Research Article

Improving Parameter Optimization in Decoy-State Quantum Key Distribution

Zijian Li and Kejin Wei

Guangxi Key Laboratory for Relativistic Astrophysics, School of Physical Science and Technology, Guangxi University, Nanning 530004, China

Correspondence should be addressed to Kejin Wei; kjwei@gxu.edu.cn

Received 9 November 2021; Accepted 19 January 2022; Published 4 February 2022

Academic Editor: Liuguo Yin

Copyright © 2022 Zijian Li and Kejin Wei. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Quantum key distribution (QKD) enables two remote users to share a string of key bits with information-theoretical security. Parameter optimization is a crucial step in achieving optimal performance in practical QKD systems. In general, such optimization is implemented using a local search algorithm (LSA). However, LSAs inevitably fail to find out the optimal values when the searched key rate function is nonconvex or has a discontinuity of first-order derivatives and a narrow parameter search space. This paper proposes a genetic algorithm-based method to overcome the limitations of LSAs for QKD parameter optimization. We tested the proposed method with various types of common QKD protocols and found that it has very high parameter optimization performance for QKD with a time consumption comparable to that using a standard LSA. We expect our method to be a valuable optimization tool for quantum information processing tasks.

1. Introduction

Quantum key distribution (QKD) [1, 2], which is guaranteed by the principles of quantum mechanics, enables two remote users to share a string of key bits with information-theoretical security. In realistic QKD systems, nonperfect single-photon sources, such as a weak coherent state laser pulses (WCPs), are widely adopted, making them susceptible to a sophisticated attack called a photon number splitting attack [3, 4]. However, the decoy-state method was proposed to solve this problem by using two or more photon intensities to evaluate the single-photon contributions [5, 6]. In recent decades, decoy-state QKD has been demonstrated in optical fiber [7–9] and free space [10, 11]. Various field-test QKD networks [12–16], assembled with decoy-state QKD systems, have been reported worldwide. Recently, combined with the idea of measurement-device-independent (MDI)-QKD [8, 17–19], which removes all detector side-channel attacks, decoy-state QKD systems have successfully overcome the repeaterless secret key capacity [20–24] and become more suitable for practical applications [25–29].

In reality, all decoy-state QKD experiments are completed within finite time. Hence, the total number of emitted signals is finite. This means that the estimation of single-photon contributions using the decoy-state method must account for statistical fluctuations. This is referred to as finite-key analysis of QKD [30–33]. When considering finite-data size, the intensity choices of signal and decoy states and the probabilities of sending these states are crucial in achieving better system performance. Thus, a full optimization must be performed to search for these parameters based on given experimental parameters, such as optical misalignment, data size, and channel loss.

In general, such full optimization is completed by a brute-force global search, which is challenging due to limited computational power. Hence, it is only suitable for searched functions with a smaller number of parameters. In 2014, Xu et al. [34] introduced a local search algorithm (LSA) called coordinate descent (CD), which significantly improves the optimization speed for symmetric MDI-QKD, which has a higher number of searched parameters. Subsequently, it has been proven to be an efficient tool for optimization problems.
in various QKD protocols, such as asymmetric MDI-QKD [35, 36] and twin-field QKD [37].

However, there are three main drawbacks when using an LSA for the optimization of parameters in a QKD protocol. First, LSAs are easily trapped in local maxima because they require the targeted function to be convex. However, quite a few QKD protocols [37, 38], such as the BB84 protocol, cannot fulfill this requirement. Second, LSAs inevitably fail to find optima, especially when searching for the optimal key rate over the maximum transmission distance of QKD systems. In this case, a positive key rate is exhibited in a very narrow search space. However, LSAs lack the ability to follow the appropriate path to the optimum on such complex search spaces. Third, LSAs require the searched key rate function to have continuous first-order derivatives with respect to the parameters. However, numerous QKD protocols [35, 36, 39, 40], such as asymmetric MDI-QKD [36], are discontinuous, which results in failure to find the optimal point. These disadvantages are discussed in more detail in Section 2.2.

A more effective method is to exploit a neural network to directly predict the optimal parameters for QKD systems [38, 41, 42]. Although this approach shows several orders of magnitude higher speed compared with LSAs, it takes several hours and requires considerable programming to train a neural network. Furthermore, it typically requires pre-generated training data sets using an LSA and can output only a portion (ranging from 80% to 99%) of the optimal secure rate of an LSA [38]. Hence, the level of optimal key rate obtained using an LSA determines the performance of the neural network.

This paper proposes a method that uses a genetic algorithm (GA) to overcome the limitations of LSAs for QKD parameter optimization. A GA is a general-purpose search algorithm capable of exploiting the information accumulated about an initially unknown search space to guide subsequent searches into useful subspaces. This key feature enables the GA to efficiently find QKD optima, even with complex, large, and poorly understood search spaces. We tested the proposed method using four common QKD protocols: BB84, symmetric and asymmetric MDI-QKD, and reference-independent (RFI)-MDI-QKD. The simulation results demonstrate that the proposed method can accurately and efficiently optimize QKD parameters.

The remainder of the paper is organized as follows. In Section 2, we describe how parameter optimization can be formulated as a function that can be searched by computer algorithms. We then further discuss the limitations of LSAs for parameter optimization. In Section 3, we introduce the proposed method with an example protocol. In Section 4, we present numerical results that show that the proposed method can accurately find optimal parameters and secure key rates for four common QKD protocols. Finally, we summarize the paper in Section 5.

2. Parameter Optimization in QKD

2.1. Optimal Parameters as a Function. Let us consider asymmetrical MDI-QKD as an example. Because Alice and Bob have nonidentical distances to the untrusted relay Charles, they must use an asymmetrical protocol [36] called the seven-intensity protocol to achieve a higher key rate. In the seven-intensity protocol, the intensities used are not only different for Alice and Bob but also different in the two bases $Z$ and $X$. Therefore, a total of 12 parameters must be optimized when considering a finite-size effect:

$$\mathbf{\gamma} = \{s_A, \mu_A, \gamma_A, p_s, p_{\mu}, p_{\gamma}, \phi_B, \gamma_B, p_s, p_{\mu}, p_{\gamma}\}.$$ Here, we represent these 12 parameters as a vector $\mathbf{\gamma}$, where $\{s_A, \mu_A, \gamma_A\}$ are the signal and decoy intensities for Alice, and $p_s, p_{\mu}, p_{\gamma}$ are the probabilities of sending them.

To calculate the key rate of a practical QKD system, we must consider the influence of experimental parameters, including dark count $Y_0$, error correction efficiency $f_e$, detection efficiency $\eta_d$, optical misalignment $e_d$, distance $L$, and total number of signals $N$ sent by Alice and Bob. These experimental parameters are united as $\mathbf{\varepsilon}$, so the key rate function of QKD can be expressed as

$$R = R(\mathbf{\varepsilon}, \mathbf{\gamma}),$$

which is a function of user parameters $\mathbf{\gamma}$ and experimental parameters $\mathbf{\varepsilon}$. The experimental parameters are typically fixed, as determined by the performance of QKD systems, and user parameters are controlled by the users. Hereinafter, we refer to the “user parameters” as simply “parameters” for ease of explanation. Therefore, to calculate the optimal key rate, we must compute

$$R_{\text{max}}(\mathbf{\varepsilon}) = \max_{\mathbf{\gamma} \in V} R(\mathbf{\varepsilon}, \mathbf{\gamma}),$$

which is the optimal key rate for a given $\mathbf{\varepsilon}$, and $V$ is the parameter search space. The key objective of parameter optimization for QKD is to acquire the optimal set of $\mathbf{\gamma}_{\text{opt}}$, given the maximum value from the target function $R(\mathbf{\varepsilon}, \mathbf{\gamma})$ with the given $\mathbf{\varepsilon}$. This can be expressed as

$$\mathbf{\gamma}_{\text{opt}}(\mathbf{\varepsilon}) = \arg \max_{\mathbf{\gamma} \in V} R(\mathbf{\varepsilon}, \mathbf{\gamma}).$$

In principle, the optimal set of $\mathbf{\gamma}_{\text{opt}}$ can be determined using a global search, which searches the optimal points by calculating $R(\mathbf{\varepsilon}, \mathbf{\gamma})$ many times with different parameters. However, for some protocols with extremely large parameter spaces, a brute-force search is impossible to complete in a reasonable time. For example, the abovementioned seven-intensity protocol has a 12-dimensional parameter space. A desktop PC would take approximately several months to search over a very rough dozens of sample resolution for each parameter.

2.2. Limitations of LSAs. LSAs enable an efficient parameter search in a reasonable time for some specific QKD protocols. However, there are some limitations of LSAs for QKD optimization if the target key rate function cannot fulfill some extreme restrictions. To explain this clearly, as in Ref. [34], we adopted the CD algorithm, which is a popular LSA. We demonstrate that the CD algorithm does not work for...
several QKD protocols. Note that these limitations do not depend on specific LSAs because they all share the same basic principles.

CD usually starts with a manually input starting point, known as the current solution, and then it descends along each axis stepwise within its neighborhood and iterates over each axis in turn. The algorithm moves from the current solution to a better closer solution, which is used in the next iteration as the current solution. This process is repeated until there is no better solution within the neighborhood of the current solution. For instance, as shown in Figure 1, we begin by iterating $s$ and finding the optimal current solution. In the next iteration, the algorithm descends along the $\mu_A$ axis. After several iterations, the optimal parameters can be found.

However, CD has several relatively strict requirements for the searched functions. First, LSAs do not work for a nonconvex key rate function because it would result in two peaks at the boundary line. Second, LSAs require the searched key rate function to have continuous first-order derivatives with respect to the parameters. Furthermore, LSAs lack the ability to follow the proper path to the optimum in narrow search spaces. Numerous QKD protocols do not fulfill all these requirements. For instance, as shown in Figure 2(a), the key rate function of the BB84 protocol is slightly nonconvex, and in Figure 2(b), the first-order derivative of the key rate function in the seven-intensity MDI-QKD protocol is discontinuous with respect to $\mu_A$ and $\mu_B$. The reason for this discontinuity is discussed in detail in Ref. [36]. Furthermore, a positive key rate is exhibited only in a very narrow search space over a long transmission distance, as shown in Figures 2(c) and 2(d), where we plot the key rate function of RFI-MDI-QKD with a distance of $L = 1\, \text{km}$ and $L = 30\, \text{km}$, respectively; it can be seen that the boundary line at $L = 30\, \text{km}$ is far lower than that at $L = 1\, \text{km}$.

Here, we further discuss LSAs for asymmetrical MDI-QKD in detail. In Ref. [36], the authors showed that LSAs are infeasible for asymmetrical MDI-QKD because the first-order derivatives of the key rate function in the seven-intensity protocol are discontinuous with respect to $\mu_A$ and $\mu_B$. As shown in Figure 3, this causes the contour map of the key rate to have some break points, as indicated by the black circles. If we used the CD algorithm, it would incorrectly stop at any break points after several iterations because no better solution could be found in orthogonal steps. As a result, it would fail to find the optima.

It is worth mentioning that, with considerable mathematical techniques, one can switch to polar coordinates of variables and remove the problem of discontinuity of derivatives in the seven-intensity protocol. Hence, a direct usage of LSA has a good performance for optimization. However, the transformation is based on the proved Theorems I and II in Ref. [36]. It is unknown whether the theorems can be applied in other protocols or not, for example, the twin-field or asymmetric reference-independent MDI-QKD protocols.

3. Methods

In this section, we describe our proposed GA-based method [43]. For simplicity, to introduce our method, we consider the seven-intensity protocol as an example protocol that has been proven not to be able to optimize parameters using the CD algorithm. In the next subsection, when detailing the numerical results, we also include the following three protocols to show that the method provides an effective search for them too: BB84, symmetric MDI-QKD, and RFI-MDI-QKD.

Our method is mainly based on a GA, which exploits a population of candidate solutions and then explores the search space by evolving the population through four biological genetic operators: parent selection, crossover, mutation, and replacement. Without any restriction on the target function, a GA is able to exploit the information collected from an initially unknown search space to guide following searches into useful subspaces. This key feature enables our method to outperform LSAs and efficiently search QKD optimal parameters.

Here, we use our method to find the optimal decoy intensities $\mu_A$ and $\mu_B$ in the seven-intensity protocol as an example. The optimization was performed using the built-in GA package in MATLAB R2016b. Figure 4 displays the procedure, and the detail is explained as follows:

(i) Initialization

After setting the population size to 30, the algorithm initializes with a random population of candidate solutions (called individuals). For the seven-intensity protocol, we used a range of common values for $s \in (0, 1)$, $\mu \in (0, 1)$, $\nu \in (0, 1)$, $P \in (0, 1)$. Here, since Alice and Bob have equal range, we omit the subscripts of $A$ and $B$. As illustrated in Figure 4(a), the whole range of possible solutions (or the search space) is allowed in the 1st generation. Though the function being
Figure 2: Examples of (a) nonconvex key rate function with respect to signal intensity $s$ and decoy intensity $\gamma$ in the BB84 protocol, (b) discontinuity of first-order derivatives of key rate function with respect to Alice’s decoy intensity $\mu_A$ and Bob’s decoy intensity $\mu_B$ in asymmetric MDI protocol, and narrow search spaces over long transmission distances of (c) $L = 1$ km and (d) $L = 30$ km with respect to each users’ signal intensity $s$ and decoy intensity $\gamma$ in RFI-MDI protocol. In each example, the boundary line was obtained by sweeping the $x$ and $y$ parameters.

Figure 3: CD algorithm for asymmetrical MDI-QKD optimization. The LSA is stopped at the current optimum because no better solution can be found in the orthogonal steps. The black circle points indicate break points. The $x$ axis is Alice’s decoy intensity $\mu_A$, and the $y$ axis is Bob’s decoy intensity $\mu_B$.
discontinuous, the solutions are possible to be inseminated in specific regions on occasion, where the potential optimal solutions will be found.

(ii) Selection

During each iteration (or successive generation), a part of the existing population is chosen to produce a new one (called the parents) through a fitness-based process. This process is conducted by a preset fitness function. Here, we choose the roulette selection function, which simulates a roulette wheel with the area of each segment proportional to its expectation. Generally, fitter individuals (measured by the roulette selection function) are at greater chance of selection for the next step, i.e., genetic operators.

(iii) Genetic operators

In this step, the main goal is to produce a second-generation population of solutions from the parents generated in the previous step. The generation is based on biologically genetic behaviors, including crossover and mutation. To generate each new solution (called child), a pair of “parent” solutions is selected from the pool generated in the step of Selection. Since the child is bred by using the combination of crossover and mutation, it typically shares many of the characteristics of its created parents. For each new child, a new pair of parents is chosen for breeding. The genetic process is repeated until a new population of solutions of preset size (we set to 30) is produced. The crossover and mutation operations are implemented with certain
probabilities, which are set as 0.8 and 0.02, respectively. In crossover operation, a multipoint crossover method is adopted, which exchanges genes of the parent at several randomly selected gene points. In mutation operation, we use a random mutation method that each gene of an individual has a probability of mutation.

As shown in Figure 4(b), this procedure will generally improve the population’s average fitness since only the best organisms from the 1st generation are chosen for breeding. Then, a fitter population of solutions is achieved after several generations, as displayed in Figure 4(c).

(iv) Stopping
The above process maintains until one of the following stopping criteria is reached: (a) the number of iterations being more than 1000; (b) function tolerance (average change in the fitness function

Table 1: Parameters for the experiment and numerical simulations for four protocols. \( \eta_d \) is the detection efficiency, \( e_d \) denotes the optical misalignment, \( Y_0 \) is the dark count rate, and \( N \) is the number of emitted signals.

<table>
<thead>
<tr>
<th>Protocol</th>
<th>( \eta_d ) (%)</th>
<th>( e_d ) (%)</th>
<th>( Y_0 )</th>
<th>( N )</th>
</tr>
</thead>
<tbody>
<tr>
<td>BB84 [44]</td>
<td>10</td>
<td>1</td>
<td>( 4 \times 10^{-6} )</td>
<td>( 1 \times 10^7 )</td>
</tr>
<tr>
<td>Asymmetric MDI [45]</td>
<td>65</td>
<td>0.5</td>
<td>( 8 \times 10^{-7} )</td>
<td>( 1 \times 10^{14} )</td>
</tr>
<tr>
<td>Symmetric MDI [29]</td>
<td>53</td>
<td>1.8</td>
<td>( 4 \times 10^{-8} )</td>
<td>( 3 \times 10^{13} )</td>
</tr>
<tr>
<td>RFI-MDI [46]</td>
<td>12.5</td>
<td>0.5</td>
<td>( 1.2 \times 10^{-6} )</td>
<td>( 3 \times 10^{12} )</td>
</tr>
</tbody>
</table>
Table 2: Comparison of time consumption estimated by a desktop PC with an Intel® Core™ i7-10700F CPU @2.90 GHz and NVIDIA GeForce GT 1030 GPU.

<table>
<thead>
<tr>
<th>Protocol</th>
<th>Our method (s)</th>
<th>Local search (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>BB84</td>
<td>0.048–0.062</td>
<td>0.031–0.041</td>
</tr>
<tr>
<td>RFI</td>
<td>1.0–1.7</td>
<td>0.5–0.92</td>
</tr>
<tr>
<td>Symmetric MDI-QKD</td>
<td>1.4–2.1</td>
<td>1.2–1.6</td>
</tr>
<tr>
<td>Asymmetric MDI-QKD</td>
<td>3.1–4.2</td>
<td>2.8–3.5</td>
</tr>
</tbody>
</table>

value within 50 successive generations) being less than $\varepsilon = 10^{-20}$. Well-optimized parameters will typically be attained once each of the two criteria is met, as illustrated in Figure 4.

4. Numerical Results and Discussion

In this section, we report the use of the proposed method to optimize parameters for four common QKD protocols: BB84 (nonconvex key rate function), seven-intensity (discontinuity of first-order derivatives), symmetric MDI-QKD (large search space), and RFI-MDI-QKD (narrow search spaces over a long transmission distance). The experimental parameter sets were taken from previous works, as summarized in Table 1. The results are plotted in Figure 5. It can be seen that the proposed method can optimize parameters well for all protocols. For comparison, we also plot the key rate using an LSA. It can be seen that our proposed method achieves very similar levels of key rate compared to the LSA for the symmetrical MDI-QKD and RFI-MDI-QKD protocols, as shown in Figures 5(c) and 5(d). In this test, the performance of LSA is mainly determined by the chosen empirical starting point that needs output a positive key rate in the first round. But this starting point would not always be the most optimal starting point. Hence, the GA is not always better than LSA for all MDI-QKD, for example, in RFI-MDI-QKD. Furthermore, our method outperforms the LSA for BB84 and asymmetrical MDI-QKD protocols: (i) LSA outputs multiple maxima, which results in a nonsmooth key rate, as indicated by the blue line in Figure 5(a), and (ii) LSA exhibits poor performance for the seven-intensity protocol, as shown in Figure 5(b).

We also performed a timing test for the two algorithms. The algorithms were run using the same compiler, and ten samples were optimized for each protocol. The average time consumption was recorded, as shown in Table 2. It is clear that our method exhibits almost the same efficiency as the LSA. Here, the LSA initialized with an empirical starting point that outputs a positive key rate in the first round. This simple modification saves some time in finding a local maximum. The time consumption of GA can further be improved by setting more proper algorithm parameters, such as population size and crossover and mutation probabilities.

5. Conclusion

In summary, we proposed an efficient way to optimize the parameters for a given decoy-state QKD protocol. We tested our method with four common QKD protocols and demonstrated that it outperforms an LSA and can accurately and efficiently find optima. Furthermore, we showed that our method outperforms LSAs, which has several limitations, in QKD parameter optimization. In this work, we apply the GA for QKD optimization and obtain a good performance. The well feature of GA would be potential to solve other problems in QKD, as Ref. [47, 48], where the authors applied the GA to find optimal QKD protocols for given user-specified restrictions and arbitrary quantum channels. Furthermore, we expect that our method could be a valuable optimization tool for other quantum information processing tasks, such as blinding quantum computing [49, 50] and quantum secure direct communication [51].

Added Note. When our work is in submission, we successfully exploit our method to improve the efficiency of parameter optimization in the quantum secure direct communication protocol [52]. In contrast to our present manuscript, the paper was focused on designing a quantum secure direct communication protocol, which is an important branch of quantum cryptography but different from QKD.

Data Availability

Data are available on request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This study was supported by the National Natural Science Foundation of China (nos. 62171144, 62031024, and 11865004) and the Guangxi Science Foundation (no. 2021GXNSFAA220011).

References