

Asymmetric Protocols for Scalable High-Rate Measurement-Device-Independent Quantum Key Distribution Networks

Wenyuan Wang^{1,*}, Feihu Xu^{2,†} and Hoi-Kwong Lo^{1,‡}

¹*Centre for Quantum Information and Quantum Control (CQIQC), Dept. of Electrical & Computer Engineering and Department of Physics, University of Toronto, Toronto, Ontario, M5S 3G4, Canada*

²*Shanghai Branch, National Laboratory for Physical Sciences at Microscale, University of Science and Technology of China, Shanghai, 201315, China*

 (Received 14 April 2018; revised manuscript received 11 July 2019; published 16 October 2019)

Measurement-device-independent quantum key distribution (MDI-QKD) can eliminate detector side channels and prevent all attacks on detectors. The future of MDI-QKD is a quantum network that provides service to many users over untrusted relay nodes. In a real quantum network, the losses of various channels are different and users are added and deleted over time. To adapt to these features, we propose a type of protocol that allows users to independently choose their optimal intensity settings to compensate for different channel losses. Such a protocol enables a scalable high-rate MDI-QKD network that can easily be applied for channels of different losses and allows users to be dynamically added or deleted at any time without affecting the performance of existing users.

DOI: [10.1103/PhysRevX.9.041012](https://doi.org/10.1103/PhysRevX.9.041012)

Subject Areas: Quantum Information

I. INTRODUCTION

Quantum computing threatens the security of conventional public key cryptography [1]. To address this increasing threat, quantum key distribution (QKD) allows two parties to share a pair of random keys with information-theoretic security and protects users from the attack of even quantum computers. Because of this, QKD is considered as one of the strong candidates for the next-generation technology for secure communications. Moreover, as we now live in the era of the Internet of Things that interconnects many users and devices, for QKD to be widely deployed in the future, an important step is to study it in a network setting, i.e., designing *quantum networks* that can connect and provide service to numerous users, who may freely join or leave a network.

The problem is, while QKD is theoretically secure, side channels still exist in a system built with practical components. Therefore, an important question in quantum cryptography is to determine how secure a system really is in practice. There have been multiple quantum hacking attacks, e.g., Refs. [2–7], that target the practical weaknesses in QKD

systems. Among the components of a QKD system, detectors are especially susceptible to attacks (and a majority of hacking attempts target the detectors), making them the Achilles' heel of QKD systems. The measurement-device-independent (MDI) QKD [8] protocol allows an untrusted third party to make measurements, thus avoiding all security breaches from detector side channels. Since its proposal, MDI-QKD has attracted worldwide interest, and there have been hundreds of follow-up theory and experimental papers. For example, some notable experimental implementations have been reported in Refs. [9–15]. An illustration of the MDI-QKD setup is shown in Fig. 1.

Up till now, multiple field implementations of point-to-point QKD networks have been reported in, e.g., Refs. [16–18]; however, they all relied on *trusted* relays (where the information stops being quantum at the relays), which are undesirable for security. MDI-QKD solves this problem by allowing *untrusted* relays in a quantum network, which is a huge advantage over previous point-to-point QKD networks, making MDI-QKD a powerful candidate for future quantum networks. For instance, the first three-user star-shaped MDI-QKD network experiment in a metropolitan setting has been reported in Ref. [19].

However, a major limitation of MDI-QKD is that it requires all users to have near-identical (i.e., symmetric) distances to the untrusted relay for the protocol to work well [20,21], and the key rate will degrade very quickly with an increased level of asymmetry between channels. [22] Because of this limitation, previous experiments of MDI-QKD either were performed in the laboratory over symmetric fiber spools [9–14] or had to deliberately add a

*wenyuan.wang@mail.utoronto.ca

†feihuxu@ustc.edu.cn

‡hklo@comm.utoronto.ca

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

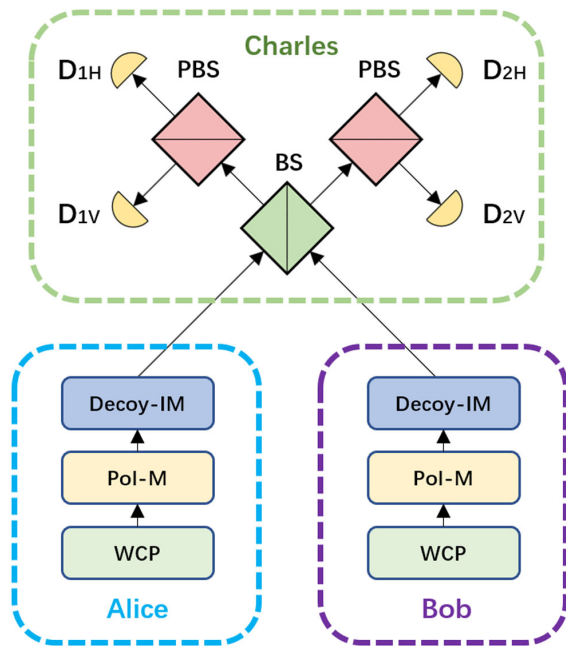


FIG. 1. An example schematic setup of MDI-QKD [8]. Alice and Bob, respectively, send signals through *two* channels, and Charles measures the signals with a Bell-state measurement [by observing the coincidence click events in detectors D_{1H} , D_{1V} , D_{2H} , D_{2V} behind the beam splitter (BS) and polarizing beam splitters (PBSs)] and announces the results. Here, weak coherent pulse (WCP) sources are used, in combination with decoy states created with intensity modulators (Decoy-IM). In this particular setup, polarization encoding is used [with polarization modulators (Pol-M)], but MDI-QKD can be performed with other degrees of freedom, such as time-bin phase encoding, too.

tailored length of fiber to the shorter channel (to introduce additional loss) in exchange for better symmetry [15]. Adding additional fibers not only is cumbersome as it requires halting the system (and not practical when there are many pairs of connections in a quantum network or when channel loss is changing), but also results in a suboptimal key rate when the channels are asymmetric. An intuitive illustration of this can be found in the Appendix A.

In a realistic setup, a quantum network will very likely have asymmetric channels due to different geographical locations of sites. For instance, the channel losses in Refs. [16,17] are largely different. Here, we select five nodes from the Vienna QKD network [16] and show them in Fig. 2(a), where the biggest difference between channels is as large as 66 km. If we want to perform MDI-QKD over these locations, although one can add additional fibers to each channel to compensate for channel differences, users have to accommodate for the lowest-transmittance channel—just like in “*Liebig’s barrel*”—and have a sub-optimal rate. Moreover, in a scalable network with large numbers of dynamically added or deleted users, it is not practical to add fibers and maintain symmetry between

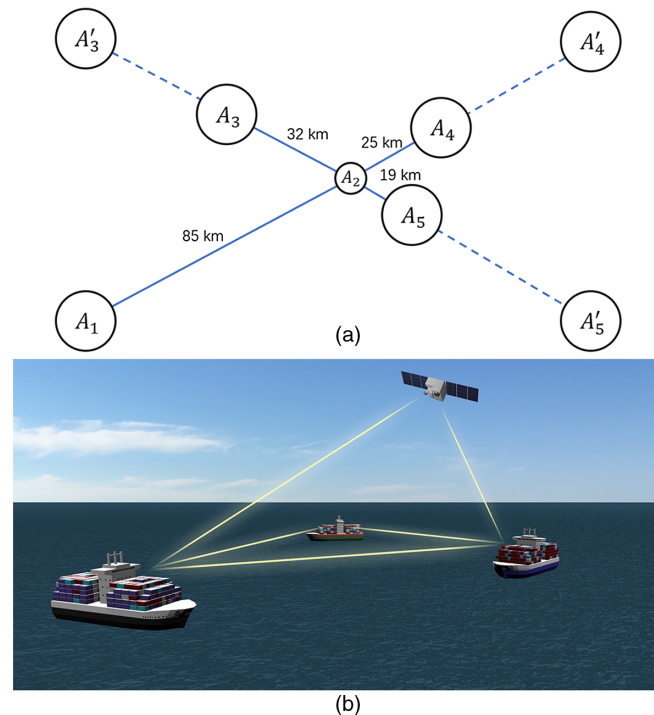


FIG. 2. (a) Part of the QKD network setup from Ref. [16]. Here as an example, we focus on the five nodes with high asymmetry (nodes A_1 , A_3 , A_4 , A_5 connected with A_2 , corresponding to nodes 1–5 in Ref. [16]), where A_2 can be set up as an untrusted relay. We keep the same topology and redraw it as a star-shaped MDI-QKD network with four users connected to a single untrusted relay. When performing MDI-QKD, all users need to accommodate for the longest channel (i.e., A_1) and add losses to their channels (e.g., extending to A'_3 , A'_4 , A'_5) if previous protocols are used. (b) Ship-to-ship communication and ground-satellite communication, where the participants’ distances to the detector are constantly changing, and the channels will thus have quickly varying asymmetry.

each pair of users all the time. Additionally, if one is to implement a MDI-QKD network over free space between mobile platforms (e.g., satellite-based MDI-QKD [25] or maritime MDI-QKD between ships), the losses in the channels are constantly changing, and the channels will often be highly asymmetric, as shown in Fig. 2(b). In summary, the requirement on symmetric channels significantly limits the key rate of previous MDI-QKD protocols in a general quantum network, thus seriously hindering the widespread deployment of MDI-QKD.

The issue of MDI-QKD with asymmetric channel losses was first considered in Ref. [4], which provided a rule of thumb on the ratio of intensities between Alice’s and Bob’s signals. However, Ref. [4] assumes infinitely large data size and was also restricted to protocols where the same set of intensities for the optical signals are used in the two bases X and Z . In this paper, we make no such assumptions.

In this work, we present a new method to overcome this crucial limitation directly and enable high-rate MDI-QKD

with arbitrary user locations. First of all, our work provides an important conceptual insight: A common folklore in the field is that MDI-QKD relies on the Hong-Ou-Mandel (HOM) dip, and therefore, it is important to use matched intensities at the beam splitter of the receiver Charles in MDI-QKD. Here, we show that such a folklore is, in fact, a misconception. We show that there is an intrinsic asymmetry between the two bases of MDI-QKD: Only the X basis relies on the indistinguishability of photons from the two beams, while the Z basis does *not*. We later show that one can make use of such an asymmetry to create protocols resilient against asymmetric channels. We also show that this is a general theoretical result applicable to many protocols, including various types of MDI-QKD protocols and potentially other protocols such as MDI quantum-digital signature [26,27] and twin-field QKD [28] in asymmetric settings.

Following this conceptual insight, we present a novel method in this paper to combat channel asymmetry. We make use of the inherent asymmetry between bases in MDI-QKD and propose a type of asymmetric MDI-QKD protocol where intensities are not only different for Alice and Bob, but also different in the X and Z bases. In this way, by decoupling the bases and also allowing Alice and Bob to independently vary their intensities, the users can effectively compensate for channel asymmetry in one basis and optimize the key generation rate in another basis, enabling a much higher key rate of our asymmetric protocols in the presence of channel asymmetry. Additionally, we present a technique that makes it possible to efficiently perform a local search for high-speed parameter optimization over the extremely large parameter space for such asymmetric protocols (which would be otherwise impossible to optimize using previous algorithms such as in Refs. [20,29]).

The protocols we propose have important practical impacts. We show that, when channels are asymmetric, our protocols can provide a much higher key rate than previous protocols [30,31] that were designed for symmetric channels [for instance, 1 to 2 orders of magnitude higher rate at mid-to-close distances, e.g., 60 km (10 km) for Alice’s (Bob’s) channels, with a 50-km difference in channel distances]. Moreover, it enables a much larger region of possible combinations of channels: For instance, even at a small data size of $N = 10^{11}$ (N is defined as the total number of pulses sent by Alice and Bob), one can generate a high secret key rate of $R = 10^{-7}$ per pulse even through an extremely asymmetric channel pair of 0 km (90 km) for Alice’s (Bob’s) channels, whereas with previous protocols no key could be generated at all. Using the type of protocol we propose, one can completely remove the requirement of symmetric channels in MDI-QKD. This makes our proposal a powerful solution that enables high-rate MDI-QKD under arbitrary asymmetry, which paves the way for practical MDI-QKD networks where users can be placed at arbitrary locations.

The structure of this paper is as follows: In Sec. II A, we point out a theoretical insight that there is an inherent asymmetry between the two bases of MDI-QKD. In Sec. II B we make use of this insight and propose a type of asymmetric protocol that simultaneously has two kinds of asymmetries: the asymmetry between Alice and Bob and the asymmetry between the X and Z bases, which, together, enable the protocol to effectively compensate for different pairs of channels and maintain a good key generation rate. We show the security of such a scheme in Sec. II C. We then describe how to optimally choose the asymmetric parameters in Sec. II D. While our proposal applies to a general type of MDI-QKD protocol, we also highlight a specific implementation, a “seven-intensity protocol,” and show that it is a good trade-off between the key rate and ease of implementation. Lastly, we present the simulation results to show the effectiveness of our protocol compared with prior protocols in Sec. III.

II. ASYMMETRIC PROTOCOLS

In this section, we present a general theoretical framework for designing protocols that can effectively compensate for channel asymmetry and provide a good key rate. Note that our method proposed here is a general result that can be applied to any decoy-state MDI-QKD protocol with WCP sources for both asymptotic and finite-size cases as long as (1) decoupled bases are used and (2) Alice and Bob have asymmetric intensities. We show in Appendixes B and C that the scaling of the key rate versus distance is determined by the signal states, so in principle, any number of decoy states (e.g., two decoys, three decoys, and four decoys) can be used so long as they can effectively estimate the single-photon contributions. In principle, such a method can potentially be applied even to other types of protocols in asymmetric settings, such as MDI quantum digital signature [26,27] and twin-field QKD [28], which are also currently limited to symmetric intensities between Alice and Bob and which also use two asymmetric bases X and Z .

A. Asymmetry between bases in MDI-QKD

Here, we start by making a key theoretical observation on MDI-QKD:

Observation 1: For MDI-QKD, there is an inherent asymmetry between the bases: Only the diagonal (X) basis requires the indistinguishability of the signals from Alice and Bob, while the rectilinear (Z) basis does not.

Such an observation is because in MDI-QKD, Charles performs a Bell-state measurement with *postselection*, making the protocol different from a simple two-photon interference in a standard HOM dip. Here, let us follow the discussions in Ref. [8] (and consider the experimental setup from Fig. 1 in Ref. [8]). Note that while Alice and Bob randomly send signals in the X and Z bases, Charles always measures in the Z basis (as defined by his PBS) and

postselects detector click events that correspond to the two Bell states $|\psi^+\rangle = 1/\sqrt{2}(|HV\rangle + |VH\rangle)$ and $|\psi^-\rangle = 1/\sqrt{2}(|HV\rangle - |VH\rangle)$. Such a postselection results in an asymmetry between the two bases. In the Z basis, only events where Alice and Bob sent opposite states (e.g., $|HV\rangle$ or $|VH\rangle$) are accepted as bits. In these cases, no photon interference takes place, and indistinguishability between the two input photon beams is not required because each of the clicking detectors, respectively, receives only a signal from either Alice or Bob but never both. For WCP sources, in the ideal case with no misalignment or dark counts, the intensities of the pulses and even their spectrum and timing need not be matching at all. In the X basis, however, the events may correspond to identical states sent by Alice and Bob [e.g., $|++\rangle$ and $|--\rangle$ corresponding to $|\psi^+\rangle = 1/\sqrt{2}(|++\rangle - |--\rangle)$], which do interfere at the beam splitter. [35] To ensure that the correct events are triggered, a good visibility of such a two-photon interference is required. Note that for WCP sources, the interference visibility is at most 50% (resulting in a 25% observed QBER in the X basis even in the ideal case—for instance, $E_{\mu\mu}^X, E_{\nu\nu}^X$ when Alice and Bob use decoy states with intensities μ_A, μ_B and ν_A, ν_B —but we can perform a decoy-state analysis to correctly estimate a low QBER among single-photon components, $e_{11}^{X,U}$), and the visibility will quickly drop when intensities are mismatched, such as observed in Ref. [21].

Therefore, a low QBER in the X basis heavily relies on the indistinguishability of the signals and the balance of incoming intensities at Charles, while such a dependence is not present in the Z basis. [37] Such a conclusion is rather general and also not dependent on the degree of freedom used for qubit encoding, such as polarization encoding or time-bin phase encoding (where $|HV\rangle$ and $|VH\rangle$ in the Z basis correspond to pairs of early and late pulses, which will similarly not interfere at the beam splitter since they have different timing).

B. Using decoupled bases and asymmetric intensities

Here, let us consider the key rate formula of MDI-QKD [8,30]:

$$R = P_{s_A} P_{s_B} \{ (s_A e^{-s_A}) (s_B e^{-s_B}) Y_{11}^{X,L} [1 - h_2(e_{11}^{X,U})] - f_e Q_{ss}^Z h_2(E_{ss}^Z) \}, \quad (1)$$

where s_A, s_B are the intensities of signal states, Q_{ss}^Z, E_{ss}^Z are the gain and QBER in the Z (signal) basis, $Y_{11}^{X,L}, e_{11}^{X,U}$ are the lower (upper) bounds of single-photon yield and QBER estimated from the decoy-state statistics in the X basis (i.e., the observed gain and QBER for decoy states Q_{ij}^X, E_{ij}^X , where i, j are decoy intensities, such as in $\{\mu_A, \nu_A, \omega\}$ and $\{\mu_B, \nu_B, \omega\}$ if Alice and Bob each choose three decoy states), h_2 is the binary entropy function, and f_e is the error-correction efficiency.

In the key rate formula, the first part corresponds to key generation (where the privacy amplification depends on the single-photon contributions estimated from decoy-state analysis), and the second part corresponds to error correction for the signal states. We can make another key observation on the intensities used in the two bases:

Observation 2: In our protocol, the intensities of the signal states $\{s_A, s_B\}$ used in the Z basis are independent from those of the decoy states used in the X basis, which means that the privacy amplification process (to bound Eve's information on the final key, i.e., estimate the phase error rate) in the X basis is completely *decoupled* from error correction in the Z basis for key generation.

This decoupling of bases means that it is possible for us to independently adjust the decoy states and the signal states in their respective bases to compensate for channel asymmetry or to optimize the key rate.

For the decoy states, their role is to estimate the single-photon contributions as accurately as possible. As we mention above, when channels are asymmetric, using the same intensities for Alice and Bob (hence different intensities arriving at Charles after the channels' attenuation) will result in poor interference visibility and high QBER in the X basis, and consequently poor estimation of $e_{11}^{X,U}$. For a good interference visibility, Alice and Bob should try to maintain similar intensities arriving at Charles, so the decoy intensities should be chosen to roughly satisfy

$$\mu_A \eta_A = \mu_B \eta_B, \quad (2)$$

where η_A and η_B are the channel transmittances in Alice's and Bob's channels. A similar equation holds true for ν_A and ν_B .

In contrast, for the signal states, they are not involved in privacy amplification. On the other hand, they affect the signal-state gain and QBER Q_{ss}^Z, E_{ss}^Z (which determine the amount of error correction) and the probability of sending single photons for key generation $s_A e^{-s_A} s_B e^{-s_B}$. The key point is, the QBER E_{ss}^Z does *not* require indistinguishability of the signals. If there is no misalignment or noise, E_{ss}^Z would be zero regardless of incoming intensities. In practice, due to imperfections such as misalignment, the QBER E_{ss}^Z [whose full expression can be found in Appendix C Eq. (C3)] still slightly depends on channel asymmetry and is also minimal if incoming intensities at Charles are balanced, but this dependence is for a much different reason (due to misalignment) than that in the X basis (mostly due to two-photon interference). Furthermore, E_{ss}^Z is much less sensitive to channel asymmetry than QBER in the X basis. We can observe this from Fig. 3.

Note that, not only do signal intensities affect the signal-state QBER, they also determine the probabilities of sending single photons, hence affecting key generation,

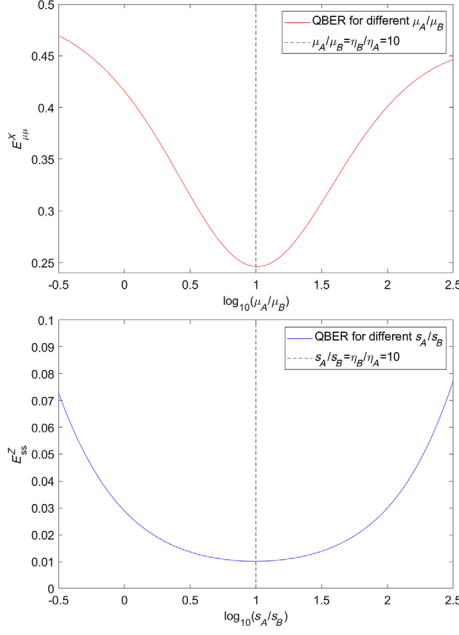


FIG. 3. An example of the respective quantum bit error rate (QBER) in the X basis and Z basis, $E_{\mu\mu}^X$ and E_{ss}^Z (we consider a pair of decoy states with intensities μ_A, μ_B in the X basis and signal-state intensities s_A, s_B in the Z basis) versus the ratio of intensities for MDI-QKD using WCP sources. Parameters from Table II are used. Here we consider the case where the respective distances from Alice and Bob to Charles are $L_A = 60$ km, $L_B = 10$ km (i.e., the ratio of transmittances in the two channels satisfies $\eta_B/\eta_A = 10$). We fix $s_B = 0.2$ (or $\mu_B = 0.2$) and scan over different s_A (or μ_A). Specifically, we also mark out the position where $s_A \eta_A = s_B \eta_B$ ($\mu_A \eta_A = \mu_B \eta_B$). Because QBER in the X basis heavily depends on the visibility of two-photon interference, it is lowest when intensities arriving at Charles's beam splitter are equal (a similar observation has been made in Ref. [21]). However, importantly, the Z basis does not require signal indistinguishability, and its QBER is determined mainly by misalignment. The misalignment makes the Z basis QBER also slightly dependent on the interference visibility and lowest when arriving intensities are equal, but such a QBER is much less sensitive to unbalanced intensities and is relatively low even if $s_A \eta_A \neq s_B \eta_B$. Therefore, by decoupling the X and Z bases, we can maintain highly balanced decoy-state intensities arriving at Charles in the X basis, while further optimizing signal intensities to obtain a higher key rate. As a quantitative example of such a difference in sensitivity, let us consider $L_A = 60$ km, $L_B = 10$ km, and $N = 10^{11}$ (Table IV, line 1) and focus on two pairs of decoy states with intensities $\mu_A, \mu_B, \nu_A, \nu_B$. An optimal key rate of $R = 3.1 \times 10^{-5}$ can be achieved, where optimal decoy-state intensities satisfy $\mu_A/\mu_B = \nu_A/\nu_B = 9 \approx \eta_B/\eta_A$, and $E_{\mu\mu}^X, E_{\nu\nu}^X$ are both close to 25% (see Table IV for the full list of intensities and probabilities). Even a relatively small deviation, such as choosing $\mu_A/\mu_B = \nu_A/\nu_B = 10^{0.5} = 3.16$ when fixing μ_B, ν_B (which results in $E_{\mu\mu}^X$ and $E_{\nu\nu}^X$ close to 32%) results in zero rate. On the other hand, the optimal signal states satisfy $s_A/s_B = 3.5$, which deviates from η_B/η_A , but $E_{\mu\mu}^Z$ is still a rather small 0.013. In fact, even if we choose $s_A = s_B = 0.2$ here, we can still get $R = 1.0 \times 10^{-5}$ while $E_{\mu\mu}^Z = 0.029$.

too. This means that, while having similar received signal intensities at Charles is surely one important criterion in achieving a good key rate, the optimal choice of signal-state intensities requires a trade-off between the single-photon probabilities and the error correction (and their optimal values can be found by numerical optimization). Generally speaking, the ratio of signal intensities s_A/s_B does not satisfy a similar relation as Eq. (2), i.e., generally,

$$s_A \eta_A \neq s_B \eta_B. \quad (3)$$

Therefore, the protocols we propose have two inherent asymmetries: an asymmetry between Alice and Bob (so that they can have different intensities and establish good two-photon interference in the X basis) and an asymmetry between the X and Z bases (which allows decoy and signal states to be independently optimized). Such inherent asymmetries in the protocols allow us to have a novel choice of parameters and maintain a good key rate of MDI-QKD, even when Alice's and Bob's channels have very different levels of loss. A more detailed discussion on how such independent choices of decoy and signal states affect the key rate can be found in Appendix D.

We discuss the security of such a scheme in Sec. II C, and in Sec. II D, we discuss how to actually choose the optimal decoy and signal intensities. We introduce the main challenge in implementing such asymmetric protocols—performing efficient parameter optimization over a huge parameter space—and how we address this problem by proposing two important theoretical results for the key rate function of asymmetric MDI-QKD and using them to design an efficient optimization algorithm.

C. Security

In this subsection, we show that the security of our protocol with decoupled bases and asymmetric intensities is not compromised compared to prior art protocols. Here, we state two *facts* that its security relies upon, both of which have been proven in established papers [8,32]:

- (1) Given the same photon number i in a pulse, Eve has no way of differentiating whether it came from the decoy states or the signal states in the same basis.
- (2) The single-photon pairs in the X and Z bases cannot be distinguished from each other.

The first fact is proven in Ref. [32], which ensures the decoy-state analysis works even with asymmetric intensities, and the second fact is proven in Ref. [8], which ensures that the decoupling of bases works.

For fact (1), note that the density matrices for the two states are the same, independent of whether they came from the decoy states or signal states. In quantum mechanics, whenever two density matrices are the same, they are indistinguishable. Similarly, for fact (2), note that the density matrices for the two cases are the same. Therefore, it is not possible to distinguish them even in principle.

In more detail, for fact (1), let us first consider a very similar process in traditional decoy-state BB84 [32–34]. Consider Alice using a laser with intensities μ or ν to send weak coherent pulses to Bob. The photon number i follows a Poissonian distribution, e.g.,

$$\begin{aligned} p(i|\mu) &= e^{-\mu} \frac{\mu^i}{i!}, \\ p(i|\nu) &= e^{-\nu} \frac{\nu^i}{i!}. \end{aligned} \quad (4)$$

The crucial point is that the conversion from such a probability distribution described by intensities μ or ν to a certain photon number i is a *Markov process*, i.e., it is memoryless, and for any given photon number i in the channel, it does not contain any information of the intensity it came from. Surely, Eve can guess with a conditional distribution, e.g., $p(\mu|i)$, the likelihood that it comes from a certain intensity, but whatever actions Eve chooses to perform on the signal (e.g., choosing different levels of yields $Y_i^{E1}, Y_i^{E2}, \dots$ with different probabilities for a given photon number i) will be completely independent of the intensities Alice chose when sending the signal. This means that in the asymptotic case with infinite data, [38] given the same photon number i in a pulse, we will still always have yield Y_i and QBER e_i satisfying

$$\begin{aligned} Y_i(\mu) &= Y_i(\nu), \\ e_i(\mu) &= e_i(\nu) \end{aligned} \quad (5)$$

when Alice uses intensities μ, ν . This is the exact observation made in Ref. [32] [in Eqs. (4) and (5)].

Similarly, for MDI-QKD, for any given pair of i -photon and j -photon pulse, there is no information on which pair of intensity settings (e.g., μ_A^1, μ_B^1 or μ_A^2, μ_B^2) they came from. That is, the yield $Y_{i,j}$ and QBER $e_{i,j}$ will satisfy

$$\begin{aligned} Y_{i,j}^X(\mu_A^1, \mu_B^1) &= Y_{i,j}^X(\mu_A^2, \mu_B^2), \\ e_{i,j}^X(\mu_A^1, \mu_B^1) &= e_{i,j}^X(\mu_A^2, \mu_B^2), \\ Y_{i,j}^Z(\mu_A^1, \mu_B^1) &= Y_{i,j}^Z(\mu_A^2, \mu_B^2), \\ e_{i,j}^Z(\mu_A^1, \mu_B^1) &= e_{i,j}^Z(\mu_A^2, \mu_B^2) \end{aligned} \quad (6)$$

for signals in each of the bases X and Z (the latter two equations are meaningful if one also uses multiple decoy intensities in the Z basis, although here we use only the first two equations as decoy-state analysis is performed only in the X basis for our protocols). This, again, is a well-established result for decoy-state MDI-QKD as used in the original MDI-QKD paper [8]. Note that this result (which simply comes from the fact that the sending of photon number i from an intensity μ is a Markov process) does not rely on the fact that Alice and Bob use the same intensities, and will remain unchanged for asymmetric intensities, too,

i.e., $\mu_A^1 \neq \mu_B^1$ and $\mu_A^2 \neq \mu_B^2$. Also, note that for successful decoy-state analysis, we do not require the symmetry between the two bases, i.e., $Y_{i,j}^X(\mu_A^1, \mu_B^1) = Y_{i,j}^Z(\mu_A^2, \mu_B^2)$ or $e_{i,j}^X(\mu_A^1, \mu_B^1) = e_{i,j}^Z(\mu_A^2, \mu_B^2)$ for multiphoton pulses are not required.

Fact (2) stems from the fact that, for single-photon components, Alice and Bob send the same density matrices $\rho_{1,1}^X = \rho_{1,1}^Z$; that is, the single-photon pairs are basis independent. This is an important result explicitly stated in the original MDI-QKD paper [8] (in the Security Analysis section in the Supplemental Material). Using decoupled bases (i.e., a different set of intensities for the X and Z bases) does not affect the single photons themselves at all but affects only the probability of sending these single-photon pairs. However, this process of sending single-photon pairs is again a Markov process. That is, although the single-photon pairs might have different probabilities of coming from either the X or Z bases (which Eve can fully be aware of, just like in “efficient BB84” [39], where basis choice probability is biased on purpose), for any given pair of single photons that are sent, they are described by exactly the same density matrix, and there is no information contained on which basis they came from; i.e., they are basis independent. Therefore, we can safely conclude that $Y_{11}^X = Y_{11}^Z$, which is the reason we can perform decoy states in the X basis only to estimate Y_{11}^X and use $Y_{11}^Z = Y_{11}^X$ to obtain the single-photon yield in the Z basis.

The security of a scheme of decoupling the bases in MDI-QKD and using $Y_{11}^Z = Y_{11}^X$ has also been theoretically studied in Ref. [30] and in the Appendix of Ref. [40,41], and the scheme has also been experimentally demonstrated in Refs. [13,40], although all these works were focused on the scenario of symmetric channels only and did not discuss the role of decoupled bases in compensating channel asymmetry, which is one of the main novelties of our work. However, physically, the only difference between the signals sent from Alice and Bob in our protocol and those in prior protocols is the different intensities on the two arms [which we know, from fact (1) $Y_{i,j}^X(\mu_A^1, \mu_B^1) = Y_{i,j}^X(\mu_A^2, \mu_B^2)$, will not affect the security of the decoy-state analysis], and for decoupled bases, we use the same result $Y_{11}^Z = Y_{11}^X$ for single-photon pairs, which is no less secure than prior works either.

D. Parameter optimization

In this section, we discuss how to perform efficient parameter optimization for such asymmetric protocols. Here, we highlight an implementation which we denote as the seven-intensity protocol (which is the case where three decoy intensities are used in the X basis). We show that it is a good trade-off between the key rate and ease of implementation and focus on this implementation when discussing parameter optimization (and the following

numerical simulations). Nonetheless, we also show the generality of our method by including the results for other protocol cases (e.g., two-intensity and four-intensity) in Appendix E.

Note that the results in the previous subsection are general and not limited to the number of decoys Alice and Bob use in the X basis. For instance, while using signal states $\{s_A, s_B\}$ in the Z basis, in the X basis Alice and Bob can each use a different set of two decoy states $\{\mu, \nu\}$, three decoy states $\{\mu, \nu, \omega\}$, or even four decoy states $\{\mu, \nu, \nu_2, \omega\}$. In principle, the concept of asymmetric intensities between Alice and Bob can also be applied to prior art protocols with nondecoupled bases, such as in Refs. [20,43] (where Alice and Bob use the same three decoy states $\{\mu, \nu, \omega\}$ for both bases, and the Z basis μ is used as the signal state for key generation)—it is just that such a protocol will have a lower key rate since μ cannot simultaneously satisfy asymmetry compensation and key rate optimization.

As an example, in Table I we list a comparison between the key rate of using different numbers of decoy states (with and without asymmetric intensities between Alice and Bob) in the presence of asymmetric channels. We include the non-decoupled-bases case [20,43], too. We can see that, regardless of the protocol, using asymmetric intensities between Alice and Bob always provides a higher key rate when the channels are asymmetric. Also, the three-decoy case provides significant performance improvement over the two-decoy case or the prior art protocol (which also has three decoy states, meaning that decoupled bases are crucial in the compensation for channel asymmetry). While the asymmetric four-decoy case can provide the highest key rate, it provides a limited performance increase (60%) over the three-decoy case, but it comes at a cost of a more complex experimental implementation as well as more difficult data collection and analysis. See Appendix E for a more detailed comparison between the protocols. Overall, we can see that the three-decoy case provides a good balance between ease of implementation and performance.

Therefore, for practicality here, in the following we focus on the three-decoy case as a concrete example (whose symmetric case is the four-intensity protocol [30]) and generalize it to the asymmetric case by allowing Alice and Bob to have independent intensities and probabilities. Such a setting enables a seven-intensity protocol (with three independent $\{s, \mu, \nu\}$ for each of Alice and Bob, and the vacuum state $\omega_A = \omega_B = \omega = 0$) in the asymmetric case.

For such a protocol, efficient and accurate parameter optimization is crucial for obtaining a good key rate (especially when considering the finite-size effects). For the seven-intensity protocol, we need to use a total of 12 parameters for a full finite-size parameter optimization:

$$\vec{v} = [s_A, \mu_A, \nu_A, P_{s_A}, P_{\mu_A}, P_{\nu_A}, s_B, \mu_B, \nu_B, P_{s_B}, P_{\mu_B}, P_{\nu_B}].$$

Here, we denote the parameters as a vector \vec{v} , and when all devices and channel parameters (e.g., channel loss, misalignment, dark count rate, detector efficiency, etc.) are fixed, the key rate is a function of the intensities and probabilities $R(\vec{v})$, and the question of intensity parameter optimization can be viewed as searching for

$$\vec{v}_{\text{opt}} = \arg \max_{\vec{v} \in V} [R(\vec{v})], \quad (7)$$

where V is the search space for the parameters.

To provide a high key rate under finite-size effects, the optimal choice of parameters is very important in implementing the protocol. However, the seven-intensity protocol has an extremely large parameter space of 12 dimensions, for which a brute-force search is next to impossible. Therefore, to efficiently search over the parameters in a reasonable time, a local search algorithm must be applied. But, as we show here, an important characteristic of asymmetric MDI-QKD is the discontinuity of first-order derivatives for the function $R(\vec{v})$ with respect to the intensity parameters in \vec{v} . This means that a straightforward local search algorithm, such as previously proposed in Ref. [29], will inevitably fail to find the optimal point, since

TABLE I. Example key rate comparison among MDI-QKD protocols where Alice and Bob use different numbers of decoy states in the X basis (and each keep one signal state in the Z basis). The protocol in Refs. [20,43] where the bases are not decoupled is also included for comparison. We use parameters from Table II, $L_A = 60$ km, $L_B = 10$ km, and $N = 10^{11}$. We can see that, regardless of the protocol, using asymmetric intensities between Alice and Bob always provides a higher key rate when the channels are asymmetric. The three-decoy protocol has a significantly higher key rate than either the prior art protocol (which also uses three decoy states but uses nondecoupled bases) or the two-decoy case. While the asymmetric four-decoy case can provide the highest key rate, it provides a limited performance increase of 60%, but comes at a cost of a more complex experimental implementation and more difficult data collection and analysis. Therefore, in the presence of channel asymmetry, the three-decoy case, whose asymmetric case corresponds to the seven-intensity protocol (marked in bold), provides a good trade-off between ease of implementation and performance.

Parameters	Prior art protocol in Refs. [20,43]	Two-decoy	Three-decoy	Four-decoy
Symmetric	6.834×10^{-10}	0	3.890×10^{-7}	1.057×10^{-5}
Asymmetric	5.378×10^{-7}	7.715×10^{-6}	3.106×10^{-5}	4.932×10^{-5}

it requires continuous first-order derivatives of the searched function.

Here, we present two important theoretical results for the key rate versus parameter function and propose a method to circumvent the problem of discontinuous derivatives and perform an efficient and correct local search in parameter space. This method helps us overcome the biggest challenge in successfully implementing the seven-intensity protocol.

First, we propose that there is an inherent symmetry constraint for the *ratio* of optimal decoy intensities.

Theorem I: For any arbitrary choice of device and channel parameters, the optimal decoy intensities μ_A^{opt} , ν_A^{opt} , μ_B^{opt} , ν_B^{opt} that maximize the key rate always satisfy the constraint

$$\frac{\mu_A^{\text{opt}}}{\mu_B^{\text{opt}}} = \frac{\nu_A^{\text{opt}}}{\nu_B^{\text{opt}}}. \quad (8)$$

Second, we make an important observation.

Theorem II: The key rate versus (μ_A, μ_B) function for any given ν_A, ν_B does not have continuous first-order derivatives.

Both of these theorems result from the fact that the lower bound for single-photon yield Y_{11}^L in the decoy-state analysis (whose expression can be found in Refs. [20,43]) is a piecewise function that depends on whether $(\mu_A/\mu_B) \leq (\nu_A/\nu_B)$, where a boundary line $(\mu_A/\mu_B) = (\nu_A/\nu_B)$ exists.

Theorem I states that the optimal parameters that maximize the key rate must lie exactly on this boundary line, while Theorem II states that the key rate does not have a continuous partial derivative with respect to μ_A or μ_B across this boundary line. This will cause the boundary line to behave like a sharp “ridge,” on which the gradient is not defined. An illustration for this ridge can be seen in Fig. 4. A rigorous proof for Theorems I and II in the asymptotic limit can be found in Appendix F.

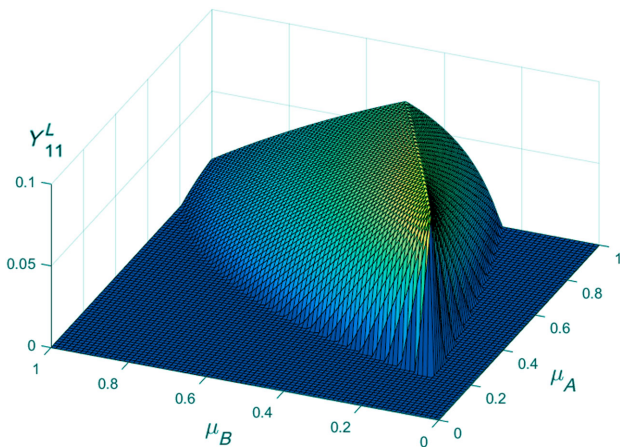


FIG. 4. An example of the discontinuity of first-order derivatives of the Y_{11}^L vs μ_A, μ_B function in decoy-state MDI-QKD for fixed values of $\nu_A = 0.2$, $\nu_B = 0.1$. Note the ridge on the line $(\mu_A/\mu_B) = (\nu_A/\nu_B) = 2$.

Using Theorems I and II, it is possible to transform the coordinates of the search variables and eliminate the undefined gradient problem of the key rate function. More specifically, instead of expressing $(\mu_A, \mu_B), (\nu_A, \nu_B)$ in Cartesian coordinates, we can express them in polar coordinates $(r_\mu^{\text{polar}}, \theta_\mu^{\text{polar}}), (r_\nu^{\text{polar}}, \theta_\nu^{\text{polar}})$, where polar angles satisfy $\theta_\mu^{\text{polar}} = \theta_\nu^{\text{polar}}$ due to Theorem I. This means we can *jointly* search for

$$\theta_{\mu\nu}^{\text{polar}} = \theta_\mu^{\text{polar}} = \theta_\nu^{\text{polar}} \quad (9)$$

with respect to which the key rate is a smooth function (graphically, this is because we are now always searching along the ridge). In Appendix G, we describe in more detail how to perform an optimization of the parameters efficiently based on local search to obtain a high secure key rate for our seven-intensity protocol. Our method allows extremely fast and highly accurate optimization for asymmetric MDI-QKD and takes below 0.1 s for each full local search (at any given distance) on a quad-core i7-4790k 4.0-GHz PC. Such computing efficiency makes it possible for real-time optimization of intensities on the field and also makes possible a dynamic MDI-QKD network that might add or delete new user nodes in real time. In addition, in Appendix G we also discuss the effect of inaccuracies and fluctuations of the intensities and probabilities on the key rate and show that our method is robust even in the presence of inaccuracies and fluctuations of the parameters.

In summary, using our two theorems and switching to polar coordinates as in Eq. (6) allow us to greatly simplify the optimization problem and allow the standard coordinate descent method to be applied here.

III. SIMULATION RESULTS

Now, we can proceed to study the performance of asymmetric MDI-QKD protocols with full parameter optimization. Again, we use the seven-intensity protocol as a concrete example, as it provides a good trade-off between performance and practicality. We also include simulation results for protocols with alternative numbers of decoy states in Appendix E.

In the main text, we focus on the practical case of having finite data size. The asymptotic case of infinite data size (and an analytical understanding of the ideal infinite-decoy case) is discussed in Appendixes B and C, and its simulation results can be found in Fig. 9.

Our finite-key analysis is described in more detail in Appendix H. For simplicity, we consider a standard error analysis in numerical simulations, but it is important to note that our theory is fully compatible with composable security. See Appendix H for discussion. In addition, compared to the “joint-bound” analysis as proposed in Ref. [21] (which jointly considers the statistical fluctuation of multiple observables; such an analysis model increases the key rate

but introduces multiple maxima undesirable for local search), in the main text here, we choose use an “independent-bound” analysis for our simulations, which considers each variable’s statistical fluctuations independently and is far more stable and faster in simulations. However, we specifically note here that all our methods are fully compatible with joint-bound analysis. We list some representative results generated with joint-bound analysis in Table IV for comparison and discuss the different finite-size analysis models in more detail in Appendix H.

First, we consider the key rate for an arbitrary combination of (L_A, L_B) and perform a simulation of the key rate over *all* possible range of Alice and Bob’s channels. This provides a bird’s-eye view of how using 7-intensities can affect the performance in asymmetric channels. We show the results in Figs. 5(a) and 5(b). From the plot, we can make three important observations:

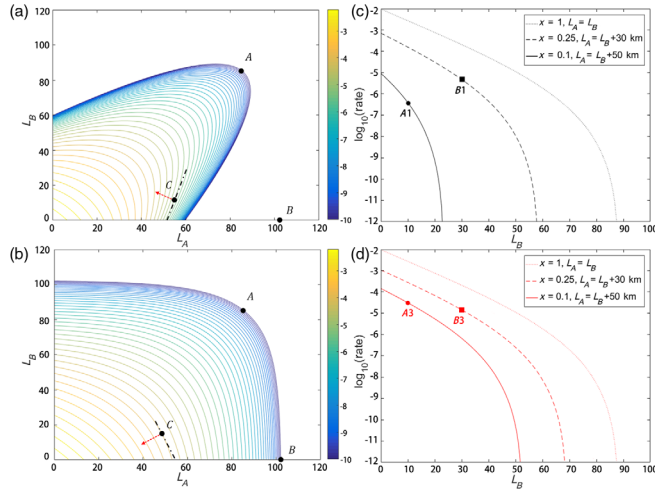


FIG. 5. Left: Comparison of rate vs (L_A, L_B) . The rates are plotted in contours in log scale from 10^{-2} to 10^{-10} . We use the parameters from Table II, and $N = 10^{11}$. (a) Using a previous four-intensity protocol; (b) using our seven-intensity protocol. As can be seen, while four-intensity MDI-QKD is limited to only high-symmetry regions, using seven-intensity can greatly increase the applicable region of MDI-QKD, even in extremely asymmetric regions such as $(L_A, 0)$ where one channel has zero distance (point B). Moreover, we see that with seven-intensity protocol, both L_A, L_B components of the gradient for the key rate (red dotted arrow) are always negative, meaning that with seven-intensity protocol, it is always optimal to adjust only the intensities and never necessary to add any fiber, while for four-intensity protocol, adding fiber (e.g., increasing L_B at point C) will sometimes increase the rate. Right: Comparison of rate vs distance (Bob to Charles) for various fixed levels of mismatch $x = (\eta_A/\eta_B)$ where η_A, η_B are the channel transmittances (c) using four-intensity protocol and (d) using seven-intensity protocol. As can be observed, the higher the mismatch, the more advantage seven-intensity protocol has (and only when the channels are symmetric will the two protocols perform identically). Data points from A1, A3, B1, and B3 from Table III are also shown in the plots.

- (1) Using seven-intensity protocol, we have a much wider applicable region for asymmetric MDI-QKD, and an acceptable key rate can be acquired even for highly asymmetric channels. In addition, seven-intensity protocol will always provide a higher key rate than four-intensity protocol, except when channels are already symmetric.
- (2) No matter what position one is at, there is never any necessity for adding loss when seven-intensity protocol is used, and optimizing on the spot always provides the highest rate. Details are provided in the Fig. 5 caption.
- (3) Using seven-intensity protocol, even extremely asymmetric scenarios, such as $(L, 0)$ where $L_B = 0$, can be used to generate a good key rate. In fact, such a scenario provides an even higher rate than with symmetric channels such as (L, L) (as the comparison between points A and B in Fig. 5).

Point 3 has an important practical implication: It can lead to a new type of “single-arm” MDI-QKD setup. More details can be found in Appendix I and Fig. 14.

Here, for points 1–3, we have a good physical understanding of why allowing different intensities for Alice and Bob can provide a larger region where the key rate is positive. As we discuss in Sec. II B, MDI-QKD requires highly balanced intensities arriving at Charles on the two arms in the X basis for good interference visibility, as well as roughly similar (but not necessarily balanced) levels of arriving intensities in the Z basis, which optimize a trade-off between error correction and the probability of sending single photons. (The optimal choice of intensities is subject to numerical optimization as described in Sec. II D). Prior methods with the same intensities for Alice and Bob will suffer from high QBER in both the X and Z bases, while our method decouples the X and Z bases and optimally chooses Alice’s and Bob’s signal and decoy intensities, respectively, to compensate for channel asymmetry in both bases, ensuing low QBER and allowing for a much higher key rate under channel asymmetry. Such an effect is present in both asymptotic and finite-key scenarios, and it is the underlying reason that the seven-intensity protocol can allow high-rate MDI-QKD regardless of channel asymmetry.

Additionally, we show that when channels are highly asymmetric, the asymptotic key rate of the seven-intensity protocol scales quadratically with the *lower* transmittance

TABLE II. Parameters for numerical simulations adopted from Ref. [13] including detector dark count rate and efficiency Y_0, η_d , optical misalignment e_d , error-correction efficiency f , and failure probability ϵ . Here the value Y_0 we use is for each *two* detectors, while the dark count rate per detector $p_d = 1 - \sqrt{1 - Y_0}$, which is approximately half of Y_0 .

Y_0	η_d	e_d	f	ϵ
8×10^{-7}	65%	0.5%	1.16	10^{-7}

TABLE III. Simulation results for asymmetric MDI-QKD in two scenarios: case *A* (10 km, 60 km) and case *B* (30 km, 60 km) using parameters from Table II and $N = 10^{11}$. We define channel mismatch as $x = (\eta_A/\eta_B)$ where η_A, η_B are the channel transmittances. Note that in reality, Alice and Bob cannot modify the physical channels, and they can either add loss to the channels or keep them as is, but they cannot decrease channel loss. Three strategies are compared here: *A1* and *B1* represent using the old four-intensity protocol directly. *A2* and *B2* (not in Fig. 5) represent adding fiber to the shorter channel to match the longer channel, i.e., making the channels (60 km, 60 km). *A3*, *B3* represent using our new seven-intensity protocol without modifying the channels. As shown here, seven-intensity protocol always returns a higher rate than both strategies using four-intensity protocol.

Protocol	Point	x	L_B	L_A	Rate	Comparison with four-intensity protocol
Four-intensity protocol	<i>A1</i>	0.1	10 km	60 km	3.891×10^{-7}	...
Four-intensity protocol + fiber	<i>A2</i>	1	60 km	60 km	1.862×10^{-6}	+379%
Our protocol	<i>A3</i>	0.1	10 km	60 km	3.106×10^{-5}	+7883%
Four-intensity protocol	<i>B1</i>	0.25	30 km	60 km	4.746×10^{-6}	...
Four-intensity protocol + fiber	<i>B2</i>	1	60 km	60 km	1.862×10^{-6}	-61%
Our protocol	<i>B3</i>	0.25	30 km	60 km	1.445×10^{-5}	+204%

among the two channels, which means that, albeit always being able to provide a higher key rate and being much more convenient than, e.g., adding fibers when channels are asymmetric (which is a relation we rigorously prove in Appendix C 2 for the asymptotic case), the seven-intensity protocol will not change the asymptotic scaling properties of the MDI-QKD key rate, which is still quadratically related to transmittance. Physically, this behavior is understandable, since although we effectively compensate for channel asymmetry with optimized intensities and allow good Hong-Ou-Mandel interference at Charles for the decoy states, MDI-QKD still fundamentally depends on two single signal photons both passing through the channels; hence, its key rate is quadratically related to transmittance, even in the asymmetric case with compensated intensities. More detailed discussions and analytical proofs of the above observations can be found in Appendixes B and C.

Now, as a concrete example, let us consider two sets of channels at ($L_B = 10$ km, $L_A = 60$ km) and ($L_B = 30$ km, $L_A = 60$ km), through which Alice and Bob would like to perform MDI-QKD. We compare strategies of using the four-intensity protocol, directly or with fibers added until the channels are symmetric, with directly using our seven-intensity protocol. As can be seen in Table III, in this specific example, using seven-intensity protocol can provide 1 or 2 magnitudes higher key rate, and its rate is also always

higher than either strategy with the four-intensity protocol. In fact, we can also show such an advantage of our protocol by plotting the key rate vs L_B under a fixed mismatch $x = (\eta_A/\eta_B)$, where η_A, η_B are the channel transmittances (i.e., a fixed difference between L_A and L_B). This is also the scenario studied by Ref. [20]. The results are shown in Figs. 5(c) and 5(d). The data points *A1*, *A3*, *B1*, and *B3* in Table III are also plotted. As can be seen, the higher the asymmetry between the channels, the more improvement we can gain from using seven-intensity protocol.

Here, we also list some examples of optimal parameters found by the optimization algorithm, which are listed in Table IV. As we can observe from the table, Alice and Bob adjust their intensities to compensate for channel asymmetry. Physically, since MDI-QKD depends on Hong-Ou-Mandel interference of two WCP sources in the *X* basis, we expect the *received* intensity for decoy state at Charles to be similar on the two arms to ensure good visibility (and consequently lower QBER) in the *X* basis; i.e., the ratio of decoy intensities μ_A/μ_B and ν_A/ν_B would roughly follow the rule of thumb of $\mu_A\eta_A = \mu_B\eta_B$, which is indeed what we can observe from Table IV and Fig. 6.

On the other hand, the ratio of signal intensities s_A/s_B deviates more from η_B/η_A . This is because, as we mention in Sec. II A, signal intensities not only affect the *Z* basis QBER, but also need to optimize a trade-off between the

TABLE IV. Examples of optimal parameters for the seven-intensity protocol using simulation parameters from Table II. The numerical values are rounded to the accuracy of 0.001 in the table here. As can be observed, Alice and Bob's intensities compensate for channel asymmetry, while their intensity probabilities are mostly identical—since the intensities have already compensated for the asymmetry—despite having some numerical noise (as the key rate is not sensitive to the probabilities near the maximum, the algorithm satisfies with them having close enough, rather than perfectly identical, values, so the optimal values found are still slightly different even when $x = 1$). As we show in Sec. II, the optimal decoy-state intensity ratios are the same, i.e., $(\mu_A/\mu_B) = (\nu_A/\nu_B)$. Moreover, we can observe that the ratio of decoy-state intensities more closely follows $1/x$ than the ratio of signal intensities.

L_A	L_B	x	s_A	μ_A	ν_A	P_{s_A}	P_{μ_A}	P_{ν_A}	s_B	μ_B	ν_B	P_{s_B}	P_{μ_B}	P_{ν_B}	R
60 km	10 km	0.1	0.662	0.522	0.100	0.600	0.033	0.255	0.202	0.058	0.011	0.600	0.031	0.256	3.106×10^{-5}
60 km	30 km	0.25	0.593	0.457	0.089	0.581	0.036	0.266	0.294	0.125	0.024	0.580	0.034	0.269	1.445×10^{-5}
60 km	60 km	1	0.402	0.305	0.063	0.478	0.047	0.330	0.402	0.305	0.063	0.480	0.047	0.329	1.862×10^{-6}

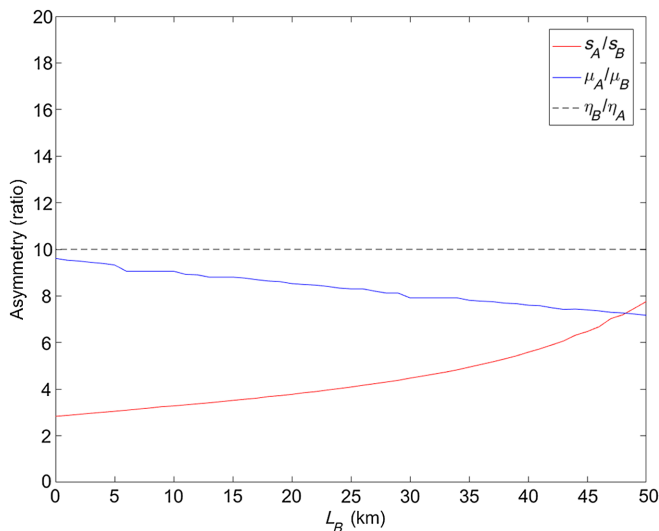


FIG. 6. Here we plot the ratios of signal intensities and decoy intensities versus distance, when the channel mismatch is fixed at $x = 0.1$ (i.e., $L_A = L_B + 50$ km). The simulation parameters are from Table II [this plot of intensities corresponds to the solid red key rate line in Fig. 5(d)]. We also include the line $(\eta_B/\eta_A) = 10$ for comparison. We can observe that the ratio of decoy states roughly follows η_B/η_A (to maintain good HOM interference visibility in the X basis), while the optimal ratio of signal intensities varies greatly between 1 (optimal for probability of sending single photons) and η_B/η_A (optimal for E_{ss}^Z). This is because signal states affect both key generation and error correction, so having similar intensities arriving at Charles after channel attenuation is not the only criteria for a good key rate, and optimal parameters do not necessarily satisfy $s_A\eta_A = s_B\eta_B$. In fact, since signal states in the Z basis are decoupled from the X basis, and E_{ss}^Z is less sensitive to unbalanced arriving intensities, s_A/s_B can be much more freely optimized between 1 and η_B/η_A , allowing seven-intensity protocol to have a higher key rate.

single-photon probabilities and error correction. This makes it usually not follow $s_A\eta_A = s_B\eta_B$. An illustration of the ratios of decoy intensities and signal intensities can be seen in Fig. 6.

Now, having demonstrated the new seven-intensity protocol, we proceed to introduce a powerful reality application for it: a scalable high-performance MDI-QKD network where any node can be dynamically added or deleted. We consider the channels from a real quantum network setup in Vienna reported in Ref. [16]. We focus here on the high-asymmetry nodes A_1, A_2, A_3, A_4, A_5 , as shown in Fig. 2(a). We find that our method leads to much higher key rates and allows easy dynamic addition or deletion of nodes. Since intensities can be independently optimized for each pair of channels, the establishment of new connections does not affect any existing connections, hence providing good scalability for the network (compared to, e.g., the case of using four-intensity protocol with the strategy of adding fibers, where each channel needs to accommodate for the *longest* link among all channels). See Appendix J for numerical results.

IV. CONCLUSION

In summary, we propose a method of effectively compensating for channel asymmetry in MDI-QKD by adjusting the two users' intensities and decoupling the two bases. Such a method can drastically increase the scenarios that MDI-QKD can be applied to while maintaining a good key rate. This study provides a powerful and robust software solution for a scalable and reconfigurable MDI-QKD network.

Our method is also a general result that can in principle be used, e.g., for various numbers of decoys or various finite-size analysis models (e.g., joint-bound analysis or composable security with Chernoff's bound). It is also potentially applicable to other types of quantum communication protocols such as twin-field QKD [28] or MDI quantum digital signature [26,27], which both use WCP sources and decoy-state analysis. We hope that our proposal can inspire more future work on the study of asymmetric protocols.

ACKNOWLEDGMENTS

This work is supported by the Natural Sciences and Engineering Research Council of Canada, U.S. Office of Naval Research, the Fundamental Research Funds for the Central Universities of China, National Natural Science Foundation of China Grant No. 61771443, and the China 1000 Young Talents Program. A key inspiration for this project comes from the study of free-space MDI-QKD. We thank the collaborators Dr. B. Qi and Professor G. Siopsis for helpful discussion and collaborative efforts in implementing free-space MDI-QKD, and we thank K. McBryde and S. Hammel from SPAWAR Systems Center, Pacific for kindly providing atmospheric data and for the helpful discussions.

Note added.—Recently, we experimentally implemented our protocol in Ref. [44] (see also Ref. [45]), thus clearly demonstrating the practicality of our work. We also recently applied our method proposed in this work to asymmetric twin-field QKD successfully; see Ref. [46]. See also Ref. [47], an independent work posted at the same time on the preprint server.

APPENDIX A: NOTE ABOUT ADDING FIBER

In this Appendix, we provide an intuitive description of why adding additional loss is suboptimal and how our method works better with asymmetric channels.

Previously, when Alice and Bob have asymmetric channels, a common solution is to add fiber (thus adding loss) to the shorter channel in exchange for better symmetry, such as in Ref. [15]. Afterwards, one selects symmetric intensities for Alice and Bob and acquires a higher rate. However, the added fiber lies in Bob's lab and is in fact securely under control of Bob. But by assuming a symmetric setup, we are effectively relinquishing its control to Eve and pessimistically estimating the key rate, as illustrated in Fig. 7. Therefore, intuitively, this is not

necessarily the optimal strategy. We show with our new protocol that when the channels are asymmetric, Alice and Bob can independently choose their optimal intensities, and that optimizing intensities and probabilities alone is sufficient to compensate for the different channel losses.

APPENDIX B: SCALING OF KEY RATE WITH TRANSMITTANCE

In this Appendix, we discuss the scaling properties of the key rate versus transmittance for prior protocols with the same parameters for Alice and Bob and our new protocol that uses different intensities for Alice and Bob. We show here and in Appendix C that the scaling of the key rate versus distance is mainly determined by the signal states (so long as we have a good single-photon estimation from the decoy states). This also means that the advantage of our method is really not dependent on the number of decoy states used or the finite-size analysis model used (or lack thereof, in the asymptotic case), and our results are in principle applicable to any protocol that decouples the signal and decoy states in the Z and X bases and allows different intensities for Alice and Bob.

The transmittance of the two channels are (η_A, η_B) , and the asymmetry (mismatch) x is defined as

$$x = \frac{\eta_A}{\eta_B}. \quad (\text{B1})$$

1. Single-photon source

Now, let us consider the single-photon case first. That is, suppose Alice and Bob both send perfect single photons only, and the key is generated from two-photon interference. If we ignore the dark counts, the asymptotic key rate can be written as [48]

$$R_{\text{SP}} = \eta_A \times \eta_B \times [1 - 2h_2(e_{11})], \quad (\text{B2})$$

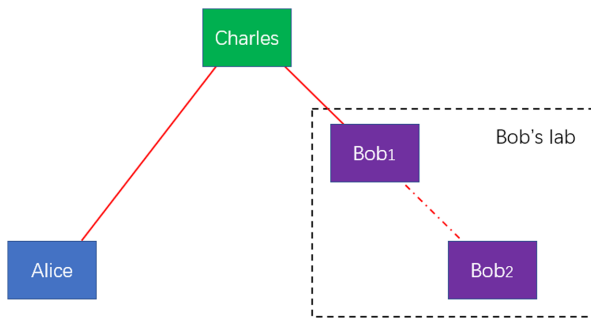


FIG. 7. Setup for asymmetric MDI-QKD. When channels are highly asymmetric (e.g., Alice and Bob1), to increase the symmetry in the channel, sometimes one adds additional loss to the system in Bob's lab [15] in exchange for better symmetry. When estimating the key rate, Bob assumes that both Charles-Bob1 and Bob1-Bob2 channels are controlled by Eve. This is therefore a pessimistic estimation of the key rate, and it is not necessarily the optimal strategy.

where h_2 is the binary entropy function, and e_{11} is the QBER (which is a quantity that, when the dark count rate is ignored is independent of the transmittance). Here we can see that in the perfect single-photon case, the key rate is proportional to $\eta_A \eta_B$, and the mismatch x does not explicitly appear in its expression:

$$R_{\text{SP}} \propto \eta_A \eta_B. \quad (\text{B3})$$

In fact, for a given total distance $L_A + L_B = L$, any positioning of the untrusted relay Charles (e.g., at the midpoint, in Alice's lab, or in Bob's lab) would not affect the key rate, since $\eta_A \eta_B$ depends only on L .

2. Weak coherent pulse source

The previous discussion for single-photon MDI-QKD suggests that, by nature, there is not really any limitation on symmetry for MDI-QKD, at least for the ideal single-photon case. Then, where does this dependence of the key rate on channel symmetry which we observe come from? In this section, we show that the scaling of the key rate depends on the signal states' trade-off between error correction and the probabilities of sending single photons when using WCP sources rather than privacy amplification (which depends on the estimation of single-photon contributions).

More concretely, (as we prove in the next section) for protocols with symmetric intensities, there are two sharp cutoff values for the mismatch x^{max} and x^{min} that prevent the protocol from acquiring any key rate when $x > x^{\text{max}}$ or $x < x^{\text{min}}$ (and optimizing identical intensities $s_A = s_B$ cannot circumvent this problem). This is why protocols such as the four-intensity protocol are limited to near-symmetric positions.

On the other hand, when a protocol allows independent intensities for Alice and Bob (such as our new seven-intensity protocol described in the main text), we show that the mismatch can always be compensated by optimizing intensities s_A and s_B (hence lifting the limitations x^{max} and x^{min}). In fact, we show that for positions with high asymmetry, the key rate no longer depends on mismatch $x = (\eta_A/\eta_B)$ at all, and the optimal key rate scales only with the *smaller* of the two channel transmittances. That is,

$$R_{\text{optimal}} \propto \min(\eta_A^2, \eta_B^2), \quad (\text{B4})$$

which means that the biggest advantage of protocols with independent intensities for Alice and Bob (e.g., seven-intensity protocol) is to completely lift the limitation on channel asymmetry. When compared with adding fiber to maintain asymmetry, we see that its scaling property is still the same, i.e., quadratically related to the (smaller of) channel transmittances, although our method will always perform better (by a constant coefficient) than adding fiber. Moreover, it provides the convenience of not needing

additional fiber, which may not be feasible in free-space channels or when channel mismatch is changing.

Proofs for the above scaling properties can be found in the next Appendix.

APPENDIX C: PROOF OF SCALING PROPERTIES OF THE KEY RATE WITH TRANSMITTANCE

In this Appendix, we outline the analytical proofs for the observations on the scaling properties of the asymptotic MDI-QKD key rate versus transmittance in the presence of asymmetry described in Appendix B. We also discuss how the finite-decoy and finite-size effects can be considered as imperfections in the infinite-decoy, infinite-data case, and that the scaling properties are still approximately the same, which are determined only by the signal states' trade-off between error correction and probabilities of sending single photons and not affected by decoy states.

To simplify the discussion, it is convenient to first use a few crucial approximations as described in Ref. [20].

- (1) We consider the asymptotic case with infinite data size.
- (2) We assume an infinite number of decoy states, i.e., Alice and Bob can perfectly estimate the single-photon gain Y_{11} and QBER e_{11} . In this case, Alice and Bob need only to choose appropriate signal intensities s_A, s_B .
- (3) We ignore the dark count rate Y_0 when studying the scaling properties with distance (as background noise affects only the maximum transmission distance where transmittance is at the same order as the dark count rate, but it does not affect the overall scaling properties of the key rate versus distance).
- (4) When describing the channel model to estimate the observable gain and QBER Q_{ss}^Z and E_{ss}^Z (which affect the error correction), we make second-order approximations to two functions:

$$\begin{aligned} I_0(x) &\approx 1 + \frac{x^2}{4} + O(x^4), \\ e^x &\approx 1 + x + \frac{x^2}{2} + O(x^3), \end{aligned} \quad (C1)$$

where I_0 is the modified Bessel function of the first kind. This approximation is relatively accurate when $s_A\eta_A\eta_d$ and $s_B\eta_B\eta_d$ are both small, where η_d is the detector efficiency.

With the above approximations, one can write the key rate conveniently as [excerpting Eqs. (C.1) and (C.2) from Ref. [20]],

$$R = \frac{\eta_B^2 \eta_d^2}{2} G(x, s_A, s_B), \quad (C2)$$

where $G(x, s_A, s_B)$ is a function determined by (s_A, s_B) and the asymmetry x only:

$$\begin{aligned} G(x, s_A, s_B) &= x s_A s_B e^{-(s_A+s_B)} \left[1 - h_2 \left(e_d - \frac{e_d^2}{2} \right) \right] \\ &\quad - \frac{2x s_A s_B + (s_B^2 + x^2 s_A^2)(2e_d - e_d^2)}{2} \\ &\quad \times f_e h_2 [E_{ss}^Z(x, s_A, s_B)] \\ E_{ss}^Z(x, s_A, s_B) &= \frac{(s_B + x s_A)^2 (2e_d - e_d^2)}{2[2x s_A s_B + (s_B^2 + x^2 s_A^2)(2e_d - e_d^2)]}, \end{aligned} \quad (C3)$$

where h_2 is the binary entropy function.

Now, having described the key rate function, we are interested in how it scales with the transmittances η_A, η_B using different optimization strategies for the intensities. We discuss two cases:

- (1) $R_{\text{symmetric}}$, where Alice and Bob use the same intensity $s = s_A = s_B$ and optimize s .
- (2) R_{optimal} , where Alice and Bob fully optimize a pair of intensities s_A, s_B , which can take different values.

1. Symmetrically optimized intensities

Let us consider the case where Alice and Bob use the same intensity $s = s_A = s_B$ and optimize s . This is the case discussed by previous protocols (such as the four-intensity protocol, although here to simplify the proof we focus on the infinite-decoy case and consider only signal intensities).

In this case, the function G is optimized over s (and is a function of x only). The rate satisfies

$$R_{\text{symmetric}} = \max_s R \propto \eta_B^2 \max_s G(x, s, s); \quad (C4)$$

therefore, $R_{\text{symmetric}}$ is proportional to η_B^2 when channel mismatch η_A/η_B is fixed.

Moreover, since $R_{\text{symmetric}}$ is also proportional to $G(x)$, we have $R_{\text{symmetric}} = 0$ if $G(x) = 0$. Note that we can rewrite the signal-state QBER E_{ss}^Z as

$$E_{ss}^Z(x) = \frac{(1+x)^2 (2e_d - e_d^2)}{2[2x + (1+x^2)(2e_d - e_d^2)]} \quad (C5)$$

since the equal intensities are canceled out; i.e., E_{ss}^Z is only a function of x . In fact, E_{ss}^Z is a function that minimizes at $x = 1$ and reaches 50% (where $R_{\text{symmetric}}$ is naturally zero) when $x \rightarrow 0$ or $x \rightarrow \infty$. Therefore, if $G(x) \neq 0$ at $x = 1$, there must exist some critical values of x^{max} and x^{min} which result in a sufficiently large QBER such that $G(x) = 0$ (and $R_{\text{symmetric}} = 0$). This means that $R_{\text{symmetric}}$ is quadratically related to η_B (or η_A) when mismatch η_A/η_B is fixed, but it also has two cutoff positions for critical levels of mismatch, beyond which no key can be generated. These two critical mismatch positions are what limit previous MDI-QKD protocols to near-symmetric positions. Also, as we have previously mentioned, we see that this critical dependence

on mismatch actually comes from the error-correction part (which involves E_{ss}^Z).

2. Fully optimized intensities

Now, let us consider the case where Alice and Bob are allowed to fully optimize their intensities s_A, s_B (such as in the seven-intensity protocol, although again, here we focus only on the signal states in the infinite-decoy case).

In this case, the function G is optimized over s_A, s_B . The rate satisfies

$$R_{\text{optimal}} = \max_{s_A, s_B} R \propto \eta_B^2 \max_{s_A, s_B} G(x, s_A, s_B). \quad (\text{C6})$$

Now, let us focus on the properties of $G(x, s_A, s_B)$. Looking at its expression Eq. (C3) in the previous section, we make the important observation that, except for the term $e^{-(s_A+s_B)}$ in the single-photon probabilities, every other term is only a function of s_B and xs_A (rather than x and s_A separately). We can rewrite $G(x, s_A, s_B)$ as

$$\begin{aligned} G'(x, s'_A, s_B) &= s'_A s_B e^{[-s'_A/x]} e^{-s_B} \left[1 - h_2 \left(e_d - \frac{e_d^2}{x} \right) \right] \\ &\quad - \frac{2s'_A s_B + (s_B^2 + s_A'^2)(2e_d - e_d^2)}{2} \\ &\quad \times f_e h_2 [E_{ss}^Z(s'_A, s_B)] \\ E_{ss}^Z(s'_A, s_B) &= \frac{(s_B + s'_A)^2 (2e_d - e_d^2)}{2[2s'_A s_B + (s_B^2 + s_A'^2)(2e_d - e_d^2)]}, \end{aligned} \quad (\text{C7})$$

where we define *equivalent intensity* s'_A as

$$s'_A = s_A \times x. \quad (\text{C8})$$

Moreover, if $\eta_A \gg \eta_B$ (i.e., mismatch $x \gg 1$), we can approximately assume that

$$e^{[-s'_A/x]} \approx 1, \quad (\text{C9})$$

which means that we can rewrite $\max_{s_A, s_B} G(x, s_A, s_B)$ as

$$G^{\text{max}} = \max_{s'_A, s_B} G'(s'_A, s_B), \quad (\text{C10})$$

which, importantly, is a constant value *not* dependent on the value of x when $x \gg 1$. The actual value of s_A equals

$$s_A = \frac{s'_A}{x}. \quad (\text{C11})$$

Physically, this means that when there is asymmetry between Alice and Bob's channels, we can compensate for this asymmetry by adjusting the intensities to keep the same "equivalent intensity" received by Charles and keep E_{ss}^Z at a low value. In this case, E_{ss}^Z is no longer limited by the mismatch x , and we can perform MDI-QKD at arbitrary values of asymmetry. Also, the key rate is now given by

$$R_{\text{optimal}} \propto \eta_B^2 G^{\text{max}}. \quad (\text{C12})$$

This means that when $\eta_A \gg \eta_B$ (e.g., the "single-arm" case previously mentioned where L_A is much shorter than L_B), the key rate of asymmetric MDI-QKD is related only to η_B and still quadratically scales with η_B . When $\eta_B \gg \eta_A$, though, we can rewrite $x' = (\eta_B/\eta_A)$ and rewrite

$$R_{\text{optimal}} \propto \eta_A^2 \max_{s'_B, s_A} G'(s'_B, s_A). \quad (\text{C13})$$

Therefore, overall,

$$R_{\text{optimal}} \propto \min(\eta_A^2, \eta_B^2). \quad (\text{C14})$$

Now, we plot the two cases (symmetric intensities and fully optimized intensities) in a contour plot. As we can observe in Fig. 8, the key rate $R_{\text{symmetric}}$ has two cutoff mismatch positions beyond which the key rate is zero. This limitation is removed when full optimization of intensities is implemented. Moreover, for R_{optimal} , we see that the contours are perpendicular to the axes in high-asymmetry regions, which means that the key rate scales only with the longer of the two channels.

Also, note that from Eqs. (C4) and (C6), we can make the observation that there is never any need to add fiber to the shorter channel when fully optimizing the intensities, and our new method always provides a higher key rate than the prior art technique of adding fiber until channels are symmetric, while using same intensities for Alice and Bob.

To show this point, consider the system having a fixed longer channel L_B (i.e., suppose η_B is fixed and $\eta_A > \eta_B$, $x = (\eta_A/\eta_B) > 1$). Adding loss to η_A is equivalent to decreasing x .

With symmetric intensities (and adding loss till $\eta_A = \eta_B$), the key rate can be written as

$$R_{\text{symmetric}} = \frac{\eta_d^2 \eta_B^2}{2} \max_s G(1, s, s). \quad (\text{C15})$$

Suppose we fully optimize the intensities for this case with added fiber, we will obtain the same key rate (since for $x = 1$, i.e., symmetric setup, the optimal choice of intensities satisfies $s_A = s_B$):

$$\max_s G(1, s, s) = \max_{s_A, s_B} G(1, s_A, s_B). \quad (\text{C16})$$

However, let us compare it with the case of using fully optimized intensities and no additional loss:

$$R_{\text{optimal}} = \frac{\eta_d^2 \eta_B^2}{2} \max_{s_A, s_B} G(x, s_A, s_B). \quad (\text{C17})$$

As we describe in Eq. (C7), we can rewrite $G(x, s_A, s_B)$ as $G'(x, s'_A, s_B)$ (recall that the equivalent intensity s'_A is defined as xs_A). We make the observation that $G'(x, s'_A, s_B)$

strictly increases with x . That is, for any two given values of s'_A, s_B , and $x > 1$,

$$G'(x, s'_A, s_B) > G'(1, s'_A, s_B); \quad (\text{C18})$$

hence, after optimization we also have

$$\max_{s'_A, s_B} G'(x, s'_A, s_B) > \max_{s'_A, s_B} G'(1, s'_A, s_B), \quad (\text{C19})$$

which means that when fully optimizing Alice and Bob's intensities (which already compensate for the mismatch between channels), it is always optimal not to add any additional loss to the channels. Moreover, combining Eqs. (C15)–(C17) and (C19), we can see that

$$\begin{aligned} R_{\text{optimal}} &= \frac{\eta_d^2 \eta_B^2}{2} \max_{s_A, s_B} G(x, s_A, s_B) \\ &> \frac{\eta_d^2 \eta_B^2}{2} \max_s G(1, s, s) = R_{\text{symmetric}}. \end{aligned} \quad (\text{C20})$$

That is, compared to the case where one adds loss to η_A until $\eta_A = \eta_B$, our new protocol always provides a higher key rate as long as the channels are asymmetric. Intuitively, this is because adding fiber while using the same intensities for Alice and Bob is in fact an suboptimal subset of the overall set of strategies Alice and Bob can take (which includes adjusting Alice and Bobs intensities independently, as well as adding any length of fibers to change x). Even when considering adding fiber as one of the valid variables, we show that the optimal point always happens when no fiber is added. Therefore, our method is a better optimized strategy than adding fiber because it considers a larger parameter space.

Note that fully optimizing Alice and Bob's intensities does not change the fundamental scaling property; the key rate is still *quadratically* related to transmittance in the longer arm. However, it always provides a better key rate than prior art techniques, and it also offers the great convenience of not having to physically add loss to the channels and being able to implement everything in software.

3. Practical imperfections

Up to here, we have analytically shown how choosing to fully optimize the intensities can affect the key rate for the asymptotic, infinite-decoy case. The behavior of contours as shown in Fig. 8 is a result of s_A, s_B compensating for the difference in channel loss. However, we have so far assumed perfect knowledge of single-photon contributions and have not yet discussed the decoy-state intensities. Moreover, nonideal experimental parameters (including dark count rate and detector efficiency) and finite-size effects will both affect the key rate. Here in this subsection, we compare the key rate under more practical assumptions and show that the above factors can be considered as *imperfections* that reduce the key rate but maintain similar contour shapes and

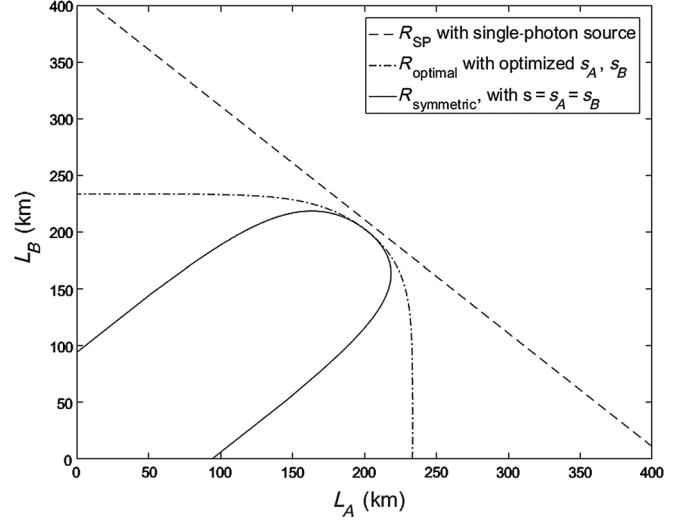


FIG. 8. Rate vs distance contours for single-photon MDI-QKD R_{SP} , decoy-state MDI-QKD with symmetric intensities $R_{\text{symmetric}}$, and with fully optimized intensities R_{optimal} . We plot the contour line of $R = 10^{-9.5}$. Here, for a better comparison with WCP sources, we arbitrarily set a probability $P_{11} = s_A s_B \times e^{-(s_A + s_B)}$ (where $s_A = s_B = 0.6533$) of single photon pairs being sent when calculating R_{SP} . For the decoy-state case, as we describe at the beginning of this appendix we assume infinite decoys, infinite data size, ignore dark count rate, and take second-order approximation when calculating gain and QBER (so that we focus only on the ideal scaling properties of the key rate with distance). As can be seen, R_{SP} is not limited by asymmetry and takes a constant value for any fixed $L_A + L_B$ (meaning that the dependence of the key rate on asymmetry does not come from single-photon contributions in the privacy amplification part when using WCP sources). For decoy-state MDI-QKD, we can clearly see $R_{\text{symmetric}}$ being limited by the two cutoff lines where $|L_A - L_B|$ takes maximum value (which corresponds to critical values of channel mismatch x^{max} and x^{min}). On the other hand, R_{optimal} is not limited by asymmetry and has contours nearly perpendicular to the axes when asymmetry is high (meaning that, when one channel is significantly longer than the other, R_{optimal} is dependent only on the longer channel).

scaling properties for the key rate; that is, we still observe a high dependence on asymmetry for protocols with identical intensities for Alice and Bob, and fully optimizing intensities can completely lift this limitation.

In practice, with a finite number of decoys (for instance, for four-intensity and seven-intensity protocols, where Alice and Bob each choose three decoy intensities μ, ν, ω), the estimation of Y_{11} and e_{11} is not perfect; therefore, the key rate will be slightly lower than the aforementioned infinite-decoy case. Moreover, to accurately estimate Y_{11} and e_{11} , the decoy intensities need to be optimized to compensate for channel loss, too. As we describe in Sec. II A in the main text, the decoy states should maintain balanced arriving intensities at Charles (e.g., $\mu_A \eta_A = \mu_B \eta_B$) to ensure good HOM visibility and low QBER in the X basis. Note that the optimization of decoy

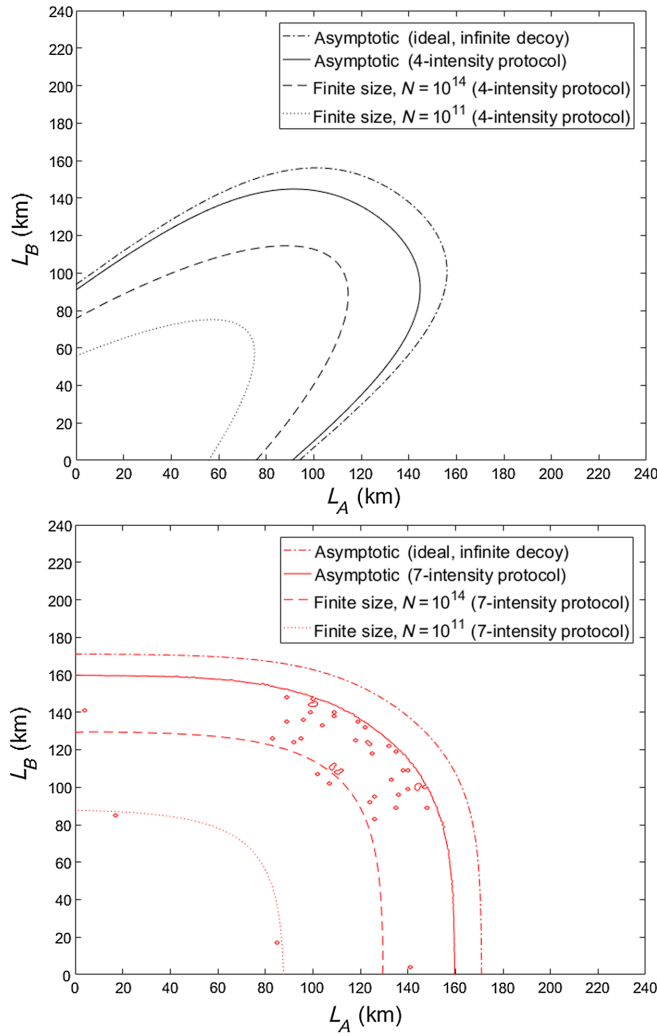


FIG. 9. Contours of rate vs distance for decoy-state MDI-QKD under different assumptions for practical imperfections, for the key rates for the asymptotic case with infinite decoys (and ideal assumption of zero dark count rate and 100% detector efficiency), asymptotic case with four-intensity or seven-intensity protocol (with practical device parameters), and finite-size case with four-intensity or seven-intensity protocol. Top: Protocols with identical intensities for Alice and Bob. Bottom: Protocols with fully optimized intensities. (Note that in the bottom plot there are some noises in the asymptotic seven-intensity protocol key rate. This is because the optimal ν can take a very small value in the ideal case where data size is infinitely large, which results in some numerical noise in computer simulations). We plot the contour lines of $R = 10^{-7}$. As can be observed here, the finite number of decoys, the nonideal experimental parameters, and the finite-size effects are all imperfections that reduce the key rate. However, the overall shapes of the contours still remain largely the same, which follow the upper bounds given by the ideal infinite-decoy case. (Except for four-intensity protocol under finite-size effects, which no longer has two clear cutoff mismatch positions, but it is still severely limited by channel asymmetry, while seven-intensity protocol lifts this constraint completely).

intensities has a very different purpose from that of the signal intensities s_A, s_B ; the signal intensities are optimized so as to reduce E_{ss}^Z (while keeping single-photon probability $s_A s_B e^{-(s_A+s_B)}$ high) and maximize the key rate, while the decoy intensities are optimized to estimate Y_{11}^L and e_{11}^U as accurately as possible, whose ideal values Y_{11} and e_{11} (used in the infinite-decoy case above) provide an upper bound for the practical key rate with a finite number of decoys. As we see in Fig. 9, the asymptotic key rate with a finite number of decoys follows a similar shape as its upper bound, the infinite-decoy case.

Additionally, the detector efficiency (which is equivalent to channel loss) contributes to a uniformly shifted key rate in both L_A and L_B directions, while dark counts reduce the key rate more significantly in the higher loss region (both of which we have so far ignored in the ideal case as we describe at the beginning of this section). However, as we observe in Fig. 9 (the solid lines consider both finite-decoys and practical parameters), these factors do not change the overall shape of the contours either.

Lastly, finite-size effects will reduce the key rate significantly. As we observe in Fig. 9 bottom plot, while the key rate is reduced, the contour shapes remain largely unchanged (meaning that even under finite-size effects, the seven-intensity protocol can still effectively compensate for channel asymmetry effectively). In Fig. 9 top plot, we can find similar observations that finite-size effects reduces the overall key rate. However, note that under finite-size effects, the shapes of the key rate contours for the four-intensity protocol are somewhat different and no longer follow the two cutoff positions $x^{\text{upper}}, x^{\text{lower}}$ for channel mismatch (which appear as straight lines in, e.g., Fig. 8). This is because, though the key rate is still limited by E_{ss}^Z (which causes the cutoff mismatch positions), it is also limited by the estimation of Y_{11}^L and e_{11}^U using the decoy states. Compared to the asymptotic case, here under finite-size effects, the increased e_{11}^U is likely a more severe limiting factor than E_{ss}^Z , and not being able to choose independent intensities for Alice and Bob prevents an accurate estimation of Y_{11}^L and e_{11}^U (due to poor HOM visibility in the X basis caused by unbalanced intensities). Therefore, here the dependence of the key rate on channel asymmetry is present in both privacy amplification and error-correction terms, and the shapes of contours are a result of both effects. (The difference in contour shape from the infinite-decoy case is more prominent for the finite-size case, likely because the key rate is more sensitive to e_{11}^U here). Importantly, under finite-size effects, the key rate for four-intensity protocol is still highly limited by channel asymmetry, while seven-intensity protocol completely removes such a constraint and allows two channels with arbitrary asymmetry between them.

APPENDIX D: NOTE ABOUT DECOUPLING SIGNAL AND DECOY INTENSITIES

In this Appendix, we provide a simple intuitive explanation for why our protocol provides a better choice of decoy and signal intensities.

Let us recall again the key rate formula of MDI-QKD [8,30]:

$$R = P_{s_A} P_{s_B} \{ (s_A e^{-s_A}) (s_B e^{-s_B}) Y_{11}^{X,L} [1 - h_2(e_{11}^{X,U})] - f_e Q_{ss}^Z h_2(E_{ss}^Z) \}. \quad (D1)$$

Here, there are three criteria that determine whether a MDI-QKD protocol generates good key rate in the presence of channel asymmetry:

- Similar arriving intensities at Charles in the X basis in order to have good HOM interference and keep QBER low in the X basis (which is important for a good estimation of $e_{11}^{X,U}$).
- Similar arriving intensities at Charles in the Z basis in order to keep QBER E_{ss}^Z low in the Z basis (which is due to misalignment), although this term is much less sensitive to difference in intensities than (a).
- A high enough probability of sending single photons $s_A e^{-s_A} s_B e^{-s_B}$. Note that both criteria (b) and (c) involve the signal states s_A, s_B , so there is a trade-off between (b) and (c).

Prior protocols require Alice and Bob to use the same set of intensities. This overly constrains the solution space [because Alice and Bob try to use the same set of intensities to satisfy (a)–(c) simultaneously] and leaves high QBER in both the X and Z bases and thus resulting in a low key rate when channels are asymmetric.

By relaxing this constraint (allowing Alice and Bob to have different intensities) and decoupling criteria (a) from criteria (b) and (c) by allowing independent decoy and signal intensities, we can satisfy (a) nicely, while simultaneously achieving a good trade-off between (b) and (c), hence ensuring a high key rate.

Remark: For more detail on the trade-off between (b) and (c), here (b) is optimal when arriving intensities are matched, i.e., $s_A/s_B = \eta_B/\eta_A$, and (c) is independent of asymmetry and is optimal when signal intensities are both 1. In fact, since E_{ss}^Z is much less sensitive to s_A/s_B , such a trade-off between two terms favors (c) more than (b); thus, the optimal s_A/s_B is often closer to 1 than η_B/η_A . The actual optimal signal intensities can be found by numerical optimization, as we describe in Sec. II D. An example of μ_A/μ_B and s_A/s_B can also be seen in Fig. 6, where we observe that μ_A/μ_B follows η_B/η_A rather closely, while s_A/s_B has relatively much more freedom in its optimization (between 1 and η_B/η_A).

APPENDIX E: GENERALITY OF OUR METHOD: MDI-QKD PROTOCOLS OTHER THAN THREE DECOY STATES

In the main text, we focus on the seven-intensity protocol, where Alice and Bob each use one signal intensity s_A (s_B), and three decoy intensities μ_A, ν_A, ω (μ_B, ν_B, ω). However, the core idea of our protocol lies in two key points: (1) the X and Z bases are decoupled, where decoy states in the X basis bound Eve’s information and the signal state in the Z basis encodes the key, and (2) Alice and Bob use different intensities to compensate for channel asymmetry. This means that our protocol is not limited to the seven-intensity protocol, but can easily be applied to other protocols, too, as long as points (1) and (2) are satisfied.

In this Appendix, we demonstrate the generality of our method by actually applying it to other kinds of MDI-QKD protocols where Alice and Bob use a different number of decoy intensities in the X basis and show that similar advantages as with the seven-intensity protocol can be observed when using asymmetric intensities. We also show with numerical results, that although these alternative protocols will certainly work, the seven-intensity protocol provides a good balance between performance and ease of experimental implementation.

Here we compare three cases:

- Alice and Bob each use two decoy intensities μ_A, ν_A (μ_B, ν_B) in the X basis. We denote this case as a six-intensity protocol (including the two signal intensities), where the parameter choices are

$$[s_A, \mu_A, \nu_A, P_{s_A}, P_{\mu_A}, P_{\nu_A}, s_B, \mu_B, \nu_B, P_{s_B}, P_{\mu_B}, P_{\nu_B}]. \quad (E1)$$

Here, $P_{\nu_A} = 1 - P_{s_A} - P_{\mu_A}$ and $P_{\nu_B} = 1 - P_{s_B} - P_{\mu_B}$. This is similar to the “one-decoy” setup that was discussed in Ref. [29]. Note that here it is not a “five-intensity” protocol because using μ_A, μ_B, ω alone is not sufficient to satisfactorily bound the single-photon contributions and will result in a low or zero key rate. Therefore, in this setup, the vacuum state is not used, and Alice and Bob each use two nonzero decoy states.

- Alice and Bob each use three decoy intensities μ_A, ν_A, ω (μ_B, ν_B, ω) in the X basis. This case is the seven-intensity protocol that we discuss in the main text, where the parameter choices are

$$[s_A, \mu_A, \nu_A, \omega, P_{s_A}, P_{\mu_A}, P_{\nu_A}, P_{\omega_A}, s_B, \mu_B, \nu_B, \omega, P_{s_B}, P_{\mu_B}, P_{\nu_B}, P_{\omega_B}]. \quad (E2)$$

Here, $P_{\omega_A} = 1 - P_{s_A} - P_{\mu_A} - P_{\nu_A}$, $P_{\omega_B} = 1 - P_{s_B} - P_{\mu_B} - P_{\nu_B}$, and ω is the vacuum state (for simplicity, we can assume $\omega = 0$).

- (3) Alice and Bob each use four decoy intensities $\mu_A, \nu_A, \nu_{2A}, \omega$ ($\mu_B, \nu_B, \nu_{2B}, \omega$) in the X basis. We denote this case as a nine-intensity protocol, where the parameter choices are

$$[s_A, \mu_A, \nu_A, \nu_{2A}, \omega, P_{s_A}, P_{\mu_A}, P_{\nu_A}, P_{\nu_{2A}}, P_{\omega_A}, \\ s_B, \mu_B, \nu_B, \nu_{2B}, \omega, P_{s_B}, P_{\mu_B}, P_{\nu_B}, P_{\nu_{2B}}, P_{\omega_B}]. \quad (\text{E3})$$

Here, $P_{\omega_A} = 1 - P_{s_A} - P_{\mu_A} - P_{\nu_A} - P_{\nu_{2A}}$, $P_{\omega_B} = 1 - P_{s_B} - P_{\mu_B} - P_{\nu_B} - P_{\nu_{2B}}$, and ω is the vacuum state.

Note that in all of these three protocols, the key rate formula stays the same as Eq. (1):

$$R = P_{s_A} P_{s_B} \{ (s_A e^{-s_A}) (s_B e^{-s_B}) Y_{11}^{X,L} [1 - h_2(e_{11}^{X,U})] \\ - f_e Q_{ss}^Z h_2(E_{ss}^Z) \}. \quad (\text{E4})$$

What is changing here is only the estimation of the single-photon contributions, namely, the yield $Y_{11}^{X,L}$ and QBER $e_{11}^{X,U}$. While we have analytical bounds for the decoy-state analysis [20] for the seven-intensity protocol, we use a linear programming approach to numerically estimate $Y_{11}^{X,L}$ and QBER $e_{11}^{X,U}$ in the six-intensity and nine-intensity cases. Such an approach has been widely discussed in the literature, as in Refs. [29,49,50].

Now, we perform numerical simulations with the six-intensity, seven-intensity, and nine-intensity protocols and show that they all have much higher performance than their symmetric intensity counterparts when channel asymmetry is present. This demonstrates the generality of our method, as using asymmetric intensities can always improve the performance of MDI-QKD with asymmetric channels.

We also compare the performances of the three protocols with each other and show that using more decoy intensities can always guarantee higher or equal performance than using fewer decoy intensities, regardless of data size and asymmetry. The seven-intensity protocol never provides a smaller key rate than the six-intensity protocol, and although the nine-intensity protocol can potentially provide an even higher key rate, the advantage is small, and the seven-intensity protocol that we use in the main text is a good balance between key rate performance and ease of experimental implementation.

Interestingly, as observed in Figs. 10(a) and 10(b), for the six-intensity and nine-intensity protocols, although the yield $Y_{11}^{X,L}$ and QBER $e_{11}^{X,U}$ are estimated numerically using linear programming, there is still a ridge (discontinuity in first-order derivatives) along $(\mu_A/\mu_B) = (\nu_A/\nu_B)$, and $(\nu_A/\nu_B) = (\nu_{2A}/\nu_{2B})$ as we see for the seven-intensity protocol in the main text. For the six-intensity protocol, the ridge is very clearly shown. For the nine-intensity protocol, the ridge exists but it is less prominent and sometimes not visible (likely because, e.g., if two pairs of proportional decoy states $(\nu_A/\nu_B) = (\nu_{2A}/\nu_{2B})$ already

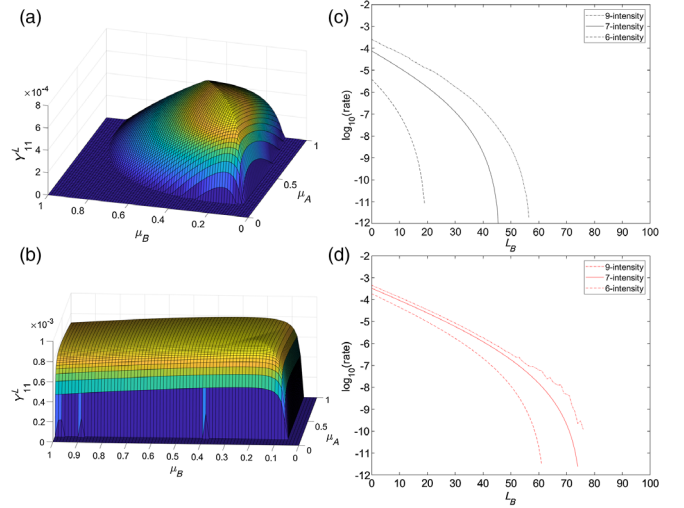


FIG. 10. Left: Examples of $Y_{11}^{X,L}$ versus μ_A and μ_B where other parameters are all fixed for (a) six-intensity protocol and (b) nine-intensity protocol. Just like for the seven-intensity protocol, we can see a ridge along $(\mu_A/\mu_B) = (\nu_A/\nu_B)$ or $(\mu_A/\mu_B) = (\nu_A/\nu_B) = (\nu_{2A}/\nu_{2B})$. Note that the ridge is a lot less obvious for the nine-intensity protocol (and sometimes is not visible), likely because two proportional pairs of decoy states can estimate single-photon contribution reasonably well, so the third pair (μ_A, μ_B) here has more freedom in the choice of intensities. Right: Comparison of rate vs L_B for six-intensity, seven-intensity, and nine-intensity protocols, where mismatch $x = 0.1$, i.e., $L_A = L_B + 50$ km (assuming fiber loss 0.2 dB/km). The rates are plotted in log scale. We use the parameters from Table II and $N = 10^{12}$. (c) Using symmetric intensities for Alice and Bob; (d) using fully optimized asymmetric parameters for Alice and Bob. As can be seen, using asymmetric intensities can greatly improve the key rate for all three protocols when channel asymmetry is present. Note that there is a higher amount of noise present for the nine-intensity case due to the numerical instability brought by linear program solvers (similar to that of joint-bound finite-size analysis, which we discuss in Appendix H), but the key points here are that the nine-intensity protocol also benefits considerably from using asymmetric intensities, and that the nine-intensity protocol does not have a significant advantage over the seven-intensity protocol despite being more complex to implement.

provide good estimation of $Y_{11}^{X,L}$, $e_{11}^{X,U}$, the third pair μ_A, μ_B has more freedom, and would not affect the decoy-state analysis or the key rate too much even if it does not provide good HOM visibility and results in high $E_{\mu\mu}^X$).

We plot the simulated key rate for the protocols in Figs. 10(c) and 10(d). We first consider a similar scenario as Figs. 5(c) and 5(d) using parameters from Table II, a channel mismatch of $(\eta_A/\eta_B) = x = 0.1$, and data size of $N = 10^{12}$ (here we use a larger data size than in the main text, since for $N = 10^{11}$ the six-intensity protocol with symmetric intensities cannot generate a key rate even at $L_B = 0$ km, so a comparison is not immediately clear in the plot). As we show in Fig. 10, for each protocol, allowing asymmetric intensities provides a much higher key rate

than using symmetric intensities only, demonstrating the general effectiveness of our method for different protocols under channel asymmetry.

We also make an important observation here: The more decoy intensities one uses, the higher key rate one can obtain after parameter optimization, even with finite-size effects considered (e.g., the nine-intensity protocol always has a higher key rate than the seven-intensity, and the seven-intensity also always has a higher rate than the six-intensity). This is because, for instance, the six-intensity protocol can in fact be considered a special case of the seven-intensity protocol, just with P_{ω_A} and P_{ω_B} infinitely close to zero, and with nine instead of four constraints, e.g., the gains Q_{ij}^X when estimating $Y_{11}^{X,L}$. With close to zero data, the five new constraints are obviously very loose (with very large finite-size fluctuation) and will not provide any useful information, but the key point is, in a linear program these loose constraints *will not decrease* the key rate. Therefore, any optimal set of parameters for the six-intensity protocol can also be considered a valid set of parameters for the seven-intensity protocol; i.e., the parameter space of the six-intensity protocol is a subset of that of the seven-intensity protocol, and the latter protocol always provides *no smaller* key rate than the former (and often the seven-intensity protocol can find a better parameter set in the larger parameter space, resulting in a higher key rate).

The same goes for the nine-intensity protocol, but as we see in Fig. 10, the advantage it provides over the seven-intensity protocol is rather small (compared to, e.g., six-intensity versus seven-intensity), while requiring more complex control of the intensity modulators in the experimental setup and more complicated data collection and analysis: The users need to collect 16 sets of gains and error gains, and the parameter optimization is also a lot slower and more unstable (evaluating the linear program is on average slower than analytical expression by about 50 times, and linear programs also introduce numerical instabilities). Similar observations have been made for the symmetric case in Ref. [29] (although in this paper, the signal states are not decoupled from decoy states, so the protocols are slightly different), that using decoy states $\{\mu, \nu, \omega\}$ provides a higher key rate than $\{\mu, \nu\}$, but adding one more decoy state ν_2 provides little further advantage.

Therefore, our conclusion is that, while our method of asymmetric intensities and decoupled bases surely works well with other protocols such as the six-intensity and nine-intensity protocols, the seven-intensity protocol we introduce in the main text strikes a good balance between key rate performance and the ease of both experimental implementation and data analysis.

APPENDIX F: DECOY-STATE INTENSITIES

In this Appendix, we describe Theorems I and II in more detail and show their theoretical proofs in the asymptotic

limit of infinite data size (moreover, numerically, we find that Theorems I and II in fact hold true even under finite-size effects).

1. Symmetry of Optimal Decoy Intensities

To prove Theorem I, here we actually propose an even stronger assumption for μ_A, μ_B :

Theorem III: For any arbitrary choice of device and channel parameters and any two given values of ν_A, ν_B , the optimal decoy intensities $\mu_A^{\text{opt}}, \mu_B^{\text{opt}}$ that maximize R always satisfy the constraint

$$\frac{\mu_A^{\text{opt}}}{\mu_B^{\text{opt}}} = \frac{\nu_A}{\nu_B}. \quad (\text{F1})$$

Remark: As we show below, Theorem I is simply a corollary of Theorem III.

Proof for Theorem III: Here for convenience, we first limit the discussion to the asymptotic case (i.e., infinite data size), and we assume that the vacuum intensity is indeed $\omega = 0$. Throughout the rest of the text, we use Q_{ij}^k and E_{ij}^k to denote the observed gain and QBER, where, if not specified, the first subscript is Alice's intensity, and the second is Bob's intensity, which can be chosen from $\{s_A, \mu_A, \nu_A, \omega\}$ and $\{s_B, \mu_B, \nu_B, \omega\}$ for Alice and Bob, respectively. The superscript k signifies the basis X or Z (although, here we only explicitly write the basis for illustration purposes, since the basis is already implied by the choice of intensities).

First, looking at the key rate expression [8,30]

$$R = P_{s_A} P_{s_B} \{ (s_A e^{-s_A}) (s_B e^{-s_B}) Y_{11}^{X,L} [1 - h_2(e_{11}^{X,U})] - f_e Q_{ss}^Z h_2(E_{ss}^Z) \}, \quad (\text{F2})$$

we can see that the term $Y_{11}^{X,L} [1 - h_2(e_{11}^{X,U})]$, i.e., the decoy-state analysis and privacy amplification, is determined by the decoy intensities (and probabilities, if finite-size effect is considered) only and not affected by the signal intensities s_A, s_B . This is an important and very convenient characteristic of the four-intensity and seven-intensity protocols, that the signal state is concerned only with key generation, while the decoy states are responsible only for privacy amplification. That is, the optimization of decoy-state intensities is decoupled from the key generation. Now, we can make an observation that, under given device parameters and channel loss, the optimization of the decoy intensities is independent of s_A, s_B , and its only goal is to maximize $Y_{11}^{X,L} [1 - h_2(e_{11}^{X,U})]$. Furthermore, to perform the decoy-state analysis, we note that the upper bound for single-photon QBER satisfies the form of

$$e_{11}^{X,U} = f(Y_{11}^{X,L}, \nu_A, \nu_B), \quad (\text{F3})$$

where $e_{11}^{X,U}$ is determined only by $Y_{11}^{X,L}$, ν_A , and ν_B . The full expression, as in Ref. [20], is

$$e_{11}^{X,U} = \frac{1}{\nu_A \nu_B Y_{11}^{X,L}} (e^{\nu_A + \nu_B} Q_{\nu\nu}^X E_{\nu\nu}^X - e^{\nu_A} Q_{\nu\omega}^X E_{\nu\omega}^X - e^{\nu_B} Q_{\omega\nu}^X E_{\omega\nu}^X + Q_{\omega\omega}^X E_{\omega\omega}^X). \quad (\text{F4})$$

Now, suppose we first fix two arbitrary values of ν_A , ν_B and try to maximize $Y_{11}^X [1 - h_2(e_{11}^X)]$ by optimizing μ_A , μ_B , we can see that maximizing $Y_{11}^{X,L}$ will suffice, since it will simultaneously minimize $e_{11}^{X,U}$, whose only component dependent on μ_A , μ_B is $Y_{11}^{X,L}$. The question now becomes simply finding

$$(\mu_A^{\text{opt}}, \mu_B^{\text{opt}}) = \arg \max [Y_{11}^{X,L}(\mu_A, \mu_B)]. \quad (\text{F5})$$

A very important characteristic of $Y_{11}^{X,L}$ is that its expression is dependent upon whether $(\mu_A/\mu_B) \leq (\nu_A/\nu_B)$; i.e., it is a piecewise function, as described in Ref. [20].

Case 1: If $(\mu_A/\mu_B) \leq (\nu_A/\nu_B)$:

$$Y_{11}^{X,L} = Y_{11}^a = \frac{1}{(\mu_A - \nu_A)} \left[\frac{\mu_A}{\nu_A \nu_B} Q_{\nu\nu}^{M1} - \frac{\nu_A}{\mu_A \mu_B} Q_{\mu\mu}^{M2} \right]. \quad (\text{F6})$$

Case 2: Otherwise, if $(\mu_A/\mu_B) > (\nu_A/\nu_B)$:

$$Y_{11}^{X,L} = Y_{11}^b = \frac{1}{(\mu_B - \nu_B)} \left[\frac{\mu_B}{\nu_A \nu_B} Q_{\nu\nu}^{M1} - \frac{\nu_B}{\mu_A \mu_B} Q_{\mu\mu}^{M2} \right], \quad (\text{F7})$$

where we denote the two expressions of $Y_{11}^{X,L}$ in the two cases as Y_{11}^a and Y_{11}^b , and the two terms $Q_{\nu\nu}^{M1}$ and $Q_{\mu\mu}^{M2}$ are linear combinations of the observable gain and are functions of (ν_A, ν_B) and (μ_A, μ_B) only, respectively. Their full expressions can be found in Appendix F 3. Also, note that if $(\mu_A/\mu_B) = (\nu_A/\nu_B)$, the two cases $Y_{11}^a = Y_{11}^b$.

Now, we can make a key observation that in case 1, for any given μ_A , the partial derivative $(\partial Y_{11}^a)/(\partial \mu_B)$ always satisfies

$$\frac{\partial Y_{11}^a}{\partial \mu_B} < 0. \quad (\text{F8})$$

(The actual expression of the partial derivative and proof of its positivity are shown in Appendix F 3). However, in case 1, μ_B is bounded by $\mu_B \geq [(\mu_A \nu_B)/\nu_A]$, so the only optimal case is to take the boundary condition

$$\mu_B^{\text{opt}} = \frac{\mu_A \nu_B}{\nu_A}. \quad (\text{F9})$$

This means that in the region of $(\mu_A/\mu_B) \leq (\nu_A/\nu_B)$, any pair of optimal values $(\mu_A^{\text{opt}}, \mu_B^{\text{opt}})$ must satisfy $(\mu_A^{\text{opt}}/\mu_B^{\text{opt}}) = (\nu_A/\nu_B)$, or else we can always decrease μ_B to get a higher rate, meaning that the previous point is not the actual maximum. We illustrate this behavior in Fig. 11.

Similarly, for case 2, the partial derivative with respect to μ_A satisfies

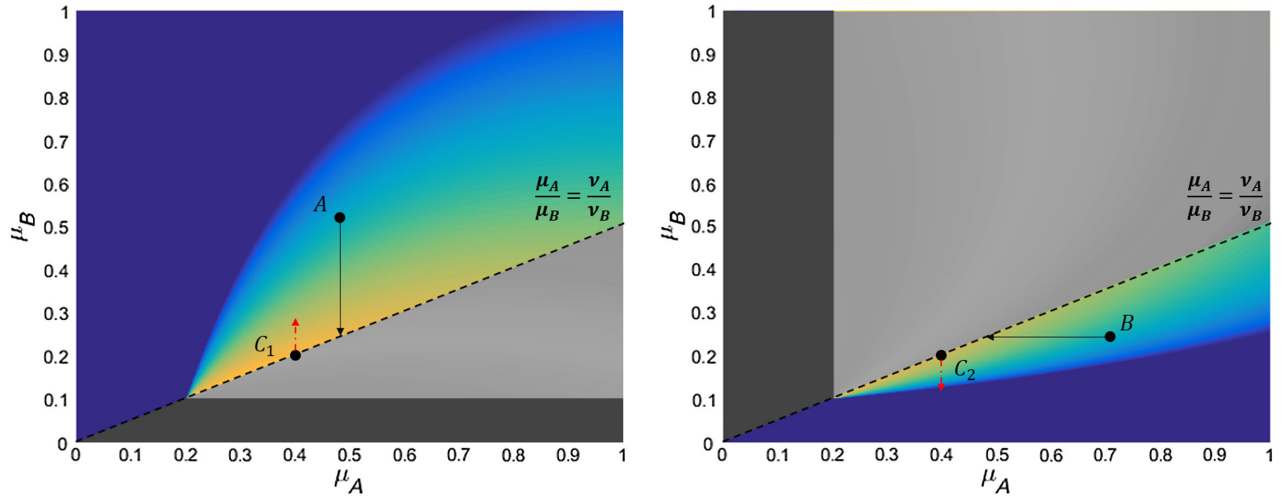


FIG. 11. An example of the two different cases of the $Y_{11}^{X,L}$ vs μ_A , μ_B function for fixed values of $\nu_A = 0.2$, $\nu_B = 0.1$. Left: Y_{11}^a , where $(\mu_A/\mu_B) \leq (\nu_A/\nu_B)$ (case 1). Right: Y_{11}^b , where $(\mu_A/\mu_B) > (\nu_A/\nu_B)$ (case 2). Allowed regions are marked in color for both cases. In case 1, we show that $[(\partial Y_{11}^a)/(\partial \mu_B)] < 0$, so any given point A can descend along the μ_B axis (the solid black arrow) to get a higher rate, until it reaches the boundary line $(\mu_A/\mu_B) = (\nu_A/\nu_B)$ where μ_B is highest. Similarly, in case 2, $[(\partial Y_{11}^b)/(\partial \mu_A)] < 0$, so any given point B can descend along the μ_A axis until $(\mu_A/\mu_B) = (\nu_A/\nu_B)$ to get the highest rate. Therefore, the optimal $(\mu_A^{\text{opt}}, \mu_B^{\text{opt}})$ that maximize the piecewise function $Y_{11}^{X,L}$ always occur on the boundary line. Moreover, for any given point $C(\mu_A, \mu_B)$ on the boundary line, the function values of Y_{11}^a , Y_{11}^b are the same. However, we show that $(\partial Y_{11}^a)/(\partial \mu_A)$ at C_1 is not equal to $(\partial Y_{11}^b)/(\partial \mu_A)$ at C_2 (along the dot-dash red lines). Therefore, the piecewise function $Y_{11}^{X,L}$ is not smooth.

$$\frac{\partial Y_{11}^b}{\partial \mu_A} < 0, \quad (\text{F10})$$

and μ_A is bounded by $\mu_A > [(\mu_B \nu_A)/\nu_B]$. In the same way, in case 2 for any given μ_B , we can acquire

$$\mu_A^{\text{opt}} = \frac{\mu_B \nu_A}{\nu_B}. \quad (\text{F11})$$

Up to here, we have proven that Theorem III is indeed correct. ■

Proof for Theorem I: Now, following the same idea, any optimal four values ($\mu_A^{\text{opt}}, \mu_B^{\text{opt}}, \nu_A^{\text{opt}}, \nu_B^{\text{opt}}$) must satisfy $(\mu_A^{\text{opt}}/\mu_B^{\text{opt}}) = (\nu_A^{\text{opt}}/\nu_B^{\text{opt}})$, or else we can always vary (μ_A, μ_B) while keeping (ν_A, ν_B) fixed and let $(\mu_A/\mu_B) = (\nu_A/\nu_B)$ to get a higher rate, meaning that the previous point is not the actual maximum. Therefore, we show that Theorem I is indeed correct and that the optimal decoy intensities always satisfy

$$\frac{\mu_A^{\text{opt}}}{\mu_B^{\text{opt}}} = \frac{\nu_A^{\text{opt}}}{\nu_B^{\text{opt}}}. \quad (\text{F12})$$

Note that the same conclusion does not hold true for traditional three-intensity MDI-QKD (i.e., using $\{\mu, \nu, \omega\}$ for both X and Z bases and using μ in the Z basis to generate the key), that is because the key rate for three-intensity depends on μ for both key generation and error correction, such that the $Q_{\mu\mu}^Z, E_{\mu\mu}^Z$ terms and the single-photon probability $\mu e^{-\mu}$ both depend on μ ; hence, optimizing only Y_{11}^L is no longer sufficient. Therefore, this independence of s from μ, ν is an additional advantage that the four-intensity or seven-intensity protocol can provide under asymmetric conditions.

Also, one thing to note is that, although the above theorem provides us with a way to constrain $\mu_A/\mu_B, \nu_A/\nu_B$, the actual values of these ratios still need to be found by optimization. In Ref. [20], the authors have proposed a rule-of-thumb formula for finding optimal intensities

$$\mu_A \eta_A = \mu_B \eta_B \quad (\text{F13})$$

for which we now have a good understanding of the reason: Such a relation keeps the arriving intensities balanced at Charles in order to maintain good HOM visibility in the X basis and low QBER. However, this formula is still only a rough approximation, and it is an exact relation only when the dark count rate Y_0 is ignored, the data size is infinite, and an infinite number of decoys are used [Ref. [20] considered the case where μ is both the signal and decoy intensity, and only proved Eq. (E13) to be exact in the ideal infinite-decoy case with no noise]. For a general case, μ_A/μ_B is not always exactly equal to η_B/η_A (and does not only depend on the mismatch $x = \eta_A/\eta_B$) but rather deviates slightly from it when (η_A, η_B) changes. But at least, one general rule is that μ_A/μ_B decreases with

$x = \eta_A/\eta_B$, or, to put in more simple words, the larger the channel loss, the higher the intensities we should choose to compensate for the loss.

2. Nonsmoothness of Key Rate vs Intensities Function

In the previous section, we show that the piecewise expression for $Y_{11}^{X,L}$ causes the optimal value to occur on the boundary line $(\mu_A/\mu_B) = (\nu_A/\nu_B)$. Here, we continue to show that Theorem II is a result of this piecewise function, too.

Proof of Theorem II: The theorem means that the key rate does not have a continuous partial derivative with respect to μ_A or μ_B at the boundary line. Such discontinuity will cause the boundary line to behave like a sharp ridge. To prove this theorem, instead of differentiating Y_{11}^a vs μ_B and Y_{11}^b vs μ_A , here we perform partial differentiation of both Y_{11}^a, Y_{11}^b vs μ_A , and observe this discontinuity of derivative.

First, we rewrite Y_{11} into

$$\begin{aligned} Y_{11}^a &= \frac{\nu_A}{(\mu_A - \nu_A)} \left[\frac{1}{\nu_A \nu_B} Q_{\nu\nu}^{M1} - \frac{1}{\mu_A \mu_B} Q_{\mu\mu}^{M2} \right] + \frac{1}{\nu_A \nu_B} Q_{\nu\nu}^{M1}, \\ Y_{11}^b &= \frac{\nu_B}{(\mu_B - \nu_B)} \left[\frac{1}{\nu_A \nu_B} Q_{\nu\nu}^{M1} - \frac{1}{\mu_A \mu_B} Q_{\mu\mu}^{M2} \right] + \frac{1}{\nu_A \nu_B} Q_{\nu\nu}^{M1}. \end{aligned} \quad (\text{F14})$$

The last term is not dependent on either μ_A or μ_B . Note that here on the boundary of $(\mu_A/\mu_B) = (\nu_A/\nu_B)$, the values of Y_{11}^a and Y_{11}^b are equal:

$$Y_{11}^a = Y_{11}^b. \quad (\text{F15})$$

Performing the partial differentiation against μ_A , we can get

$$\begin{aligned} \frac{\partial Y_{11}^a}{\partial \mu_A} &= -\frac{\nu_A}{\mu_A - \nu_A} \frac{\partial}{\partial \mu_A} \left(\frac{Q_{\mu\mu}^{M2}}{\mu_A \mu_B} \right) \\ &\quad + \frac{\nu_A}{(\mu_A - \nu_A)^2} \left(\frac{Q_{\mu\mu}^{M2}}{\mu_A \mu_B} - \frac{Q_{\nu\nu}^{M1}}{\nu_A \nu_B} \right), \\ \frac{\partial Y_{11}^b}{\partial \mu_A} &= -\frac{\nu_B}{\mu_B - \nu_B} \frac{\partial}{\partial \mu_A} \left(\frac{Q_{\mu\mu}^{M2}}{\mu_A \mu_B} \right). \end{aligned} \quad (\text{F16})$$

We can see that on the boundary of $(\mu_A/\mu_B) = (\nu_A/\nu_B)$, the first terms are again equal; however, the second term in $(\partial Y_{11}^a)/(\partial \mu_A)$ is strictly larger than 0 (A detailed proof by expanding $Q_{\mu\mu}^{M2}, Q_{\nu\nu}^{M1}$ is shown in Appendix F3). Therefore,

$$\frac{\partial Y_{11}^a}{\partial \mu_A} \neq \frac{\partial Y_{11}^b}{\partial \mu_A}. \quad (\text{F17})$$

The derivatives of $Y_{11}^{X,L}$ vs μ_A on the two sides of the ridge are not equal, causing the rate function R to have a nondefined gradient. A similar proof can be applied to μ_B , and it leads to the same result. ■

An illustration can be seen in the main text Fig. 2, which chooses a given set of values ($\nu_A = 0.2, \nu_B = 0.1$) and plots the key rate over (μ_A, μ_B) . As can be clearly observed, there is a sharp ridge on the line $(\mu_A/\mu_B) = (\nu_A/\nu_B) = 2$, meaning the key rate function versus intensities is not smooth.

3. Proof of Negativity of Partial Derivatives for Decoy Intensities

As we describe above, the expression for the single-photon yield $Y_{11}^{X,L}$ depends on whether $(\mu_A/\mu_B) \leq (\nu_A/\nu_B)$. For case 1, if $(\mu_A/\mu_B) \leq (\nu_A/\nu_B)$, we would like to prove that

Lemma I: $(\partial Y_{11}^a)/(\partial \mu_B)$ and $(\partial Y_{11}^b)/(\partial \mu_A)$ are both always negative.

Proof of Lemma I: Here, we use a simplified model of the gains Q_{ij}^X as in Ref. [20], which ignores the dark count rate Y_0 and takes a second-order approximation for the modified Bessel function:

$$Q_{\mu\mu}^X = \frac{\eta_B^2 \eta_d^2}{4} [2x\mu_A\mu_B + (\mu_B^2 + x^2\mu_A^2)(2e_d - e_d^2)], \quad (\text{F18})$$

where η_B is the transmittance in Bob-Charles channel, $x = (\eta_A/\eta_B)$ is the channel mismatch, η_d is the detector efficiency, and e_d is the misalignment. Here for convenience, we can further define

$$\epsilon = 2e_d - e_d^2, \quad T = \frac{\eta_B^2 \eta_d^2}{4}, \quad (\text{F19})$$

such that

$$Q_{\mu\mu}^X = T(2x\mu_A\mu_B + \epsilon\mu_B^2 + \epsilon x^2\mu_A^2). \quad (\text{F20})$$

Now, let us consider $(\partial Y_{11}^a)/(\partial \mu_B)$, where

$$Y_{11}^{X,L} = Y_{11}^a = \frac{1}{(\mu_A - \nu_A)} \left[\frac{\mu_A}{\nu_A \nu_B} Q_{\nu\nu}^{M1} - \frac{\nu_A}{\mu_A \mu_B} Q_{\mu\mu}^{M2} \right]. \quad (\text{F21})$$

To calculate the single-photon gain, the two terms

$$\begin{aligned} Q_{\nu\nu}^{M1} &= e^{\nu_A + \nu_B} Q_{\nu\nu}^X - e^{\nu_A} Q_{\nu\omega}^X - e^{\nu_B} Q_{\omega\nu}^X + Q_{\omega\omega}^X, \\ Q_{\mu\mu}^{M2} &= e^{\mu_A + \mu_B} Q_{\mu\mu}^X - e^{\mu_A} Q_{\mu\omega}^X - e^{\mu_B} Q_{\omega\mu}^X + Q_{\omega\omega}^X \end{aligned} \quad (\text{F22})$$

are linear combinations of the observable gains Q_{ij}^Z .

We can make the observation that, only the term

$$-\frac{\nu_A}{(\mu_A - \nu_A)\mu_A} \left(\frac{Q_{\mu\mu}^{M2}}{\mu_B} \right) \quad (\text{F23})$$

contains μ_B , so, we only need to prove the *positivity* of $(\partial/\partial \mu_B)(Q_{\mu\mu}^{M2}/\mu_B)$, where

$$\begin{aligned} Q_{\mu\mu}^{M2} &= e^{\mu_A + \mu_B} Q_{\mu\mu}^X - e^{\mu_A} Q_{\mu\omega}^X - e^{\mu_B} Q_{\omega\mu}^X + Q_{\omega\omega}^X \\ &= e^{\mu_A + \mu_B} Q_{\mu\mu}^X - e^{\mu_A} Q_{\mu\omega}^X - e^{\mu_B} Q_{\omega\mu}^X \end{aligned} \quad (\text{F24})$$

substituting with Eq. (E20),

$$\begin{aligned} \frac{1}{T} Q_{\mu\mu}^{M2} &= (2x\mu_A\mu_B + x^2\epsilon\mu_A^2 + \epsilon\mu_B^2)e^{\mu_A + \mu_B} - x^2\epsilon\mu_A^2 e^{\mu_A} - \epsilon\mu_B^2 e^{\mu_B} \\ &= 2x\mu_A\mu_B e^{\mu_A + \mu_B} \\ &\quad + x^2\epsilon\mu_A^2 e^{\mu_A}(e^{\mu_B} - 1) + \epsilon\mu_B^2 e^{\mu_B}(e^{\mu_A} - 1). \end{aligned} \quad (\text{F25})$$

Therefore,

$$\begin{aligned} \frac{Q_{\mu\mu}^{M2}}{\mu_B} &= T \times [2x\mu_A e^{\mu_A + \mu_B} + x^2\epsilon\mu_A^2 e^{\mu_A} \frac{e^{\mu_B} - 1}{\mu_B} \\ &\quad + \epsilon(e^{\mu_A} - 1)\mu_B e^{\mu_B}]. \end{aligned} \quad (\text{F26})$$

Note that here, as $\mu_A, \mu_B > 0$, we have $e^{\mu_A}, e^{\mu_B} > 1$, and each of the three functions satisfy

$$\begin{aligned} \frac{\partial}{\partial \mu_B} (e^{\mu_A + \mu_B}) &> 0, \\ \frac{\partial}{\partial \mu_B} \left(\frac{e^{\mu_B} - 1}{\mu_B} \right) &> 0, \\ \frac{\partial}{\partial \mu_B} (\mu_B e^{\mu_B}) &> 0. \end{aligned} \quad (\text{F27})$$

Therefore, we have proven that $(\partial/\partial \mu_B)(Q_{\mu\mu}^{M2}/\mu_B) > 0$ and that $\partial Y_{11}^a/\partial \mu_B < 0$. Similarly, we can also prove that $\partial Y_{11}^b/\partial \mu_A < 0$. Thus, the optimal point $(\mu_A^{\text{opt}}, \mu_B^{\text{opt}})$ must happen on the boundary, i.e.,

$$\frac{\mu_A^{\text{opt}}}{\mu_B^{\text{opt}}} = \frac{\nu_A}{\nu_B}. \quad (\text{F28})$$

4. Proof of Discontinuity of Partial Derivatives for Decoy Intensities

Now, to prove the discontinuity of the first-order derivatives of the key rate function, we need to show that

Lemma II: Partial derivatives of Y_{11}^a and Y_{11}^b with respect to μ_A , i.e., $(\partial Y_{11}^a)/(\partial \mu_A)$ and $(\partial Y_{11}^b)/(\partial \mu_A)$, are not equal.

Proof of Lemma II:

$$\begin{aligned} \frac{\partial Y_{11}^a}{\partial \mu_A} &= -\frac{\nu_A}{\mu_A - \nu_A} \frac{\partial}{\partial \mu_A} \left(\frac{Q_{\mu\mu}^{M2}}{\mu_A \mu_B} \right) \\ &\quad + \frac{\nu_A}{(\mu_A - \nu_A)^2} \left(\frac{Q_{\mu\mu}^{M2}}{\mu_A \mu_B} - \frac{Q_{\nu\nu}^{M1}}{\nu_A \nu_B} \right), \\ \frac{\partial Y_{11}^b}{\partial \mu_A} &= -\frac{\nu_B}{\mu_B - \nu_B} \frac{\partial}{\partial \mu_A} \left(\frac{Q_{\mu\mu}^{M2}}{\mu_A \mu_B} \right). \end{aligned} \quad (\text{F29})$$

On the boundary of $(\mu_A/\mu_B) = (\nu_A/\nu_B)$, the first terms are equal, i.e.,

$$-\frac{\nu_A}{\mu_A - \nu_A} \frac{\partial}{\partial \mu_A} \left(\frac{Q_{\mu\mu}^{M2}}{\mu_A \mu_B} \right) = -\frac{\nu_B}{\mu_B - \nu_B} \frac{\partial}{\partial \mu_A} \left(\frac{Q_{\mu\mu}^{M2}}{\mu_A \mu_B} \right). \quad (\text{F30})$$

However, here we have to show that the second term in $(\partial Y_{11}^a)/(\partial \mu_A)$ is strictly larger than zero:

$$\frac{\nu_A}{(\mu_A - \nu_A)^2} \left(\frac{Q_{\mu\mu}^{M2}}{\mu_A \mu_B} - \frac{Q_{\nu\nu}^{M1}}{\nu_A \nu_B} \right) > 0, \quad (\text{F31})$$

or, since $\mu_A > \nu_A$ and $\mu_A, \nu_A > 0$, simply

$$\frac{Q_{\mu\mu}^{M2}}{\mu_A \mu_B} - \frac{Q_{\nu\nu}^{M1}}{\nu_A \nu_B} > 0. \quad (\text{F32})$$

Just like in Appendix F 3, we can expand the gains Q_{ij}^X using Eq. (E20):

$$\begin{aligned} \frac{Q_{\nu\nu}^{M1}}{\nu_A \nu_B} &= 2x e^{\nu_A + \nu_B} + x^2 \epsilon \frac{\nu_A}{\nu_B} e^{\nu_A} (e^{\nu_B} - 1) \\ &\quad + \epsilon \frac{\nu_B}{\nu_A} e^{\nu_B} (e^{\nu_A} - 1), \\ \frac{Q_{\mu\mu}^{M2}}{\mu_A \mu_B} &= 2x e^{\mu_A + \mu_B} + x^2 \epsilon \frac{\mu_A}{\mu_B} e^{\mu_A} (e^{\mu_B} - 1) \\ &\quad + \epsilon \frac{\mu_B}{\mu_A} e^{\mu_B} (e^{\mu_A} - 1). \end{aligned} \quad (\text{F33})$$

Subtracting them, we can acquire

$$\begin{aligned} \frac{Q_{\mu\mu}^{M2}}{\mu_A \mu_B} - \frac{Q_{\nu\nu}^{M1}}{\nu_A \nu_B} &= 2x (e^{\mu_A + \mu_B} - e^{\nu_A + \nu_B}) \\ &\quad + x^2 \epsilon \left(\mu_A e^{\mu_A} \frac{e^{\mu_B} - 1}{\mu_B} - \nu_A e^{\nu_A} \frac{e^{\nu_B} - 1}{\nu_B} \right) \\ &\quad + \epsilon \left(\mu_B e^{\mu_B} \frac{e^{\mu_A} - 1}{\mu_A} - \nu_B e^{\nu_B} \frac{e^{\nu_A} - 1}{\nu_A} \right). \end{aligned} \quad (\text{F34})$$

Note that when a given variable $x > 0$, the functions

$$\begin{aligned} \frac{d}{dx} (e^x) &> 0, \\ \frac{d}{dx} \left(\frac{e^x - 1}{x} \right) &> 0, \\ \frac{d}{dx} (x e^x) &> 0. \end{aligned} \quad (\text{F35})$$

Therefore, these three functions strictly increase with their variable x , i.e., for any $x_1 > x_2$, $f(x_1) > f(x_2)$. Now, we can use the conditions $\mu_A > \nu_A$, $\mu_B > \nu_B$, and acquire

$$\begin{aligned} e^{\mu_A + \mu_B} &> e^{\nu_A + \nu_B}, \\ \mu_A e^{\mu_A} &> \nu_A e^{\nu_A}, \\ \mu_B e^{\mu_B} &> \nu_B e^{\nu_B}, \\ \frac{e^{\mu_A} - 1}{\mu_A} &> \frac{e^{\nu_A} - 1}{\nu_A}, \\ \frac{e^{\mu_B} - 1}{\mu_B} &> \frac{e^{\nu_B} - 1}{\nu_B}. \end{aligned} \quad (\text{F36})$$

Therefore, we have proven that $Q_{\mu\mu}^{M2}/(\mu_A \mu_B) - Q_{\nu\nu}^{M1}/(\nu_A \nu_B) > 0$, i.e.,

$$\frac{\partial Y_{11}^a}{\partial \mu_A} \neq \frac{\partial Y_{11}^b}{\partial \mu_A}. \quad (\text{F37})$$

Similarly, one can show that

$$\frac{\partial Y_{11}^b}{\partial \mu_B} \neq \frac{\partial Y_{11}^a}{\partial \mu_B}. \quad (\text{F38})$$

Therefore, for any given intensities (ν_A, ν_B) , the rate function $R(\mu_A, \mu_B)$ is not smooth against the two intensities (μ_A, μ_B) . ■

Remark: Also, though not explicitly proven here, since ν_A, ν_B also not only $Y_{11}^{X,L}$ but also $e_{11}^{X,U}$, their derivatives will be a lot more complex than μ_A, μ_B . Numerically, we observe that for any given (μ_A, μ_B) , the rate function $R(\nu_A, \nu_B)$ is actually not smooth against the two intensities (ν_A, ν_B) either, and the ridge still appears at $(\mu_A/\mu_B) = (\nu_A/\nu_B)$.

APPENDIX G: LOCAL SEARCH ALGORITHM

In this Appendix, we describe how to perform the optimization for the parameters, which is an indispensable process in obtaining the optimal key rate. In addition, we also discuss the effect of inaccuracies and fluctuations of the intensities and probabilities on the key rate and show that our method is robust even in the presence of inaccuracies and fluctuations of the parameters.

To provide a good key rate under finite-size effects, an optimal choice of parameters is an extremely important factor in implementing the protocol. However, the seven-intensity protocol has an extremely large parameter space of 12 dimensions, for which a brute-force search is next to impossible. To put it into context, a desktop PC (quad-core i7-4790k 4.0 GHz) can evaluate the function $R(\vec{v})$ at approximately 10^5 parameter combinations \vec{v} per second. Simply searching over a very crude ten-sample resolution for each parameter would take over four months, and a 100-sample resolution for each parameter would take 3×10^{11} years, a time longer than the age of the Universe! Therefore, a local search algorithm must be used to efficiently search the parameters in a reasonable time.

There have been studies to apply convex optimization to QKD, e.g., in Ref. [29] to find the optimal set of parameters and in Refs. [51–53] to bound the information leakage and secure key rate. Here, we start by adopting a local search algorithm for parameter optimization proposed in Ref. [29] called ‘‘coordinate descent’’ (CD), which requires drastically less time than using an exhaustive search. Instead of performing an exhaustive search over the parameter space, we can descend along each axis at a time and iterate over each axis in turn. For instance, suppose we currently iterate s_A ,

$$\begin{aligned}
R^{i+1} &= \max_{s_A \in (s_A^{\min}, s_A^{\max})} R(s_A, \mu_A^i, \nu_A^i, P_{s_A}^i, P_{\mu_A}^i, P_{\nu_A}^i, \\
&\quad s_B^i, \mu_B^i, \nu_B^i, P_{s_B}^i, P_{\mu_B}^i, P_{\nu_B}^i) \\
&= R(s_A^{i+1}, \mu_A^i, \nu_A^i, P_{s_A}^i, P_{\mu_A}^i, P_{\nu_A}^i, \\
&\quad s_B^i, \mu_B^i, \nu_B^i, P_{s_B}^i, P_{\mu_B}^i, P_{\nu_B}^i), \tag{G1}
\end{aligned}$$

which freezes the other coordinates and replaces s_A with the optimal position on the current coordinate axis s_A . In the next iteration, the algorithm will descend along axis μ_A , etc., hence the name coordinate descent. The search space satisfies that the probabilities lie within (0,1), and while the intensities could be larger than 1 in principle, typically that does not provide a good key rate, so here we also define the domain for all intensities as (0,1). The decoy intensities also follow two additional constraints $\mu_A > \nu_A$ and $\mu_B > \nu_B$. The CD algorithm is able to reach the same optimal position as a gradient descent algorithm (which descends along the gradient vector), the commonly used approach for parameter optimization.

However, a significant limitation of coordinate descent is that it does not work correctly over functions that have discontinuous first-order derivatives (which cause the gradient to be nondefined). For instance, in the presence of a sharp ridge as in Fig. 2 in the main text, any arbitrary point P on the ridge will cause the CD algorithm to terminate incorrectly and fail to find the maximum point. Mathematically, this is caused by the gradient being not clearly defined at a position where derivatives are discontinuous. Therefore, coordinate descent does not work anymore for asymmetric MDI-QKD.

As we discuss above, such a discontinuity of derivatives comes from the ridge $(\mu_A/\mu_B) = (\nu_A/\nu_B)$. Moreover, we know that the optimal parameters must satisfy $(\mu_A^{\text{opt}}/\mu_B^{\text{opt}}) = (\nu_A^{\text{opt}}/\nu_B^{\text{opt}})$. Therefore, here we propose to use polar coordinates instead of Cartesian coordinates to perform coordinate descent and *jointly* search μ_A/μ_B and ν_A/ν_B . In this way, we can make the rate vs parameter function smooth. We redefine \vec{v} as

$$\vec{v}_{\text{polar}} = [s_A, s_B, r_\mu, r_\nu, \theta_{\mu\nu}, P_{s_A}, P_{\mu_A}, P_{\nu_A}, P_{s_B}, P_{\mu_B}, P_{\nu_B}], \tag{G2}$$

where

$$\begin{aligned}
r_\mu &= \sqrt{\mu_A^2 + \mu_B^2}, & r_\nu &= \sqrt{\nu_A^2 + \nu_B^2}, \\
\theta_{\mu\nu} &= \tan^{-1}(\mu_A/\mu_B) = \tan^{-1}(\nu_A/\nu_B). \tag{G3}
\end{aligned}$$

In this way, the expression of Y_{11}^L always takes the boundary value (and has only a single expression). Therefore, when other parameters are fixed, $R(\theta_{\mu\nu})$ is actually a smooth function; therefore, by searching over

the parameters \vec{v}_{polar} , we can successfully find the optimal parameters and maximum rate.

After converting to polar coordinates and jointly searching $\theta_{\mu\nu}$, the coordinate descent algorithm becomes

$$\begin{aligned}
R^{i+1} &= \max_{s_A \in (s_A^{\min}, s_A^{\max})} R(s_A, s_B^i, r_\mu^i, r_\nu^i, \theta_{\mu\nu}^i, \\
&\quad P_{s_A}^i, P_{\mu_A}^i, P_{\nu_A}^i, P_{s_B}^i, P_{\mu_B}^i, P_{\nu_B}^i) \\
&= R(s_A^{i+1}, s_B^i, r_\mu^i, r_\nu^i, \theta_{\mu\nu}^i, \\
&\quad P_{s_A}^i, P_{\mu_A}^i, P_{\nu_A}^i, P_{s_B}^i, P_{\mu_B}^i, P_{\nu_B}^i). \tag{G4}
\end{aligned}$$

Additionally, when searching along each coordinate (for instance, fixing other parameters and searching s_A), we employ an iterative searching technique to further accelerate the algorithm, which starts out with a coarse resolution and iteratively narrows the search region while increasing the resolution (this is a similar technique as introduced in Ref. [29] but efficiently parallelized to utilize multithreading on modern PCs). For instance, we can start out with, e.g., 100 samples within the (0,1) region and evaluate them in parallel. After the maximal point is found, we can then choose two neighboring samples on the left and right of the maximal point and start a finer search among ten more samples between them. This process can be iterated until the maximum value no longer changes significantly or until the maximum depth is reached. Such a technique allows a search resolution that dynamically changes as needed (from 10^{-2} down to even 10^{-5} , although in practice often 10^{-3} is sufficient), and it efficiently uses, e.g., all eight threads on a quad-core CPU, enabling fast and accurate optimization below 0.1 s.

One more note is that the key rate obtained by our method is in fact robust against small inaccuracies in the parameters. For instance, for point A3 (10 km, 60 km) in Table III, if we round all parameters to an accuracy of 0.001 (as shown in Table IV) and use it for simulation, we can still get 99.5% of the optimal key rate 3.106×10^{-5} , while rounding the parameters to 0.01 will still give us 93.0% of the optimal key rate. In fact, even if we just keep one significant digit of each parameter, we can still get 47.6% of the optimal key rate. This would make it much easier for an experimental implementation of our method, as the key rate is very forgiving of inaccuracies in the parameters, which makes a much less stringent requirement on the intensity modulators and random number generators.

Note that the above ‘‘accuracy’’ discusses how strict the requirement is for us to generate an intensity or probability with mean value close to the desired optimal value (e.g., limited by bits in the random number generator or the accuracy of the intensity modulator), but we are still assuming we have perfect knowledge of the variables we generate. In addition, here we would like to point out that our conclusions remain unchanged, even in the presence of

intensity fluctuations or imprecision in the intensity probabilities.

First, the system is not very sensitive to the probabilities (since the partial derivatives with respect to them are zero at the optimal points), so even if all signal and decoy probabilities are simultaneously set 5% away from optimal value (and we take the global worst-case key rate value among all possible combinations of positive or negative deviation for each variable), the key rate will not significantly drop; for instance, for the (10 km, 60 km) case, one can still obtain 92.3% the ideal key rate (2.869×10^{-5} versus 3.106×10^{-5}) even with a 5% deviation for the probabilities.

Similarly, for intensity fluctuations, even if we add a 5% deviation to *all* intensities [again, taking the (10 km, 60 km) case as an example], we can still get 73.1% the ideal key rate (2.270×10^{-5} versus 3.106×10^{-5}). Moreover, one important point to note is that intensity fluctuation is not a problem unique to asymmetric MDI-QKD (or the new asymmetric protocol that we propose in this work). Even if one uses prior protocols (such as the four-intensity protocol), one would still obtain a significantly lower key rate if taking intensity fluctuation into consideration, such as 39.9% the key rate (3.671×10^{-5} versus 9.206×10^{-6} with no fluctuation) at (0 km, 50 km), and zero key rate (versus 3.891×10^{-7} with no fluctuation) at (10 km, 60 km). Therefore, the advantage of our method remains unchanged, even if intensity fluctuations are considered.

APPENDIX H: FINITE-SIZE ANALYSIS

In this Appendix, we describe the finite-key analysis used in our simulations.

The analytical proofs in Appendix F are shown for the asymptotic case. Numerically, we show that the seven-intensity protocol works effectively in the finite-key regime too, as can be observed in the main text Fig. 5.

To account for finite-size effects, we perform a standard error analysis [29,30] and estimate the expected value $\langle n \rangle$ of an observable n by

$$\underline{n} = n - \gamma\sqrt{n} \leq \langle n \rangle \leq n + \gamma\sqrt{n} = \bar{n}, \quad (\text{H1})$$

where we define the upper and lower bound for an observable n as \bar{n} and \underline{n} . Here, γ is the number of standard deviations the confidence interval of the observed value is from the expected value (for instance, for a required failure probability of no more than $\epsilon = 10^{-7}$, we should set $\gamma = 5.3$).

We can denote the observed counts as n_{μ_i, μ_j}^X , and error counts as m_{μ_i, μ_j}^X , where $\mu_i \in \{\mu_A, \nu_A, \omega\}$, $\mu_j \in \{\mu_B, \nu_B, \omega\}$. Then, the observed gain and error can be acquired from

$$\begin{aligned} Q_{\mu_i, \mu_j}^X &= \frac{n_{\mu_i, \mu_j}^X}{NP_{\mu_i}P_{\mu_j}}, \\ T_{\mu_i, \mu_j}^X &= \frac{m_{\mu_i, \mu_j}^X}{NP_{\mu_i}P_{\mu_j}}, \\ E_{\mu_i, \mu_j}^X &= \frac{T_{\mu_i, \mu_j}^X}{Q_{\mu_i, \mu_j}^X}, \end{aligned} \quad (\text{H2})$$

where N is the total number of signals sent, and P_{μ_i}, P_{μ_j} are the probabilities for Alice and Bob to send the respective intensities. Note that here we define the QBER in terms of error gains:

$$T_{\mu_i, \mu_j}^X = Q_{\mu_i, \mu_j}^X E_{\mu_i, \mu_j}^X. \quad (\text{H3})$$

As we describe in Appendix F, we can define the key rate expression as [8,30]

$$\begin{aligned} R &= P_{s_A} P_{s_B} \{s_A s_B e^{-(s_A + s_B)} Y_{11}^{X,L} [1 - h_2(e_{11}^{X,U})] \\ &\quad - f_e Q_{ss}^Z h_2(E_{ss}^Z)\}, \end{aligned} \quad (\text{H4})$$

and the single-photon gain and error estimated by [20]

$$\begin{aligned} Y_{11}^{X,L} &= \frac{1}{\mu_A - \nu_A} \left(\frac{\mu_A}{\nu_A \nu_B} Q_{\nu\nu}^{M1} - \frac{\nu_A}{\mu_A \mu_B} Q_{\mu\mu}^{M2} \right) \\ e_{11}^{X,U} &= \frac{1}{\nu_A \nu_B Y_{11}^{X,L}} (e^{\nu_A + \nu_B} T_{\nu\nu} - e^{\nu_A} T_{\omega\nu} \\ &\quad - e^{\nu_B} T_{\omega\nu} + T_{\omega\omega}), \end{aligned} \quad (\text{H5})$$

where $Q_{\nu\nu}^{M1}, Q_{\mu\mu}^{M2}$ are linear combination terms of the observables

$$\begin{aligned} Q_{\nu\nu}^{M1} &= e^{\nu_A + \nu_B} Q_{\nu\nu}^X - e^{\nu_A} Q_{\nu\omega}^X - e^{\nu_B} Q_{\omega\nu}^X + Q_{\omega\omega}^X, \\ Q_{\mu\mu}^{M2} &= e^{\mu_A + \mu_B} Q_{\mu\mu}^X - e^{\mu_A} Q_{\mu\omega}^X - e^{\mu_B} Q_{\omega\mu}^X + Q_{\omega\omega}^X. \end{aligned} \quad (\text{H6})$$

Now, with standard error analysis, we can define the upper and lower bounds for the gain and error gain:

$$\begin{aligned} \overline{Q_{\mu_i, \mu_j}^X} &= Q_{\mu_i, \mu_j}^X + \gamma \sqrt{\frac{Q_{\mu_i, \mu_j}^X}{NP_{\mu_i}P_{\mu_j}}}, \\ \underline{Q_{\mu_i, \mu_j}^X} &= Q_{\mu_i, \mu_j}^X - \gamma \sqrt{\frac{Q_{\mu_i, \mu_j}^X}{NP_{\mu_i}P_{\mu_j}}}, \\ \overline{T_{\mu_i, \mu_j}^X} &= T_{\mu_i, \mu_j}^X + \gamma \sqrt{\frac{T_{\mu_i, \mu_j}^X}{NP_{\mu_i}P_{\mu_j}}}, \\ \underline{T_{\mu_i, \mu_j}^X} &= T_{\mu_i, \mu_j}^X - \gamma \sqrt{\frac{T_{\mu_i, \mu_j}^X}{NP_{\mu_i}P_{\mu_j}}}. \end{aligned} \quad (\text{H7})$$

Therefore, we have

$$\begin{aligned}
\underline{Q}_{\nu\nu}^{M1} &= e^{\nu_A + \nu_B} \underline{Q}_{\nu\nu}^X - e^{\nu_A} \overline{Q}_{\nu\omega}^X - e^{\nu_B} \overline{Q}_{\omega\nu}^X + \underline{Q}_{\omega\omega}^X, \\
\underline{Q}_{\mu\mu}^{M2} &= e^{\mu_A + \mu_B} \overline{Q}_{\mu\mu}^X - e^{\mu_A} \underline{Q}_{\mu\omega}^X - e^{\mu_B} \underline{Q}_{\omega\mu}^X + \underline{Q}_{\omega\omega}^X, \\
Y_{11}^{X,L} &= \frac{1}{\mu_A - \nu_A} \left(\frac{\mu_A}{\nu_A \nu_B} \underline{Q}_{\nu\nu}^{M1} - \frac{\nu_A}{\mu_A \mu_B} \overline{Q}_{\mu\mu}^{M2} \right), \\
e_{11}^{X,U} &= \frac{1}{\nu_A \nu_B Y_{11}^{X,L}} (e^{\nu_A + \nu_B} \overline{T}_{\nu\nu} - e^{\nu_A} \underline{T}_{\nu\omega} \\
&\quad - e^{\nu_B} \underline{T}_{\omega\nu} + \overline{T}_{\omega\omega}), \tag{H8}
\end{aligned}$$

which we can use to substitute into Eq. (G4) to obtain the key rate under finite-size effects. (Note that here, $\underline{Q}_{\omega\omega}^X$ takes the lower bound in both $\underline{Q}_{\nu\nu}^{M1}$ and $\overline{Q}_{\mu\mu}^{M2}$, because its overall coefficient is positive in $Y_{11}^{X,L}$.)

Note that in Ref. [30], in addition to proposing the four-intensity protocol, Zhou *et al.* has proposed a ‘‘joint-bound’’ finite-key analysis which jointly considers the statistical fluctuations of observable gain and QBER. It is a tighter bound and can provide a higher rate than considering each observable’s fluctuation independently as we discuss above in this section (i.e., using ‘‘independent bounds’’). To illustrate this, we perform a simple simulation of the key rate versus distance plot using independent bounds and joint bounds (as well as using traditional three-intensity protocol [29] for comparison). As can be seen in Fig. 12, the four-intensity protocol with joint-bound analysis provides a higher rate than independent bounds (and both have a higher rate than the three-intensity protocol). However, the joint-bound analysis is based on linear optimization and sometimes brings multiple maxima for $R(\vec{v})$, which is undesirable for local search and will result in unpredictable behaviors (such as sudden ‘‘jitters’’ in the resulting rate versus distance plot, as can be observed in the joint-bound plot in Fig. 12; similar behavior is observed in Ref. [30], too).

Here, just for comparison, we list in Table V some example data points where we apply both independent-bound and joint-bound analysis. As can be seen, using joint bounds, we can indeed gain a further improved key rate. However, this comes at the expense of not knowing whether we are indeed at the global maximum or not, due to the existence of multiple maxima (and it is not ideal for comparing asymmetric and symmetric protocols, as the key rate estimated could be just local maxima for both of them). Therefore, as the purpose of this work is studying asymmetric MDI-QKD, we focus on independent bounds throughout the main text.

Also, note that although we have so far used standard error analysis for simplicity, our method here can in principle be applied to finite-key analysis with composable security, too, such as using the Chernoff bound [50]. The key point is that (as explicitly demonstrated in Appendixes B and C), the scaling of the asymmetric MDI-QKD key rate versus distance depends on the signal states (which perform

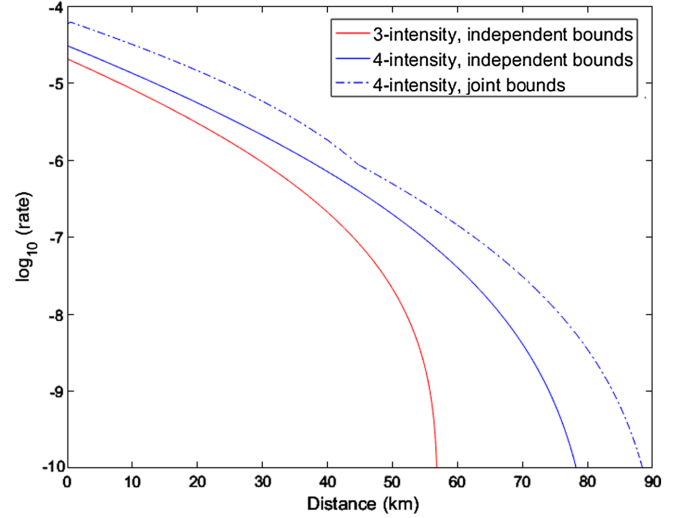


FIG. 12. Rate vs distance (Alice to Bob) for the symmetric case, for $N = 10^{11}$ using parameters $Y_0 = 6.02 \times 10^{-6}$, $\eta_d = 14.5\%$, $e_d = 1.5\%$, a parameter set in Zhou *et al.*’s paper [30]. Here, we compare the traditional three-intensity protocol as proposed in Ref. [29] (red solid line) and the four-intensity protocol [30] with independent-bound (blue solid line) and joint-bound analysis (blue dot-dash line).

a trade-off between error correction and single-photon probability). The decoy states need to maintain balanced arriving intensities at Charles but serve only to estimate the single-photon contributions as accurately as possible, whose asymptotic bounds are given by the infinite-data, infinite-decoy case. Adopting a different finite-key analysis (or no analysis at all, as in the asymptotic case) affects the bounds on single-photon yield and QBER Y_{11}^L and e_{11}^U . The finite-size case can be seen as the asymptotic case with correction terms (i.e., imperfections) added to the privacy amplification, but its key rate will have a similar scaling property as the asymptotic case. This means that the advantage of our method is independent of the finite-size analysis model used (or lack thereof, in the asymptotic case).

TABLE V. Simulation results of the key rate estimated with independent bounds versus joint bounds using the parameters in Table II. The data points for independent bounds correspond to the solid red curve in Fig. 5(d). As can be seen, using joint bounds for finite-size estimation can improve the key rate significantly. However, this will result in multiple maxima and cause instabilities in simulations. Therefore, we use independent bounds throughout the main text.

L_A	L_B	$R_{\text{independent}}$	R_{joint}
60 km	10 km	3.106×10^{-5}	6.714×10^{-5}
100 km	50 km	4.677×10^{-11}	7.568×10^{-8}
113 km	63 km	0	7.311×10^{-10}

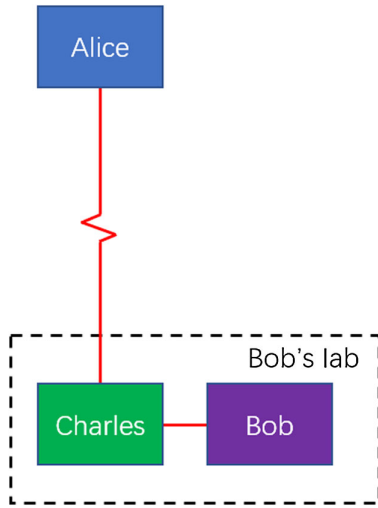


FIG. 13. Single-arm MDI-QKD where Bob and Charles are both in the same lab, with Bob’s channel having as little loss as possible. By optimizing intensities, we can achieve the maximum distance (loss) in the single channel between Alice and Charles, while enjoying the security of MDI-QKD.

APPENDIX I: SINGLE-ARM MDI-QKD

In the main text, we propose a new type of “single-arm” MDI-QKD setup, which is the extremely asymmetric case where one channel has high loss while the other channel has close to zero loss. In this Appendix, we describe it in more detail and outline its potential applications.

Suppose we have one crucial channel (e.g., a free-space channel, say, in a satellite-ground connection, or a ship-to-ship connection) through which we would like to send quantum signals. We would like to prevent all attacks on the detector and improve the security with MDI-QKD, but we cannot add a third party in the middle of the free-space channel. In this case, it is possible to add another source Bob in the laboratory (alongside Charles’s detectors, with as small loss as possible in the Bob-Charles channel) and use it to interfere with the signals coming from Alice over the longer free-space channel, as shown in Fig. 13. With seven-intensity protocol, a high key rate can be generated

from this extremely asymmetric case, providing the security of MDI-QKD to a single channel where relays cannot be added while still maintaining good performance.

If one uses four-intensity protocol, Bob has to add a fiber similar in loss to that of the free-space channel (to maintain the symmetry), while as we show with seven-intensity protocol, Bob can simply choose as small a loss as possible and obtain maximum acceptable loss in Alice’s channel. Not only does seven-intensity protocol make such a highly asymmetric MDI-QKD possible, it actually provides a higher rate compared to the symmetric case (if Bob adds a fiber). Moreover, since Alice’s channel loss might be constantly changing, it can be very difficult to adjust an added fiber and maintain the symmetry; thus, the convenience of not having to add any loss with seven-intensity protocol is a significant factor, too.

As we can observe in the main text Figs. 5(a) and 5(b), for the same required minimum rate, rather than performing an experiment at (L_{\max}, L_{\max}) , if we are free to adjust one channel (and want maximum distance in the other channel), we can set the shorter channel to zero and obtain a longer distance in the other channel, e.g., $(L'_{\max}, 0)$ with $L'_{\max} > L_{\max}$. For instance, in the main text Figs. 5(a) and 5(b), choosing point B (102 km, 0 km) can extend the longer arm from 85 to 102 km, from the symmetric point A (85 km, 85 km) for the same $R = 10^{-10}$.

Here, we list the simulations results for single-arm MDI-QKD. To demonstrate the advantage, we study three cases: using four-intensity (but being able to add fiber until the channels are symmetric), using four-intensity (however, due to being, e.g., in a free-space channel or a dynamic network, without the luxury to add fibers and compensate for the channels), and using seven-intensity directly on the asymmetric channels. As can be seen in Fig. 14, seven-intensity protocol provides better performance than both strategies using four-intensity and increases maximum distance from 56.8 and 72.5 km (respectively, for adding or not adding fiber) to 90 km. Thus, our new protocol can enable a unique new application of providing the security of MDI-QKD to a single channel where relays cannot be added (e.g., a free-space link), while still maintaining a high key rate.

TABLE VI. Simulation results of the key rate between each pair of nodes in a MDI-QKD network, using parameters from Table II, $N = 10^{11}$, and channels in main text Fig. 1(a). As can be seen, using seven-intensity protocol always provides a higher rate than either using four-intensity directly (which fails to establish some connections) or using four-intensity after adding fiber to each channel to accommodate the longest channel (which results in an identical low rate for every connection since every channel equals the longest channel after adding fiber). The seven-intensity protocol therefore enables high scalability and reconfigurability because each link is independent of other links and no added fiber is needed.

Method	A_1-A_3	A_1-A_4	A_1-A_5	A_3-A_4	A_3-A_5	A_4-A_5
Four-intensity, add fiber	1.28×10^{-10}	1.28×10^{-10}	1.28×10^{-10}	1.28×10^{-10}	1.28×10^{-10}	1.28×10^{-10}
Four-intensity, direct	0	0	0	2.41×10^{-4}	3.22×10^{-4}	5.77×10^{-4}
Seven-intensity, direct	1.97×10^{-7}	2.42×10^{-7}	2.77×10^{-7}	2.48×10^{-4}	3.53×10^{-4}	5.87×10^{-4}

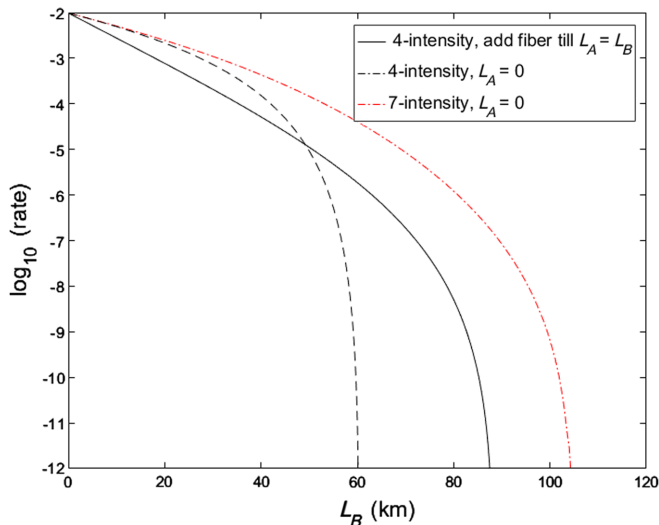


FIG. 14. Simulations of single-arm MDI-QKD. We use parameters from Table II and set $N = 10^{11}$. The three lines are generated using four-intensity protocol and adding fiber until $L_A = L_B$ (black solid line), using four-intensity protocol but without being able to add fiber (black dashed line), and using seven-intensity protocol directly (red dot-dash line). As can be seen, using seven-intensity protocol tremendously increases the key rate and maximum distance for the longer single arm. At $R = 10^{-7}$, using seven-intensity protocol (having maximum distance at 90 km) increases the maximum distance by 17.5 or 33.2 km (or 3.5 to 6.6 dB of loss) compared to four-intensity with or without fiber, respectively.

APPENDIX J: MDI-QKD NETWORK NUMERICAL RESULTS

In this Appendix, we consider the channels from a real quantum network setup in Vienna reported in Ref. [16] and numerically show that using seven-intensity protocol can provide high-rate communication between each pair of users, while previous protocols either fail to establish some connections in the network or suffer from low key rate for all connections.

Here, we focus on the high-asymmetry nodes in Ref. [16] A_1, A_2, A_3, A_4, A_5 plotted in the main text Fig. 1(a) and consider the case where an untrusted relay is placed at A_2 . The topology here is a commonly studied model of a star-type network, which is considered for a QKD network in Refs. [54,55], and it is also the model for the MDI-QKD network experiment in Ref. [19]. Such a network can provide a complete graph of connections between any two users, but it requires only one physical connection from each user. We show the simulation results in Table VI, where using seven-intensity protocol consistently provides high-rate connections even for nodes with very high asymmetry and maintains the same (in fact, moderately higher) key rate for nodes that are near-symmetric; i.e., including a long channel does not affect the rate between pairs of existing shorter channels.

Being able to establish connections with arbitrarily placed new nodes without affecting existing nodes is a very important property for a protocol to be used in a scalable and reconfigurable network, whose links will obviously be, more often than not, asymmetric. For the four-intensity protocol, to accommodate the highest-loss channel, all connections will suffer from a nonoptimal key rate. Moreover, since new users might be added or deleted dynamically, such an adding-fiber strategy will have poor scalability, since each new node affects the performance of all existing nodes and also causes interruption of service when users update their fibers. With seven-intensity protocol, we are completely free of the worries of asymmetry and can directly use the protocol on any channel combination optimally, so each node can be added or deleted without affecting the rest. This greatly improves not only the key rate but also the scalability of a MDI-QKD network.

- [1] P. W. Shor, *Algorithms for Quantum Computation: Discrete Logarithms and Factoring*, in *Proceedings of the 35th Annual IEEE Symposium on Foundations of Computer Science* (IEEE, New York, 1994).
- [2] Y. Zhao, C. H. F. Fung, B. Qi, C. Chen, and H. K. Lo, *Quantum Hacking: Experimental Demonstration of Time-Shift Attack against Practical Quantum-Key-Distribution Systems*, *Phys. Rev. A* **78**, 042333 (2008).
- [3] L. Lydersen, C. Wiechers, C. Wittmann, D. Elser, J. Skaar, and V. Makarov, *Hacking Commercial Quantum Cryptography Systems by Tailored Bright Illumination*, *Nat. Photonics* **4**, 686 (2010).
- [4] F. Xu, B. Qi, and H. K. Lo, *Experimental Demonstration of Phase-Remapping Attack in a Practical Quantum Key Distribution System*, *New J. Phys.* **12**, 113026 (2010).
- [5] I. Gerhardt, Q. Liu, A. Lamas-Linares, J. Skaar, C. Kurtsiefer, and V. Makarov, *Full-Field Implementation of a Perfect Eavesdropper on a Quantum Cryptography System*, *Nat. Commun.* **2**, 349 (2011).
- [6] H. Weier, H. Krauss, M. Rau, M. Fürst, S. Nauerth, and H. Weinfurter, *Quantum Eavesdropping without Interception: An Attack Exploiting the Dead Time of Single-Photon Detectors*, *New J. Phys.* **13**, 073024 (2011).
- [7] N. Jain, C. Wittmann, L. Lydersen, C. Wiechers, D. Elser, C. Marquardt, V. Makarov, and G. Leuchs, *Device Calibration Impacts Security of Quantum Key Distribution*, *Phys. Rev. Lett.* **107**, 110501 (2011).
- [8] H. K. Lo, M. Curty, and B. Qi, *Measurement-Device-Independent Quantum Key Distribution*, *Phys. Rev. Lett.* **108**, 130503 (2012).
- [9] T. F. da Silva, D. Vitoreti, G. B. Xavier, G. C. do Amaral, G. P. Temporão, and J. P. von der Weid, *Proof-of-Principle Demonstration of Measurement-Device-Independent Quantum Key Distribution Using Polarization Qubits*, *Phys. Rev. A* **88**, 052303 (2013).
- [10] Y. Liu *et al.*, *Experimental Measurement-Device-Independent Quantum Key Distribution*, *Phys. Rev. Lett.* **111**, 130502 (2013).

- [11] Z. Tang, Z. Liao, F. Xu, B. Qi, L. Qian, and H. K. Lo, *Experimental Demonstration of Polarization Encoding Measurement-Device-Independent Quantum Key Distribution*, *Phys. Rev. Lett.* **112**, 190503 (2014).
- [12] Y. L. Tang *et al.*, *Measurement-Device-Independent Quantum Key Distribution over 200 km*, *Phys. Rev. Lett.* **113**, 190501 (2014).
- [13] H. L. Yin *et al.*, *Measurement-Device-Independent Quantum Key Distribution over a 404 km Optical Fiber*, *Phys. Rev. Lett.* **117**, 190501 (2016).
- [14] L. C. Comandar, M. Lucamarini, B. Fröhlich, J. F. Dynes, A. W. Sharpe, S. W.-B. Tam, Z. L. Yuan, R. V. Penty, and A. J. Shields, *Quantum Key Distribution without Detector Vulnerabilities Using Optically Seeded Lasers*, *Nat. Photonics* **10**, 312. (2016).
- [15] A. Rubenok, J. A. Slater, P. Chan, I. Lucio-Martinez, and W. Tittel, *Real-World Two-Photon Interference and Proof-of-Principle Quantum Key Distribution Immune to Detector Attacks*, *Phys. Rev. Lett.* **111**, 130501 (2013).
- [16] M. Peev *et al.*, *The SECOQC Quantum Key Distribution Network in Vienna*, *New J. Phys.* **11**, 075001 (2009).
- [17] M. Sasaki *et al.*, *Field Test of Quantum Key Distribution in the Tokyo QKD Network*, *Opt. Express* **19**, 10387 (2011).
- [18] R. Courtland, *China's 2,000-km Quantum Link Is Almost Complete*, *IEEE Spectrum* **53**, 11 (2016).
- [19] Y. L. Tang *et al.*, *Measurement-Device-Independent Quantum Key Distribution over Untrustful Metropolitan Network*, *Phys. Rev. X* **6**, 011024 (2016).
- [20] F. Xu, M. Curty, B. Qi, and H. K. Lo, *Practical Aspects of Measurement-Device-Independent Quantum Key Distribution*, *New J. Phys.* **15**, 113007 (2013).
- [21] E. Moschandreou, J. I. Garcia, B. J. Rollick, B. Qi, R. Pooser, and G. Siopsis, *Experimental Study of Hong-Ou-Mandel Interference Using Independent Phase Randomized Weak Coherent States*, arXiv:1804.02291.
- [22] Note that there have also been proposals for continuous-variable MDI-QKD [23,24], which provides a high key rate for short distances but is typically limited to distances < 25 km even when assuming a high detector efficiency of 98%. In this work, we focus on only discrete-variable MDI-QKD.
- [23] Z. Li, Y. C. Zhang, F. Xu, X. Peng, and H. Guo, *Continuous-Variable Measurement-Device-Independent Quantum Key Distribution*, *Phys. Rev. A* **89**, 052301 (2014).
- [24] S. Pirandola, C. Ottaviani, G. Spedalieri, C. Weedbrook, S. L. Braunstein, S. Lloyd, T. Gehring, C. S. Jacobsen, and U. L. Andersen, *High-Rate Measurement-Device-Independent Quantum Cryptography*, *Nat. Photonics* **9**, 397 (2015).
- [25] J. Yin *et al.*, *Satellite-Based Entanglement Distribution over 1200 Kilometers*, *Science* **356**, 1140 (2017).
- [26] I. V. Puthoor, R. Amiri, P. Wallden, M. Curty, and E. Andersson, *Measurement-Device-Independent Quantum Digital Signatures*, *Phys. Rev. A* **94**, 022328 (2016).
- [27] G. L. Roberts, M. Lucamarini, Z. L. Yuan, J. F. Dynes, L. C. Comandar, A. W. Sharpe, A. J. Shields, M. Curty, I. V. Puthoor, and E. Andersson, *Experimental Measurement-Device-Independent Quantum Digital Signatures*, *Nat. Commun.* **8**, 1098 (2017).
- [28] M. Lucamarini, Z. L. Yuan, J. F. Dynes, and A. J. Shields, *Overcoming the Rate-Distance Limit of Quantum Key Distribution without Quantum Repeaters*, *Nature (London)* **557**, 400 (2018).
- [29] F. Xu, H. Xu, and H. K. Lo, *Protocol Choice and Parameter Optimization in Decoy-State Measurement-Device-Independent Quantum Key Distribution*, *Phys. Rev. A* **89**, 052333 (2014).
- [30] Y. H. Zhou, Z. W. Yu, and X. B. Wang, *Making the Decoy-State Measurement-Device-Independent Quantum Key Distribution Practically Useful*, *Phys. Rev. A* **93**, 042324 (2016).
- [31] A previous protocol of interest for MDI-QKD is the four-intensity protocol proposed by Zhou *et al.* [30]. In this protocol, Alice and Bob each use three intensities $\{\mu, \nu, \omega\}$ in the X basis to perform a decoy-state analysis [32–34] and uses one signal intensity $\{s\}$ in the Z basis to generate the secret key. The four-intensity protocol can greatly improve MDI-QKD performance under limited data size. However, it limits its discussions to the symmetric case only (Alice and Bob using identical parameters), which is suboptimal in an asymmetric setting. Although Ref. [30] mentioned in passing the possibility of using different intensities of optical signals for Alice and Bob, little analysis of this important case was performed there. So, up till now, it has not been clear how exactly Alice and Bob could compensate for asymmetric channel losses with different signal intensities.
- [32] H. K. Lo, X. Ma, and K. Chen, *Decoy State Quantum Key Distribution*, *Phys. Rev. Lett.* **94**, 230504 (2005).
- [33] W. Y. Hwang, *Quantum Key Distribution with High Loss: Toward Global Secure Communication*, *Phys. Rev. Lett.* **91**, 057901 (2003).
- [34] X. B. Wang, *Beating the Photon-Number-Splitting Attack in Practical Quantum Cryptography*, *Phys. Rev. Lett.* **94**, 230503 (2005).
- [35] Another case where Alice and Bob sent $|+-\rangle$ or $| - + \rangle$ corresponds to the other Bell state $|\psi^-\rangle$. A two-photon interference happens not at the beam splitter but at the PBS instead. This setup is slightly different from HOM interference but similar to that of Ref. [36] and also requires indistinguishability of, e.g., spectrum, timing, and matching intensities. For simplicity, here we use the term “two-photon interference” for both cases.
- [36] J. Beugnon, M. P. A. Jones, J. Dingjan, B. Darquié, G. Messin, A. Browaeys, and P. Grangier, *Quantum Interference between Two Single Photons Emitted by Independently Trapped Atoms*, *Nature (London)* **440**, 779 (2006).
- [37] In the nonideal case with basis misalignment, there may be a slight dependence in the Z basis, too, as we see in Fig. 3, because misalignment results in crosstalk between signals from the two bases, but it will be a much smaller dependence than that in the X basis.
- [38] For the finite-size case, the yield and QBER in Eq. (5) are replaced with their *expected values* $\langle Y_i \rangle$ and $\langle e_i \rangle$, and the result still holds true, i.e., $\langle Y_i \rangle(\mu) = \langle Y_i \rangle(\nu)$, $\langle e_i \rangle(\mu) = \langle e_i \rangle(\nu)$. Similar applies for the yield and QBER in MDI-QKD. More details on finite-size analysis can be found in Appendix H.
- [39] H. K. Lo, H. F. Chau, and M. Ardehali, *Efficient Quantum Key Distribution Scheme and a Proof of Its Unconditional Security*, *J. Cryptol.* **18**, 133 (2005).

- [40] L. C. Comandar, M. Lucamarini, B. Fröhlich, J. F. Dynes, A. W. Sharpe, S. W.-B. Tam, Z. L. Yuan, R. V. Penty, and A. J. Shields, *Quantum Key Distribution without Detector Vulnerabilities Using Optically Seeded Lasers*, *Nat. Photonics* **10**, 312 (2016).
- [41] The idea of decoupling the bases was first studied for BB84 in Ref. [42]. There, the relation $Y_1^Z = Y_1^X$ was used for single photons instead of single-photon pairs.
- [42] Z. Wei, W. Wang, Z. Zhang, M. Gao, Z. Ma, and X. Ma, *Decoy-State Quantum Key Distribution with Biased Basis Choice*, *Sci. Rep.* **3**, 2453 (2013).
- [43] X. B. Wang, *Three-Intensity Decoy-State Method for Device-Independent Quantum Key Distribution with Basis-Dependent Errors*, *Phys. Rev. A* **87**, 012320 (2013).
- [44] H. Liu *et al.*, *Experimental Demonstration of High-Rate Measurement-Device-Independent Quantum Key Distribution over Asymmetric Channels*, *Phys. Rev. Lett.* **122**, 160501 (2019).
- [45] http://2018.qcrypt.net/scientific-program/#Wednesday_1110.
- [46] W. Wang and H. K. Lo, *Simple Method for Asymmetric Twin Field Quantum Key Distribution*, [arXiv:1907.05291](https://arxiv.org/abs/1907.05291).
- [47] F. Grasselli, A. Navarrete, and M. Curty, *Asymmetric Twin-Field Quantum Key Distribution*, [arXiv:1907.05256](https://arxiv.org/abs/1907.05256).
- [48] P. Shor and J. Preskill, *Simple Proof of Security of the BB84 Quantum Key Distribution Protocol*, *Phys. Rev. Lett.* **85**, 441 (2000).
- [49] X. Ma, C. H. F. Fung, and M. Razavi, *Statistical Fluctuation Analysis for Measurement-Device-Independent Quantum Key Distribution*, *Phys. Rev. A* **86**, 052305 (2012).
- [50] M. Curty, F. Xu, C. C. W. Lim, K. Tamaki, and H. K. Lo, *Finite-Key Analysis for Measurement-Device-Independent Quantum Key Distribution*, *Nat. Commun.* **5**, 3732 (2014).
- [51] P. Coles, E. Metodiev, and N. Lütkenhaus, *Numerical Approach for Unstructured Quantum Key Distribution*, *Nat. Commun.* **7**, 11712 (2016).
- [52] N. Islam, C. C. W. Lim, C. Cahall, J. Kim, and D. Gauthier, *Securing Quantum Key Distribution Systems Using Fewer States*, *Phys. Rev. A* **97**, 042347 (2018).
- [53] A. Winick, N. Lütkenhaus, and P. J. Coles, *Reliable Numerical Key Rates for Quantum Key Distribution*, *Quantum* **2**, 77 (2018).
- [54] E. Y. Zhu, C. Corbari, A. V. Gladyshev, P. G. Kazansky, H. K. Lo, and L. Qian, *Multi-Party Agile Quantum Key Distribution Network with a Broadband Fiber-Based Entangled Source*, *J. Opt. Soc. Am. B* **36**, B1 (2019).
- [55] S. Wengerowsky, S. K. Joshi, F. Steinlechner, H. Hübel, and R. Ursin, *Entanglement-Based Wavelength Multiplexed Quantum Communication Network*, *Nature (London)* **564**, 225 (2018).