\hat{\mathbf{k}} - \mathbf{k}_{\text{avg}}) \cdot \hat{\mathbf{k}}_{\text{ex}} \sin(\mu t)} e^{-\eta^2 |\hat{\mathbf{k}} - \mathbf{k}_{\text{avg}}|^2 (\bar{n}_{th} + \frac{1}{2})} d^2 \hat{\mathbf{k}}. \quad (\text{E4})$$

The collection efficiency in the denominator remains unchanged under $K(t) \mapsto M_{\text{corr}}(t)K(t)$ for any unitary M_{corr} : physically, a postherald correction cannot modify the collection efficiency. To obtain the fidelity curves in the figures, we numerically optimize $\mathbf{k}_{\text{avg}}$ to maximize the fidelity for the given parameters. The magnitude of the optimal $\mathbf{k}_{\text{avg}}$ is weakly dependent on the initial $\bar{n}_{th}$.

More generally, if ρ_M is a pure state $|\zeta\rangle_M$, the motional part of the $|1\rangle_I$ component of the postherald state will also be pure ($K(t)|\zeta\rangle_M$ before normalization). As such, it is in principle possible to design a more complex correction $M_{\text{corr}}(t)$ that perfectly inverts the kick operator (up to normalization) for this particular input state, such that $M_{\text{corr}}(t)K(t)|\zeta\rangle_M \propto |\zeta\rangle_M$ (for which higher-order spin-dependent interactions [86] will be required). We leave the question of how well this can be accomplished for other classes of states (e.g., thermal states) to future work.

Unless stated otherwise, we choose a numerical aperture of 0.6 for the numerical results presented in the manuscript. This is characteristic of the optical systems employed in several photon-mediated entanglement demonstrations in trapped ions (e.g., Refs. [22,65]), as it is readily realized with well-controlled aberrations using fairly standard out-of-vacuum objectives, though it by no means represents a fundamental limit (other photon-mediated entanglement demonstrations have used objectives up to NA 0.8 [21], and diffraction-limited objectives of even higher NA have been constructed at the expense of a shorter working distance, e.g., NA 0.92 in Ref. [87]). The effective numerical aperture of typical single-mode fibers is significantly lower than this; a Gaussian beam in the collimated plane of the imaging optics as assumed in Appendix B, truncated by the aperture stop setting the system NA, is a reasonable approximation. If desired, a more accurate description of the mode guided in the fiber (e.g., an LP_{01} mode for an ideal step-index fiber) and the fiber-side $\sqrt{\cos\theta}$ apodization factor [see Appendix B and Chap. 4.2 of Ref. [85]] could be inserted for improved quantitative predictions of the coupling efficiency, though the resulting corrections to the fidelity will be fairly small.

The first exemplary choice of confinement parameters, $\eta = 0.07$, $\mu/\Gamma = 0.1$, corresponds to the 422 nm $S_{1/2} \leftrightarrow P_{1/2}$ transition of $^{88}\text{Sr}^+$ (lifetime $1/\Gamma = 7.39$ ns) confined at a secular frequency of $\mu = 2\pi \times 2$ MHz, which is in a range typical of trapped-ion experiments such as those of Refs. [13,22,68]. The second example set of parameters, $\eta = 0.5$, $\mu/\Gamma = 0.002$, was chosen to correspond to the considerably weaker confinement found in optical tweezer experiments, here in particular the 795 nm $S_{1/2} \leftrightarrow P_{1/2}$ transition of ^{87}Rb (lifetime $1/\Gamma = 26.2$ ns) and the longitudinal trap frequency of $\mu = 2\pi \times 13$ kHz from the experiments in Ref. [8].

For the curves showing thermal states resulting from Doppler cooling in Figs. 3(e) and 4(e), we take the Doppler limit to be $\hbar\Gamma/(2k_B)$, corresponding to a mean occupation number of $\bar{n}_{\text{Doppler}} = 1/(e^{2\mu/\Gamma} - 1)$. We show $2\bar{n}_{\text{Doppler}}$ as a representative value to account for some inefficiency such as from the use of a single cooling beam for multiple modes, or additional heating before the entanglement attempt in a scheme where blocks of attempts are periodically interrupted by cooling. The expression for $\bar{n}_{\text{Doppler}}$ only strictly applies in the weak confinement regime $\mu \ll \Gamma$ [35]. However, as the system transitions into the resolved-sideband regime and $\bar{n}_{\text{Doppler}} \ll 1$, the ion wavepacket size is in any case dominated by its ground-state extent, such that the asymptotic behavior remains valid.

APPENDIX F: GEOMETRIC EFFECTS IN HIGH-NA COLLECTION FROM CROSSED DIPOLES

In the discussion of two-photon protocols in the main text, we focused on errors due to the spread in detection times (Sec. VA), and considered errors from the collection geometry only for the case of two decay channels collected from different directions (Sec. VB). However, when photons from decays corresponding to two orthogonal dipoles are collected from large solid angles (high NAs), there are error mechanisms that affect even collection along

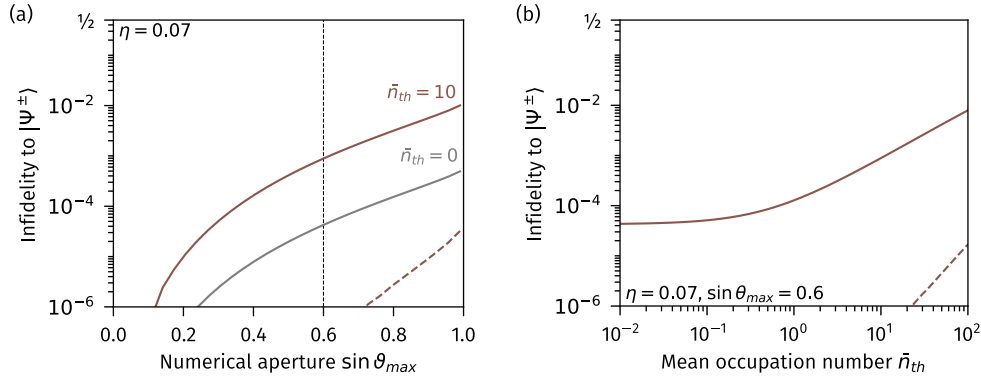


FIG. 8. Loss in fidelity from geometric effects in two-photon entanglement from π/σ decays (ignoring temporal effects, i.e., $\Gamma \rightarrow \infty$), with $\eta = 0.07$ typical of trapped-ion experiments. (a) As the NA is increased, the error from σ emission, though its longitudinal dipole component, giving rise to polarization aligned with the quantization axis (solid lines), while the error just from the mismatch of the emission patterns of the two transversal dipoles remains negligible (3×10^{-5} at NA 1.0 and $\bar{n}_{th} = 10$; dashed line). (b) For fixed NA 0.6, dependence of the fidelity on the mean occupation number $\bar{n}_{th}$ of the initially thermal state of motion (the solid line again showing the full mechanism, the dashed line only the effect from the crossed transversal dipoles).

the same optical path (e.g., into the same single-mode fibre).

Two such mechanisms commonly exist. First, consider the case where the emitting dipoles are $\mathbf{d}_{e0} = \hat{x}/\sqrt{2}$ and $\mathbf{d}_{e1} = \hat{y}/\sqrt{2}$ and collection is along $\hat{z}$. The collection weights for the two decays differ away from the optical axis due to the asymmetry of the dipole radiation patterns (e.g., $w_{0H}(\hat{\mathbf{k}})$ will fall off more quickly in the x than the y direction, as the radiation pattern from an x -oriented dipole has a node along $\hat{x}$). Second, in the case of ion-photon entanglement generation between $J = 1/2 \leftrightarrow J' = 1/2$ levels perpendicular to the magnetic field direction [22,88], technically advantageous due to its simplicity, the two decay channels correspond to $\sigma^\pm$ and π dipole transitions with $\mathbf{d}_{e0} = -(\hat{x} \pm i\hat{z})/\sqrt{3}$ and $\mathbf{d}_{e1} = \hat{y}/\sqrt{3}$ (magnetic field along $\hat{y}$). When the emitter is exactly placed on the optical axis, emission from the longitudinal $\hat{z}$ dipole that is a component of the circular σ dipole does not couple into an azimuthally symmetric fiber mode [22,67], enabling high-fidelity entanglement generation. In the presence of motion, this cancellation is no longer exact, leading to errors.

Our framework naturally covers arbitrary collection geometries in the full vector-optics picture through appropriate calculation of the $w_{i\sigma}(\hat{\mathbf{k}})$ weights. Instead of Eq. (D1) with the diagonal decay channel/polarization mapping, the full wavepackets from Eq. (5) with the sum over polarizations need to be used, but the rest of the calculation proceeds analogously [giving more nonzero terms in the equivalent of Eq. (D3)]. A numerical evaluation for parameters typical of trapped-ion experiments ($\eta = 0.07$), coupled into an optimally sized Gaussian mode, is shown in Fig. 8. The first mechanism discussed (mismatch of the transversal dipole emission patterns) remains

negligible, reaching only an error of 3×10^{-5} in the limit of NA 1.0 at $\bar{n}_{th} = 10$ (and $< 10^{-6}$ for all NAs at the motional ground state). Similar to the effect discussed in Sec. VB, the fact that the superposition of displacement operators is symmetric around the plane perpendicular to the dipole suppresses the associated errors; they could be feasibly reduced somewhat further by the application of state-dependent squeezing (and higher-order terms). The second mechanism, otherwise-suppressed longitudinal dipole coupling, dominates the errors and is non-negligible at moderate temperatures; it reaches about 0.1% at NA 0.6, $\eta = 0.07$ and $\bar{n}_{th} = 10$.

While these effects do not dominate current experimental error budgets, we conclude that they are not negligible even for typical trapped-ion experiments operating close to the Doppler limit, and certainly require scrutiny for more weakly trapped emitters or at higher temperatures. In such cases, working with, e.g., pure $\sigma^\pm$ emission imaged along the magnetic field, frequency encoding, etc. where there is no longitudinal dipole component to reject, may yield better results. In typical atomic structures, if two such $\sigma^\pm$ decays are possible, a third π decay channel will also be present that must then be rejected via symmetry, suffering from the same thermal dependence if such decays are not explicitly selected out in a postthermal check (our calculations allow for an arbitrary number of lower states and can thus be straightforwardly extended to such scenarios). Excitation into a superposition of levels with only one allowed decay channel each (e.g., σ^+ and σ^- on a $J = 1/2 \leftrightarrow J' = 3/2$ transition), or a time-bin encoding scheme making repeated use of the same transition, may avoid the polarization mixing effects, though the timing (phase-modulation) effects of large thermal states will still lead to errors (as discussed in Secs. VA and VI).

- [1] N. Gisin, G. Ribordy, W. Tittel, and H. Zbinden, Quantum cryptography, *Rev. Mod. Phys.* **74**, 145 (2002).
- [2] H. J. Kimble, The quantum internet, *Nature* **453**, 1023 (2008).
- [3] J. I. Cirac, A. K. Ekert, S. F. Huelga, and C. Macchiavello, Distributed quantum computation over noisy channels, *Phys. Rev. A* **59**, 4249 (1999).
- [4] L. Jiang, J. M. Taylor, A. S. Sørensen, and M. D. Lukin, Distributed quantum computation based on small quantum registers, *Phys. Rev. A* **76**, 062323 (2007).
- [5] N. H. Nickerson, J. F. Fitzsimons, and S. C. Benjamin, Freely scalable quantum technologies using cells of 5-to-50 qubits with very lossy and noisy photonic links, *Phys. Rev. X* **4**, 041041 (2014).
- [6] C. Monroe, R. Raussendorf, A. Ruthven, K. R. Brown, P. Maunz, L. M. Duan, and J. Kim, Large-scale modular quantum-computer architecture with atomic memory and photonic interconnects, *Phys. Rev. A* **89**, 022317 (2014).
- [7] B. Hensen, H. Bernien, A. E. Dréau, A. Reiserer, N. Kalb, M. S. Blok, J. Ruitenbergh, R. F. L. Vermeulen, R. N. Schouten, C. Abellán, W. Amaya, V. Pruneri, M. W. Mitchell, M. Markham, D. J. Twitchen, D. Elkouss, S. Wehner, T. H. Taminiau, and R. Hanson, Loophole-free Bell inequality violation using electron spins separated by 1.3 kilometres, *Nature* **526**, 682 (2015).
- [8] W. Rosenfeld, D. Burchardt, R. Garthoff, K. Redeker, N. Ortegel, M. Rau, and H. Weinfurter, Event-ready bell test using entangled atoms simultaneously closing detection and locality loopholes, *Phys. Rev. Lett.* **119**, 010402 (2017).
- [9] E. T. Khabiboulline, J. Borregaard, K. De Greve, and M. D. Lukin, Optical interferometry with quantum networks, *Phys. Rev. Lett.* **123**, 070504 (2019).
- [10] D. Gottesman and I. L. Chuang, Demonstrating the viability of universal quantum computation using teleportation and single-qubit operations, *Nature* **402**, 390 (1999).
- [11] S. Olmschenk, D. N. Matsukevich, P. Maunz, D. Hayes, L.-M. Duan, and C. Monroe, Quantum teleportation between distant matter qubits, *Science* **323**, 486 (2009).
- [12] W. Pfaff, B. J. Hensen, H. Bernien, S. B. van Dam, M. S. Blok, T. H. Taminiau, M. J. Tiggelman, R. N. Schouten, M. Markham, D. J. Twitchen, and R. Hanson, Unconditional quantum teleportation between distant solid-state quantum bits, *Science* **345**, 532 (2014).
- [13] D. Main, P. Drmota, D. P. Nadlinger, E. M. Ainley, A. Agrawal, B. C. Nichol, R. Srinivas, G. Araneda, and D. M. Lucas, Distributed quantum computing across an optical network link, *Nature* **638**, 383 (2025).
- [14] M. Żukowski, A. Zeilinger, M. A. Horne, and A. K. Ekert, “Event-ready-detectors” Bell experiment via entanglement swapping, *Phys. Rev. Lett.* **71**, 4287 (1993).
- [15] D. L. Moehring, P. Maunz, S. Olmschenk, K. C. Younge, D. N. Matsukevich, L.-M. Duan, and C. Monroe, Entanglement of single-atom quantum bits at a distance, *Nature* **449**, 68 (2007).
- [16] S. Ritter, C. Nölleke, C. Hahn, A. Reiserer, A. Neuzner, M. Uphoff, M. Mücke, E. Figueroa, J. Bochmann, and G. Rempe, An elementary quantum network of single atoms in optical cavities, *Nature* **484**, 195 (2012).
- [17] H. Bernien, B. Hensen, W. Pfaff, G. Koolstra, M. S. Blok, L. Robledo, T. H. Taminiau, M. Markham, D. J. Twitchen, L. Childress, and R. Hanson, Heralded entanglement between solid-state qubits separated by three metres, *Nature* **497**, 86 (2013).
- [18] M. K. Bhaskar, R. Riedinger, B. Machielse, D. S. Levonian, C. T. Nguyen, E. N. Knall, H. Park, D. Englund, M. Lončar, D. D. Sukachev, and M. D. Lukin, Experimental demonstration of memory-enhanced quantum communication, *Nature* **580**, 60 (2020).
- [19] A. Delteil, Z. Sun, W.-b. Gao, E. Togan, S. Faelst, and A. Imamoglu, Generation of heralded entanglement between distant hole spins, *Nat. Phys.* **12**, 218 (2016).
- [20] A. Ruskuc, C.-J. Wu, E. Green, S. L. N. Hermans, W. Pajak, J. Choi, and A. Faraon, Multiplexed entanglement of multi-emitter quantum network nodes, *Nature* **639**, 54 (2025).
- [21] J. O’Reilly, G. Toh, I. Goetting, S. Saha, M. Shalaev, A. L. Carter, A. Risinger, A. Kalakuntla, T. Li, A. Verma, and C. Monroe, Fast photon-mediated entanglement of continuously cooled trapped ions for quantum networking, *Phys. Rev. Lett.* **133**, 090802 (2024).
- [22] L. J. Stephenson, D. P. Nadlinger, B. C. Nichol, S. An, P. Drmota, T. G. Ballance, K. Thirumalai, J. F. Goodwin, D. M. Lucas, and C. J. Ballance, High-rate, high-fidelity entanglement of qubits across an elementary quantum network, *Phys. Rev. Lett.* **124**, 110501 (2020).
- [23] C. Monroe and J. Kim, Scaling the ion trap quantum processor, *Science* **339**, 1164 (2013).
- [24] J. P. Covey, H. Weinfurter, and H. Bernien, Quantum networks with neutral atom processing nodes, *npj Quantum Inf.* **9**, 90 (2023).
- [25] J. Ramette, J. Sinclair, N. P. Breuckmann, and V. Vuletić, Fault-tolerant connection of error-corrected qubits with noisy links, *npj Quantum Inf.* **10**, 58 (2024).
- [26] Y. Li and J. D. Thompson, High-rate and high-fidelity modular interconnects between neutral atom quantum processors, *PRX Quantum* **5**, 020363 (2024).
- [27] P. Drmota, D. Main, D. P. Nadlinger, B. C. Nichol, M. A. Weber, E. M. Ainley, A. Agrawal, R. Srinivas, G. Araneda, C. J. Ballance, and D. M. Lucas, Robust quantum memory in a trapped-ion quantum network node, *Phys. Rev. Lett.* **130**, 090803 (2023).
- [28] J. M. Pino, J. M. Dreiling, C. Figgatt, J. P. Gaebler, S. A. Moses, M. S. Allman, C. H. Baldwin, M. Foss-Feig, D. Hayes, K. Mayer, C. Ryan-Anderson, and B. Neyenhuis, Demonstration of the trapped-ion quantum CCD computer architecture, *Nature* **592**, 209 (2021).
- [29] A. Reiserer and G. Rempe, Cavity-based quantum networks with single atoms and optical photons, *Rev. Mod. Phys.* **87**, 1379 (2015).
- [30] H. Goto, S. Mizukami, Y. Tokunaga, and T. Aoki, Figure of merit for single-photon generation based on cavity quantum electrodynamics, *Phys. Rev. A* **99**, 053843 (2019).
- [31] J. Schupp, V. Krcmarsky, V. Krutyanskiy, M. Meraner, T. E. Northup, and B. P. Lanyon, Interface between trapped-ion qubits and traveling photons with close-to-optimal efficiency, *PRX Quantum* **2**, 020331 (2021).
- [32] C. Cabrillo, J. I. Cirac, P. Garcia-Fernandez, and P. Zoller, Creation of entangled states of distant atoms by interference, *Phys. Rev. A* **59**, 1025 (1999).
- [33] S. Saha, M. Shalaev, J. O’Reilly, I. Goetting, G. Toh, A. Kalakuntla, Y. Yu, and C. Monroe, High-fidelity remote

- entanglement of trapped atoms mediated by time-bin photons, *Nat. Commun.* **16**, 2533 (2025).
- [34] S. Sunami, S. Tamiya, R. Inoue, H. Yamasaki, and A. Goban, Scalable networking of neutral-atom qubits: Nanofiber-based approach for multiprocessor fault-tolerant quantum computer, *PRX Quantum* **6**, 010101 (2025).
- [35] D. J. Wineland, W. M. Itano, J. C. Bergquist, and R. G. Hulet, Laser-cooling limits and single-ion spectroscopy, *Phys. Rev. A* **36**, 2220 (1987).
- [36] C. Hempel, B. P. Lanyon, P. Jurcevic, R. Gerritsma, R. Blatt, and C. F. Roos, Entanglement-enhanced detection of single-photon scattering events, *Nat. Photonics* **7**, 630 (2013).
- [37] A. J. Rasmusson, I. Jung, F. G. Schroer, A. Kyprianidis, and P. Richerme, Measurement-induced heating of a trapped ion, *J. Phys. B: At. Mol. Opt. Phys.* **57**, 225002 (2024).
- [38] D. Leibfried, R. Blatt, C. Monroe, and D. Wineland, Quantum dynamics of single trapped ions, *Rev. Mod. Phys.* **75**, 281 (2003).
- [39] I. I. Ryabtsev, I. I. Beterov, D. B. Tretyakov, V. M. Entin, and E. A. Yakshina, Doppler- and recoil-free laser excitation of Rydberg states via three-photon transitions, *Phys. Rev. A* **84**, 053409 (2011).
- [40] L. Van Damme, Z. Zhang, A. Devra, S. J. Glaser, and A. Alberti, Motion-insensitive time-optimal control of optical qubits, *Quantum Sci. Technol.* **10**, 035025 (2025).
- [41] To elucidate the connection to usual concepts in atomic physics, we have defined the Lamb-Dicke parameter η_j for a constant photon wavevector $k \approx \omega_{ei}/c$ (with $\omega_{ei} := \omega_e - \omega_i$), which is an excellent approximation as spontaneous emission from atoms is narrowband ($\Gamma \ll \omega_{ei}$), and the lower states i of interest typically do not differ much in energy, $\omega_{ei} \approx \omega_{ei'}$. This choice merely aids exposition in the remote entanglement generation context; the model is straightforwardly generalized to lower levels with significantly different energies by not stating displacements in terms of a common η_j .
- [42] V. Weisskopf and E. Wigner, Berechnung der natürlichen Linienbreite auf Grund der Dirac'schen Lichttheorie, *Proc. R. Soc. Lond. A* **114**, 710 (1927).
- [43] P. W. Milonni, *The Quantum Vacuum: An Introduction to Quantum Electrodynamics* (Academic, Boston, 1994).
- [44] K. Rzaewski and W. Zakowicz, Spontaneous emission from an extended wavepacket, *J. Phys. B: At. Mol. Opt. Phys.* **25**, L319 (1992).
- [45] M. V. Fedorov, M. A. Efremov, A. E. Kazakov, K. W. Chan, C. K. Law, and J. H. Eberly, Spontaneous emission of a photon: Wave-packet structures and atom-photon entanglement, *Phys. Rev. A* **72**, 032110 (2005).
- [46] The combined effect of excitation and decay from repeated unsuccessful attempts of a photon-mediated entanglement protocol can lead to a build-up of motional excitation, so on a conceptual level, the ρ_{IM}^{rest} fraction cannot entirely be ignored. However, in this manuscript, we are concerned with the analysis of the qubit-qubit state resulting from the final attempt leading to success of such heralded protocols, for which the exact form of ρ_{IM}^{rest} turns out not to be relevant, as long as the initial motional state ρ_M is known (this includes single-photon protocols; see Appendix C). To obtain the precise ρ_M for a given entanglement attempt, the effect of prior noncollected photon decays [given by ρ_{IM}^{rest} representing the residual fraction of Eq. (4)], but also of other protocol steps, such as state preparation or recooling, may need to be considered.
- [47] To factor out the transversal modes, i.e., to write the various quantities as functions of $\hat{k}$ only, independent of $k = |\hat{k}|$, we have assumed the action of the imaging system to be independent of the photon frequency, which will be true in excellent approximation given the narrowband nature of atomic spontaneous emission. As with the definition of η_j , these assumptions are made for expository purposes only and are not essential to the model.
- [48] U. Eichmann, J. C. Bergquist, J. J. Bollinger, J. M. Gilligan, W. M. Itano, D. J. Wineland, and M. G. Raizen, Young's interference experiment with light scattered from two atoms, *Phys. Rev. Lett.* **70**, 2359 (1993).
- [49] R. Stockill, M. J. Stanley, L. Huthmacher, E. Clarke, M. Hugues, A. J. Miller, C. Matthiesen, C. Le Gall, and M. Atatüre, Phase-tuned entangled state generation between distant spin qubits, *Phys. Rev. Lett.* **119**, 010503 (2017).
- [50] P. C. Humphreys, N. Kalb, J. P. J. Morits, R. N. Schouten, R. F. L. Vermeulen, D. J. Twitchen, M. Markham, and R. Hanson, Deterministic delivery of remote entanglement on a quantum network, *Nature* **558**, 268 (2018).
- [51] S. L. N. Hermans, M. Pompili, L. D. S. Martins, A. R.-P. Montblanch, H. K. C. Beukers, S. Baier, J. Borregaard, and R. Hanson, Entangling remote qubits using the single-photon protocol: An in-depth theoretical and experimental study, *New J. Phys.* **25**, 013011 (2023).
- [52] A. J. Stolk *et al.*, Metropolitan-scale heralded entanglement of solid-state qubits, *Sci. Adv.* **10**, eadp6442 (2024).
- [53] E. T. Campbell and S. C. Benjamin, Measurement-based entanglement under conditions of extreme photon loss, *Phys. Rev. Lett.* **101**, 130502 (2008).
- [54] N. Kalb, A. A. Reiserer, J. J. W. Bakermans, S. J. Kamberling, N. H. Nickerson, S. C. Benjamin, D. J. Twitchen, M. Markham, and R. Hanson, Entanglement distillation between solid-state quantum network nodes, *Science* **356**, 928 (2017).
- [55] L. Slodička, G. Hétet, N. Röck, P. Schindler, M. Hennrich, and R. Blatt, Atom-atom entanglement by single-photon detection, *Phys. Rev. Lett.* **110**, 083603 (2013).
- [56] Here, we assume impulsive excitation to $|e\rangle$ from an approximate plane wave field, as describes well a pulse from a mode-locked laser that is not exceedingly strongly focused, as often employed in high-rate, high-fidelity remote entanglement experiments. Given that in the trapped ion/atom context, the linewidth of the dipole transition used to quickly extract photons usually significantly exceeds the motional frequencies ($\Gamma \gg \mu_j$ for each mode j), high-fidelity excitation to $|e\rangle$ implies this impulsive regime, though the model could be adapted to different excitation schemes by an appropriate choice of L .
- [57] J. Mizrahi, C. Senko, B. Neyenhuis, K. G. Johnson, W. C. Campbell, C. W. S. Conover, and C. Monroe, Ultrafast spin-motion entanglement and interferometry with a single atom, *Phys. Rev. Lett.* **110**, 203001 (2013).
- [58] The approximate equality in Eq. (9) is due to a minor approximation made which holds as long as the overall collection/detection efficiency of the apparatus is not too

- large, as discussed in Appendix C. If the collection efficiency is high, the erroneous double-excitation component can sometimes result in a distinct herald signal that can be rejected as invalid (two detector clicks). This leads to a slight upward correction of the fidelity, but is immaterial for practical parameters.
- [59] S. M. Barnett and P. M. Radmore, *Methods in Theoretical Quantum Optics*, Oxford Series in Optical and Imaging Sciences No. 15 (Clarendon, Oxford, 2005).
 - [60] D. J. Wineland, C. Monroe, W. M. Itano, D. Leibfried, B. E. King, and D. M. Meekhof, Experimental issues in coherent quantum-state manipulation of trapped atomic ions, *J. Res. Natl. Inst. Stand. Technol.* **103**, 259 (1998).
 - [61] C. Simon and William T. M. Irvine, Robust long-distance entanglement and a loophole-free bell test with ions and photons, *Phys. Rev. Lett.* **91**, 110405 (2003).
 - [62] L.-M. Duan and H. J. Kimble, Efficient engineering of multiatom entanglement through single-photon detections, *Phys. Rev. Lett.* **90**, 253601 (2003).
 - [63] J. Hofmann, M. Krug, N. Ortegel, L. Gérard, M. Weber, W. Rosenfeld, and H. Weinfurter, Heralded entanglement between widely separated atoms, *Science* **337**, 72 (2012).
 - [64] C. Nölleke, A. Neuzner, A. Reiserer, C. Hahn, G. Rempe, and S. Ritter, Efficient teleportation between remote single-atom quantum memories, *Phys. Rev. Lett.* **110**, 140403 (2013).
 - [65] D. Hucul, I. V. Inlek, G. Vittorini, C. Crocker, S. Deb-nath, S. M. Clark, and C. Monroe, Modular entanglement of atomic qubits using both photons and phonons, *Nat. Phys.* **11**, 37 (2015).
 - [66] V. Krutyanskiy, M. Galli, V. Krcmarsky, S. Baier, D. A. Fioretto, Y. Pu, A. Mazloom, P. Sekatski, M. Canteri, M. Teller, J. Schupp, J. Bate, M. Meraner, N. Sangouard, B. P. Lanyon, and T. E. Northup, Entanglement of trapped-ion qubits separated by 230 meters, *Phys. Rev. Lett.* **130**, 050803 (2023).
 - [67] T. Kim, P. Maunz, and J. Kim, Efficient collection of single photons emitted from a trapped ion into a single-mode fiber for scalable quantum-information processing, *Phys. Rev. A* **84**, 063423 (2011).
 - [68] D. P. Nadlinger, P. Drmota, B. C. Nichol, G. Araneda, D. Main, R. Srinivas, D. M. Lucas, C. J. Ballance, K. Ivanov, E. Y.-Z. Tan, P. Sekatski, R. L. Urbanke, R. Renner, N. Sangouard, and J.-D. Bancal, Experimental quantum key distribution certified by Bell's theorem, *Nature* **607**, 682 (2022).
 - [69] K. K. Mehta, C. D. Bruzewicz, R. McConnell, R. J. Ram, J. M. Sage, and J. Chiaverini, Integrated optical addressing of an ion qubit, *Nat. Nanotechnol.* **11**, 1066 (2016).
 - [70] K. K. Mehta, C. Zhang, M. Malinowski, T.-L. Nguyen, M. Stadler, and J. P. Home, Integrated optical multi-ion quantum logic, *Nature* **586**, 533 (2020).
 - [71] F. W. Knollmann, E. Clements, P. T. Callahan, M. Gehl, J. D. Hunker, T. Mahony, R. McConnell, R. Swint, C. Sorace-Agaskar, I. L. Chuang, J. Chiaverini, and D. Stick, Integrated photonic structures for photon-mediated entanglement of trapped ions, *Opt. Quantum* **2**, 230 (2024).
 - [72] O. Smedley, V. Natarajan, O. Jaramillo, H. M. Rivy, and K. K. Mehta, Atomic fluorescence collection into planar photonic devices, *Opt. Quantum* **3**, 28 (2025).
 - [73] E. M. Ainley, A. Agrawal, D. Main, P. Drmota, D. P. Nadlinger, B. C. Nichol, R. Srinivas, and G. Araneda, Multipartite entanglement for multi-node quantum networks, [arXiv:2408.00149](https://arxiv.org/abs/2408.00149).
 - [74] A. R. Vasquez, C. Mordini, C. Vernière, M. Stadler, M. Malinowski, C. Zhang, D. Kienzler, K. K. Mehta, and J. P. Home, Control of an atomic quadrupole transition in a phase-stable standing wave, *Phys. Rev. Lett.* **130**, 133201 (2023).
 - [75] E. Clements *et al.*, Sub-Doppler cooling of a trapped ion in a phase-stable polarization gradient, *Phys. Rev. Lett.* (2025).
 - [76] S. D. Barrett and P. Kok, Efficient high-fidelity quantum computation using matter qubits and linear optics, *Phys. Rev. A* **71**, 060310(R) (2005).
 - [77] M. T. Uysal and J. D. Thompson, Rephasing spectral diffusion in time-bin spin-spin entanglement protocols, *Phys. Rev. A* **110**, 052606 (2024).
 - [78] S. Kikura, R. Inoue, H. Yamasaki, A. Goban, and S. Sunami, Taming the recoil effect in cavity-assisted quantum interconnects, *PRX Quantum* **6**, 040351 (2025).
 - [79] Y. Yu, S. Saha, M. Shalaev, G. Toh, J. O'Reilly, I. Goetting, A. Kalakuntla, and C. Monroe, Entanglement fidelity limits of photonically-networked atomic qubits from recoil and timing, [arXiv:2503.19818](https://arxiv.org/abs/2503.19818).
 - [80] M. O. Scully and M. S. Zubairy, *Quantum Optics* (Cambridge University, Cambridge, 1997).
 - [81] B. Richards and E. Wolf, Electromagnetic diffraction in optical systems, II. Structure of the image field in an aplanatic system, *Proc. R. Soc. Lond. A* **253**, 358 (1959).
 - [82] L. Novotny, *Principles of Nano-Optics* (Cambridge University, Cambridge, 2012), 2nd ed.
 - [83] C. M. Bender and S. A. Orszag, *Advanced Mathematical Methods for Scientists and Engineers*, International Series in Pure and Applied Mathematics (McGraw-Hill, New York, 1978), Chap. 6.
 - [84] A. Erdelyi, Asymptotic representations of Fourier integrals and the method of stationary phase, *J. Soc. Ind. Appl. Math.* **3**, 17 (1955).
 - [85] D. P. Nadlinger, Device-independent key distribution between trapped-ion quantum network nodes, DPhil thesis, University of Oxford, 2022.
 - [86] O. Băzăvan, S. Saner, D. J. Webb, E. M. Ainley, P. Drmota, D. P. Nadlinger, G. Araneda, D. M. Lucas, C. J. Ballance, and R. Srinivas, Squeezing, trisqueezing, and quadsqueezing in a spin-oscillator system, [arXiv:2403.05471](https://arxiv.org/abs/2403.05471).
 - [87] C. Robens, S. Brakhane, W. Alt, F. Kleiβler, D. Meschede, G. Moon, G. Ramola, and A. Alberti, High numerical aperture (NA = 0.92) objective lens for imaging and addressing of cold atoms, *Opt. Lett.* **42**, 1043 (2017).
 - [88] B. B. Blinov, D. L. Moehring, L.-M. Duan, and C. Monroe, Observation of entanglement between a single trapped atom and a single photon, *Nature* **428**, 153 (2004).