Quantum Science and Technology

PAPER • OPEN ACCESS

Memory-assisted quantum key distribution resilient against multiple-excitation effects

To cite this article: Nicolò Lo Piparo *et al* 2018 *Quantum Sci. Technol.* **3** 014009

View the article online for updates and enhancements.

Related content

- Memory-assisted measurement-device-independent quantum key distribution
Christiana Panayi, Mohsen Razavi, Xiongfeng Ma et al.

- Optical quantum memory based on electromagnetically induced transparency
Lijun Ma, Oliver Slattery and Xiao Tang

- Engineered quantum dot single-photon sources
Sonia Buckley, Kelley Rivoire and Jelena Vukovi

This content was downloaded from IP address 129.11.22.210 on 11/01/2018 at 12:07

Quantum Science and Technology

OPEN ACCESS

RECEIVED
19 July 2017

REVISED
16 November 2017

ACCEPTED FOR PUBLICATION
24 November 2017

PUBLISHED
29 December 2017

Original content from this work may be used under the terms of the Creative Commons Attribution 3.0 licence.

Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

PAPER

Memory-assisted quantum key distribution resilient against multiple-excitation effects

Nicolò Lo Piparo[1,2,4], Neil Sinclair[3] and Mohsen Razavi[2]

[1] National Institute of Informatics, 2-1-2 Hitotsubashi, Chiyoda, Tokyo 101-0003, Japan
[2] School of Electronic and Electrical Engineering, University of Leeds, Leeds, United Kingdom
[3] Institute for Quantum Science and Technology, and Department of Physics & Astronomy, University of Calgary, Calgary, Alberta T2N 1N4, Canada
[4] Author to whom any correspondence should be addressed.

E-mail: nicopale@gmail.com

Keywords: quantum key distribution (QKD), quantum memory, single-photon source, quantum networks, quantum cryptography, quantum communications

Abstract

Memory-assisted measurement-device-independent quantum key distribution (MA-MDI-QKD) has recently been proposed as a technique to improve the rate-versus-distance behavior of QKD systems by using existing, or nearly-achievable, quantum technologies. The promise is that MA-MDI-QKD would require less demanding quantum memories than the ones needed for probabilistic quantum repeaters. Nevertheless, early investigations suggest that, in order to beat the conventional memory-less QKD schemes, the quantum memories used in the MA-MDI-QKD protocols must have high bandwidth-storage products and short interaction times. Among different types of quantum memories, ensemble-based memories offer some of the required specifications, but they typically suffer from multiple excitation effects. To avoid the latter issue, in this paper, we propose two new variants of MA-MDI-QKD both relying on single-photon sources for entangling purposes. One is based on known techniques for entanglement distribution in quantum repeaters. This scheme turns out to offer no advantage even if one uses ideal single-photon sources. By finding the root cause of the problem, we then propose another setup, which can outperform single memory-less setups even if we allow for some imperfections in our single-photon sources. For such a scheme, we compare the key rate for different types of ensemble-based memories and show that certain classes of atomic ensembles can improve the rate-versus-distance behavior.

1. Introduction

Providing secure key exchange at long distances is a yet-to-be achieved objective for quantum key distribution (QKD) systems. While some recent demonstrations have managed to exchange secret keys at 307 km [1] and 404 km [2], the key rate achieved at such distances is extremely low. The limitation in going to further distances is dictated by the exponentially-growing loss factor in optical fibers [3]. Probabilistic quantum repeaters offer a solution to extend the communication distance to over thousands of kilometers [4–10]. However, such quantum repeaters rely on quantum memory modules [11] with characteristics that are hard to achieve with the current technology. This does not necessarily mean that the existing quantum memories cannot offer any advantages. In fact, it has been shown that by using imperfect memories in measurement-device-independent QKD (MDI-QKD) systems, one may beat the memory-less QKD systems in rate and range to enable inter-city QKD operation [12, 13]. Although, unlike quantum repeaters, they are not scalable, such memory-assisted (MA) MDI-QKD setups can relax some of the demanding constraints on quantum memories, leading to more feasible implementations. Early investigations suggest that quantum memories with large storage-bandwidth products as well as short access and entangling times are necessary for MA-MDI-QKD [13]. These requirements may be achieved by quantum memories based on atomic ensembles [11], with the added benefit of strong light–matter

© 2017 IOP Publishing Ltd

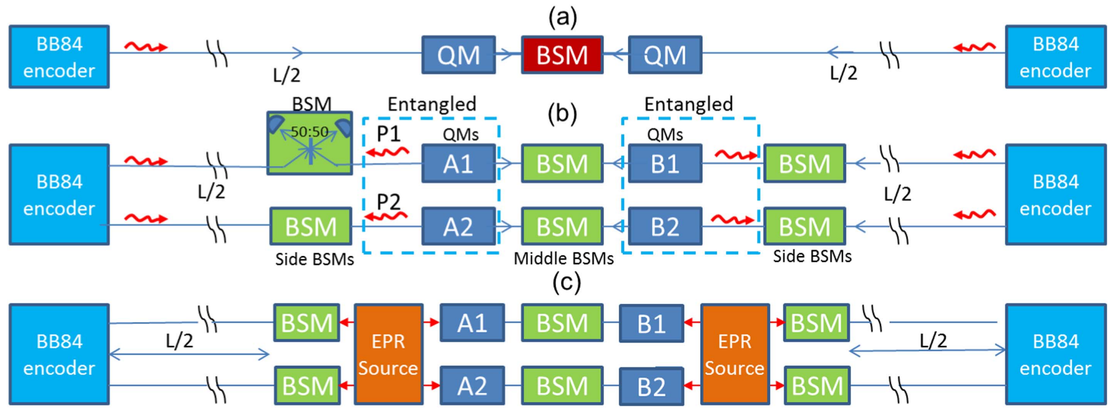

Figure 1. MA-MDI-QKD schemes with (a) heralding and (b) non-heralding quantum memories (QM boxes). In (a), we assume that, using certain mechanisms, the transmitted photon by the user can be written into the quantum memory and the memory can herald its successful loading [13]. In (b), the dual-rail configuration for ensemble-based quantum memories is shown. Here, in each round, one entangles quantum memories A_1 and A_2, and, similarly, B_1 and B_2, with two optical modes in the vacuum or single-photon state. At the transmitters, users encode their bits using phase encoded BB84 as explained in [16]. The BSM is performed using two single-photon detectors and a 50:50 beam splitter on each rail; see the BSM box for memory A_1. All other BSM boxes in (b) and (c) are the same. A click on only one detector would herald success for the corresponding BSM. Once both BSMs on one side are successful, we assume that the user's state has been teleported to the corresponding quantum memories. One then continues with loading the other two quantum memories, and, once done, they will proceed to perform the middle BSMs. (c) MA-MDI-QKD with EPR sources. At each round, one generates an entangled state in the form $|\psi_{\text{entg}}\rangle_{AP}$, use half of it to do the side BSM, and, if successful, attempt to store the other half in the quantum memories. Note that the dual-rail configuration in (b) and (c) is for illustration purposes only. In practice, one can use the equivalent single-rail time-bin encoding techniques.

coupling offering the possibility for efficient implementations. Ensemble-based quantum memories may, however, allow for storage of multiple excitations [14], which have been shown to be deleterious to their performance [15]. Here, we propose two MA-MDI-QKD schemes, both relying on single-photon sources, in an attempt to rectify the multiple-excitation problem. We begin with what seems to be the more obvious choice for our setup, but that turns out to not fully solve the problem, even if one uses ideal single-photon sources. We then fix the problem in our second setup, and show that it can outperform memory-less counterparts even if the employed single-photon sources are not ideal.

MA-MDI-QKD is a simple, but effective, extension of MDI-QKD, which inherits its resilience against detector attacks, and enhances its rate scaling. In MA-MDI-QKD, the photons transmitted by the users are stored each in a quantum memory before the entanglement-swapping Bell-state measurement (BSM) in the middle; see figure 1(a). This setup resembles a quantum repeater link with nesting level one, because of which rate enhancement follows, but without any quantum memories at the end users. The users instead need to have a BB84 encoder, which not only makes the implementation of the system easier, but also has an additional operational advantage: now, the repetition rate of the protocol is not determined by the distance, or the transmission delay, between the end users. Instead, one can in-principle run the protocol as fast as our quantum memories and optical sources in the setup allow without the need to wait for classical signals to acknowledge the success of entanglement distribution. If one employs quantum memories that feature short light–matter interaction times, then one may improve the total key generation rate per unit of time. One still, however, requires that the storage of the photon in the quantum memory to be *heralding*. Direct heralding mechanisms for writing photons into quantum memories are often slow, because of which the authors in [13] suggested to use the teleportation idea. That is, by first entangling a photon with the quantum memory, and performing a side BSM on this photon and the photon sent by the user, one can indirectly herald the transfer of the user's state to the corresponding quantum memory.

One of the first investigations [15] of the above technique utilized atomic-ensemble based quantum memories in conjunction with a heralding scheme based on off-resonant Raman interactions [4]. By using such a scheme [4] for interaction between weak pump signals and atomic ensembles, one can generate states with dominant terms in the form (neglecting normalization factors throughout this section) $|0\rangle_P|0\rangle_A + \sqrt{p_c}\,|1\rangle_P|1\rangle_A$, where $|0\rangle_P$ and $|1\rangle_P$ are, respectively, vacuum and single-photon states, $|0\rangle_A$ represents an ensemble with all atoms in their ground states, and $|1\rangle_A$ is an ensemble with only one atom, randomly, in a meta-stable excited state, while the rest are in the ground state. Using two of such states, see figure 1(b), plus post-selection succeeding with a probability proportional to p_c, one can then end up with an entangled state between two ensembles A_1 and A_2, and their corresponding photonic modes P_1 and P_2, in the form of $|\psi_{\text{entg}}\rangle_{AP} = |0\rangle_{P_1}|0\rangle_{A_1}|1\rangle_{P_2}|1\rangle_{A_2} + |1\rangle_{P_1}|1\rangle_{A_1}|0\rangle_{P_2}|0\rangle_{A_2}$, provided that p_c, the excitation probability, is much lower than one. The setup in figure 1(b) was investigated in [15] and it turned out that primarily the $|1\rangle_{P_1}|1\rangle_{A_1}|1\rangle_{P_2}|1\rangle_{A_2}$ state, which would be generated with probability

p_c^2, could result in such an amount of error that would prevent this system from outperforming memory-less systems. We refer to this issue by the two excited quantum memory (TEQM) problem. Note that reducing p_c would also reduce the success rate of the post-selection mechanism, and, on balance, would not result in an overall rate advantage.

There are several solutions to the TEQM problem. First, one may consider quasi-single-atom quantum memories, such as nitrogen-vacancy (NV) centers in diamond, as proposed in [17, 18]. In order to obtain a significant improvement in the key rate however, the NV centers must be embedded into microcavities [17]. While it is shown that the required cavity cooperativity is not necessarily high, their entangling protocol requires an appropriate single-photon source to entangle a photon with the electron spin of the NV center [17], a combination of which has yet to be demonstrated. Another remedy to TEQM, proposed in [15], is to use nearly ideal entangled-photon (EPR) sources for creating the initial memory-photon entanglement; see figure 1(c). The idea is that if one has an EPR source that ideally generates only one pair of photons per trigger, one of these photons can be used for the side-BSM operation, whereas the other can be stored (efficiently) in the quantum memory. Now, the latter photon can, in principle, drive only *one* transition and that would mitigate the multiple-excitation issue. In [15], the authors show that conventional EPR sources relying on parametric down-conversion would not solve the problem, but suggest to instead use quantum-dot based EPR sources, which have been shown to have very low second-order coherence properties [19]. Similarly, this solution would benefit quantum repeater implementations [20]. The other benefit of the EPR-based approach is that one only needs to write into quantum memories if the corresponding side-BSM is successful. We refer to this technique as 'delayed writing', which further reduces the requirements on the access times of quantum memories as they do not need to be initialized in every round.

Our proposed solutions consider using single-photon sources as a replacement for EPR sources for implementing the above ideas. Single-photon sources are at a more advanced stage of development than EPR sources, which opens up the possibility of a proof-of-principle experiment to be accomplished in the short term. Among different approaches to develop ideal single-photon sources, those that employ solid-state structures are one of the most attractive due to their potential for scalability and ease-of-use (see [21, 22] for recent reviews). Ideal single-photon sources for our proposal correspond to those that emit bright, stable, and high-rate streams of pure and indistinguishable single photons. Pure photon streams feature a near-zero second-order intensity correlation function. Wavelength tunability or operating temperature are often other considerations when evaluating single-photon sources, however these may be of reduced importance for our scheme if efficient wavelength conversion and low-temperature quantum memories, respectively, are employed.

There are different options for the single-photon sources needed in our scheme. One system that may be the most suitable is semiconductor quantum dots. Continuously-improving in quality [21], quantum dots offer high single photon generation rates (GHz, in principle) and second-order coherence ($g^{(2)}(0)$) on the order of 10^{-3} [23]. Recent work has increased the chance of generating and collecting a single photon per trigger pulse (up to 0.79 at the first lens [24]) and the degree of indistinguishability (up to 0.99 [23], in which the ideal photon corresponds to one), as well as the production yield [24] of dots embedded into micropillar cavities. Other sources of single photons that could be considered are heralded single photons that are produced by spontaneous parametric down-conversion or four-wave mixing processes. They feature exceptional indistinguishabilities (effectively unity), wide-band tunability (from the infrared to the ultraviolet), and room-temperature operation. Nonetheless these sources suffer from a low brightness (few percent) due to their statistical nature of emission [21, 22] and multiple-photon components that could be detrimental to our scheme. Multiplexing strategies have been studied to increase their emission rates [25, 26] without compromising their properties, e.g. purity. Crystal color centers in diamond [27], silicon carbide [28], or other inorganic crystals [29] have been investigated but currently suffer from low brightness due to the presence of other decay channels or other crystal defects [22]. Note that approaches using single molecules [30] and, recently, two-dimensional monolayers [22] have shown promise for high brightness. In our numerical results, we have taken system parameters from the recent quantum-dot implementations to verify the practicality of our proposed schemes.

Based on the state of the art for single-photon sources and memories, here, we propose two setups. The first of which resembles a noiseless linear amplifier (NLA) [31] and involves an entangling procedure that is based on the method described in [8, 32]. Simply, the user's photons are passed through a NLA before storing them into the quantum memories. For this setup, we optimize over the NLA parameters to maximize the key rate, finding that, while the TEQM issue is resolved, the rate scaling does not improve. Our second, improved, solution consists of a 'quasi-EPR' source relying on two single-photon sources. This setup provides the required entanglement after post-selection (via the side BSMs), solves the TEQM issue, improves the rate, and is compatible with some non-ideal single-photon sources [33].

The key rate of our system not only depends on the entangling procedure but also on the characteristics of the quantum memories that are employed [15]. Thus, we calculate the secret key rate of the proposed MA-MDI-QKD protocols considering different types of ensemble-based quantum memories. The latter may differ in coherence time, efficiency, bandwidth and access time, reading and writing procedures, and operating wavelength for example. In particular, we consider a selection of state-of-the-art memories based on warm vapors at room temperature [34–37], cold ensembles of rubidium atoms [38–40], and cryogenically-cooled rare-earth-ion-doped crystals [41–45]. In the latter case, we utilize the possibility of spectral multi-mode storage [46] in such memories.

We consider all major sources of errors in each MA-MDI-QKD setup, such as, channel loss, efficiency and background noise due to photodetection and frequency conversion, as well as coherence time, and writing-reading efficiencies of quantum memories. Based on our calculations, we find existing and near-future candidates of MA-MDI-QKD systems that offer better performance than existing QKD links.

Note that there are other schemes that offer a similar key rate scaling versus distance to the schemes proposed here. In particular, in [47], the authors present a memory-less structure, which can asymptotically achieve the square root scaling with the channel loss similar to a unity-nesting-level repeater. The key enabling ideas are (a) running multiple links in parallel; (b) performing quantum non-demolition measurements on arriving photons from the users; and (c) using an optical switch that directs the photons surviving the channel loss on each side to the middle BSM. The non-demolition measurement can, for instance, be done using the EPR source and the teleportation idea in figure 1(c). It is, however, shown that once we account for the growing insertion loss with size in optical switches, the rate scaling may not keep up with the desired scaling, and overall the system may perform worse than the MA-MDI-QKD systems [17]. In particular, because, in long distances, the chance of photon arrival decreases, in the scheme of [47], we need to run a larger number of parallel links, hence requiring a larger size for the optical switch resulting in higher insertion losses. Moreover, the fact that we need a large number of parallel links makes the implementation of such systems non-trivial as compared to our schemes, which just require a single physical link. There are also other MA-QKD schemes [48, 49] that start with entangling a photon with the quantum memory but the users, instead of being equipped with BB84 encoders, perform BB84 measurements similar to an entanglement-based QKD scheme [50]. Such schemes are not immune to detector attacks, but, in terms of key rate per transmitted pulse, offer again a similar scaling to MA-MDI-QKD systems. The downside is, however, in their repetition rate, which, like a repeater setup, is distance dependent. That requires quantum memories with long coherence times, but, more importantly, it makes it very hard for such systems to offer a *total* key rate comparable to high-clock-rate memory-less systems. Overall, all these systems provide us with interesting avenues to pursue in the near future term. MA-MDI-QKD, in particular, offers a high potential to be used as a competent commercial product in the QKD market.

The paper is structured as follows. In section 2 we describe our two proposed setups. In section 3, we study the performance of these setups by calculating their secret key rates. In section 4, we present our numerical results by comparing the key rate with the fundamental rate bounds for the distribution of secure keys over a lossy channel found in [51]. We also determine the secret key rate of the quasi-EPR-based setup for different types of ensemble-based quantum memories and we compare the rate with that of no-memory systems. In section 5, we draw our conclusions.

2. System description

In this section, we describe our two MA-MDI-QKD setups that rely on single-photon sources. The setups we present here both use the EPR source structure in figure 1(c) except that, instead of the actual EPR source, we use modules that employ single-photon sources to offer a similar functionality. We run the protocol with a repetition rate $R_S = 1/T$, where T is the repetition period, which is mainly specified by the single-photon source. In both cases we assume that the delayed writing procedure is used. That is, one attempts to write the photons into the quantum memories A_1 and A_2 only if both corresponding side BSMs are successful, and do similarly for B_1 and B_2. The delay required for this step can be on the order of nano seconds, corresponding to the measurement time at the BSMs [52], and should not incur much additional loss or complexity. The benefit is that the potentially time-consuming initialization of the quantum memories shall only be done once the memories have been loaded and read instead of in every round. The loading/reading step occurs at a much lower rate especially at long distances. In the following, we first explain our NLA- and quasi-EPR-based setups, and then give a precise description of all components used in these systems.

2.1. NLA-based MA-MDI-QKD

The key requirement of the setups of figures 1(b) and (c) is to generate entanglement between photons and quantum memories. Our aim is to achieve the same objective by using single-photon sources. A solution that may by envisioned utilizes entanglement distribution techniques that rely on single-photon sources. Not surprisingly, there is a class of probabilistic quantum repeaters that have such a property. In the scheme proposed in [32], the authors use a single-photon source and an imbalanced beam splitter to create spin-photon entanglement. They interfere two such photonic modes at a BSM to entangle the corresponding quantum

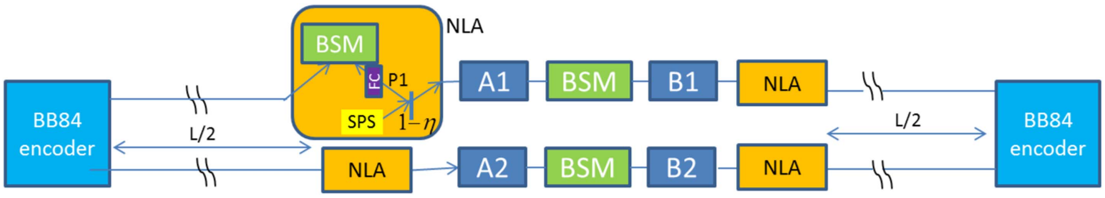

Figure 2. NLA-based MA-MDI-QKD. Users' photons will be effectively amplified before being stored in the quantum memories. A successful loading would be declared if the two NLAs corresponding to each user are both successful, that is, their corresponding BSM modules have a single click. We use the same BSM modules as in figure 1(b). The FC and SPS boxes, respectively, represent frequency converters and single-photon sources.

memories. In figure 2, we have used a similar idea to create our desired entangled state in the form $|\psi_{\text{entg}}\rangle_{AP}$, for A_1 and A_2 and their corresponding photonic modes P_1 and P_2. The same can be done for B_1 and B_2 in figure 2. Here, we use a beam splitter with reflectivity η to split a single photon into two paths: one would be stored into a quantum memory, and the other interferes at a BSM with the signal sent by the user. This structure, as shown for memory A_1 in figure 2, then resembles a NLA module based on quantum scissors [31]. We consider the quantum memories to be loaded with the user's transmitted state (within a known rotation) if both NLAs on one side are successful, meaning that their BSM module generates exactly one click,

As shown below, the above NLA structure can provide us with the required entanglement. Suppose the writing efficiency into quantum memories is unity, and, without loss of generality, let us focus on quantum memories A_1 and A_2. Assuming ideal on-demand single-photon sources, the joint state of the quantum memories and their corresponding optical modes P_1 and P_2 is given by

$$\begin{aligned}|\psi\rangle_{AP} = &\sqrt{\eta(1-\eta)}(|10\rangle_{P_1P_2}|01\rangle_{A_1A_2} + |01\rangle_{P_1P_2}|10\rangle_{A_1A_2})\\ &+ \eta|11\rangle_{P_1P_2}|00\rangle_{A_1A_2} + (1-\eta)|00\rangle_{P_1P_2}|11\rangle_{A_1A_2},\end{aligned} \tag{1}$$

where the first term, in brackets, is the desired entangled state. After the postselection by the two BSMs, which requires exactly one click in each module, the last term in 1 would be ideally removed. This last term is what could cause the TEQM problem. Therefore, this scheme resolves the TEQM issue. There is, however, a remaining term in the form $|11\rangle_{P_1P_2}|00\rangle_{A_1A_2}$, which is unwanted but can result in successful BSMs with a probability proportional to η^2, *whether or not* the transmitted photons have survived the path loss. That is, because of one background photon in each leg, in the asymptotic limit, when the distance L is large, the success rate of the side BSMs is nonzero. Let us give a name to this issue and call it the 'two loss-independent click' (TLIC) problem. We will show in section 3 how this problem prevents us from getting any rate advantage over memory-less setups. The scheme of [15] as shown in figure 1(b) also suffers from the TLIC issue. Note again that reducing η alone may not solve the problem, as our desired term occurs with a probability proportional to η. In principle, dark counts could also cause the TLIC problem, but, we may ignore it for now if it is small in comparison with other sources of background photons. We comment on the effect of dark counts later in this section and fully account for it in our key rate analysis. Next, we examine another solution that resolves the TLIC problem as well as the TEQM one.

2.2. MA-QKD with quasi-EPR sources

Figure 3(a) shows the MA-MDI-QKD setup with quasi-EPR sources for entangled photons. Our proposed quasi-EPR module is shown in figure 3(b), which may be built using integrated optics. It produces the desired entangled states, by interfering two single photons at different balanced beam splitters. It also generates additional spurious terms, which we aim to select out after successful side BSMs. Analyzing the circuit in figure 3(b), and using ideal A_1 and A_2 memories, the joint state of A and P modes can be written as follows:

$$\begin{aligned}|\psi\rangle_{AP} = &1/2(|10\rangle_{P_1P_2}|10\rangle_{A_1A_2} - |01\rangle_{P_1P_2}|01\rangle_{A_1A_2})\\ &+ \frac{1}{2\sqrt{2}}(|20\rangle_{P_1P_2} + |02\rangle_{P_1P_2})|00\rangle_{A_1A_2}\\ &+ \frac{1}{2\sqrt{2}}(|20\rangle_{A_1A_2} + |02\rangle_{A_1A_2})|00\rangle_{P_1P_2},\end{aligned} \tag{2}$$

where, again, the first term, in brackets, on the right-hand side, represents the desired entangled state. The last term represents the no-photon term, hence, unless for negligible dark count effects, cannot result in successful side BSMs, and it would be selected out. The term in the middle could result in successful BSMs, provided that the user's photon survives the path loss and/or because of dark counts. But, then, the quantum memories are both in their ground states, and except for a probability proportional to the dark count rate, they will not produce successful results at the middle BSM, and will be selected out at that stage.

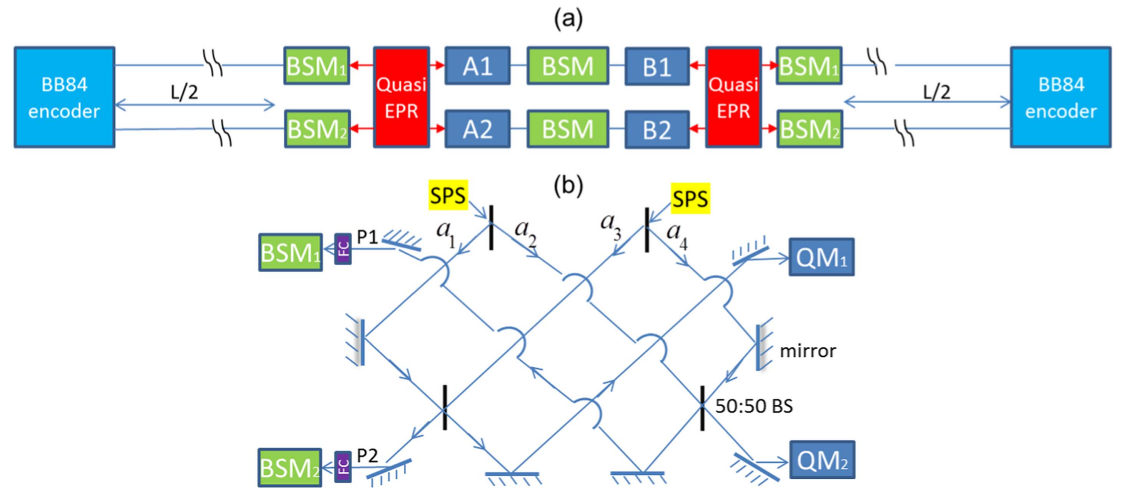

Figure 3. (a) MA-MDI-QKD with quasi-EPR sources. The quasi-EPR source, shown in (b), relies on two single-photon sources and a network of 50:50 beam splitters and mirrors that interfere the two photons via different paths. If both side BSMs are successful, the post-selected state resembles an entangled state in the desired form. The FC and SPS boxes, respectively, represent frequency converters and single-photon sources, and QM stands for quantum memory.

What happens in the module of figure 3(b) is that, by proper use of the quantum interference effect, we manage to group the unwanted states into terms in which both photons appear at the same output port. This creates only *one* background-induced click, making it easier to remove them by postselection. In the case of ideal single-photon sources, the above solution does resolve both the TEQM and TLIC problems. Even in the case of the second term, in order to get two successful side BSMs, one needs to have a user's photon arriving at the receiver, whose probability goes to zero at large distances. All of the previous discussion is based on the assumption that dark counts are negligible. The situation would be different if we have non-ideal single-photon sources with non-zero probabilities for emitting more than one photon, or when we have substantial dark count or background noise. We will consider theses scenarios later in our paper.

We made some idealistic assumptions in explaining how our proposed entanglement generation processes work. In the next section, we properly model major non-idealities in the system from which a realistic account of the key rate performance can be obtained.

2.3. Device modeling

We model different components of our system as follows.

BB84 encoders. We use phase encoding in the dual rail setup, or, equivalently, and if allowed by the quantum memory setups, time-bin encoding in a single-rail setup, in bases Z and X [16]. We assume that efficient QKD protocols are in use [53], where basis Z is chosen most often. We also assume that both users employ ideal single-photon sources for their BB84 encoding. In principle, they can use the decoy-state version of BB84, but, for the sake of our comparison, it would be sufficient to assume that both memory-assisted and memory-less systems use single photons to encode their bits. The multi-photon terms in a decoy-state protocol can be characterized by statistical analysis and they will not impose a change in the rate scaling [16]. The pulse duration is denoted by τ_p, and it is assumed to be equal to T in our numerical analysis.

Channel. We denote the total channel length by L, and its attenuation length by L_{att}. That is, the total channel transmissivity will be given by $\exp(-L/L_{\mathrm{att}})$. We assume that the channel does not impose any phase or polarization distortions. In practice, such effects can be compensated by classical-feedback mechanisms. The error in such compensating mechanisms can then be analytically modeled via misalignment parameters. In this work, we neglect such errors as they are not major error bearing components of our system, and they are common for both memory-assisted and memory-less systems. One can use the methods proposed in [13, 15] to account for such imperfections.

Single-photon detectors (SPDs). All our employed SPDs are assumed to be non-resolving detectors with efficiency η_D. The dark count rate is denoted by γ_{dc}, which results in a dark count probability $d_c = \gamma_{\mathrm{dc}}\tau_p$ per pulse. Here we assume that photodetectors are gated with an opening time that is identical to the pulse duration. The time that it takes to detect a photon and prepare the detector for next measurement is denoted by τ_M. Using self-difference techniques [52], τ_M can be on the order of nanoseconds.

Quantum memory. We consider several characteristics of quantum memories pertinent to our setups. The writing efficiency into quantum memories, i.e., the probability of successfully transferring the qubit-state encoded on a single photon to the quantum memory, is denoted by η_w. The probability of successfully reading the quantum memory, i.e., transferring the qubit-state stored in a quantum memory (back) onto a single photon is denoted by η_r. The latter will be affected by amplitude decay with time constant T_r. The reading efficiency at time t after the loading is then assumed to be given by $\eta_r = \eta_{r0} \exp(-t/T_r)$ [13, 15], where η_{r0} is the reading efficiency right after the loading. The exponential decay is not necessarily the case for all memories studied in this work. For instance, the decay is Gaussian for atomic frequency comb-based quantum memories that do not compensate for the dephasing induced by ground-level inhomogeneous broadening. In the regime of interest, where the relevant system time parameters are shorter than T_r, the exponential decay assumption will then be a pessimistic one for such quantum memories and it would not alter the overall conclusions made in our work. We also denote the required time to initialize the memory by τ_{init} and the time needed to interact with single photons by τ_{int}.

Single-photon source. We assume that the single-photon sources used in the middle site of figures 2 and 3 are identical but probabilistic. That is, upon trigger, there is a likelihood η_{SPS} that they generate the following normalized state

$$\rho_{\text{SPS}} = p_1|1\rangle\langle 1| + p_2|2\rangle\langle 2|, \tag{3}$$

where $|2\rangle$ is the two-photon state, and $p_1\eta_{\text{SPS}}$ and $p_2\eta_{\text{SPS}}$ are, respectively, the single-photon and double-photon probabilities. For most of this paper, we assume that $p_2 = 0$. We will later examine the range of values for p_2 that are tolerable for our setups.

Frequency converter. Given that many quantum memories do not operate at the telecom wavelengths, we may need to convert the frequency of some of the generated photons to match that of the quantum memory or the telecom channel used. We consider three scenarios: (1) use single-photon sources that generate photons at telecom wavelengths. One may then need an upconverter right before the quantum memories. The advantage is that the side BSM can be done more efficiently. On the downside, however, all the errors in the upconversion will affect the quantum memory as well; (2) generate photons that are matched to the quantum memory, but we downconvert the photons that enter the side BSM. Here, the advantage is that one can possibly use a matched single-photon source, in terms of the quantum memory bandgap and its bandwidth, to maximize the writing efficiency, but one will have noisier side-BSMs in this case; and (3), which is similar to (2), but one upconverts the photons sent by the user before the side BSM. In this work, we adopt the second scenario and assume that the wavelength and the bandwidth of the single-photon sources match that of quantum memories. In order to do side-BSMs, one may need to use a down-converter to match the wavelength of the two interfering photons [54–56]. We account for the conversion efficiency of such devices in our analysis. We also assume that the additional background Raman photons generated by the down-converter would modify the dark count of the side-BSM detectors.

In all devices, the sources of inefficiency are modeled by fictitious beam splitters with proper transmissivities.

3. Key rate analysis

In this section, we find the secret key generation rate for our proposed schemes shown in figures 2 and 3. We assume that there is no eavesdropping and we are only affected by device imperfections of the system as modeled in section 2.3. For convenience, we assume that both setups are symmetric. Under these conditions, in the infinite-key setting, the secret key generation rate in the setups of figures 2 and 3 is lower bounded by

$$R_{\text{QM}} = R_S Y_{11}^{\text{QM}}[1 - h(e_{11;X}^{\text{QM}}) - fh(e_{11;Z}^{\text{QM}})], \tag{4}$$

where $e_{11;X}^{\text{QM}}$ and $e_{11;Z}^{\text{QM}}$, respectively, represent the quantum bit error rate (QBER) between Alice and Bob in the X and Z basis, and $R_S Y_{11}^{\text{QM}}$ is the rate at which one generates raw key bits; the index 11 means that single photons are used at BB84 encoders; f denotes the inefficiency of the error correction scheme, and $h(q) = -q\log_2(q) - (1-q)\log_2(1-q)$ is the Shannon binary entropy function [13, 57].

We use the techniques of [13, 15] to calculate the above terms in the scenarios of interest. Without fully repeating the detailed calculations, here we just highlight the key steps in the derivation that are important in our understanding of the key-rate behavior of setups in figures 2 and 3. The key idea behind calculating Y_{11}^{QM} is to decompose the problem into two parts: (1) how often one loads the quantum memories on both sides, and (2) once loaded, how often the middle BSMs succeeds. Let us denote the former by P_{SBSM} and the latter by P_{MBSM}, to give

$$Y_{11}^{\text{QM}} = P_{\text{SBSM}} P_{\text{MBSM}}. \tag{5}$$

Here, P_{SBSM} partly depends on the probability to obtain two successful side-BSMs on one side, and partly on memory reading and writing times. Once both quantum memories are loaded, one has to spend a time equivalent to $\tau_r = \tau_{\text{int}} + \tau_M + \tau_{\text{init}}$ to obtain a measurement outcome for the middle BSM, and prepare the quantum memories for the next round [17]. Accounting for $\tau_w = \tau_{\rm int} + \tau_M$ to write into the quantum memory, there is a minimum time of $\tau_w + \tau_r$ to get one raw key bit. The inverse of this parameter then sets a bound on the maximum key rate achievable from our delayed-writing schemes. At long distances, however, the challenge of ensuring both sides to be loaded would take precedent, hence we would expect that $P_{\rm SBSM} \approx \frac{2}{3}{\rm Pr(Successful\ side\text{-}BSMs\ on\ one\ side)}$ [13]. As for $P_{\rm MBSM}$, the difficult part is to account for the decay of the quantum memories that may be loaded earlier. This requires us to average over the statistics of loading as has been detailed in [13, 15]. The same averaging is required in the calculation of $e_{11;X}^{\rm QM}$ and $e_{11;Z}^{\rm QM}$. Note that when T_r is sufficiently large, we can ignore the averaging, and we have $P_{\rm MBSM} \approx {\rm Pr(two\ successful\ middle\ BSMs)}$.

All above terms are found by calculating the relevant output density matrices in the setups of interest. We analytically obtain the pre-measurement state of the system by applying a series of transformations on the input density matrix considering channel transit, the entangling circuits, and the BSM modules. After applying relevant measurement operators, we then find the post-measurement states and the relevant probability terms This has been implemented using a generic Maple code developed for such setups.

Next, we examine the key rate scaling of the NLA-based and the quasi-EPR MA-QKD setups.

3.1. Key rate scaling: NLA-based setup

In this section, we investigate how the secret key rate of the scheme shown in figure 2 behaves at long distances. Here we ignore all inefficiencies except for the channel loss for simplicity. We also assume that T_r is sufficiently large. Under these conditions, from equation (1), there are two major terms that correspond to successful side-BSMs. The first term in brackets on the right-hand side of equation (1), corresponding to the desired entangled state, would result in successful side-BSMs provided that the user's photon has survived the path loss. This happens with a probability proportional to $(1-\eta)\eta\exp(-L/2/L_{\rm att})$ for which the quantum memories are left in the desired state. The other term that could result in successful side-BSMs is $|11\rangle_{P_1P_2}|00\rangle_{A_1A_2}$, which succeeds with probability η^2 and would leave the quantum memories in their ground state. At long distances, the post-measurement state of the quantum memories would be roughly given by

$$\rho_{A_1A_2}^{\rm (PM)} = \frac{\eta^2}{P_{\rm click}}|00\rangle_{A_1A_2}\langle 00| + \frac{(1-\eta)\eta {\rm e}^{-L/2/L_{\rm att}}}{P_{\rm click}}\rho_A^{\rm (TX)}, \tag{6}$$

where

$$P_{\rm click} = \eta^2 + (1-\eta)\eta{\rm e}^{-L/2/L_{\rm att}} \approx P_{\rm SBSM} \tag{7}$$

and $\rho_K^{\rm (TX)}$ represents the transmitted state (up to a known rotation) by user $K = A, B$. Starting with $\rho_{A_1A_2}^{\rm (PM)} \otimes \rho_{B_1B_2}^{\rm (PM)}$, then, for the middle BSM, one has

$$P_{\rm MBSM} \approx \frac{1}{P_{\rm SBSM}^2}[\eta^4 d_c^2 + 2\eta^3(1-\eta){\rm e}^{-L/2/L_{\rm att}}d_c + \eta^2(1-\eta)^2{\rm e}^{-L/L_{\rm att}}]. \tag{8}$$

In the regime of operation where $\eta \gg (1-\eta){\rm e}^{-L/2/L_{\rm att}} \gg d_c$, we then obtain

$$Y_{11}^{\rm QM} = P_{\rm SBSM}P_{\rm MBSM} \propto (1-\eta)^2{\rm e}^{-L/L_{\rm att}}, \tag{9}$$

that is, the key rate scales with the loss in the entire channel as is the case for a conventional memory-less system, and one should not expect any benefit from the NLA-based setup. As mentioned before, the distance-independent terms in equations (6) and (7) are the root causes of the TLIC problem. The dependence of both desired and undesired terms on η is another factor that results in such a rate scaling, even if one ignores the error terms The condition $\eta \gg (1-\eta){\rm e}^{-L/2/L_{\rm att}} \gg d_c$ represents operating regime of interest when long distances are considered as we show in section 4.

3.2. Key rate scaling: quasi-EPR setup

Using a similar analysis as in the previous section, we calculate how the secret key rate scales for the quasi-EPR setup of figure 3. In this case, for an ideal single-photon source with $p_1 = 1$ and $p_2 = 0$ and ignoring dark counts, from equation (2), one has $P_{\rm SBSM} \propto \exp(-L/2/L_{\rm att})$, as the arrival of the user's photon is necessary to have two successful side-BSMs. This implies that $\rho_{A_1A_2}^{\rm (PM)}$ is in the form $\alpha|00\rangle_{A_1A_2}\langle 00| + \beta\rho_A^{\rm (TX)}$ for some comparable constant probabilities α and β, adding up to one. $P_{\rm MBSM}$ would then be given by

$$P_{\rm MBSM} \approx \frac{1}{P_{\rm SBSM}^2}[\alpha^2 d_c^2 + 2\alpha\beta{\rm e}^{-L/2/L_{\rm att}}d_c + \beta^2{\rm e}^{-L/L_{\rm att}}], \tag{10}$$

which, when $d_c \ll {\rm e}^{-L/2/L_{\rm att}}$, results in

$$Y_{11}^{\rm QM} \propto {\rm e}^{-L/2/L_{\rm att}}. \tag{11}$$

From equation (11), we infer that the key rate for the setup of figure 3 scales similarly as a single-node quantum repeater system. Although this rough analysis does not account for the possible errors, the quasi-EPR setup promises to outperform the no-memory schemes. We examine this conjecture in the next section.

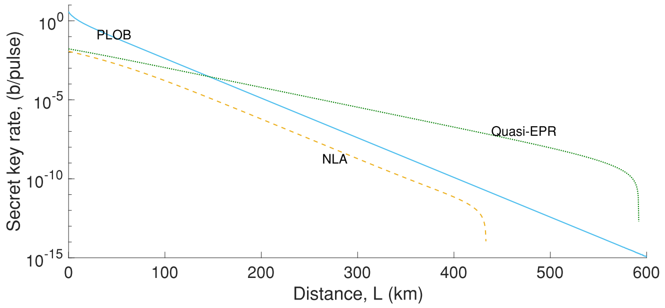

Figure 4. Secret key rate per pulse corresponding to the setups of figures 2 and 3. We have used $\eta = 0.2$ for the NLA-based system, which maximizes the rate. Memories are assumed to be ideal. The only sources of nonideality are listed in table 1 for a detection gate of 1 ns. We compare the rate with the PLOB bound obtained in [51].

The above conclusion relies on the assumption that $p_2 = 0$, which results in $P_{\rm SBSM}$ being proportional to the channel loss. If the probability to obtain the two-photon terms of the single-photon source is non-zero, then the distance-independent terms in $P_{\rm SBSM}$ are on the order of p_2, similar to η^2 in equation (7). Such terms would result in the TLIC problem as before once p_2/p_1 is comparable to $e^{-L/(2L_{\rm att})}$. The same holds if d_c is on the order of $e^{-L/(2L_{\rm att})}$, which could happen if the frequency converters generate a large background noise. In the following section, we explore the requirements on the employed devices in practical setups.

4. Numerical results

In this section we calculate the secret key rate that can be achieved using the schemes illustrated in figures 2 and 3. Specifically, we first calculate the secret key rate with the assumption that *ideal* quantum memories, meaning those that feature no limitations in performance, are employed. We compare the secret key rate per pulse of both schemes with the maximum rate achievable over a lossy channel, as obtained in [51]. We refer to this bound by the PLOB acronym. We find that the quasi-EPR scheme can outperform the PLOB bound, while the NLA scheme, due to the TLIC problem, fails to surpass it. Next we calculate the secret key rate, in bits per second, corresponding to the quasi-EPR scheme in conjunction with experimentally-measured properties of state-of-the-art warm and cold atomic ensembles as well as solid-state quantum memories based on rare-earth-ion-doped crystals. For comparison we also plot the secret key rate for a no-memory MDI-QKD implementation driven at 1 GHz repetition rate; we use the 'no-memory' label to refer to this system. We find that, under certain assumptions, some cold atom memories can surpass the no-memory bound due to their favorable coherence properties.

We also calculate the secret key rate of the quasi-EPR scheme with the assumption that we employ quantum memories that feature properties with *modest* improvements over the state-of-the-art memories. We refer to these as 'near-future' quantum memories, and find that almost all near-future memories can outperform the no-memory system. We conclude with a discussion around other possible sources of imperfection, such as multi-photon events and background noise, and explore how these impact the quasi-EPR scheme.

4.1. Ideal quantum memories

Let us first consider the case of ideal quantum memories. Specifically, these memories feature unity reading and writing efficiencies and fidelities, infinitely long coherence times, unlimited bandwidth, and zero interaction and initialization times. We calculate the secret key rate for the NLA and quasi-EPR schemes and that provided by the PLOB bound. The results are shown in figure 4, where we have used the values in table 1 for the relevant parameters. The NLA-based scheme clearly cannot surpass the PLOB bound, running below and parallel to it at long distances. Our results validate the calculations of section 3.1, which show that, at long distances, the rate-scaling with distance of the NLA-based scheme is the same as a no-repeater system. If one accounts for imperfections in quantum memories, the NLA scheme can only perform worse, which is not promising. This observation can, however, shed some light on the question of whether NLAs can help discrete-variable QKD systems, as compared to the continuous-variable QKD schemes, where, for the latter, some improvement is expected [59].

Due to its improved rate-versus-distance scaling, the quasi-EPR scheme can, however, beat the PLOB bound at distances roughly greater than 150 km. Note that this scheme improves the key rate by nearly 5 orders of magnitude over the PLOB bound at a distance of 700 km. Based on this performance, in the following sections, we only focus on the quasi-EPR scheme for practical and near-future quantum memories.

Table 1. List of common parameters and their nominal values used in our simulations. The channel loss corresponds to 0.25 dB km^{-1}. The detector efficiency and dark count at around 1550 nm correspond to the superconducting telecom-wavelength detectors reported in [58], which can be used at side BSMs. The parameters at around 800 nm correspond to commercially available silicon SPDs needed for the middle BSM. This is justified by the fact that the largest wavelength of operation for the quantum memories we consider is 850 nm, which corresponds to cesium-based quantum memories, see section 4.2. Similar efficiencies for single-photon sources and frequency converters are reported in [33, 56], respectively.

Attenuation length, L_{att}	17.3 km
Detection efficiency, η_D	0.93 at 1550 nm; 0.6 at 800 nm
Dark count rate, γ_{dc}	1 cps at 1550 nm; 1000 cps at 800 nm
Error correction inefficiency, f	1.16
Frequency conversion efficiency	0.68
Single-photon source efficiency, η_{SPS}	0.72

4.2. State-of-the-art quantum memories

Here we evaluate the performance of the quasi-EPR scheme using a selection of state-of-the-art ensemble-based memories. There are a variety of systems that have been utilized for optical quantum memories; see [11] for a recent overview. We consider ensemble-based memories due to their strong light–matter coupling and, in several cases, the possibility of long coherence times (up to seconds [38]) and high bandwidths (up to several GHz [35]). Furthermore, they offer the possibility of multi-mode storage [7, 11]. By multi-mode we are referring to memories that can simultaneously store more than one qubit during a single storage event by encoding many qubits each into a different mode. This feature has been exploited to enhance secret key generation rates in certain quantum repeater schemes [7, 10, 46]. It is important to stress that the definition of multi-mode storage differs from our reference to (the detrimental) storage of multiple excitations. The former involves many excitations, in which each individual excitation occupies a single distinguishable mode (or a pair as required for a qubit), while the latter concerns many excitations that occupy a single mode and thus each excitation may not be distinguished. Motivated by their impressive, and continually-improving, experimental record, we specifically consider warm vapor (Cs and Rb atomic gas) and cold atom (Rb atoms in a magneto-optical trap or atomic lattice) systems [11] that rely on the so-called Raman quantum memory protocol [60] as well as cryogenically-cooled rare-earth-ion-doped crystals that utilize atomic frequency combs [61].

Raman memory schemes [60] rely on three energy levels, usually a Λ-level system that features long-lived ground levels. A strong control pulse maps a propagating off-resonant photon onto the ground level. This is called the 'writing' step and at this point the photon is 'stored'. To retrieve the excitation, a control pulse is applied again, in which the excitation is mapped back onto a propagating photon. This is referred to as the 'reading' step. Note that the Raman protocol has been applied to photons that encode qubits with respect to various degrees of freedom (see [11, 39] and references therein). Along with the convenience of operation at room temperature, warm vapor Raman quantum memories feature the possibility to efficiently store GHz-bandwidth photons with microsecond-long coherence times (with up to 100 μs being possible [36]) [11, 34, 35]. Cold atoms reduce the impact of collisional or motional-induced decoherence, and, if magnetic-field-insensitive states are used, they offer very long coherence times reaching hundreds of ms and possibly more [11, 38, 40].

In a similar way, on-demand atomic frequency comb quantum memories also require a Λ-level system except here an optical inhomogeneously-broadened transition is tailored into a series of narrow absorption lines (the 'comb'), each of which are detuned from each other by an integer multiple of a fixed detuning [61]. A photon is absorbed by the comb, creating a delocalized atomic excitation and, using an optical control pulse, the excitation is reversibly-mapped onto a long-lived spin level. The photon is emitted due to a quantum interference effect between each absorption line of the comb. Ensembles of rare-earth-ions are particularly suited for atomic frequency comb quantum memories due to the long coherence times of both the optical (100s of microseconds [62, 63]) and spin (up to milliseconds [62, 63] or even seconds [64]) transitions in conjunction with level structures that allow for efficient atomic frequency combs over ~MHz bandwidths [11, 62, 63].

In the following, we study the performance of certain representatives from each group of memories. In this subsection and next, we focus mainly on the memory characteristics and neglect two-photon emissions from the source (i.e., $p_2 = 0$), or other issues that may arise in the photonic part of the system. We address the latter issues

Table 2. Properties of a selection of demonstrated warm vapor memories. All values are derived from the corresponding references given in the table. We denote the warm vapor (WV) memories of [34–37] as WV1 through WV4, respectively.

	WV1 [34]	WV2 [35]	WV3 [36]	WV4 [37]
Efficiency, $\eta_w \eta_{r0}$	0.1	0.3	0.05	0.15
Coherence time, T_r	100 ns	1.5 μs	120 μs	5 ns
Interaction time, τ_{int}	320 ps	300 ps	~1.43 ns	440 ps
Repetition rate, R_S	1.2 GHz	1.25 GHz	518 MHz	~667 MHz

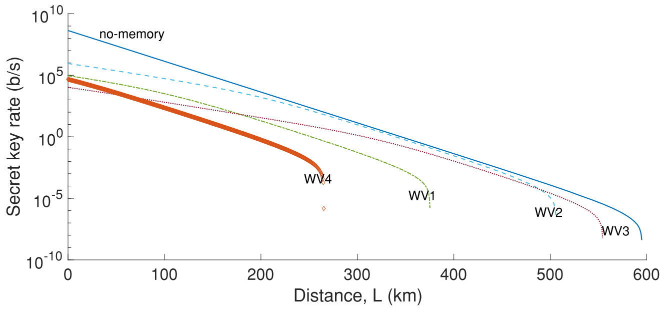

Figure 5. Secret key rate for the setups of figure 3 using the parameters of table 1 and the warm vapor memories featured in table 2.

in section 4.4. We also assume that memories feature no additional noise for the purpose of our simulations except for the decoherence effect and coupling issues already accounted for. This assumption is supported by several recent rare-earth atomic frequency comb [41, 45], as well as cold and warm Raman experiments [35, 37, 39] that have shown storage of non-classical light. We have ensured that the repetition rate of each quantum memory does not exceed the corresponding memory bandwidth. Furthermore, the choice of $\tau_p = T$ would minimize any inefficiency due to bandwidth mismatch between the source and the quantum memory. In practice, one may need to choose τ_p to be shorter than T, in which case its effect on the coupling efficiency must be considered. For all memories considered, we also assume that $\tau_{\text{init}} = 0$ given that these quantum memories would ideally go back to the desired initial state after being read out. In practice, memory re-initialization may be occasionally needed to avoid the spread of error. We assume that the frequency at which the initialization is needed is sufficiently low that it would not affect our key rate analysis.

Warm vapor. We consider the Raman memory demonstrations of [34–37] for our calculations. Each of the experiments use Cs vapor, except for [36] which uses ^{85}Rb, and feature memories of varying performance. See table 2 for a list of relevant memory properties. Specifically, the quantum memory demonstrated in [35] exhibits a reasonable combination of efficiency and coherence time as well as low noise, while [34] uses an anti-resonance of a Fabry–Perot cavity to suppress four-wave-mixing-induced noise that is present in [35]. The limitations of coherence time in these demonstrations are largely due to imperfect magnetic shielding, allowing magnetic-field-induced dephasing. The experiment of [36] employs exceptional magnetic shielding, but does not feature storage of non-classical light. Finally, [37] uses a ladder energy-level system to achieve storage in an excited level, which opens the possibility of storage of light pulses of less than ~100 ps duration. The storage time is, however, restricted to 5 ns in this experiment. Considering the excited nature of the level used for storage, the coherence time can be limited to around 100 ns. Figure 5 shows the secret key rate of the quasi-EPR scheme using the memories listed in table 2 as compared to the no-memory case. It can be seen that none of the considered quantum memories can surpass the no-memory curve. Nonetheless, the quantum memory of [35] allows the rate to become very close to that of the no-memory case, and could surpass the no-memory curve if the quantum memory coherence time was a bit longer or its coupling efficiency was a bit higher. Because of insufficient coherence time, the slope of the curve corresponding to memory WV2 starts changing around 200 km of distance. The lower slope corresponds to rate scaling with $\exp[-L/(2L_{\text{att}})]$, whereas the higher slope corresponds to $\exp[-L/(L_{\text{att}})]$ scaling, similar to the no-memory case. The change in slope happens later for WV3, which has the highest coherence time, and much earlier for the other two quantum memories. In the case of WV4, the coherence time is so low that the entire curve is parallel to that of the no-memory curve, indicating

Table 3. Properties of a selection of demonstrated cold atom memories and the corresponding interaction times and repetition rates used for the numerical calculation of the secret key rate of the quasi-EPR scheme. We denote the cold atom (CA) memories of [38–40] as CA1 through CA3, respectively.

	CA1 [38]	CA2 [39]	CA3 [40]
Efficiency, $\eta_w\eta_{r0}$	0.14	0.27	0.76
Coherence time, T_r	16 s	1.4 μs	220 ms
Interaction time, τ_{int}	82 ns	7 ns	240 ns
Repetition rate, R_S	12 MHz	133 MHz	4.2 MHz

Table 4. Properties of a selection of demonstrated rare-earth-ion-doped memories and the corresponding interaction times and repetition rates used for the numerical calculation of the secret key rate of the quasi-EPR scheme. We denote the rare-earth-ion-doped crystal memories of [41–45] as REIC1 through REIC5, respectively.

	REIC1 [41]	REIC2 [42]	REIC3 [43]	REIC4 [44]	REIC5 [45]
Efficiency, $\eta_w\eta_{r0}$	0.06	0.53	0.56	0.04	0.11
Coherence time, T_r	0.7 ms	37 μs	3 μs	38 μs	50 μs
Repetition rate, R_S	2 MHz	5 MHz	3 MHz	3.5 MHz	1 MHz

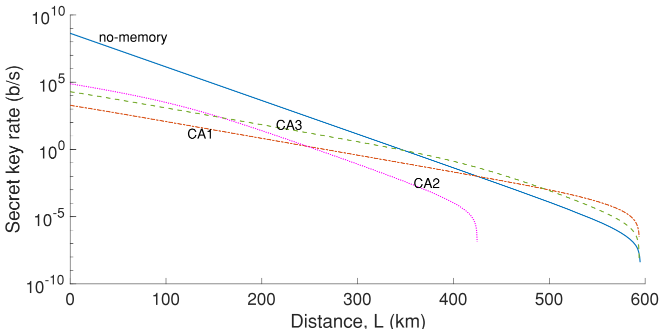

Figure 6. Secret key rate for the setups of figure 3 using the parameters of table 1 and the cold atom memories featured in table 3.

the same rate-versus-distance scaling. In section 3, we show that the no-memory bound may be overcome with some minor improvement in these quantum memories.

Cold atoms. We consider the three experiments described in [38–40]. Reference [39] utilizes ^{85}Rb in a magneto-optical trap while [38, 40] feature atomic lattices of ^{87}Rb. The coherence times of the magneto-optical trap implementations are limited by, among many factors, atomic diffusion in comparison to those of the atomic lattice [38–40]. The exceptional coherence time of [38] is due to compensation of light shifts, the insensitivity of the spin states to magnetic field fluctuations, and use of dynamical decoupling. We note that even though [38] does not explicitly show storage of non-classical light, the experiment of [65] importantly shows that no noise is introduced by dynamical decoupling. Our simulations, which are presented in figure 6, show that the atomic lattice experiments of [38, 40] can allow rates that surpass the no-memory bound. Both memories have such long coherence times that, in both cases, the maximum security distance has been dictated by the dark count noise, and not the memory decoherence. However, these memories are only useful if a low secret key rate is acceptable. Towards the possibility of higher rates and shorter-distance operation, we consider small improvements to memory properties (e.g. bandwidth) in section 3. Note that the experiments of [38, 40] employ off-resonant Raman scattering to achieve memory-photon entanglement and have not explicitly performed storage of an externally-provided photon as is required for the quasi-EPR scheme. We assume that the quantum memory parameters derived from these experiments may be translated to a Raman memory demonstration (as is achieved in [39]). We also mention that there is a theoretical proposal [66] for Raman memory using an optical lattice.

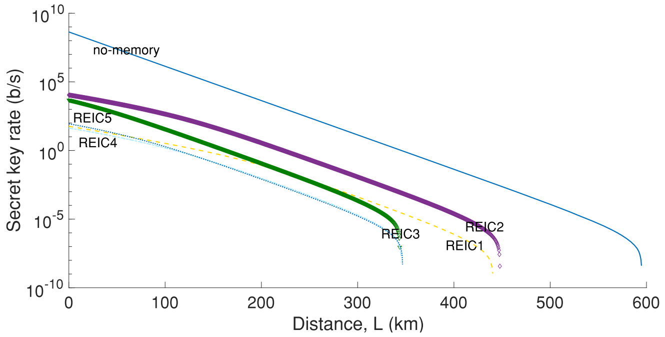

Figure 7. Secret key rate for the setups of figure 3 using the parameters of table 1 and the rare-earth-ion-doped crystal memories featured in table 4.

Table 5. Parameters for near-future warm vapor memories, and the corresponding interaction times and repetition rates, used for our numerical calculation of the secret key rate assuming the setup of figure 3. Memory abbreviations are explained in the main text.

	ExC	EnC	EnE
Efficiency, $\eta_w \eta_{r0}$	0.30	0.30	0.60
Coherence time, T_r	120 μs	10 μs	1.5 μs
Repitition rate, R_S	1.25 GHz	1.2 GHz	1.2 GHz

Rare-earth-ion-doped crystals. We consider the five atomic frequency comb experiments described in [41–45]. Europium-doped Y_2SiO_5 crystals are employed in the investigations of [41, 42] while the well-studied $Pr:Y_2SiO_5$ is featured in [44, 45]. On-demand storage at the single photon level is shown in [41], in which dynamical decoupling techniques are also used to overcome dephasing due to spin inhomogeneous broadening. Reference [42] utilizes a low-finesse cavity to show (up to 50%) efficient and on-demand storage of strong pulses. Efficient storage using a low-finesse cavity is achieved in [43], while on-demand storage of qubits and heralded single photons are shown in [44, 45], respectively. As shown in figure 7, again we simulate the key rate of the quasi-EPR scheme and find that none of the rare-earth-ion-doped crystal implementations will surpass the no-memory performance. The best performance is offered by REIC2, which has a high efficiency and a decent coherence time. Taking into consideration the technical challenges of obtaining both high efficiency and low noise in a rare-earth-ion-doped crystal-based atomic frequency comb system, in section 3 we explore the possibility of using several (spectral) modes to overcome the no-memory bound. Note that coherence times of 6 hours [64] and one minute [67] have been measured using magnetically-insensitive ground-level transitions of $^{151}Eu:Y_2SiO_5$ and $Pr:Y_2SiO_5$, respectively. However, it has yet to be shown that these coherence times can be combined with the possibility of efficient and broadband storage, hence these transitions may not be suitable for MA-MDI-QKD.

4.3. Near-future quantum memories

In this section we evaluate the performance of the quasi-EPR scheme using near-future quantum memories. Specifically, we suggest memory parameters that could be obtained with realistic experimental improvements to the memories of [34–45]. We attempt to be conservative with our suggested parameters, in particular with those of efficiency and coherence time, and acknowledge that there are fundamental limitations of some parameters, e.g. the restriction of bandwidth due to a certain energy level structure. Our enhanced memory parameters may represent a short-term goal for developing quantum memories.

Warm vapor. Here we consider three potential quantum memories with properties displayed in table 5. The corresponding quasi-EPR key rates shown in figure 8. The first we refer to as 'excellent coherence' (ExC) in which improved magnetic shielding will eliminate inhomogeneous spin dephasing such that a coherence time of [36] is achieved. Furthermore, we assume that a cavity is used to ensure low noise operation [34] and an enhancement of efficiency to that of [35], either by the cavity or control field tailoring [60]. We find that this memory enables surpassing the bound at just over 200 km and obtains maximal advantage at 400–500 km. This is a promising result given that MDI-QKD has been demonstrated over 400 km [2]—a distance for which

Table 6. Parameters of near-future cold atom memories, and the corresponding interaction times and repetition rates, used for our numerical calculation of the secret key rate assuming the setup of figure 3. Memory abbreviations are explained in the main text.

	CA2+BW	CA2+MI	CA3+BW	CA1+BW
Efficiency, $\eta_w\eta_{r0}$	0.90	0.50	0.30	0.10
Coherence time, T_r	1.4 μs	1 ms	220 ms	16 s
Repetition rate, R_S	~667 MHz	95 MHz	95 MHz	95 MHz

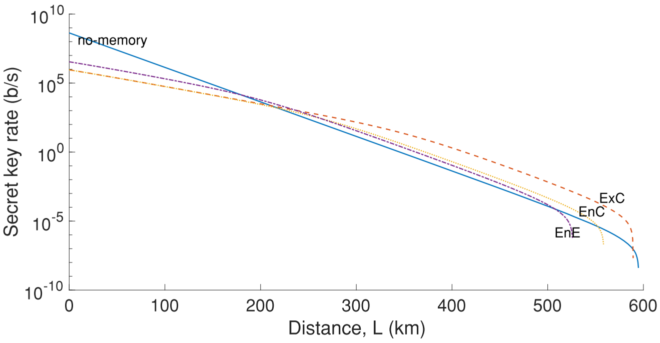

Figure 8. Secret key rate for the setups of figure 3 using the parameters of table 1 and the near-future warm vapor memories featured in table 5.

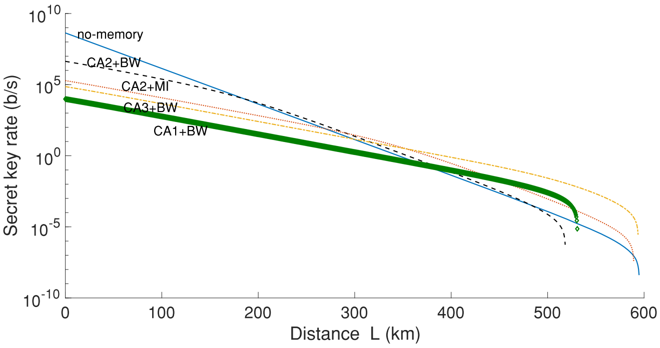

Figure 9. Secret key rate for the setups of figure 3 using the parameters of table 1 and the near-future cold atom memories featured in table 6.

channel stabilization has been realized. The second we refer to as 'enhanced coherence' (EnC) in which we keep all parameters the same as ExC except the coherence time, of which corresponds to the minimum required to surpass the no-memory bound. Interestingly, we find that a (reasonable) coherence time of approximately 10 μs will beat the bound at around 200 km, while the difference with memory ExC lies in the rate-distance scaling at longer distances. The last quantum memory we refer to as 'enhanced efficiency' (EnE) in which we keep the parameters the same as in [35] except we find the minimum efficiency to beat the bound, this being an efficiency of 60% at a distance of less than 200 km. Although it is likely that the EnE memory is challenging to achieve without added noise, improvements in experimental geometry in conjunction with control field optimization may reach this requirement without any compromise to coherence time. The quantum memory of [37] is not useful for MA-MDI-QKD due to the limited coherence time (up to 100 ns) of the (excited) level used for storage.

Cold atoms. Here we consider the quantum memories outlined in table 6, with the corresponding key rates shown in figure 9. We consider the memory of [39] with a bandwidth expanded to 1 GHz (CA2+BW), which results in $R_S \sim 667$ MHz. Note that the bandwidth must be less than half of the 3 GHz ground-state splitting of ^{85}Rb to ensure minimum impact of noise. Unfortunately, we find that, due its low coherence time, this quantum memory will only (just) beat the no-memory bound if it is ~90% efficient. Next we assume that a magnetically-insensitive ground-state transition is employed for the investigation of [39] (CA2+MI), finding that about 50% efficiency is needed to beat the bound, which can be realized by control pulse shaping or backward retrieval [60]. We also consider the quantum memory of [40] except we allow the bandwidth to be expanded to 100 MHz (CA3+BW), resulting in

Table 7. Parameters of near-future rare-earth-ion-doped memories, and the corresponding interaction times and repetition rates, used for our numerical calculation of the secret key rate assuming the setup of figure 3. Memory abbreviations are explained in the main text.

	Eu+DD	Eu+MM	Pr+DD	Pr+MM
Efficiency, $\eta_w\eta_{r0}$	1	0.53	1	0.56
Coherence time, T_r	15 ms	15 ms	500 μs	500 μs
Repitition rate, R_S	2 MHz	2 MHz	1 MHz	1 MHz
Number of spectral modes, N	1	30	1	90

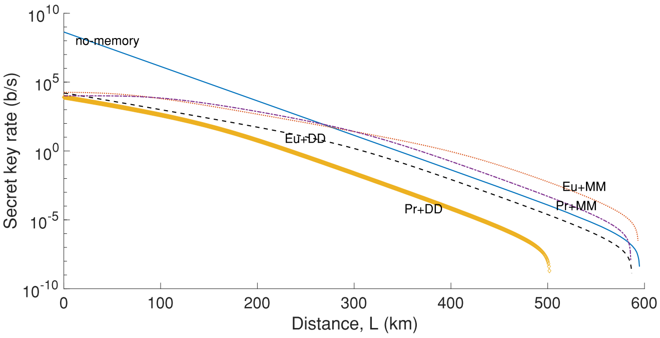

Figure 10. Secret key rate for the setups of figure 3 using the parameters of table 1 and the rare-earth-ion-doped crystal memories featured in table 7.

$R_S = 95$ MHz. This is well below the limitations given by the ground-state structure, but may pose a challenge if a cavity setup is employed. Encouragingly, we find that this quantum memory easily overcomes the bound if it is 30% efficient. Finally, if the highly-coherent memory of [38] is employed and its bandwidth is expanded from 12.2 to 100 MHz (CA1+BW), only 10% efficiency is required to be useful for MA-MDI-QKD for distances greater than 600 km, albeit at a low key rate. Note that, in majority of cases, the cross-over distance is around 300 km.

Rare-earth-ion-doped crystals. The corresponding quantum memory properties and key rates are shown in table 7 and figure 10, respectively. We employ the ^{151}Eu:Y$_2$SiO$_5$ memory of [41], except that we assume perfect dynamical decoupling is in use to achieve a coherence time that is entirely limited by the ground-level homogeneous broadening (Eu+DD), and we employ the Pr:Y$_2$SiO$_5$ crystal of [44, 45] (Pr+DD) in a similar way. Even with perfect efficiency, we find that neither of the quantum memories overcome the bound, mainly due to their limited bandwidth in comparison to the Raman quantum memories. To gain an advantage, we first assume the cavity enhanced setups of [42, 43] in conjunction with memories Eu+DD and Pr+DD, respectively. We then consider the possibility of multi-mode storage, which we refer to as memories Eu+MM and Pr+MM for Eu- and Pr-doped Y$_2$SiO$_5$, respectively. A multi-mode setup is less sensitive to decoherence issues as now we just need to have a successful side-BSM for one, out of many, modes on each side, which happens more often. This increases the maximum security distance as can be seen in figure 10.

The choice of degree of freedom over which the multi-mode quantum memory is designed needs further considerations. Since our implementation is already intrinsically temporally multi-mode, a convenient degree of freedom to use for multiplexing could be that of frequency. This is especially true of rare-earth-ion-doped crystals where their sub-level structure limits the atomic frequency comb bandwidth, but their inhomogeneously-broadened lines offer simultaneous storage of many, in some cases up to 1000 [46], spectral modes [62, 63]. In the case of example memories considered here, praseodymium-doped Y$_2$SiO$_5$ offers the possibility to store up to ~100 spectral modes given its hyperfine structure and its ~5 GHz inhomogeneous linewidth [44], while ^{151}Eu:Y$_2$SiO$_5$ only offers the possibility of storing a single spectral mode [41]. Nonetheless, one could employ spatial multiplexing, or explore the possibility to increase the inhomogneous linewidth by co-doping methods [68]. In order to use the multi-mode feature of the memory we may need to employ an array of single-photon sources, each generating single photons at different wavelengths or spatial modes. A normalized rate per channel use would then be of interest.

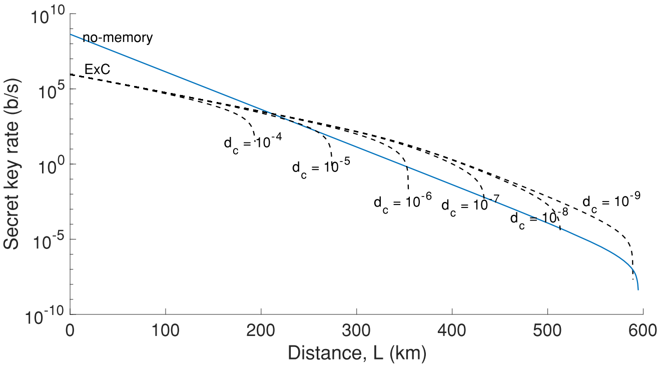

Figure 11. Comparison of the secret key rates of the setup of figure 3 with near-future warm vapor atomic ensembles at different values of dark count for the side-BSM detectors. The dark count term here accounts for not only the detector dark count, but also the background noise due to frequency converters and possibly the quantum memories. The other values are as in memory ExC in tables 1 and 5.

4.4. Near-future memories with additional system imperfections

The implication that some (possibly enhanced) atomic ensembles could outperform the memory-less QKD is based on several assumptions. One of the key assumption is that the two single-photon sources in the quasi-EPR setup can generate identical single photons that are (bandwidth-) matched to the quantum memories. We have also thus far ignored the additional background noise coming from the frequency converters. Any deviation from these assumptions may change the rate scaling and add to the QBER of the system. Below, we use our rough calculations of section 3.2 to investigate how resilient our setup is to the following imperfections.

- *Multi-photon terms*. We now test the resilience of our setup against possible multiple-photon components in the single-photon source. In fact, one can say that so long as $p_2/p_1 \ll \exp(-L/(2L_{\text{att}}))$, our system is immune against the two-photon terms generated by the source. At $L = 200$ km, that would require $p_2/p_1 \ll 0.003$, which is almost achievable with today's quantum dot technology for generating entangled and/or single photons [19], and possibly even those that rely on parametric down-conversion. In the latter case, a bank of downconverters is needed to boost the trigger rate of the system [69]. The additional QBER due to two-photon terms is on the order of p_2, which is negligible.
- *Photons distinguishibility*. If the two single photons generated by the two single-photon sources in figure 3(b) do not fully couple to each other at 50:50 beam splitters, then some TLIC-related issues occur at the side BSMs. Yet, similar to the two-photon terms, our system can tolerate the same order of magnitude (0.1%–1%) mismatch between the corresponding modes of the two single photons, which is again achievable by the current technology [23]. The additional QBER is also expected to be on the same order. The overlap between the user's photon and the single photons generated in the middle node is important, but not as vital as the overlap between that of the two single-photon sources. The former issue could increase the QBER to some extent but given that long-distance MDI-QKD has been demonstrated, this issue can be dealt with using existing technologies.
- *Bandwidth mismatch*. If the bandwidth of the single-photon source and the quantum memory do not match, one may end up with a large loss factor in the writing efficiency. For instance, the bandwidth of cold atomic ensembles is on the order of 10–100 MHz, which does not match that of many quantum-dot sources. If a quantum-dot source is used with CA1–CA3 memories, a drop of one to two orders of magnitude may be expected in their corresponding key rates in figures 6 and 9. The situation is more promising for warm vapor quantum memories, as their bandwidths are compatible with that of quantum-dot sources.
- *Background photons*. Finally, we have thus far ignored the effect of additional background noise generated by the frequency converters in our numerical analysis. In principle, at $L = 200$ km, based on the condition $d_c \ll \exp(-L/(2L_{\text{att}}))$, one expects to tolerate a dark count on the order of 10^{-4} per pulse, which is an order of magnitude higher than the typical background noise from frequency converters [55].

In order to test the above expectations, in figure 11, we have plotted the effect of dark counts from the side-BSM modules on the key rate of the MA-QKD system that uses memory ExC from near-future warm vapor atomic ensembles. Since warm vapor quantum memories are employed, no loss due to bandwidth mismatch is considered. The results show that, at $d_c = 10^{-6}$, the rate is nearly one order of magnitude above the MDI-QKD curve at $L = 300$ km, which leaves room for losses due to other experimental imperfections. Note that such a study of dark noise also guides the development of future Raman quantum memories based on warm vapor, which, without special considerations, are plagued by four-wave-mixing-induced noise [34].

5. Conclusions

In this paper we explored the possibility of using ensemble-based quantum memories in MA-MDI-QKD setups. Such quantum memories promise high efficiencies due to their strong light–matter coupling, large time-bandwidth products, and the ability to store multiple modes. By using single-photon sources, which are at an advanced stage of development, we proposed setups that could remove or alleviate the (single-mode) multiple-excitation problem. We identified the key problems in previously-proposed setups or the ones that resembled NLAs, and proposed a quasi-EPR setup that could outperform single no-memory QKD links. We showed that our solution is resilient against main imperfections in the source, the quantum memory module, and other required devices such as frequency converters and single-photon detectors. Based on our calculations, warm vapor atomic ensembles have the best chance to improve the rate-versus-distance behavior at channel distances above 200 km provided their efficiencies and coherence times can be improved. Cold atomic ensembles also offer a good performance provided that the bandwidth mismatch between the quantum memories and the driving single-photon sources can be reduced. Certain atomic frequency comb memories, such as $^{153}\mathrm{Eu}^{3+}$: $\mathrm{Y_2SiO_5}$, were also able to get close to the memory-less systems, but they need improvement in their coupling efficiency, coherence time and multi-mode capacity in order to offer a stable improvement. Our analysis ensures that a proof-of-principle experiment for our proposed setup would be within reach of the current technology and sets the stage for larger quantum repeater links to be implemented in the future.

Acknowledgments

This work was partly funded by the UK's EPSRC Grant EP/M013472/1 and EPSRC Grant EP/M506951/1, and the EU's H2020 Programme under the Marie Skłodowska-Curie project QCALL (GA 675662). NLP was also funded by the JSPS International Research Fellowship. NS acknowledges funding from Alberta Innovates Technology Futures (AITF) and thanks Wolfgang Tittel for his encouragement and support. All data generated in this paper can be reproduced by the provided methodology and equations.

ORCID iDs

Mohsen Razavi https://orcid.org/0000-0003-4172-2125

References

[1] Korzh B, Lim C C W, Houlmann R, Gisin N, Li D, Nolan M J, Sanguinetti B, Thew D and Zbinden H 2015 *Nat. Photon.* **9** 163
[2] Yin H-L *et al* 2016 *Phys. Rev. Lett.* **117** 190501
[3] Gisin N 2015 *Front. Phys.* **10** 100307
[4] Duan L-M, Lukin M D, Cirac J I and Zoller P 2001 *Nature* **414** 413–8
[5] Razavi M, Piani M and Lütkenhaus N 2009 *Phys. Rev.* A **80** 032301
[6] Amirloo J, Razavi M and Majedi A H 2010 *Phys. Rev.* A **82** 032304
[7] Sangouard N, Simon C, de Riedmatten H and Gisin N 2011 *Rev. Mod. Phys.* **83** 33
[8] Lo Piparo N and Razavi M 2013 *Phys. Rev.* A **88** 012332
[9] Lo Piparo N and Razavi M 2015 *IEEE J. Sel. Top. Quantum Electron.* **21** 6600508
[10] Guha S, Krovi H, Fuchs C A, Dutton Z, Slater J A, Simon C and Tittel W 2015 *Phys. Rev.* A **92** 022357
[11] Heshami K, England D G, Humphreys P C, Bustard P J, Acosta V M, Nunn J and Sussman B J 2016 *J. Mod. Opt.* **63** 2005
[12] Abruzzo S, Kampermann H and Bruß D 2014 *Phys. Rev.* A **89** 012301
[13] Panayi C, Razavi M, Ma X and Lütkenhaus N 2013 *New. J. Phys.* **16** 043005
[14] Razavi M and Shapiro J H 2006 *Phys. Rev.* A **73** 042303
[15] Lo Piparo N, Razavi M and Panayi C 2015 *IEEE J. Sel. Top. Quantum Electron.* **21** 6601010
[16] Ma X, Fung C-H F and Razavi M 2012 *Phys. Rev.* A **86** 052305
[17] Lo Piparo N, Razavi M and Munro W J 2017a *Phys. Rev.* A **95** 022338
[18] Lo Piparo N, Razavi M and Munro W J 2017b *Phys. Rev.* A **96** 052313
[19] Müller M, Bounouar S, Jöns K D, Glässl M and Michler P 2014 *Nat. Photon.* **8** 224
[20] Krovi H, Guha S, Dutton Z, Slater J A, Simon C and Tittel W 2016 *Appl. Phys.* B **122** 52
[21] Senellart P, Solomon G and White A 2017 *Nat. Nanotechnol.* **12** 1026
[22] Aharonovich I, Englund D and Toth M 2016 *Nat. Photon.* **10** 631
[23] Somaschi N *et al* 2016 *Nat. Photon.* **10** 340
[24] Gazzano O, Michaelis de Vasconcellos S, Arnold C, Nowak A, Galopin E, Sagnes I, Lanco L, Lemaître A and Senellart P 2013 *Nat. Commun.* **4** 1425
[25] Kaneda F, Christensen B G, Wong J J, Park H S, McCusker K T and Kwiat P G 2015 *Optica* **2** 1010

[26] Grimau Puigibert M, Aguilar G H, Zhou Q, Marsili F, Shaw M D, Verma V B, Nam S W, Oblak D and Tittel W 2017 *Phys. Rev. Lett.* **119** 083601

[27] Aharonovich I and Neu E 2014 *Adv. Opt. Mater.* **2** 911

[28] Castelletto S, Johnson B C, Ivády V, Stavrias N, Umeda T, Gali A and Ohshima T 2013 *Nat. Mater.* **13** 151

[29] Kolesov R, Xia K, Reuter R, Stöhr R, Zappe A, Meijer J, Hemmer P R and Wrachtrup J 2012 *Nat. Commun.* **3** 1029

[30] Chu X-L, Götzinger S and Sandoghdar V 2016 *Nat. Photon.* **11** 58

[31] Ralph T C and Lund A P 2009 *Quantum Communication Measurement and Computing Proc. 9th Int. Conf.* ed A Lvovsky p 155

[32] Sangouard N, Simon C, Minář J, Zbinden H, de Riedmatten H and Gisin N 2007 *Phys. Rev.* A **76** 050301

[33] Claudon J, Bleuse J, Malik N S, Bazin M, Jaffrennou P, Gregersen N, Sauvan C, Lalanne P and Gerard J M 2010 *Nat. Photon.* **4** 174

[34] Saunders D J, Munns J H D, Champion T F M, Qiu C, Kaczmarek K T, Poem E, Ledingham P M, Walmsley I A and Nunn J 2016 *Phys. Rev. Lett.* **116** 090501

[35] Reim K F, Michelberger P, Lee K C, Nunn J, Langford N K and Walmsley I A 2011 *Phys. Rev. Lett.* **107** 053603

[36] Camacho R M, Vudyasetu P K and Howell J C 2009 *Nat. Photon.* **3** 103

[37] Kaczmarek K T *et al* 2017 arXiv:1704.00013v1

[38] Dudin Y O, Li L and Kuzmich A 2013 *Phys. Rev.* A **87** 031801

[39] Ding D-S, Zhang W, Zhou Z-Y, Shi S, Shi B S and Guo G-C 2015 *Nat. Photon.* **9** 332

[40] Yang S-J, Wang X-J, Bao X-H and Pan J-W 2016 *Nat. Photon.* **10** 381

[41] Jobez P, Laplane C, Timoney N, Gisin N, Ferrier A, Goldner P and Afzelius M 2015 *Phys. Rev. Lett.* **114** 230502

[42] Jobez P, Usmani I, Timoney N, Laplane C, Gisin N and Afzelius M 2014 *New J. Phys.* **16** 083005

[43] Sabooni M, Li Q, Kröll S and Rippe L 2013 *Phys. Rev. Lett.* **110** 133604

[44] Gündoğan M, Ledingham P M, Kutluer K, Mazzera M and de Riedmatten H 2015 *Phys. Rev. Lett.* **114** 230501

[45] Seri A, Lenhard A, Rieländer D, Gündoğan M, Ledingham P M, Mazzera M and de Riedmatten H 2017 *Phys. Rev.* X **7** 021028

[46] Sinclair N *et al* 2014 *Phys. Rev. Lett.* **113** 053603

[47] Azuma K, Tamaki K and Munro W J 2015 *Nat. Commun.* **6** 10171

[48] Luong D, Jiang L, Kim J and Lütkenhaus N 2016 *Appl. Phys.* B **122** 96

[49] Rozpedek F, Goodenough K, Ribeiro J, Kalb N, Vivoli V C, Reiserer A, Hanson R, Wehner S and Elkouss D 2017 Realistic parameter regimes for a single sequential quantum repeater, arXiv:1705.00043

[50] Bennett C H, Brassard G and Mermin N D 1992 *Phys. Rev. Lett.* **68** 557

[51] Pirandola S, Laurenza R, Ottaviani C and Banchi L 2017 **8** 15043

[52] Yuan Z, Kardynal B, Sharpe A and Shields A 2007 *Appl. Phys. Lett.* **91** 041114

[53] Lo H-K, Chau H F and Ardehali M 2005 *J. Cryptol.* **18** 133

[54] Tang R, Li X, Wu W, Pan H, Zeng H and Wu E 2015 *Opt. Express* **23** 9796

[55] Pelc J S, Ma L, Phillips C R, Zhang Q, Langrock C, Slattery O, Tang X and Fejer M M 2011 *Opt. Express* **19** 21445

[56] Zaske S, Lenhard A and Becher C 2011 *Opt. Express* **19** 12825

[57] Shor P W and Preskill J 2000 *Phys. Rev. Lett.* **85** 441

[58] Marsili F *et al* 2013 *Nat. Photon.* **7** 210

[59] Blandino R, Leverrier A, Barbieri M, Etesse J, Grangier P and Tualle-Brouri R 2012 *Phys. Rev.* A **86** 012327

[60] Nunn J, Walmsley I A, Raymer M G, Surmacz K, Waldermann F C, Wang Z and Jaksch D 2007 *Phys. Rev.* A **75** 011401

[61] Afzelius M, Simon C, de Riedmatten H and Gisin N 2009 *Phys. Rev.* A **79** 052329

[62] Macfarlane R and Shelby R 1987 *Spectroscopy of Solids Containing Rare Earth Ions* (*Modern Problems in Condensed Matter Sciences* vol 21) (Amsterdam: Elsevier Science Publishers) p 51

[63] Thiel C, Bottger T and Cone R 2011 *J. Lumin.* **131** 353

[64] Zhong M, Hedges M P, Ahlefeldt R L, Bartholomew J G, Beavan S E, Wittig S M, Longdell J J and Sellars M J 2015 *Nature* **517** 177

[65] Rui J, Jiang Y, Yang S-J, Zhao B, Bao X-H and Pan J-W 2015 *Phys. Rev. Lett.* **115** 133002

[66] Nunn J, Dorner U, Michelberger P, Reim K F, Lee K C, Langford N K, Walmsley I A and Jaksch D 2010 *Phys. Rev.* A **82** 022327

[67] Heinze G, Hubrich C and Halfmann T 2014 *Phys. Rev.* A **89** 053825

[68] Sun Y C 2005 Rare earth materiais in optical storage and data processing applications *Spectroscopic Properties of Rare Earths in Optical Materials* ed R Hull *et al* (Berlin: Springer) pp 379–429

[69] Migdall A L, Branning D and Castelletto S 2002 *Phys. Rev.* A **66** 053805