Realistic parameter regimes for a single sequential quantum repeater

F. Rozpędek,[1,*] K. Goodenough,[1,*] J. Ribeiro,[1] N. Kalb,[1,2] V. Caprara Vivoli,[1] A. Reiserer,[1,2,3] R. Hanson,[1,2] S. Wehner,[1] and D. Elkouss[1]

[1] *QuTech, Delft University of Technology, Lorentzweg 1, 2628 CJ Delft, The Netherlands*
[2] *Kavli Institute of Nanoscience, Delft University of Technology, Lorentzweg 1, 2628 CJ Delft, The Netherlands*
[3] *Quantum Networks Group, Max-Planck-Institute of Quantum Optics, Hans-Kopfermann-Str. 1, 85748 Garching, Germany*

Quantum key distribution allows for the generation of a secret key between distant parties connected by a quantum channel such as optical fibre or free space. Unfortunately, the rate of generation of a secret key by direct transmission is fundamentally limited by the distance. This limit can be overcome by the implementation of so-called quantum repeaters. Here, we assess the performance of a specific but very natural setup called a single sequential repeater for quantum key distribution. We offer a fine-grained assessment of the repeater by introducing a series of *benchmarks*. The benchmarks, which should be surpassed to claim a working repeater, are based on finite-energy considerations, thermal noise and the losses in the setup. In order to boost the performance of the studied repeaters we introduce two methods. The first one corresponds to the concept of a *cut-off*, which reduces the effect of decoherence during storage of a quantum state by introducing a maximum storage time. Secondly, we supplement the standard classical post-processing with an *advantage distillation* procedure. Using these methods, we find realistic parameters for which it is possible to achieve rates greater than each of the benchmarks, guiding the way towards implementing quantum repeaters.

I. INTRODUCTION

Quantum communication enables the implementation of tasks with qualitative advantages with respect to classical communication, including secret key distribution [5, 14], clock synchronization [16] and anonymous state transfer [11]. Unfortunately, the transmission of both classical and quantum information over optical fibres decreases exponentially with the distance. While the problem of losses applies both to classical and quantum communication, classical information can be amplified at intermediate nodes, preventing the signal from dying out and thus increasing the rate of transmitted information. At the same time, the existence of a quantum analogue of a classical amplifier is prohibited by the no-cloning theorem [50]. Fortunately, in principle it is possible to construct a *quantum repeater* to increase the rate of transmission without having to amplify the signal [8, 29]. Hence, the construction of a quantum repeater would represent a fundamental milestone towards long distance quantum communications.

Many quantum repeater proposals require significant resources and are thus not within experimental reach. However, the recent experimental progress in the development of quantum memories [27, 36, 40] has brought the realisation of a quantum repeater closer to reality. In this paper, we evaluate a realistic setup of a so-called single sequential quantum repeater for the specific task of quantum key distribution. The setup considers two parties which we call Alice and Bob who are spatially separated, and want to generate a shared secret key. They use a single sequential quantum repeater [26] located between Alice and Bob, where both of them are connected to the quantum repeater by optical fibre. The repeater is composed of two quantum memories, both of which have the ability to become entangled with a photon, see FIG. 1. However, the repeater has a single photonic interface which means that it can only address Alice and Bob in a sequential fashion. Examples where only one of the qubit memories has an interface to the photonic channel include modular ion traps [22] and nitrogen-vacancy centres in diamond [6, 15, 36]. The situation is similar for atoms or ions trapped in a single cavity [35]. In this case, both memories can have a photonic interface. However, typically only one of the interfaces can be active at a given moment.

The figure of merit that we have chosen to evaluate the repeater is the secret-key rate. That is, the ratio between the number of generated secret bits and the number of uses of the quantum channel connecting the two parties. We compare the secret-key rate achievable with the repeater with a set of benchmarks that we introduce here. The most strict of these benchmarks is the capacity of the channel [47]. That is, the optimal secret-key rate achievable over optical fibre unassisted by a quantum repeater. The other benchmarks correspond to the optimal rates achievable with additional restrictions. In consequence, these benchmarks form a set of stepping stones towards the first quantum repeater able to produce a secure key over large distances.

* These authors contributed equally; filiproz@gmail.com

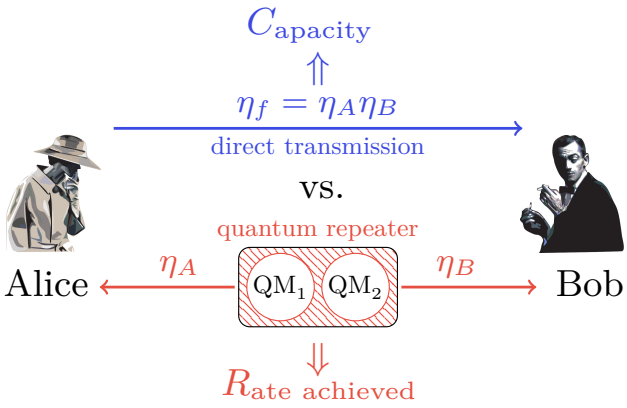

FIG. 1. The quantum repeater will send photons entangled with the QM_1 to Alice through the optical fibre of transmissivity η_A. After receiving one photon she will perform a BB84 or six-state measurement. After Alice has measured a photon and communicated her success to the quantum repeater, the quantum repeater tries to send a photon entangled with the QM_2 to Bob through the optical fibre of transmissivity η_B. If Bob does not receive a photon within some pre-defined amount of trials (i.e. the cut-off), Alice and Bob will abort the round. This is done to prevent the state in the QM_1 from decohering too much. If Bob does succeed, the quantum repeater performs a Bell state measurement on the two quantum memories.

The idea of assessing quantum repeaters by comparing with the optimal unassisted rates [3, 10, 18, 32, 33, 41, 48, 49] has spurred a significant amount of research devoted to developing sophisticated repeater proposals. Analysis of practical systems that utilise only parametric down-conversion sources and optical measurement setups [23] has shown that such systems do not allow for overcoming the channel capacity, which hints at the importance of quantum memories in repeater architectures. Specific architectures that utilise entangled-photon pair sources together with multimode quantum memories have also been considered in this context [20, 24]. Their analysis suggests that the required efficiency of those entangled-photon pair sources and number of storage modes might be experimentally very challenging for implementation in the very near future. Finally, the so called all-optical repeaters that do not require quantum memories but allow to overcome the channel capacity have been proposed [31]. However, they necessitate the ability to create large photonic cluster states which are beyond current experimental capabilities.

A detailed analysis of a realistic, single-node proof of principle repeater that includes all the specific system imperfections has been recently performed [26]. In particular, the analysis identified parameter regimes where it would be possible to surpass the optimal direct transmission rates with a repeater scheme that is close to experimental implementation. We build upon the analysis of [26] by introducing two methods that allow us to achieve higher rates. The first of these methods is the introduction of a maximum storage time for the memories in the quantum repeater. This restriction effectively reduces the effect of decoherence. We derive tight analytical bounds for the secret-key rate as a function of the maximum storage time. In this way we can perform efficient optimisation of the secret-key rate over the maximum storage time. The second of these methods is advantage distillation [19], a two-way classical post-processing technique that allows for distilling secret key at a higher rate than achievable with only one-way post-processing.

The structure of the paper is as follows. In Section II we detail our key distribution protocol. The sources of errors, such as losses in the apparatus and noisy operations and storage, are discussed in Section III. In Section IV, we calculate the secret-key rate that the single sequential quantum repeater would achieve. We define the benchmarks in Section V, and in Section VI we numerically explore the parameter regimes for which the quantum repeater implementation overcomes each benchmark and how the proposed protocol scales as a function of the distance. We end in Section VII with some concluding remarks.

II. PROTOCOL FOR A SINGLE SEQUENTIAL QUANTUM REPEATER

A quantum key distribution protocol consists of two main parts. First, Alice and Bob exchange quantum signals over a quantum channel and measure them to obtain a raw key that is post-processed in a second, purely classical part into a secure key [39]. Here, we focus our interest on the entanglement-based version of the BB84 [5] and the six-state [9] protocols. In this section, we describe the first part of both key distribution protocols.

The physical setup consists of two spatially separated parties Alice and Bob connected to an intermediate repeater via fibre optical channels. We note that such a repeater does not need to be positioned exactly half-way between Alice and Bob. The repeater is composed of two qubit quantum memories which we denote by QM_1 and QM_2. The repeater is then able to generate memory-photon entanglement, where the photonic degree of freedom in which the qubits are encoded could depend on the physical system, e.g. time-bin or polarisation encoding. Alice and Bob each have an optical detector setup that performs a BB84 or a six-state measurement. For technical reasons (see Section III), we consider slightly different setups for BB84 and six-state. More concretely, for BB84 we consider an active setup that switches randomly between the two measurement bases, while in the six-state protocol we consider a passive setup that chooses between the three measurement bases by a passive optical construction [17].

Let us now describe a first version of the protocol without a maximum storage time. First, the quantum repeater attempts to generate an entangled qubit-qubit state between a photon and the first quantum memory QM_1, after which the photon is sent through a fibre to Alice. Such a *trial* is attempted repeatedly until a photon arrives at Alice's side, after which Alice performs either a BB84 or a six-state measurement. Second, the quantum repeater attempts to do the same on Bob's side with the second quantum memory QM_2 while the state in QM_1 is kept stored. We denote the number of trials performed until a photon arrives at Alice's and Bob's sides n_A and n_B respectively. After Bob has received and measured a photon, a Bell state measurement is performed on the two states in QM_1 and QM_2. We denote by p_{bsm} the probability that the measurement succeeds. The classical outcome of the Bell state measurement is communicated to Bob. This concludes a single *round* of the protocol. We note that in this protocol every round ends with a successful generation of one bit of raw key.

One of the main problems in a quantum repeater implementation is that a quantum state will decohere when it is stored in a quantum memory. This means that if it takes Bob a large amount of trials to receive a photon, the state in the quantum memory QM_1 will have significantly decohered, preventing the generation of secret key. This motivates the introduction of a *cut-off*. A cut-off is a limit on the amount of trials that Bob can attempt to receive a photon. We denote this maximum number by $n^\star$.

The protocol that we consider here, modifies the protocol described above as follows: if in a given round Bob reaches the cut-off without success, the round is interrupted and a new round starts from the beginning with the quantum repeater again attempting to send a photon to Alice. In this scheme a large number of rounds might be required until a single bit of raw key is successfully generated. See Algorithm 1 for a description of the modified protocol with the cut-off.

Algorithm 1 Generation of a bit of raw key with a single sequential quantum repeater

Initialize:
  $n_A \leftarrow 0,\ n_B \leftarrow 0,\ k \leftarrow 0$
loop
  $k \leftarrow k + 1$ ▷ Increment the number of rounds
  repeat
    $n_A \leftarrow n_A + 1$ ▷ Increment the number of Alice's channel uses
    Generate entangled photon-QM_1 pair
    Send entangled photon through fibre towards Alice
  until Alice receives photon
  Alice performs a BB84 or a six-state measurement, stores result
  repeat
    $n_B \leftarrow n_B + 1$ ▷ Increment the number of Bob's channel uses
    Generate entangled photon-QM_2 pair
    Send entangled photon through fibre towards Bob
  until Bob receives photon or $n_B = kn^\star$
  if Bob received photon **then**
    Bob performs a BB84 or a six-state measurement, stores result
    Perform the Bell state measurement on the memories, communicate result
    Store $\max(n_A, n_B)$ ▷ Store channel uses
    return

III. SOURCES OF ERRORS

In this section, we model the different elements in the setup to identify the sources of losses and noise. The losses in the system are not only due to the transmissivity of the fibre; depending on the implementation a significant amount of photons is lost before they enter the fibre or due to the non-unit detector efficiency. The causes of noise are the experimental imperfections of the operations, measurements and quantum memories.

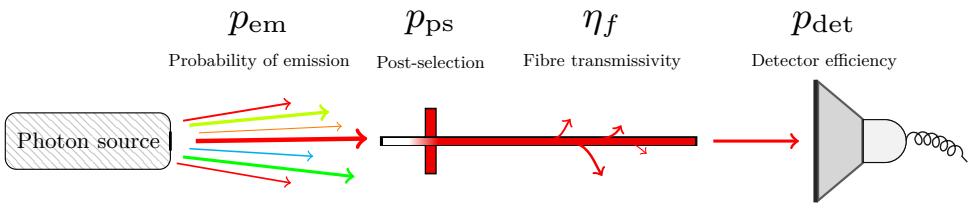

FIG. 2. General model of all photon losses occurring in the repeater setup. p_{em} is the probability of generating and capturing a photon into the fibre. For experimental reasons a fraction $(1 - p_{\text{ps}})$ of photons are additionally filtered out. The fibre has a transmissivity η_f. After exiting the fibre, the photons produce a click in the detector with probability p_{det}. The total efficiency of the apparatus is described by one parameter, $p_{\text{app}} = p_{\text{em}} p_{\text{det}}$.

Losses

We model the process of generating and sending an entangled photon through a fibre as follows (see FIG. 2). First, the photon has to be generated at some photon source and be captured in the fibre. This process happens with probability p_{em}. Depending on the experimental implementation, only a fraction p_{ps} of the photons entering the fibre can be used for secret key generation. This can occur for any number of reasons, for instance photons might be filtered according to frequency or a certain time-window [15, 35]. The filtering can happen either before or after the transmission through the fibre. The fibre losses are modelled as an exponential decay of the transmissivity η_f with the distance L, i.e. $\eta_f = \exp\left(-\frac{L}{L_0}\right)$ for some fibre attenuation length L_0. We denote by η_A the fibre losses on Alice's side and by η_B the fibre losses on Bob's side. Finally, the arriving photons will be captured by the detectors with an efficiency p_{det}. This probability of detecting a photon will be increased by the presence of dark counts (which will also inevitably add noise to the system), see the discussion of the dark counts at the bottom of this section and in Appendix A. We define the quantity $p_{\text{app}} = p_{\text{em}} p_{\text{det}}$ describing the total efficiency of our apparatus.

Noise

We model all noise processes either by the action of a dephasing channel

$$\mathcal{D}^{\lambda_1}_{\text{dephase}}(\rho) = \lambda_1 \rho + (1 - \lambda_1)\, Z \rho Z \tag{1}$$

or that of a depolarising channel

$$\mathcal{D}^{\lambda_2}_{\text{depol}}(\rho) = \lambda_2 \rho + (1 - \lambda_2)\, \frac{\mathbb{I}}{2} \tag{2}$$

where the parameters λ_1 and λ_2 quantify the noise, Z is the qubit gate $\left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$ and $\mathbb{I}/2$ is the maximally mixed state. The noise processes occur due to imperfect operations, decoherence of the state while stored in QM_1 and dark counts in the detectors.

The noise from imperfect quantum operations is captured by two parameters: F_{prep} and F_{gm}. F_{prep} is a dephasing parameter which corresponds to the preparation fidelity of the memory-photon entangled state [44]. F_{gm} is a depolarising parameter that describes the noise introduced by the imperfect gates and measurements performed on the two quantum memories during the protocol [12]. Hence, the noise can be modelled by a dephasing and a depolarising channel with $\lambda_1 = F_{\text{prep}}$ and $\lambda_2 = F_{\text{gm}}$.

The decoherence is modelled by a decay of the fidelity in the number of trials n. This decoherence is caused by two distinct effects. Firstly, there is the decoherence due to the time that the quantum repeater has to wait between sending photons. This time is the time it takes to confirm whether the photon got lost plus the time it takes to generate a photon entangled with the memory. We model this effect through an exponential decay of fidelity with time, which is expected whenever excess dephasing is suppressed (e.g. by dynamical decoupling [13]). However, we note that this is not the only possible model of decay, in several experiments a Gaussian decay has been observed [22, 38, 40, 43]. Secondly, attempting to generate an entangled photon-memory pair at QM_2 might also decohere the state stored in the QM_1. The latter effect is the most prominent decoherence mechanism in nitrogen-vacancy implementations [36], for example, where an exponential decay of fidelity with number of trials was observed. This is also how we model that effect here.

The quantum state ρ that is subjected to those effects undergoes an evolution given by the dephasing and depolarising channels with $\lambda_1 = (1 + e^{-an})/2$ and $\lambda_2 = e^{-bn}$. The two parameters a and b are given by

$$a = a_0 + a_1 \left(\frac{2n_{\text{ri}} L_B}{c} + t_{\text{prep}} \right) , \tag{3}$$

$$b = b_0 + b_1 \left(\frac{2n_{\text{ri}} L_B}{c} + t_{\text{prep}} \right) , \tag{4}$$

where $n_{\text{ri}} \approx 1.44$ is the refractive index of the fibre, c is the speed of light in vacuum, L_B the distance from the quantum repeater to Bob and t_{prep} is the time it takes to prepare for the emission of an entangled photon. Here a_0 and b_0 quantify the noise due to a single attempt at generating an entangled state and a_1 and b_1 quantify the noise during storage per second. Finally, the dark counts in the detectors introduce depolarising noise, this model is justified for the two quantum key distribution protocols that we consider, see [4, 17]. We let $\alpha_{A/B}$ denote the corresponding depolarising parameter on Alice's/Bob's side. The details of this model are presented in Appendix A.

IV. SECRET-KEY RATE OF A SINGLE SEQUENTIAL QUANTUM REPEATER

The secret-key rate R is defined as the amount of secret-key bits generated by a protocol divided by the number of channel uses and the number of optical modes. In the particular case of our sequential quantum repeater, the secret-key rate is given by

$$R = \frac{Y}{2} r \ . \tag{5}$$

The yield Y of the protocol is defined as the rate of raw bits per channel use. The secret-key fraction r is defined as the average amount of secret key that can be extracted from a single raw bit. The factor of a half is due to the fact that the encoding uses two optical modes. Since we consider two possible quantum key distribution protocols we take

$$r = \max\{r_{\text{BB84}}, r_{\text{six-state}}\} . \tag{6}$$

where r_{BB84} and $r_{\text{six-state}}$ are the secret-key fractions of BB84 and the six-state protocols respectively (see Eq. (12) and Appendix C).

Yield

The yield can be calculated as p_{bsm} (i.e. the success probability of the Bell state measurement) divided by the (average) number of channel uses needed for the successful detection of a photon by both Alice and Bob in the same round. With a single sequential quantum repeater it is not obvious how to count the number of channel uses. As in [26], we count the *maximum* of the two channel uses on Alice's and Bob's sides respectively,

$$Y = \frac{p_{\text{bsm}}}{\mathbb{E}\left[N\right]} = \frac{p_{\text{bsm}}}{\mathbb{E}\left[\max(N_A, N_B)\right]} \ . \tag{7}$$

where N, N_A and N_B are the random variables that model the number of channel uses, the number of channel uses at Alice's side and the number of channel uses at Bob's side, respectively.

Without the cut-off, it is possible to obtain an analytical formula for the average number of channel uses [26, 30],

$$\mathbb{E}\left[\max(N_A, N_B)\right] = \frac{1}{p_A} + \frac{1}{p_B} - \frac{1}{p_A + p_B - p_A p_B} \ , \tag{8}$$

where p_A and p_B depend on the quantum key distribution protocol and are given by the following equations (see Appendix A),

$$p_{A/B,\text{BB84}} = 1 - (1 - p_{\text{app}} p_{\text{ps}} \eta_{A/B})(1 - p_d)^2 \ , \tag{9}$$

$$p_{A/B,\text{six-state}} = 1 - (1 - p_{\text{app}} p_{\text{ps}} \eta_{A/B})(1 - p_d)^6 \ . \tag{10}$$

Every time that Bob reaches $n^\star$ trials, Alice and Bob restart the round and start over again. The cut-off thus increases the average number of channel uses. We have developed an analytic approximation of $\mathbb{E}[N]$ which is essentially tight (see Appendix D for the derivation and error bounds)

$$\mathbb{E}\left[\max(N_A, N_B)\right] \approx \begin{cases} \frac{1}{p_A\left(1-(1-p_B)^{n^\star}\right)} & \frac{1}{p_A} > n^\star \\ \frac{1}{p_A} + \frac{1}{p_B} - \frac{1}{p_A+p_B-p_Ap_B} & \frac{1}{p_A} \leq n^\star \end{cases} . \tag{11}$$

Secret-key fraction

Here we consider the secret-key fraction of the BB84 and six-state protocols. As we discussed previously, we consider the BB84 protocol with an active measuring scheme and the six-state protocol with a passive one. Moreover, we consider a fully asymmetric version of BB84 and a fully symmetric version of six-state. Fully symmetric means that all bases are used with equal probability while fully asymmetric means that the ratio at which one of the bases is used is arbitrarily close to one. Finally, we consider a one-way key distillation scheme for BB84 [39] while for the six-state protocol we consider the advantage distillation scheme in [45]. Advantage distillation [19] is a classical post-processing technique that allows to increase the secret-key fraction at all levels of noise.

The reasons for not analysing the BB84 protocol with advantage distillation and the fully asymmetric six-state with advantage distillation are technical. In the case of BB84, computing the rate with advantage distillation requires the optimisation over a free parameter. The combination of the optimisation over the cut-off together with the extra free parameter was computationally too intensive.

For the six-state protocol there is, to our knowledge, no security proof that can deal with the fully asymmetric six-state protocol with photonic qubits without introducing extra noise [2, 17]. However, these protocol choices do not have a strong impact on our analysis. Advantage distillation does not significantly increase the amount of distillable key for low error rates. Hence, asymmetric BB84 without advantage distillation is only slightly suboptimal. For higher error rates, where advantage distillation plays a role, the symmetric six-state protocol with advantage distillation is a factor of three away from the asymmetric version.

The expression for the secret-key fraction of both protocols depends on the error rates in the X, Y and Z bases, which we denote by e_X, e_Y and e_Z. In the case of the BB84 protocol, [25, 39] it is given by

$$r_{\text{BB84}} = 1 - h(e_Z) - h(e_X) , \tag{12}$$

where $h(p) = -p\log_2 p - (1-p)\log_2(1-p)$ is the binary entropy function. The expression for $r_{\text{six-state}}$ is more complex. We leave its discussion to Appendix C.

We can directly evaluate the error rates in each basis as a function of the general parameters of Section III. For our single sequential quantum repeater these average errors are

$$e_X = e_Y = e_{XY} = \frac{1}{2} - \frac{1}{2}F_{\text{gm}}\alpha_A\alpha_B\left(2F_{\text{prep}} - 1\right)^2\left\langle e^{-(a+b)n}\right\rangle , \tag{13}$$

$$e_Z = \frac{1}{2} - \frac{1}{2}F_{\text{gm}}\alpha_A\alpha_B\langle e^{-bn}\rangle . \tag{14}$$

where $\langle e^{-cn}\rangle$ is the average of the exponential e^{-cn} over a geometric distribution over the first $n^\star$ trials. The detailed derivation of the error expressions is presented in Appendix B.

V. BENCHMARKS FOR THE ASSESSMENT OF QUANTUM REPEATERS

We introduce a set of benchmarks to assess the performance of a quantum repeater implementation.

The first benchmark that we consider is the rate that would be achieved with the same parameters for the system losses and dark counts and for the same protocol but without a quantum repeater. Overcoming this benchmark gives the first indication that the repeater setup is useful; it means that the repeater setup outperforms the setup without repeater. We call this benchmark the direct transmission benchmark.

The remaining benchmarks represent the optimal secret-key rate that Alice and Bob could achieve if they were to communicate over the same quantum channel without a repeater under some constraints.

The optimal secret-key rate without a repeater highly depends on the channel model. The first modelling decision is the placement of the boundary between Alice's and Bob's laboratories and the quantum channel. This is because it is not *a priori* clear where the channel begins and ends. However, this decision has a strong impact on the optimal achievable rate; if the channel includes most of Alice's and Bob's laboratories, then the channel is more lossy and

noisy and the benchmark is easier to overcome. If, on the other hand, the channel is just the fibre optical cable the benchmark becomes more difficult to overcome.

We consider three cases in terms of the individual lossy components of our setup (see FIG. 1, FIG. 2 and their captions):

Case 1: Fibre only, in this case the transmissivity is: $\eta = \eta_{\mathrm{f}} = \eta_A \eta_B$.

Case 2: Fibre and different filters, then the channel transmissivity becomes: $\eta = \eta_{\mathrm{f}} p_{\mathrm{ps}}$.

Case 3: Fibre, filters and Alice's and Bob's apparatus, then the transmissivity becomes: $\eta = \eta_{\mathrm{f}} p_{\mathrm{ps}} p_{\mathrm{app}}$.

Note that although in the experimental implementation of the repeater the terms p_{ps} and p_{app} appear twice in the expression of the transmissivity, they appear only once in the benchmarks which include them. The reason is that in a scenario without a repeater the emission inefficiency and the filters only affect the transmissivity once.

The second design parameter for these benchmarks is the type of channel. Transmission of photons through fibres is modelled as a pure-loss channel [46], where only a fraction η of the input photons reach the end of the channel. The first type of channel that we consider is the pure-loss channel without any additional restriction. The optimal achievable rate over one mode of the pure-loss channel is given by the secret-key capacity [33]

$$-\log_2 (1-\eta) \ . \tag{15}$$

Note that for high losses the scaling of this capacity with distance is proportional to $\eta_f = \exp\left(-\frac{L}{L_0}\right)$. At the same time with an ideal (noiseless) single quantum repeater placed half-way between Alice and Bob, the expected secret-key rate would scale proportionally to $\sqrt{\eta_f} = \exp\left(-\frac{L}{2L_0}\right)$ [26].

The second type of channel that we consider is the pure-loss channel when the transmitter has a limitation in the energy that can be introduced into the channel. There has been some recent work studying the optimal rate per mode of the finite-energy pure-loss channel [18, 42, 48]. However, the optimal rate remains unknown. The bound that we consider here [42] is given by

$$g\left((1+\eta)\, P/2\right) - g\left((1-\eta)\, P/2\right) \ , \tag{16}$$

where $g(x) := (x+1)\log_2(x+1) - x \log_2 x$ and P is the mean photon number. In our repeater setup, the finite energy restriction arises from the fact that, on average, only a fraction of a photon enters the fibre in each trial. More precisely, the average photon number satisfies $P = p_{\mathrm{em}}$ in cases 1 and 2 above and $P = 1$ in case 3. Unfortunately, since Eq. (16) is an upper bound, it is only strictly smaller than the capacity of the pure-loss channel for small mean photon number. Expanding the bounds from equations (15) and (16) around $\eta = 0$ shows that the cross-over between the two bounds occurs when $p_{\mathrm{em}} \log_2 \left(\frac{p_{\mathrm{em}}+2}{p_{\mathrm{em}}}\right) = \frac{1}{\ln 2}$. In other words, for high losses the finite-energy bound is tighter when $p_{\mathrm{em}} \lesssim 0.796$. This implies that the finite-energy bound does not yield an interesting benchmark in case 3.

The third type of channel that we consider is the thermal-loss channel. An upper bound on the capacity of the thermal-loss channel is

$$-\log_2[(1-\eta)\, \eta^{\overline{n}}] - g\left(\overline{n}\right), \tag{17}$$

if $\overline{n} < \frac{\eta}{1-\eta}$ and zero otherwise [33]. Here, $\overline{n}$ is the average number of thermal photons per channel use [46]. This is an interesting channel because the effect of dark counts can be seen as caused by the thermal photons. Hence this type of channel becomes relevant for case 3, where detectors, and therefore also the dark counts, are regarded as part of the channel. The details of the dark count model are presented in Appendix A. There we also show how to easily convert the experimentally relevant dark count rate of the detector and the duration of the detection window t_{int} into $\overline{n}$ and p_d, the probability of getting a dark count within the given time window.

The combinations of a channel boundary together with a channel type give us a set of benchmarks. Not all combinations yield interesting benchmarks. In Table I, we summarise the benchmarks that we consider.

	Infinite	Finite	Thermal	Direct transmission
Case 1: η_{f}	1a	1b	–	–
Case 2: $\eta_{\mathrm{f}} p_{\mathrm{ps}}$	2a	2b	–	–
Case 3: $\eta_{\mathrm{f}} p_{\mathrm{ps}} p_{\mathrm{app}}$	–	–	3c	3d

TABLE I. Labels of the benchmarks that we use to assess the performance of a quantum repeater. These labels are frequently referred to in the numerical results. Each row corresponds to a different channel boundary, which translates into an effective channel transmissivity. Each column corresponds to a different type of channel: pure loss, pure loss with energy constraint and thermal channel, and the final column corresponds to the direct transmission benchmark.

VI. NUMERICAL RESULTS

In this section, we perform a numerical analysis of our model applied to a specific physical system. In particular, we have chosen a setup based on nitrogen-vacancy centres in cavities kept at cryogenic temperatures. All numerical results have been obtained using a Mathematica notebook [1]. Unless specified otherwise, we use the following parameters that we call "expected parameters":

- a_0 (dephasing due to interaction) $= \frac{1}{2000}$ per attempt [36],
- a_1 (dephasing with time) $= \frac{1}{3}$ per second [28],
- b_0 (depolarisation due to interaction) $= \frac{1}{5000}$ per attempt [36],
- b_1 (depolarisation with time) $= \frac{1}{3}$ per second [28],
- t_{prep} (memory-photon entanglement preparation time) $= 6\,\mu$s [21],
- F_{gm} (depolarising parameter for gates and measurements) $= 0.9$ [12],
- F_{prep} (dephasing parameter for the memory-photon state preparation) $= 0.97$ [21, 44],
- p_{em} (probability of emission) $= 0.49$ [7, 21],
- p_{ps} (post-selection) $= 0.46$ [37],
- p_{det} (detector efficiency) $= 0.8$ [21],
- p_{bsm} (Bell state measurement success probability) $= 1$,
- Dark count rate $= 10$ per second [21],
- t_{int} (detection window) $= 30$ ns [21],
- L_0 (attenuation length) $= 0.542$ km [21].

Before we present the results, we note that the emission frequency of the nitrogen-vacancy centres results in a relatively low L_0 which in turn does not allow to achieve large distances. In practical quantum key distribution networks, this problem might be overcome using the frequency conversion of the emitted photons into the telecom frequency for which $L_0 \approx 21$ km [34]. However, this conversion process is usually probabilistic which results in lower rates. Since the main objective of this paper is the study of the regimes where the various benchmarks could be overcome, the frequency conversion is not included in our numerical results.

Tightness of the error bounds for the secret-key rate. We have derived upper and lower bounds on the yield, and thus also on the secret-key rate, for the two studied protocols.

In FIG. 3, we plot both the upper and the lower bound on the achieved rate with the current and improved parameters ($p_{\text{ps}} = p_{\text{em}} = 0.6$ and $F_{\text{gm}} = 0.97$) and optimised cut-off as a function of the distance in units of L_0. There are two regimes visible on the plot. This is a consequence of the fact that our bounds have a different analytical form in the two regimes (see Appendix D).

Since for practical purposes our bounds are essentially tight, from now on we will refer to the upper bound as the expected secret-key rate, and will omit the lower bound for the legibility of the plots.

The impact of the cut-off on the secret-key rate. In FIG. 4 we plot the secret-key rate versus the cut-off for different sets of parameters. The repeater is assumed to be positioned half-way between Alice and Bob. We observe a strong dependency of the secret-key rate on the cut-off. In particular, for large cut-off the secret-key rate drops to zero. This is due to the inclusion of rounds where the state has significantly decohered. This implies that the cut-off is essential for generating a key at large distances. Moreover, we observe that the optimal cut-off highly depends on the explored parameter regime.

Optimal positioning of the repeater. The asymmetry of the studied sequential protocol raises the question of whether it is best to position the repeater half-way between Alice and Bob. In fact, in the absence of a cut-off this is not the case [26]. For sufficiently large distances, shifting the repeater towards Bob can increase both the secret-key rate and the distance over which the secret-key rate is non-zero in the presence of dark counts. Specifically, the optimal positioning remains a fixed distance away from Bob independently of the actual total distance. Here, we find that with the cut-off and for the parameters considered this phenomenon disappears. We see in FIG. 5 that the optimal position with the cut-off optimisation appears to be exactly in the middle of Alice and Bob. Nevertheless, we note

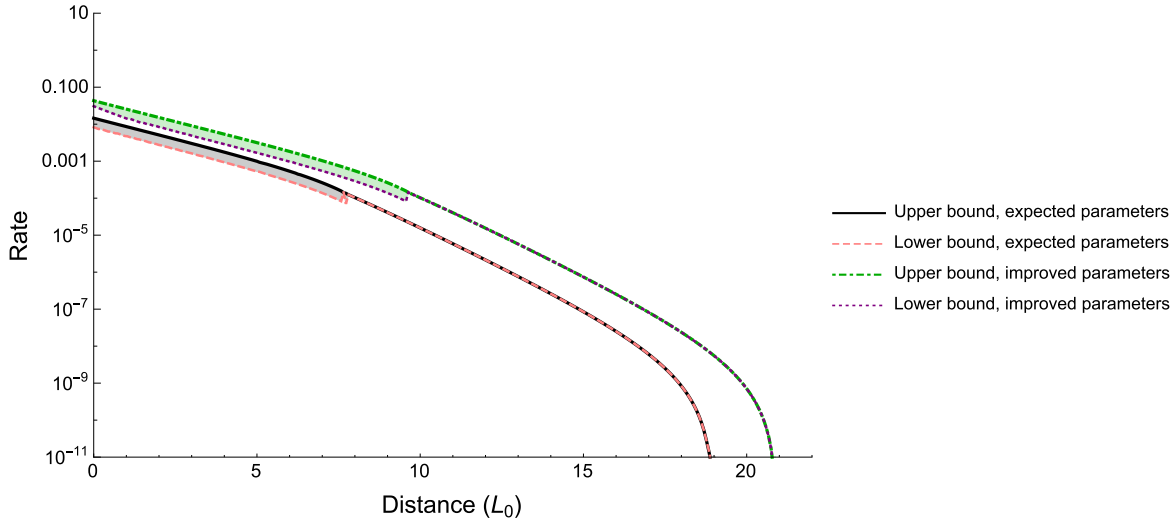

FIG. 3. Upper- and lower bounds on the secret-key rate rate with a quantum repeater as a function of the distance in units of L_0. The repeater is positioned half-way between Alice and Bob. The curves correspond to the expected and improved parameters with optimised cut-off. The improved parameters correspond to setting $p_{\mathrm{ps}} = p_{\mathrm{em}} = 0.6$ and $F_{\mathrm{gm}} = 0.97$. For high losses, the upper- and lower bounds become essentially tight. For this reason, the upper bound on the achieved rate forms a reliable estimate of the secret-key rate.

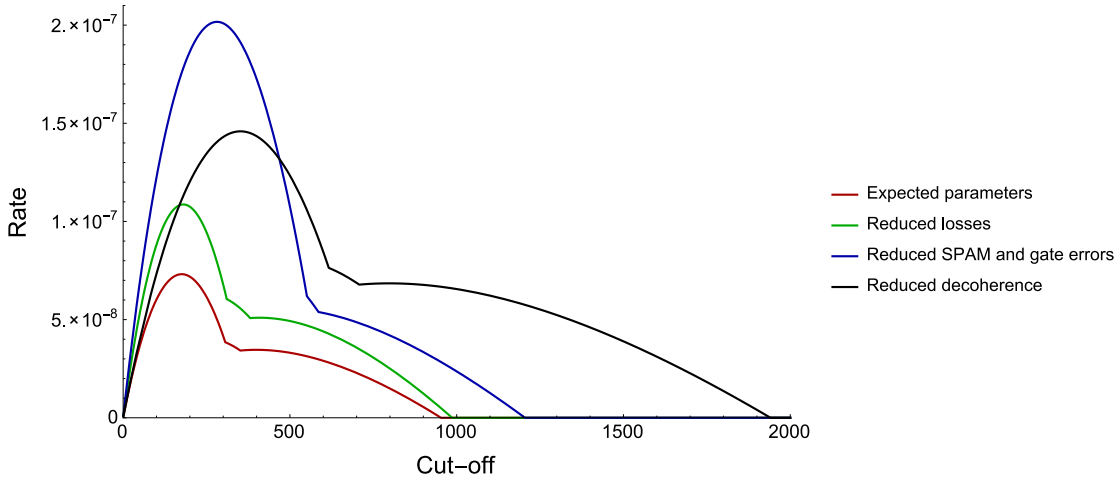

FIG. 4. Secret-key rate as a function of the cut-off for the expected parameters with the repeater positioned half-way between Alice and Bob. The reduced losses are for $p'_{\mathrm{app}} = (p_{\mathrm{app}})^{0.9}$ and $p'_{\mathrm{ps}} = (p_{\mathrm{ps}})^{0.9}$, the reduced SPAM (state preparation and measurement) and gate errors are for $F'_{\mathrm{gm}} = (F_{\mathrm{gm}})^{0.7}$ and $F'_{\mathrm{prep}} = (F_{\mathrm{prep}})^{0.7}$ and the improved coherence is for $a' = a/2$ and $b' = b/2$. The optimal $n^\star$ shifts depending on the parameters. The kinks arise due to the fact that we optimise over two protocols: fully asymmetric BB84 and symmetric six-state protocol with advantage distillation which itself consists of two subprotocols. The optimal protocol depends on the bit error rates. The data have been plotted for the distance of 8.2 km ($\approx 15.1\, L_0$).

that the bounds for the yield derived in Appendix D are valid under the condition $\eta_B \geq \eta_A$. This means that we can only study the effect of moving the repeater towards Bob. However, we do not expect any benefit in shifting the repeater towards Alice as this could only increase the noise due to decoherence. From now on for the scenarios with the cut-off optimisation, we always consider the repeater to be placed half-way between Alice and Bob. Interestingly, in FIG. 5 we also see that the rates for the two scenarios with and without the cut-off start to coincide after the quantum repeater is shifted within a certain distance of Bob. Intuitively this happens when the probability of Bob getting a photon is large enough so that the significance of the cut-off becomes marginal.

Cut-off versus no cut-off. Having established the optimal positioning of the repeater, we can now compare the two scenarios: optimised cut-off with middle positioning of the repeater and no cut-off with optimised positioning. We find

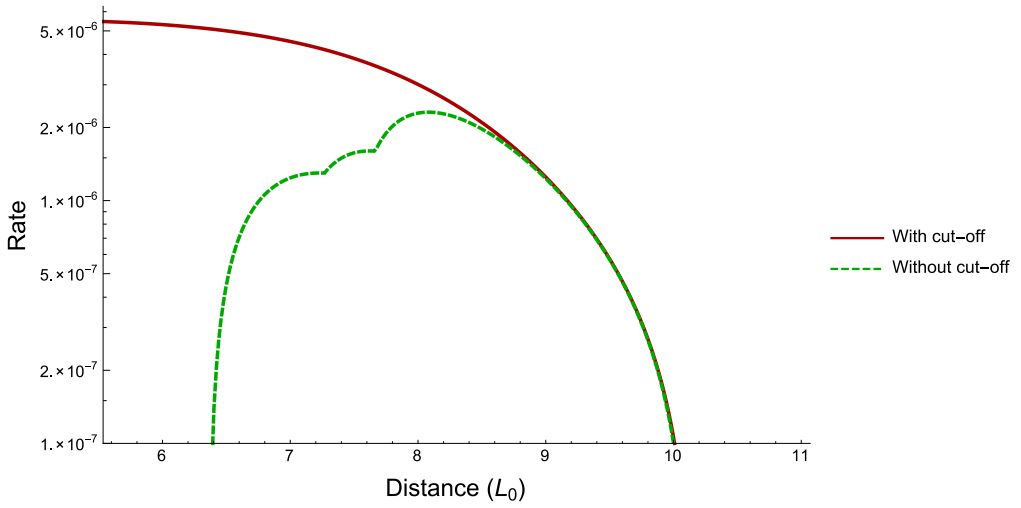

FIG. 5. Secret-key rate with and without the cut-off as a function of the distance in units of L_0 between Alice and quantum repeater. The total distance between Alice and Bob is fixed to 6 kilometres ($\approx 11.1\,L_0$). We see that with the cut-off optimisation, positioning the repeater half-way between Alice and Bob is optimal. This behaviour was also observed for other parameter regimes. This result contrasts with the optimal positioning for the no cut-off scenario, for which we see that shifting the repeater towards Bob is beneficial. We also note that the two rates overlap when the repeater is shifted towards Bob.

that in the absence of dark counts the scaling with distance of both schemes is the same, with a small advantage of the cut-off scheme. However, the cut-off is more robust against dark counts. Hence, for imperfect detectors the cut-off allows distributing keys at larger distances. These results can be seen in FIG. 6 and FIG. 7, which show the secret-key rate as a function of distance for detectors without and with dark counts, together with the channel capacity of the optical fibre (i.e. benchmark 1a). We plot the data for the expected and improved parameters ($p_{\text{ps}} = p_{\text{em}} = 0.6$ and $F_{\text{gm}} = 0.97$).

In FIG. 6 where we assume no dark counts, we see that for small distances the rate scales approximately with the square root of the transmissivity for both scenarios. That is, they are proportional to the theoretical optimum [26] of $\sqrt{\eta_f} = e^{L/2L_0}$. For sufficiently large distances time-dependent decoherence of the memory QM_1 becomes a problem. Both schemes overcome it at the expense of reducing the yield. As a result, the scaling becomes proportional to $\eta_f = e^{L/L_0}$ for both schemes. In FIG. 7 however we see that the presence of dark counts affects the two schemes quite differently. While for both schemes the effect of dark counts becomes the dominant source of noise after a certain distance, this distance is shorter for the no cut-off scheme than for the scheme with the cut-off. In other words, we see that the cut-off is more robust towards dark counts than the repositioning method. This fact can be explained by noting that shifting the repeater towards Bob increases the losses on Alice's side and as a result makes the link Alice-repeater vulnerable to dark counts. With the cut-off however, the repeater remains in the middle making both of the individual links Alice-repeater and repeater-Bob shorter than the Alice-repeater link in the no cut-off scheme. As a result the setup with the cut-off and with the improved parameters allows us to clearly overcome the channel capacity (1a), which is barely surpassed without it.

Comparison with the proposed benchmarks. Let us now investigate the secret key rate achievable with the expected parameters and how it compares with the proposed benchmarks. The comparison is depicted in FIG. 8. The benchmarks corresponding to direct transmission (3d) and the thermal-loss channel (3c) are outperformed. The other benchmarks are not overcome but are within experimental reach.

Parameter trade-off. Let us now give a general overview of how good the improved parameters need to be in order to overcome individual benchmarks. This information is presented on two contour plots. In FIG. 9, we study the parameter regions for which it is possible to beat the benchmarks in Table I as a function of p_{ps} and p_{em}.

A similar plot as a function of F_{gm} and p_{em} can be seen in FIG. 10. We omit here the direct transmission benchmark which, as we have already seen, can be easily surpassed with the expected parameters. Moreover, we note that the capacity of the thermal channel in the benchmark (3c) goes to zero for very low p_{ps} and p_{em} for which it is still possible to generate key with the quantum repeater. Hence it is trivially easy to beat this benchmark for low p_{ps} and p_{em}. In that sense this benchmark is not so interesting in that regime. It is for this reason that this regime is not depicted on the contour plots. In both FIG. 9 and FIG. 10 we observe a crossing between the finite energy benchmarks (1b) and (2b) and their infinite energy counterparts (1a) and (2a) at $p_{\text{em}} \approx 0.796$ as discussed in Section V.

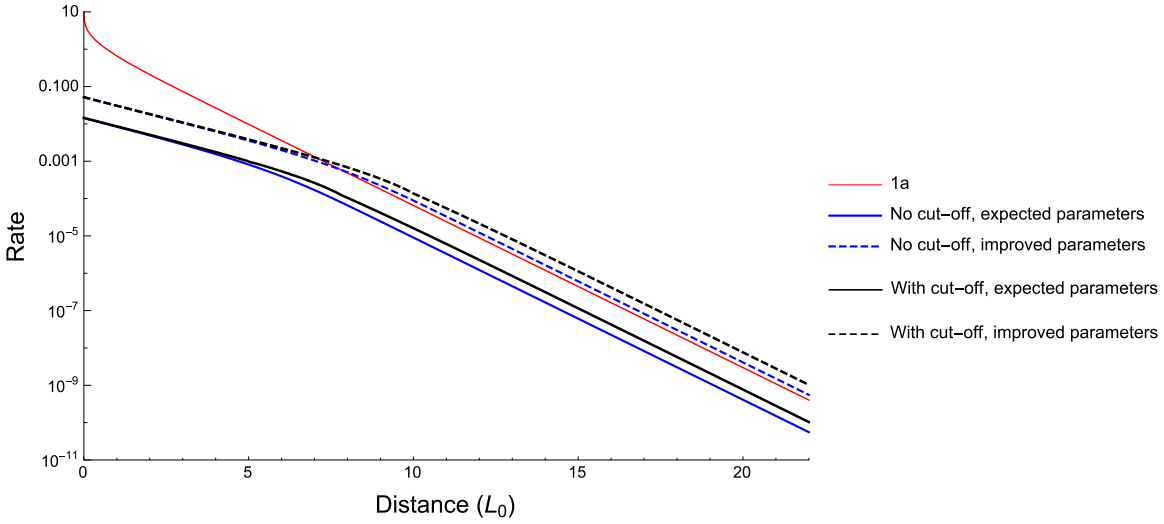

FIG. 6. Secret-key rate as a function of the distance in units of L_0, assuming detectors without dark counts. The black lines correspond to the protocol with cut-off and the blue lines to the protocol without the cut-off but with optimised positioning of the repeater. We plot the data for both the expected and improved parameters. The improved parameters correspond to setting $p_{\mathrm{ps}} = p_{\mathrm{em}} = 0.6$ and $F_{\mathrm{gm}} = 0.97$. Finally, the channel capacity (1a) is also included for comparison. It can be seen that both the cut-off and repositioning of the repeater allows to generate key for all distances.

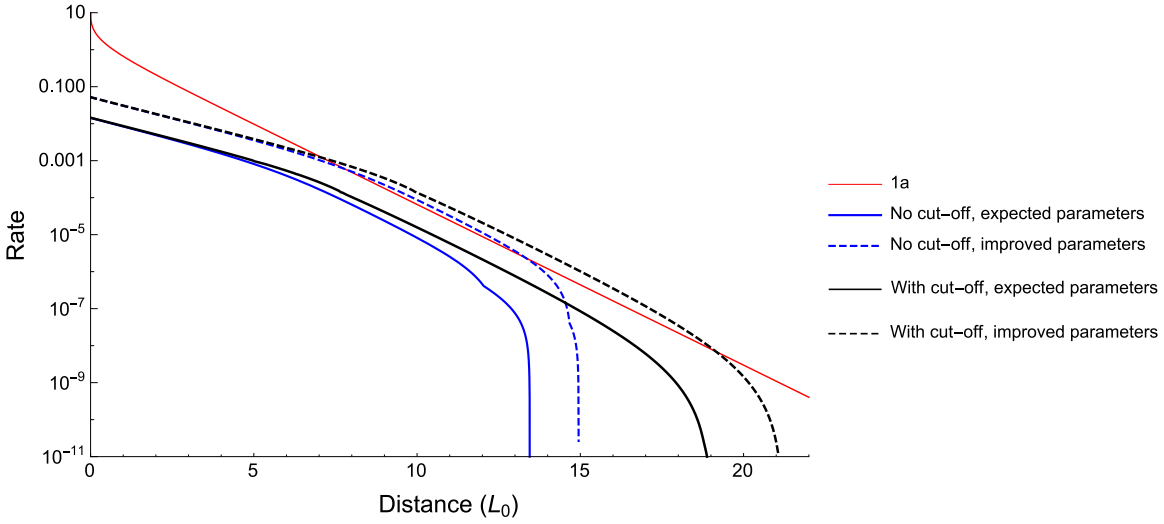

FIG. 7. Secret-key rate as a function of the distance in units of L_0 with dark counts. The black lines correspond to the protocol with cut-off and the blue lines to the protocol without the cut-off but with optimised positioning of the repeater. We plot the data for both the expected and improved parameters. The improved parameters correspond to setting $p_{\mathrm{ps}} = p_{\mathrm{em}} = 0.6$ and $F_{\mathrm{gm}} = 0.97$. Finally, the channel capacity (1a) is also included for comparison. It can be seen that the protocol with the cut-off is more robust against dark counts than the protocol without the cut-off.

VII. CONCLUSIONS

In this work, we have analysed numerically a realistic quantum repeater implementation for quantum key distribution. We have introduced two methods for improving the rates of the repeater with respect to previous proposals: advantage distillation and the cut-off. Advantage distillation is a classical post-processing method that increases the secret-key rate at all levels of noise. The cut-off on the other hand allows for a trade-off between the channel uses required and the secret-key fraction. Utilising the cut-off results in three benefits with respect to the previous scheme for the single sequential quantum repeater [26]. Firstly, the cut-off method achieves a higher rate for all distances. Secondly, the protocol is more robust against dark counts, in the sense that non-zero secret key can be generated over larger distances. Finally, the cut-off can be adjusted on the fly, unlike the repositioning of the repeater [26]. This is

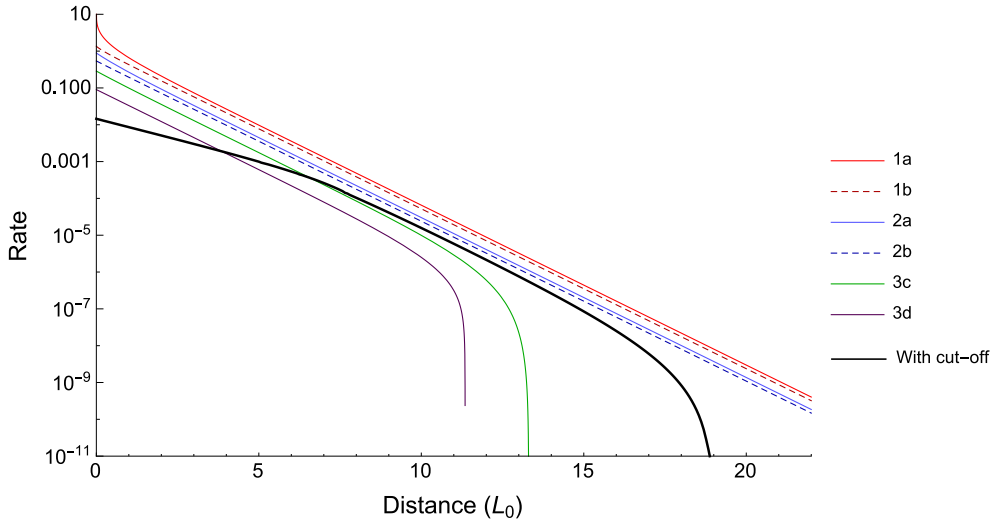

FIG. 8. Secret-key rate with the quantum repeater implementation for the expected parameters with optimised cut-off as a function of the distance in units of L_0. The rate is compared to all the benchmarks defined in Table I.

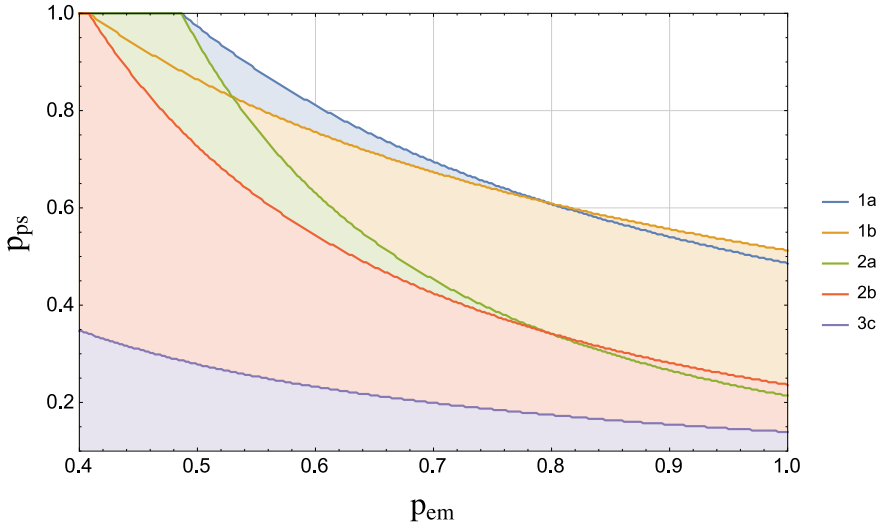

FIG. 9. Contour plot of regions of p_{em} versus p_{ps} with the expected parameters where the benchmarks listed in Table I can be surpassed. The contour lines correspond to the parameters that achieve the corresponding benchmarks while the parameter regimes above the curves allow us to surpass them. The data is plotted for the distance of 5.2 kilometres ($\approx 9.6\, L_0$).

especially convenient in the scenario where the experimental setup might be modified. With the previous scheme for example, improving the coherence times of the memories would lead to a new optimal position. The repositioning of the repeater node would be both costly and time-inefficient, while modifying the cut-off corresponds to a simple change in the programming of the devices.

By optimising over the cut-off, we have found realistic parameter regions where it is possible to surpass several different benchmarks including the secret-key capacity. These benchmarks are relevant milestones towards claiming a quantum repeater, and thus form an important step in the creation of the first large-scale quantum networks. To make our arguments concrete, we have chosen a specific parameter set induced by some recent experimental results. However, other platforms or technological advances might allow to improve upon our results and predict particularly simple setups for performing the first quantum repeater experiment. We leave the investigation of other parameter regimes open. In this respect our model has a very broad functionality, as it allows us to perform efficient optimisation of the secret-key rate over the cut-off for any set of parameters. We achieve this functionality by finding tight analytical bounds for the number of channel uses needed to generate one bit of raw key as a function of the cut-off. Our numerical package is freely available for further exploration [1].

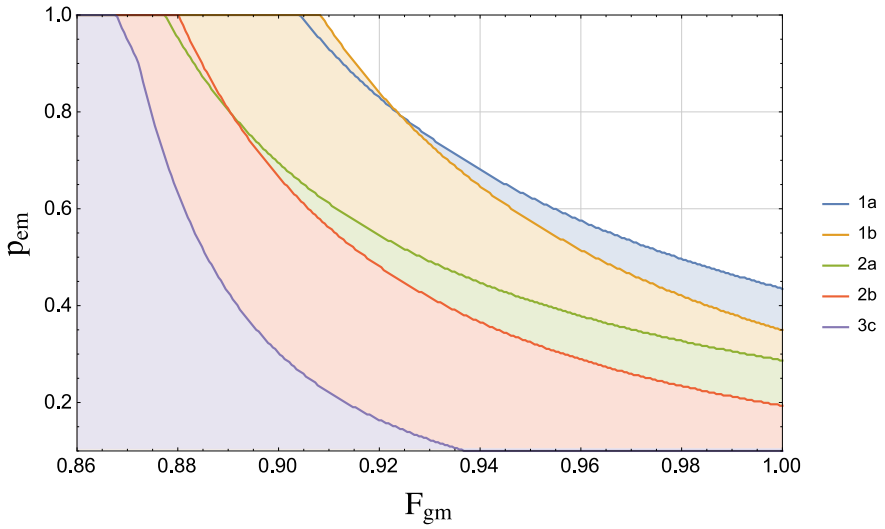

FIG. 10. Contour plot of regions of F_{gm} versus p_{em} with the expected parameters where the benchmarks listed in Table I can be surpassed. The contour lines correspond to the parameters that achieve the corresponding benchmarks while the parameter regimes above the curves allow us to surpass them. The data is plotted for the distance of 5.2 kilometres ($\approx 9.6\, L_0$).

VIII. ACKNOWLEDGEMENTS

The authors would like to thank Suzanne van Dam, Peter Humphreys, Thinh Le Phuc and Mark Steudtner for helpful discussions and feedback, and Dmytro Vasylyev for the illustrations of Alice and Bob. This work was supported by the Dutch Organization for Fundamental Research on Matter (FOM), Dutch Technology Foundation (STW), the Netherlands Organization for Scientific Research (NWO) through a VICI grant (RH), a VIDI grant (SW) and the European Research Council through a Starting Grant (RH and SW).

[1] Available on demand.
[2] Ballester, M. A., Wehner, S., and Winter, A. State discrimination with post-measurement information. *IEEE Transactions on Information Theory*, 54(9):4183–4198, 2008.
[3] Bardhan, B. R. and Wilde, M. M. Strong converse rates for classical communication over thermal and additive noise bosonic channels. *Physical Review A*, 89(2):022302, 2014.
[4] Beaudry, N. J., Moroder, T., and Lütkenhaus, N. Squashing models for optical measurements in quantum communication. *Physical review letters*, 101(9):093601, 2008.
[5] Bennett, C. H. and Brassard, G. Quantum cryptography: Public key distribution and coin tossing. In *International Conference on Computer System and Signal Processing, IEEE, 1984*, pages 175–179, 1984.
[6] Blok, M., Kalb, N., Reiserer, A., Taminiau, T., and Hanson, R. Towards quantum networks of single spins: analysis of a quantum memory with an optical interface in diamond. *Faraday discussions*, 184:173–182, 2015.
[7] Bogdanovic, S., van Dam, S. B., Bonato, C., Coenen, L. C., Zwerver, A., Hensen, B., Liddy, M. S., Fink, T., Reiserer, A., Loncar, M., and Hanson, R. Design and low-temperature characterization of a tunable microcavity for diamond-based quantum networks. *Applied Physics Letters*, 110(17):171103, 2017.
[8] Briegel, H.-J., Dür, W., Cirac, J. I., and Zoller, P. Quantum repeaters: The role of imperfect local operations in quantum communication. *Physical Review Letters*, 81(26):5932, 1998.
[9] Bruß, D. Optimal eavesdropping in quantum cryptography with six states. *Physical Review Letters*, 81(14):3018, 1998.
[10] Christandl, M. and Müller-Hermes, A. Relative entropy bounds on quantum, private and repeater capacities. *arXiv preprint arXiv:1604.03448*, 2016.
[11] Christandl, M. and Wehner, S. Quantum anonymous transmissions. In *International Conference on the Theory and Application of Cryptology and Information Security*, pages 217–235. Springer, 2005.
[12] Cramer, J., Kalb, N., Rol, M. A., Hensen, B., Blok, M. S., Markham, M., Twitchen, D. J., Hanson, R., and Taminiau, T. H. Repeated quantum error correction on a continuously encoded qubit by real-time feedback. *Nature communications*, 7, 2016.

[13] De Lange, G., Wang, Z., Riste, D., Dobrovitski, V., and Hanson, R. Universal dynamical decoupling of a single solid-state spin from a spin bath. *Science*, 330(6000):60–63, 2010.

[14] Ekert, A. K. Quantum cryptography based on Bell's theorem. *Physical review letters*, 67(6):661, 1991.

[15] Gao, W., Imamoglu, A., Bernien, H., and Hanson, R. Coherent manipulation, measurement and entanglement of individual solid-state spins using optical fields. *Nature Photonics*, 9(6):363–373, 2015.

[16] Giovannetti, V., Lloyd, S., and Maccone, L. Quantum-enhanced positioning and clock synchronization. *Nature*, 412(6845): 417–419, 2001.

[17] Gittsovich, O., Beaudry, N. J., Narasimhachar, V., Alvarez, R. R., Moroder, T., and Lütkenhaus, N. Squashing model for detectors and applications to quantum-key-distribution protocols. *Physical Review A*, 89(1):012325, 2014.

[18] Goodenough, K., Elkouss, D., and Wehner, S. Assessing the performance of quantum repeaters for all phase-insensitive gaussian bosonic channels. *New Journal of Physics*, 18(6):063005, 2016.

[19] Gottesman, D. and Lo, H.-K. Proof of security of quantum key distribution with two-way classical communications. *IEEE Transactions on Information Theory*, 49(2):457–475, 2003.

[20] Guha, S., Krovi, H., Fuchs, C. A., Dutton, Z., Slater, J. A., Simon, C., and Tittel, W. Rate-loss analysis of an efficient quantum repeater architecture. *Physical Review A*, 92(2):022357, 2015.

[21] Hensen, B., Bernien, H., Dréau, A., Reiserer, A., Kalb, N., Blok, M., Ruitenberg, J., Vermeulen, R., Schouten, R., Abellán, C., et al. Loophole-free Bell inequality violation using electron spins separated by 1.3 kilometres. *Nature*, 526 (7575):682–686, 2015.

[22] Hucul, D., Inlek, I., Vittorini, G., Crocker, C., Debnath, S., Clark, S., and Monroe, C. Modular entanglement of atomic qubits using photons and phonons. *Nature Physics*, 11(1):37–42, 2015.

[23] Khalique, A. and Sanders, B. C. Practical long-distance quantum key distribution through concatenated entanglement swapping with parametric down-conversion sources. *JOSA B*, 32(11):2382–2390, 2015.

[24] Krovi, H., Guha, S., Dutton, Z., Slater, J. A., Simon, C., and Tittel, W. Practical quantum repeaters with parametric down-conversion sources. *Applied Physics B*, 122(3):1–8, 2016.

[25] Lo, H.-K., Chau, H. F., and Ardehali, M. Efficient quantum key distribution scheme and a proof of its unconditional security. *Journal of Cryptology*, 18(2):133–165, 2005.

[26] Luong, D., Jiang, L., Kim, J., and Lütkenhaus, N. Overcoming lossy channel bounds using a single quantum repeater node. *Applied Physics B*, 122(4):1–10, 2016.

[27] Lvovsky, A. I., Sanders, B. C., and Tittel, W. Optical quantum memory. *Nature photonics*, 3(12):706–714, 2009.

[28] Maurer, P. C., Kucsko, G., Latta, C., Jiang, L., Yao, N. Y., Bennett, S. D., Pastawski, F., Hunger, D., Chisholm, N., Markham, M., et al. Room-temperature quantum bit memory exceeding one second. *Science*, 336(6086):1283–1286, 2012.

[29] Munro, W. J., Azuma, K., Tamaki, K., and Nemoto, K. Inside quantum repeaters. *Selected Topics in Quantum Electronics, IEEE Journal of*, 21(3):1–13, 2015.

[30] Panayi, C., Razavi, M., Ma, X., and Lütkenhaus, N. Memory-assisted measurement-device-independent quantum key distribution. *New Journal of Physics*, 16(4):043005, 2014.

[31] Pant, M., Krovi, H., Englund, D., and Guha, S. Rate-distance tradeoff and resource costs for all-optical quantum repeaters. *Physical Review A*, 95(1):012304, 2017.

[32] Pirandola, S. and Laurenza, R. General benchmarks for quantum repeaters. *arXiv preprint arXiv:1512.04945*, 2015.

[33] Pirandola, S., Laurenza, R., Ottaviani, C., and Banchi, L. Fundamental limits of repeaterless quantum communications. *Nature Communications*, 8:15043 EP –, 04 2017. URL `http://dx.doi.org/10.1038/ncomms15043`.

[34] Radnaev, A., Dudin, Y., Zhao, R., Jen, H., Jenkins, S., Kuzmich, A., and Kennedy, T. A quantum memory with telecom-wavelength conversion. *Nature Physics*, 6(11):894–899, 2010.

[35] Reiserer, A. and Rempe, G. Cavity-based quantum networks with single atoms and optical photons. *Reviews of Modern Physics*, 87(4):1379, 2015.

[36] Reiserer, A., Kalb, N., Blok, M. S., van Bemmelen, K. J., Taminiau, T. H., Hanson, R., Twitchen, D. J., and Markham, M. Robust quantum-network memory using decoherence-protected subspaces of nuclear spins. *Physical Review X*, 6(2): 021040, 2016.

[37] Riedel, D., Söllner, I., Shields, B. J., Starosielec, S., Appel, P., Neu, E., Maletinsky, P., and Warburton, R. J. Deterministic enhancement of coherent photon generation from a nitrogen-vacancy center in ultrapure diamond. *arXiv preprint arXiv:1703.00815*, 2017.

[38] Sangouard, N., Simon, C., De Riedmatten, H., and Gisin, N. Quantum repeaters based on atomic ensembles and linear optics. *Reviews of Modern Physics*, 83(1):33, 2011.

[39] Scarani, V., Bechmann-Pasquinucci, H., Cerf, N. J., Dušek, M., Lütkenhaus, N., and Peev, M. The security of practical quantum key distribution. *Reviews of modern physics*, 81(3):1301, 2009.

[40] Specht, H. P., Nölleke, C., Reiserer, A., Uphoff, M., Figueroa, E., Ritter, S., and Rempe, G. A single-atom quantum memory. *Nature*, 473(7346):190–193, 2011.

[41] Takeoka, M., Guha, S., and Wilde, M. M. Fundamental rate-loss tradeoff for optical quantum key distribution. *Nature communications*, 5, 2014.

[42] Takeoka, M., Guha, S., and Wilde, M. M. The squashed entanglement of a quantum channel. *Information Theory, IEEE Transactions on*, 60(8):4987–4998, 2014.

[43] Thiel, C., Böttger, T., and Cone, R. Rare-earth-doped materials for applications in quantum information storage and signal processing. *Journal of luminescence*, 131(3):353–361, 2011.

[44] Togan, E., Chu, Y., Trifonov, A., Jiang, L., Maze, J., Childress, L., Dutt, M. G., Sørensen, A. S., Hemmer, P., Zibrov, A., et al. Quantum entanglement between an optical photon and a solid-state spin qubit. *Nature*, 466(7307):730–734, 2010.

[45] Watanabe, S., Matsumoto, R., Uyematsu, T., and Kawano, Y. Key rate of quantum key distribution with hashed two-way classical communication. *Physical Review A*, 76(3):032312, 2007.
[46] Weedbrook, C., Pirandola, S., Garcia-Patron, R., Cerf, N. J., Ralph, T. C., Shapiro, J. H., and Lloyd, S. Gaussian quantum information. *Reviews of Modern Physics*, 84(2):621, 2012.
[47] Wilde, M. M. *Quantum information theory*. Cambridge University Press, 2013.
[48] Wilde, M. M. and Qi, H. Energy-constrained private and quantum capacities of quantum channels. *arXiv preprint arXiv:1609.01997*, 2016.
[49] Wilde, M. M., Tomamichel, M., and Berta, M. Converse bounds for private communication over quantum channels. *IEEE Transactions on Information Theory*, 63(3):1792–1817, 2017.
[50] Wootters, W. K. and Zurek, W. H. A single quantum cannot be cloned. *Nature*, 299(5886):802–803, 1982.

Appendix A: Dark counts

In this section we detail the effect of dark counts in the detectors of Alice and Bob on our protocol. In particular, we briefly go over the concept of so-called *squashing models* [4, 17], after which we will be able to calculate the induced depolarising noise. We conclude with detailing how dark counts increase the yield.

Quantum states of light are naturally described by operators on an infinite-dimensional Hilbert space. However, a significant number of optical experiments have been performed where the infinite-dimensional states and operations are approximated by a lower dimensional description. An example of this is where the state of light is assumed to lie within a two-dimensional subspace spanned by the vacuum state and a single-photon excitation. Such an approximation is valid in the sense that the theoretical predictions of measurement statistics correspond accurately to those that are observed experimentally.

However, in cryptographic contexts one usually has to make unconditional statements about the information held by a third party. This third party might be malicious and all-powerful, and her measurement statistics are, by definition, unknown. This implies that there is not necessarily a bound on the information held by a malicious third party, despite the fact that the truncation of the Hilbert space is a good approximation for experimental statistics.

Since the theoretical analysis in a infinite-dimensional Hilbert space is difficult, one would prefer to be able to bound the held information by a third party, while at the same time applying a truncation to the finite-dimensional Hilbert space. This can be done if a so-called squashing model exists, which is a way of relating measurements performed on a high-dimensional state to a truncated space. Squashing models exist for both the fully asymmetric BB84 protocol and the symmetric six-state protocol (with only passive measurements), implying that one can, without loss of generality, perform the fully asymmetric BB84 and symmetric (passive) six-state protocol with photons [4, 17]. The squashing model also dictates how multiple clicks in different detectors give rise to noise in the truncated space. In the next section, we discuss how to map the dark counts in the detectors to depolarising noise according to the corresponding squashing model.

The parameters typically used to quantify detectors are the dark counts per second and the detection window t_{int}, which is the duration of the integration period of the detectors. The number of thermal photons $\overline{n}$ relevant for the thermal benchmark is given by t_{int} times the dark counts per second. Assuming a Poisson distribution of the dark counts, it follows that the probability p_d of getting at least a single dark count click within the time window of awaiting the signal photon is given by $p_d = 1 - \exp(-\overline{n}) \approx \overline{n}$ for small $\overline{n}$.

The noise caused by the dark counts at Alice's or Bob's detector can then be modelled by a depolarising channel, where the depolarising parameter $\alpha_{A/B}$ depends on the protocol implemented,

$$\alpha_{A/B,\,\text{BB84}} = \frac{p_{\text{app}} p_{\text{ps}} \eta_{A/B}(1-p_d)}{1-(1-p_{\text{app}} p_{\text{ps}} \eta_{A/B})(1-p_d)^2} \ , \tag{A1}$$

$$\alpha_{A/B,\,\text{six-state}} = \frac{p_{\text{app}} p_{\text{ps}} \eta_{A/B}(1-p_d)^5}{1-(1-p_{\text{app}} p_{\text{ps}} \eta_{A/B})(1-p_d)^6} \ . \tag{A2}$$

That is, conditioned on a click in at least one of the detectors, Alice or Bob receive the desired state if they receive the signal photon and no other detector was triggered. Due to the squashing map all other events can be mapped onto a maximally mixed state [4, 17]. To explain the exponents, we note that the active BB84 protocol requires an optical measurement setup with two detectors, while for the six-state protocol such a measurement setup will consist of six detectors.

Furthermore, independent of the existence of a squashing map, the dark counts increase the total probability that Alice or Bob gets a click. This probability depends on whether the BB84 or six-state protocol is implemented, and is

given by

$$p_{A/B,\ BB84} = 1 - (1 - p_{\text{app}} p_{\text{ps}} \eta_{A/B})(1 - p_d)^2 \ , \tag{A3}$$

$$p_{A/B,\ \text{six-state}} = 1 - (1 - p_{\text{app}} p_{\text{ps}} \eta_{A/B})(1 - p_d)^6 \ . \tag{A4}$$

Appendix B: Quantum bit error rate

In this Appendix we derive the expressions for the average quantum bit error rate as a function of the experimental parameters, such as the losses. These are given by

$$\langle e_X \rangle = \langle e_Y \rangle = \frac{1}{2} - \frac{1}{2} F_{\text{gm}} \alpha_A \alpha_B \left(2F_{\text{prep}} - 1\right)^2 \left\langle e^{-(a+b)n} \right\rangle \ , \tag{B1}$$

$$\langle e_Z \rangle = \frac{1}{2} - \frac{1}{2} F_{\text{gm}} \alpha_A \alpha_B \langle e^{-b \cdot n} \rangle \ . \tag{B2}$$

where the average is performed over the geometric distribution with only the first $n^\star$ trials. That is, the average of the exponential e^{-cn} is given by

$$\begin{aligned} \langle e^{-cn} \rangle &= \frac{\sum_{n=1}^{n^\star} p_B \left(1 - p_B\right)^{n-1} e^{-cn}}{\sum_{n=1}^{n^\star} p_B \left(1 - p_B\right)^{n-1}} \\ &= \frac{p_B e^{-c}}{1 - (1-p)^{n^\star}} \frac{1 - (1 - p_B)^{n^\star} e^{-cn^\star}}{1 - (1 - p_B) e^{-c}} \ . \end{aligned} \tag{B3}$$

To derive these quantum bit error rates, let us firstly define the two-qubit bell states as

$$|\psi(\mathsf{x}, \mathsf{z})\rangle = \frac{1}{\sqrt{2}} (|0\rangle |0 + \mathsf{x}\rangle + (-1)^{\mathsf{z}} |1\rangle |1 + \mathsf{x} \ (\text{mod}\, 2)\rangle), \tag{B4}$$

for $\mathsf{x}, \mathsf{z} \in \{0, 1\}$. The noise in the preparation can be modelled as dephasing noise [44]. The initially generated entangled state between the quantum memory and the state of the photon flying to Alice is then

$$\rho_{AR} = F_{\text{prep}} |\psi(1,0)\rangle\langle\psi(1,0)| + (1 - F_{\text{prep}}) |\psi(1,1)\rangle\langle\psi(1,1)| \ , \tag{B5}$$

where F_{prep} is the preparation fidelity of this state. The state in the first quantum memory is now kept stored there. During this time, a second entangled photon-memory is attempted to be generated at the second quantum memory. During these attempts, the state stored in the first quantum memory decoheres through time-dependent dephasing and depolarising noise acting on it. This means that at the time when the second copy is generated, the first copy will have decohered. This second copy will be of the same form as the first one. The decohered first copy is of the form

$$\begin{aligned} \rho'_{AR} = F_{T_1} [F_{\text{prep}} (F_{T_2} |\psi(1,0)\rangle\langle\psi(1,0)| + (1 - F_{T_2}) |\psi(1,1)\rangle\langle\psi(1,1)|) \\ + (1 - F_{\text{prep}}) \left(F_{T_2} |\psi(1,1)\rangle\langle\psi(1,1)| + (1 - F_{T_2}) |\psi(1,0)\rangle\langle\psi(1,0)|\right)] + (1 - F_{T_1}) \frac{\mathbb{I}}{4} \ , \end{aligned} \tag{B6}$$

where F_{T_1}, F_{T_2} are respectively the depolarising and dephasing parameters due to the decoherence processes on the stored state in the first memory. The fidelity decays exponentially with the number of attempts [36] and hence these parameters be written as

$$F_{T_1} = e^{-b \cdot n} \ , \tag{B7}$$

$$F_{T_2} = \frac{1 + e^{-a \cdot n}}{2} \ . \tag{B8}$$

Here n is the number of attempts that have been performed on the second memory to successfully generate the repeater-Bob entanglement and the decay rates a and b are defined in the main text. Hence we can rewrite the state of ρ'_{AR} as

$$\rho'_{AR} = F_{T_1} (F_{\text{deph},AR} |\psi(1,0)\rangle\langle\psi(1,0)| + (1 - F_{\text{deph},AR}) |\psi(1,1)\rangle\langle\psi(1,1)|) + (1 - F_{T_1}) \frac{\mathbb{I}}{4} \ . \tag{B9}$$

where

$$F_{\text{deph},AR} = \frac{1 + (2F_{\text{prep}} - 1)e^{-an}}{2} \ . \tag{B10}$$

The entanglement swapping is performed at the two memories at the repeater node. Since the situation is symmetric for all the four measurement outcomes, without loss of generality we can consider the resulting state on AB as if the repeater measured $|\psi(1,0)\rangle$. If a different Bell state was measured, a Pauli rotation could be used to bring the state to this form. The state that we obtain is

$$\begin{aligned}\rho''_{AB} = F_{T_1}\big(&[F_{\text{deph},AR}F_{\text{prep}} + (1 - F_{\text{deph},AR})(1 - F_{\text{prep}})]\,|\psi(1,0)\rangle\langle\psi(1,0)| \\ &+ [F_{\text{deph},AR}(1 - F_{\text{prep}}) + (1 - F_{\text{deph},AR})F_{\text{prep}}]\,|\psi(1,1)\rangle\langle\psi(1,1)|\big) + (1 - F_{T_1})\,\frac{\mathbb{I}}{4} \ .\end{aligned} \tag{B11}$$

Finally we note that the operations such as Bell state measurements or any other required gates performed on the memories are also noisy. We will model them by the depolarising channel here [12]. The depolarising channel commutes with the dephasing channel. For the two copies of the Bell-diagonal state, it also commutes with the entanglement swapping, in the sense that applying it to one of our memory qubits is mathematically equivalent to applying the same channel to one of the photons flying to Alice or Bob. Hence independently of when exactly in the protocol those gates or measurements on the memories are applied, we can add the resulting depolarisation to the final state shared between Alice and Bob, so that we obtain

$$\begin{aligned}\rho''_{AB} = F_{\text{gm}}\alpha_A\alpha_B F_{T_1}\big(&[F_{\text{deph},AR}F_{\text{prep}} + (1 - F_{\text{deph},AR})(1 - F_{\text{prep}})]\,|\psi(1,0)\rangle\langle\psi(1,0)| \\ &+ [F_{\text{deph},AR}(1 - F_{\text{prep}}) + (1 - F_{\text{deph},AR})F_{\text{prep}}]\,|\psi(1,1)\rangle\langle\psi(1,1)|\big) + (1 - F_{\text{gm}}\alpha_A\alpha_B F_{T_1})\,\frac{\mathbb{I}}{4} \ .\end{aligned} \tag{B12}$$

Here by F_{gm} we denote the product of all the depolarising parameters corresponding to all noisy gates and measurements and $\alpha_{A/B}$ corresponds to the noise caused by the dark counts on Alice's/Bob's side. From the final state it follows that

$$\langle e_X\rangle = \langle e_Y\rangle = \frac{1}{2} - \frac{1}{2}F_{\text{gm}}\alpha_A\alpha_B\,(2F_{\text{prep}} - 1)^2\left\langle e^{-(a+b)n}\right\rangle \ , \tag{B13}$$

$$\langle e_Z\rangle = \frac{1}{2} - \frac{1}{2}F_{\text{gm}}\alpha_A\alpha_B\langle e^{-b\cdot n}\rangle \ . \tag{B14}$$

where the average is over the geometric distribution with only the first $n^\star$ trials. This is due to the fact that, by construction, the state is never allowed to decohere more than $n^\star$ trials.

Appendix C: Secret-key fraction and advantage distillation

In this section the secret-key fraction formula for the six-state protocol with advantage distillation of [45] is briefly reviewed. We note here that while the analysis in Appendix B has the state $|\psi(1,0)\rangle$ as the target state, here we follow the analysis of [45] for which $|\psi(0,0)\rangle$ is the target state. This doesn't affect the overall analysis as the final state from Appendix B can be rotated locally such that $|\psi(0,0)\rangle$ could be made the target state. The secret key fraction can be expressed in terms of the Bell coefficients of the Bell diagonal state

$$\rho_{AB} = \sum_{\mathsf{x},\mathsf{z}\in\{0,1\}} P_{\mathsf{XZ}}(\mathsf{x},\mathsf{z})\,|\psi(\mathsf{x},\mathsf{z})\rangle\,\langle\psi(\mathsf{x},\mathsf{z})| \ . \tag{C1}$$

Here P_{XZ} is a probability distribution and we will abbreviate $P_{\mathsf{XZ}}(\mathsf{x},\mathsf{z})$ as p_{xz}. For the description of the advantage distillation protocol we refer the reader to [45]. It is shown there that the secret-key fraction can be written as

$$r_{\text{six-state}} = \frac{1}{3}\max\left[1 - H(P_{\mathsf{XZ}}) + \frac{P_{\tilde{\mathsf{X}}}(1)}{2}h\left(\frac{p_{00}p_{10} + p_{01}p_{11}}{(p_{00}+p_{01})(p_{10}+p_{11})}\right), \frac{P_{\tilde{\mathsf{X}}}(0)}{2}(1 - H(P'_{\mathsf{XZ}}))\right], \tag{C2}$$

where

$$P_{\tilde{\mathsf{X}}}(0) = (p_{00}+p_{01})^2 + (p_{10}+p_{11})^2 \ , \tag{C3}$$

$$P_{\tilde{\mathsf{X}}}(1) = 2(p_{00}+p_{01})(p_{10}+p_{11}) \ , \tag{C4}$$

and

$$P'_{\mathsf{XZ}}(0,0) = \frac{p_{00}^2 + p_{01}^2}{(p_{00}+p_{01})^2 + (p_{10}+p_{11})^2}, \tag{C5}$$

$$P'_{\mathsf{XZ}}(1,0) = \frac{2p_{00}p_{01}}{(p_{00}+p_{01})^2 + (p_{10}+p_{11})^2}, \tag{C6}$$

$$P'_{\mathsf{XZ}}(0,1) = \frac{p_{10}^2 + p_{11}^2}{(p_{00}+p_{01})^2 + (p_{10}+p_{11})^2}, \tag{C7}$$

$$P'_{\mathsf{XZ}}(1,1) = \frac{2p_{10}p_{11}}{(p_{00}+p_{01})^2 + (p_{10}+p_{11})^2} \ . \tag{C8}$$

and $H(P_{\mathsf{XZ}})$ is the Shannon entropy of the distribution P_{XZ}. The factor of a third arises from the fact that for a symmetric six-state protocol only a third of the measurements will be performed in the same basis by Alice and Bob.

In our model we only consider depolarising noise and dephasing noise in standard basis. Hence for the six-state protocol the error rates in X and Y basis will be the same. Therefore

$$p_{10} + p_{11} = e_Z \ , \tag{C9}$$

$$p_{01} + p_{11} = e_{XY} \ , \tag{C10}$$

$$p_{01} + p_{10} = e_{XY} \ , \tag{C11}$$

$$p_{00} + p_{01} + p_{10} + p_{11} = 1 \ . \tag{C12}$$

Hence

$$p_{00} = 1 - \frac{e_Z}{2} - e_{XY} \ , \tag{C13}$$

$$p_{01} = e_{XY} - \frac{e_Z}{2} \ , \tag{C14}$$

$$p_{10} = p_{11} = \frac{e_Z}{2} \ . \tag{C15}$$

And so

$$P_{\bar{\mathsf{X}}}(0) = 1 - 2e_Z + 2e_Z^2 \ , \tag{C16}$$

$$P_{\bar{\mathsf{X}}}(1) = 2(1-e_Z)e_Z \ . \tag{C17}$$

Appendix D: Yield

In this Appendix we derive the analytical approximation for the yield with the cut-off $n^\star$. The yield Y is given by

$$Y = \frac{p_{\text{bsm}}}{\mathbb{E}\left[N\right]} = \frac{p_{\text{bsm}}}{\mathbb{E}\left[\max(N_A, N_B)\right]} \ . \tag{D1}$$

The approximation used for $\mathbb{E}\left[\max(N_A, N_B)\right]$ is

$$\mathbb{E}\left[\max(N_A, N_B)\right] \approx \begin{cases} \frac{1}{p_A\left(1-(1-p_B)^{n^\star}\right)} & \frac{1}{p_A} \geq n^\star \\ \frac{1}{p_A} + \frac{1}{p_B} - \frac{1}{p_A + p_B - p_A p_B} & \frac{1}{p_A} < n^\star, \end{cases} \tag{D2}$$

where p_A and p_B are defined in Eq. (A3) for BB84 and in Eq. (A4) for the six-state protocol. In the rest of this Appendix, we will motivate this approximation by finding tight analytical lower and upper bounds on the expression above.

We note that we consider separately two parameter regimes. One of them is the regime where on average the dominant number of channel uses per round is on Alice's side $\left(\frac{1}{p_A} > n^\star\right)$. This corresponds to the high-loss regime since the number of channel uses per round on Bob's side is upper bounded by the cut-off. The other regime is the low-loss regime $\left(\frac{1}{p_A} \leq n^\star\right)$. In this regime we will show that the cut-off does not play any significant role, so that in this regime the formula for the yield with no cut-off [26, 30] can be used. Moreover, for our derivation to be valid we require an additional constraint to be satisfied, namely $p_B \geq p_A$. This means that we cannot consider scenarios when the repeater is positioned closer to Alice than to Bob. Such a constraint is well-justified since the time-dependent decoherence in quantum memory QM_1 would only increase by shifting the repeater towards Alice.

High-loss regime

The high-loss regime is the regime where the losses on Alice's side together with the cut-off on Bob's side ensure that the predominant number of channel uses is almost always on Alice's side, i.e. $\mathbb{E}[N] = \mathbb{E}[\max(N_A, N_B)] \approx \mathbb{E}[N_A]$. This regime is described by the condition $p_A n^\star < 1$. More specifically, as we will show in this section, if

$$\frac{1}{p_A} := \mu = \beta n^\star, \ \beta > 1 \ , \tag{D3}$$

then

$$\mathbb{E}[N_A] \leq \mathbb{E}[N] \leq (g_{\text{err}}(p_A, p_B, n^\star) + 1)\, \mathbb{E}[N_A] \ , \tag{D4}$$

where $\mathbb{E}[N_A] = \frac{1}{p_A(1-(1-p_B)^{n^\star})}$ (see Eq. (D12)) and $g_{\text{err}}(p_A, p_B, n^\star) = \mathcal{O}\left(\frac{1}{\beta^2}\right)$ is a function defined in Eq. (D31). This implies that for β large enough, $\mathbb{E}[N]$ can be accurately approximated by $\frac{1}{p_A(1-(1-p_B)^{n^\star})}$.

We start the proof of equation (D4) by first noticing that $\mathbb{E}[N_A] \leq \mathbb{E}[N]$. It is, thus, only necessary to find an upper bound for $\mathbb{E}[N]$. Now, let $p(K = k) = (1 - p_r)^{k-1} p_r$ be the probability that Bob succeeds in round k. Here $p_r = 1 - (1 - p_B)^{n^\star}$ is the probability that Bob succeeds in a given round. Then

$$\mathbb{E}[N] = \mathbb{E}[\max(N_A, N_B)] = \sum_{k=1}^{\infty} p(K = k) \left(\sum_{n_A = k}^{\infty} \left(\sum_{n_B = (k-1)n^\star + 1}^{kn^\star} p(N_A = n_A \wedge N_B = n_B | K = k) \max(n_A, n_B) \right) \right) . \tag{D5}$$

One can split the sum over n_A in two, depending on whether n_A is greater than n_B or vice versa. We get

$$\mathbb{E}[N] = \sum_{k=1}^{\infty} p(k) \left(\sum_{n_B = (k-1)n^\star + 1}^{kn^\star} \left(\sum_{n_A = k}^{n_B} p(n_A \wedge n_B | k) n_B \right) + \sum_{n_B = (k-1)n^\star + 1}^{kn^\star} \left(\sum_{n_A = n_B + 1}^{\infty} p(n_A \wedge n_B | k) n_A \right) \right) , \tag{D6}$$

where $p(k) = p(K = k)$, and $p(n_A \wedge n_B | k) = p(N_A = n_A \wedge N_B = n_B | K = k)$. The first term of Eq. (D6) can be upper bounded noticing that $n_B \leq kn^\star$, i.e.

$$\sum_{k=1}^{\infty} p(k) \left(\sum_{n_B = (k-1)n^\star + 1}^{kn^\star} \left(\sum_{n_A = k}^{n_B} p(n_A \wedge n_B | k) n_B \right) \right) \leq \sum_{k=1}^{\infty} p(k) p\left(N_A \leq N_B | K = k\right) kn^\star. \tag{D7}$$

The second term of Eq. (D6) can be upper bounded in the following way

$$\sum_{k=1}^{\infty} p(k) \left(\sum_{n_B = (k-1)n^\star + 1}^{kn^\star} \left(\sum_{n_A = n_B + 1}^{\infty} p(n_A \wedge n_B | k) n_A \right) \right) \leq \sum_{k=1}^{\infty} p(k) \left(\sum_{n_A = k}^{\infty} p(n_A | k) n_A \right) \tag{D8}$$

$$= \sum_{k=1}^{\infty} p(k) \sum_{n_A = 1}^{\infty} p(n_A | k) n_A \tag{D9}$$

$$= \sum_{n_A = 1}^{\infty} p(n_A) n_A = \mathbb{E}[N_A] \ . \tag{D10}$$

Inputting Eq. (D7) and Eq. (D10) back into Eq. (D6), we obtain

$$\mathbb{E}[N] \leq \left(\frac{n^\star}{\mathbb{E}[N_A]} \sum_{k=1}^{\infty} p(k) p\left(N_A \leq N_B | k\right) k + 1 \right) \mathbb{E}[N_A] \ . \tag{D11}$$

Let N_A^i be the random variable describing the number of trials on Alice's side in round i. Since $p(N_A^i = n_A^i) = (1 - p_A)^{n_A^i - 1} p_A$, we clearly have that $\mathbb{E}[N_A^i] = \frac{1}{p_A} = \mu$. Then we note that

$$\mathbb{E}[N_A] = \sum_{k=1}^{\infty} p(k) \sum_{i=1}^{k} \sum_{n_A^i = 1}^{\infty} p(n_A^i) n_A^i = \sum_{k=1}^{\infty} p(k) \sum_{i=1}^{k} \mathbb{E}[N_A^i] = \mu \sum_{k=1}^{\infty} p(k) k = \mathbb{E}[K]\mu = \frac{1}{p_A p_r} = \frac{1}{p_A(1 - (1 - p_B)^{n^\star})} \ . \tag{D12}$$

Here, we first express $\mathbb{E}[N_A]$ by calculating the average number of trials in each of the k rounds. Then, we sum the k averages together, and finally, we average over the total number of rounds k. Since all the rounds are independent, we replace each $\mathbb{E}[N_A^i]$ by μ as stated above. By inputting Eq. (D12) into Eq. (D11), we get

$$\mathbb{E}\left[N_A\right] \leq \mathbb{E}\left[N\right] \leq \left(\frac{1}{\mathbb{E}[K]\beta}\sum_{k=1}^{\infty} p(k)p\left(N_A \leq N_B|k\right)k + 1\right)\mathbb{E}[N_A] \ . \tag{D13}$$

We now upper bound the $p\left(N_A \leq N_B|k\right)$ term. Note that

$$p\left(N_A \leq N_B|k\right) = p\left(\sum_{i=1}^{k} N_A^i \leq \sum_{i=1}^{k} N_B^i \Big| k\right) \ . \tag{D14}$$

We note that conditioned on $K = k$, we have that $\sum_{i=1}^{k} N_B^i = (k-1)n^\star + N_B^k$. It then follows that

$$p\left(N_A \leq N_B|k\right) = p\left(\sum_{i=1}^{k} N_A^i \leq (k-1)n^\star + N_B^k \Big| k\right) \leq p\left(\sum_{i=1}^{k} N_A^i \leq kn^\star \Big| k\right) \ . \tag{D15}$$

Condition (D3) and $-\sum_{i=1}^{k} N_A^i \geq -kn^\star$ is equivalent to $k\mu - \sum_{i=1}^{k} N_A^i \geq k(\beta-1)n^\star$. Hence,

$$p\left(\sum_{i=1}^{k} N_A^i \leq kn^\star \Big| k\right) = p\left(k\mu - \sum_{i=1}^{k} N_A^i \geq k(\beta-1)n^\star \Big| k\right) \ . \tag{D16}$$

We can use the Chernoff bound to upper bound this probability. The Chernoff bound for a random variable X is

$$p(X \geq a) \leq \frac{\mathbb{E}[e^{tX}]}{e^{ta}}, \ t > 0 \ . \tag{D17}$$

Let X be the sum of k random variables $X_1, X_2, \ldots, X_k$, where

$$X_i = \mu - N_A^i \ , \tag{D18}$$

i.e. $X = \sum_{i=1}^{k} X_i = k\mu - \sum_{i=1}^{k} N_A^i$. From this we can now bound the desired probability using (D17) and $a = k(\beta-1)n^\star$, we have the inequality

$$p\left(k\mu - \sum_{i=1}^{k} N_A^i \geq k(\beta-1)n^\star \Big| k\right) \leq \frac{\mathbb{E}\left[\exp\left(t\left(k\mu - \sum_{i=1}^{k} N_A^i\right)\right)\Big| k\right]}{e^{tk(\beta-1)n^\star}} \tag{D19}$$

$$= \exp\left[tk\left(\mu - (\beta-1)n^\star\right)\right]\mathbb{E}\left[\Pi_{i=1}^{k} e^{-tN_A^i}|k\right] \ . \tag{D20}$$

Let us now focus on $\mathbb{E}\left[\prod_{i=1}^{k} e^{-tN_A^i}|k\right]$,

$$\mathbb{E}\left[\prod_{i=1}^{k} e^{-tN_A^i}|k\right] = \prod_{i=1}^{k}\mathbb{E}\left[e^{-tN_A^i}|k\right] = \prod_{i=1}^{k}\left(\sum_{n_A^i=1}^{\infty} p_A(1-p_A)^{n_A^i-1}\, e^{-tn_A^i}\right) = \left(\frac{p_A e^{-t}}{1-(1-p_A)e^{-t}}\right)^k \ . \tag{D21}$$

Here, after the first equality sign we have used the fact that the random variables N_A^i are independent for different i's. After the second equality we note that all of them have exactly the same geometric distribution over the k rounds. Specifically, it is now important to note that this holds provided that k is the value of K on which we have conditioned, i.e., the success on Bob's side occurs exactly in the k'th round. Furthermore, the common ratio $(1-p_A)e^{-t}$ satisfies the convergence condition $|(1-p_A)e^{-t}| < 1$ for all $t > 0$. This yields

$$p\left(N_A \leq N_B|K = k\right) \leq \left(\exp\left[t\left(\frac{1}{p_A} - (\beta-1)n^\star\right)\right]\frac{p_A e^{-t}}{1-(1-p_A)e^{-t}}\right)^k \ . \tag{D22}$$

Let's define the function $f(t)$ as

$$f(t) := \exp\left[t\left(\frac{1}{p_A} - (\beta-1)n^\star\right)\right]\frac{p_A e^{-t}}{1-(1-p_A)e^{-t}} \ . \tag{D23}$$

This function should be minimised subject to $t > 0$ to obtain the tightest bound. A single stationary point is analytically found at

$$t_0 = \ln\left(\frac{(1-p_A)(p_A(\beta-1)n^\star - 1)}{p_A(\beta-1)n^\star + p_A - 1}\right) . \tag{D24}$$

We now want to make sure that t_0 always satisfies the condition $t > 0$, necessary for applying the Chernoff bound. By condition (D3), the denominator of the above expression inside the logarithm is $p_A(\beta-1)n^\star + p_A - 1 = 1 - p_A n^\star + p_A - 1 = p_A(1-n^\star) < 0$ as long as $n^\star > 1$. From this it follows that $t_0 > 0$ if and only if

$$(1-p_A)(p_A(\beta-1)n^\star - 1) < p_A(\beta-1)n^\star + p_A - 1 . \tag{D25}$$

Clearly this condition is equivalent to $-p_A^2(\beta-1)n^\star < 0$ which is satisfied for $\beta > 1$. This means that $t_0 > 0$ is always satisfied. Now note that $f(t=0) = 1$. Moreover, one can also easily verify that $f'(t=0) = n^\star(1-\beta) < 0$ for $\beta > 1$, and that $\lim_{t\to\infty} f(t) \to \infty$ as long as $n^\star > 1$. These properties of $f(t)$, together with the continuity of $f(t)$, prove that $t = t_0$ corresponds to the global minimum of this function in the regime $t > 0$ and that $f(t_0) < 1$. Hence, we can now calculate $f(t_0)$ which gives

$$f(t_0) = \left(\frac{(p_A(\beta-1)n^\star - 1)(1-p_A)}{p_A(\beta-1)n^\star + p_A - 1}\right)^{\frac{1}{p_A} - (\beta-1)n^\star - 1} (1 - p_A(\beta-1)n^\star) . \tag{D26}$$

This formula can be simplified by substituting the condition (D3) to eliminate β

$$f(t_0) = p_A n^\star \left(\frac{n^\star(1-p_A)}{n^\star - 1}\right)^{n^\star - 1} . \tag{D27}$$

$\mathbb{E}[N]$ can now be upper bounded by an expression that depends by $f(t_0)$, that is

$$\mathbb{E}[N] \leq \left(\frac{1}{\mathbb{E}[K]\beta}\sum_{k=1}^{\infty} p(K=k) f(t_0)^k k + 1\right)\mathbb{E}[N_A] . \tag{D28}$$

We can now average over the number of rounds k,

$$\sum_{k=1}^{\infty} \frac{p_r}{(1-p_r)} \left[(1-p_r)f(t_0)\right]^k k = \frac{p_r f(t_0)}{\left[1-(1-p_r)f(t_0)\right]^2} . \tag{D29}$$

Moreover, $\mathbb{E}[K] = \frac{1}{p_r}$ and again removing β through condition (D3) yields to

$$\mathbb{E}[N] \leq \left(\frac{p_r^2 p_A n^\star f(t_0)}{\left[1-(1-p_r)f(t_0)\right]^2} + 1\right)\mathbb{E}[N_A] = \left(\frac{(1-(1-p_B)^{n^\star})^2 p_A n^\star f(t_0)}{\left[1-(1-p_B)^{n^\star} f(t_0)\right]^2} + 1\right)\mathbb{E}[N_A] . \tag{D30}$$

Now by taking the number of channel uses to be $\mathbb{E}[N_A]$, we can define the relative error $g_{\text{err}}(p_A, p_B, n^\star)$,

$$g_{\text{err}}(p_A, p_B, n^\star) := \frac{(1-(1-p_B)^{n^\star})^2 p_A n^\star f(t_0)}{\left[1-(1-p_B)^{n^\star} f(t_0)\right]^2} , \tag{D31}$$

with $f(t_0)$ given in Eq. (D27), so that

$$\mathbb{E}[N_A] \leq \mathbb{E}[N] \leq (g_{\text{err}}(p_A, p_B, n^\star) + 1)\,\mathbb{E}[N_A] , \tag{D32}$$

where the conditions required to satisfy the above formula are $n^\star > 1$ and $p_A n^\star < 1$. Finally, we can now show how $g_{\text{err}}(p_A, p_B, n^\star)$ scales with β. Note that

$$f(t_0) \leq p_A n^\star \left(1 + \frac{1}{n^\star - 1}\right)^{n^\star - 1} \leq p_A n^\star e . \tag{D33}$$

This together with $f(t_0) < 1$ gives

$$g_{\text{err}}(p_A, p_B, n^\star) < \frac{p_r^2 (p_A n^\star)^2 e}{p_r^2} = \frac{e}{\beta^2} . \tag{D34}$$

Therefore $g_{\text{err}}(p_A, p_B, n^\star) = \mathcal{O}\left(\frac{1}{\beta^2}\right)$, implying that the bounds in the high-loss regime are good enough to tightly bound the achieved yield.

Low-loss regime

Now we consider the complementary low-loss regime characterised by the condition $p_A n^\star \geq 1$. Firstly, since in our protocol there is never any benefit in placing the repeater closer to Alice than to Bob, we also have that $p_B \geq p_A$. This implies that $\frac{1}{p_B} \leq \frac{1}{p_A} = \mathbb{E}[N_A^i] \leq n^\star$. This is the regime where the cut-off is large in comparison with the average number of channel uses required to detect a single photon on Bob's side. That is,

$$\frac{\beta'}{p_B} = n^\star,\ n^\star \geq \beta' \geq 1\ . \tag{D35}$$

As we will show in this section, in this region we can approximate $\mathbb{E}[N] = \mathbb{E}[\max(N_A, N_B)]$ by N_{NC}, where

$$N_{NC} = \frac{1}{p_A} + \frac{1}{p_B} - \frac{1}{p_A + p_B - p_A p_B}\ , \tag{D36}$$

is the average number of channel uses in the no cut-off (NC) scenario [26, 30]. Intuitively, this is because Alice and Bob almost never have to restart due to Bob reaching the cut-off. More specifically, we show that

$$N_{NC} \leq \mathbb{E}[N] \leq (\tilde{g}_{\text{err}}(p_A, p_B, n^\star) + 1)\, N_{NC}\ , \tag{D37}$$

where $\tilde{g}_{\text{err}}(p_A, p_B, n^\star)$ is defined in (D48). Since $\tilde{g}_{\text{err}}(p_A, p_B, n^\star) = \mathcal{O}\left(\beta' e^{-\beta'}\right)$, for sufficiently large β' the expectation value $\mathbb{E}[N]$ can be accurately approximated by N_{NC}.

Here we detail a proof of Eq. (D37). We note that the presence of the cut-off increases the number of needed channel uses with respect to the no cut-off scenario, i.e. $N_{NC} \leq \mathbb{E}[N]$. For the upper bound we can write now

$$\mathbb{E}[N] = \mathbb{E}[\max(N_A, N_B)] = \sum_{k=1}^{\infty} p(K=k) \left(\sum_{n_A=k}^{\infty} \left(\sum_{n_B=(k-1)n^\star+1}^{kn^\star} p(n_A \wedge n_B | K=k) \max(n_A, n_B) \right) \right) \tag{D38}$$

$$\begin{aligned} &= p(K=1) \sum_{n_B=1}^{n^\star} \sum_{n_A=1}^{\infty} p(n_A|K=1) p(n_B|K=1) \max(n_A, n_B) \\ &\qquad + \sum_{k=2}^{\infty} p(K=k) \left(\sum_{n_B=(k-1)n^\star+1}^{kn^\star} \left(\sum_{n_A=k}^{\infty} p(n_A \wedge n_B|k) \max(n_A, n_B) \right) \right)\ . \end{aligned} \tag{D39}$$

In (D39) we split the sum over k into two terms, one with $k=1$ and the other with $k>1$. Since the first term has fixed $k=1$, the variables N_A and N_B are independent here (there is only one round in which Bob for sure succeeds, so the value of n_B doesn't affect the value of n_A). Moreover, the geometric distribution of N_B is normalised over the interval $[1, \ldots, n^\star]$.

$$\mathbb{E}[N] \leq p(K=1) N_{NC} + \sum_{k=2}^{\infty} p(K=k) \left(\sum_{n_B=(k-1)n^\star+1}^{kn^\star} \left(\sum_{n_A=k}^{\infty} p(n_A \wedge n_B|k) \max(n_A, kn^\star) \right) \right)\ . \tag{D40}$$

We have upper bounded the first term of Eq. (D39) by upper bounding the sum $\sum_{n_B=1}^{n^\star}$ with $\sum_{n_B=1}^{\infty}$. In this case the expression after $p(K=1)$ in the first term becomes N_{NC}. In the second term we upper bound n_B by $kn^\star$. Since the second term does not depend on n_B anymore we upper bound it by removing the constraints on N_B completely from the probabilities $p(n_A \wedge n_B|K=k)$, i.e.

$$\mathbb{E}[N] \leq p(K=1) N_{NC} + \sum_{k=2}^{\infty} p(K=k) \sum_{n_A=k}^{\infty} p(n_A|K=k) \max(n_A, kn^\star) \tag{D41}$$

$$= p(K=1) N_{NC} + \sum_{k=2}^{\infty} p(K=k) \left(\sum_{n_A=k}^{kn^\star} p(n_A|K=k) kn^\star + \sum_{n_A=kn^\star+1}^{\infty} p(n_A|K=k) n_A \right)\ , \tag{D42}$$

where in the last line of (D42) we split the second term into two terms corresponding to the regime where $kn^\star$ is larger than n_A and vice versa. Since $kn^\star$ does not depend on n_A, we upper bound this term by removing the constraints on n_A,

$$\mathbb{E}\left[N\right] \leq p(K=1)N_{NC} + \sum_{k=2}^{\infty} p(K=k)kn^\star + \sum_{k=2}^{\infty} p(K=k) \sum_{n_A=k}^{\infty} p(n_A|K=k)n_A \ . \tag{D43}$$

Eq. (D43) can be greatly simplified. We can perform the sum over n_A in the third term obtaining $k\mu$. Then the sums over k can also be easily evaluated so that the right hand side of Eq. (D43) can be rewritten as

$$p(K=1)N_{NC} + \sum_{k=2}^{\infty} p(K=k)kn^\star + \sum_{k=2}^{\infty} p(K=k)k\mu = p(K=1)N_{NC} + (n^\star+\mu)(\mathbb{E}(K) - p(K=1)) \tag{D44}$$

$$= \left(p_r + \frac{n^\star+\mu}{N_{NC}}\left(\frac{1}{p_r} - p_r\right)\right) N_{NC} \tag{D45}$$

$$= \left(p_r + \left(\frac{n^\star+\mu}{N_{NC}}\right)\left(\frac{1-p_r^2}{p_r}\right)\right) N_{NC} \ . \tag{D46}$$

Hence we have that

$$N_{NC} \leq \mathbb{E}\left[N\right] \leq \left(\tilde{g}_{\text{err}}(p_A, p_B, n^\star) + 1\right) N_{NC} \ , \tag{D47}$$

where $\tilde{g}_{\text{err}}(p_A, p_B, n^\star)$ is defined as

$$\tilde{g}_{\text{err}}(p_A, p_B, n^\star) := (1-p_B)^{n^\star}\left[\left(\frac{n^\star+\mu}{N_{NC}}\right)\left(\frac{2-(1-p_B)^{n^\star}}{1-(1-p_B)^{n^\star}}\right) - 1\right] \ . \tag{D48}$$

We now show that $\tilde{g}_{\text{err}}(p_A, p_B, n^\star)$ is small compared to the other quantities in Eq. (D47). Observe that

$$(1-p_B)^{n^\star} = \left(1 - \frac{\beta'}{n^\star}\right)^{n^\star} \leq e^{-\beta'} \ . \tag{D49}$$

From Eq. (D48) it follows that

$$\tilde{g}_{\text{err}}(p_A, p_B, n^\star) \leq e^{-\beta'}\left[\frac{n^\star + \frac{1}{p_A}}{N_{NC}}\left(\frac{2}{1-e^{-\beta'}}\right) - 1\right] \ . \tag{D50}$$

We can upper bound the relative error starting by upper bounding a term inside it, namely

$$\frac{n^\star + \frac{1}{p_A}}{N_{NC}} = \frac{n^\star + \frac{1}{p_A}}{\frac{1}{p_A} + \frac{1}{p_B} - \frac{1}{p_A+p_B-p_Ap_B}} \leq \frac{n^\star + \frac{1}{p_A}}{\frac{1}{p_A} + \frac{1}{p_B} - \frac{1}{p_A+p_B-p_A}} = p_An^\star + 1 \ . \tag{D51}$$

$\tilde{g}_{\text{err}}(p_A, p_B, n^\star)$, then, is upper bounded by

$$\tilde{g}_{\text{err}}(p_A, p_B, n^\star) \leq e^{-\beta'}\left[(p_An^\star+1)\left(\frac{2}{1-e^{-\beta'}}\right) - 1\right] \tag{D52}$$

$$= \frac{e^{-\beta'}}{1-e^{-\beta'}}(2p_An^\star + 1 + e^{-\beta'}) \tag{D53}$$

$$\leq \frac{e^{-\beta'}}{1-e^{-\beta'}}(2\beta' + 1 + e^{-\beta'}) \tag{D54}$$

$$= e^{-\beta'}\left(\frac{2\beta'}{1-e^{-\beta'}} + \coth\left(\frac{\beta'}{2}\right)\right) \tag{D55}$$

$$< e^{-\beta'}\left(\frac{2\beta'}{1-e^{-1}} + \coth\left(\frac{1}{2}\right)\right) \tag{D56}$$

$$< e^{-\beta'}\coth\left(\frac{1}{2}\right)(2\beta'+1) \tag{D57}$$

$$< 3\coth\left(\frac{1}{2}\right)\beta' e^{-\beta'} \ . \tag{D58}$$

Therefore $\tilde{g}_{\text{err}}(p_A, p_B, n^\star) = \mathcal{O}\left(\beta' e^{-\beta'}\right)$.