MDPI

Article

Reliable Optimization of Arbitrary Functions over Quantum Measurements

Jing Luo and Jiangwei Shang *

Key Laboratory of Advanced Optoelectronic Quantum Architecture and Measurement of Ministry of Education, School of Physics, Beijing Institute of Technology, Beijing 100081, China
* Correspondence: jiangwei.shang@bit.edu.cn

Abstract: As the connection between classical and quantum worlds, quantum measurements play a unique role in the era of quantum information processing. Given an arbitrary function of quantum measurements, how to obtain its optimal value is often considered as a basic yet important problem in various applications. Typical examples include but are not limited to optimizing the likelihood functions in quantum measurement tomography, searching the Bell parameters in Bell-test experiments, and calculating the capacities of quantum channels. In this work, we propose reliable algorithms for optimizing arbitrary functions over the space of quantum measurements by combining the so-called Gilbert's algorithm for convex optimization with certain gradient algorithms. With extensive applications, we demonstrate the efficacy of our algorithms with both convex and nonconvex functions.

Keywords: quantum measurement; Gilbert's algorithm; convex optimization; nonconvex optimization

1. Introduction

In quantum information science, numerous complex mathematical problems remain to be solved. Since the set of quantum states as well as quantum measurements form convex sets, various important tasks in this field, such as the calculation of ground state energy, violation of the Bell inequality, and the detection and quantification of quantum entanglement [1,2], conform to the framework of convex optimization theory. The primary tool in convex optimization is semidefinite programming (SDP) [3,4], which can be used to derive relaxed constraints and provide accurate solutions for a large number of computationally challenging tasks. However, serious drawbacks also exist for SDP including its slow computation speed and low accuracy. For instance, SDP can only compute up to four qubits in quantum state tomography (QST), while improved superfast algorithms [5] can quickly go up to eleven qubits with a higher precision. Consequently, developing more efficient algorithms in convex optimization is becoming more and more crucial as quantum technologies rapidly advance.

Recently, an efficient convex optimization algorithm [6] was proposed by Brierley et al. based on the so-called Gilbert's algorithm [7]. Concurrently, Ref. [8] used Gilbert's algorithm to investigate whether nonlocal relationships can be distinguished in polynomial time. In Ref. [9], Gilbert's algorithm was employed as a tool to satisfy certain constraints, based on which two reliable convex optimization schemes over the quantum state space were proposed. In addition, some nonconvex optimization algorithms were also brought out for QST; for instance, the one in Ref. [10] is faster and more accurate as compared to previous approaches. One notices that all these studies concern only the optimization over quantum state space, with the consideration over quantum measurement space rarely being mentioned.

In fact, various important and meaningful problems related to quantum measurements exist in convex optimization, including, for example, searching the Bell parameters in

Bell-test experiments [11], optimizing the correlation of quantum measurements under different measurement settings [12–15], and maximizing the likelihood functions in quantum measurement tomography. Meanwhile, characterization of quantum measurements forms the basis for quantum state tomography [16–18] and quantum process tomography [19–21]. Therefore, convex optimization over the quantum measurement space stands as an independent yet important problem in quantum information theory. However, the space of quantum measurements is much more complex as compared to the quantum state space since it is possible to produce an infinite variety of different measurement outcomes as long as the probabilities for these outcomes sum to one. Recently, Ref. [22] proposed a method to optimize over the measurement space based on SDP, but it fails to solve complex tasks due to the intrinsic problem with SDP. Worst of all, nonconvex functions [23] easily appear in the space of quantum measurements. Unlike convex functions, local optima might be found during the process of optimization. Hence, nonconvex optimization is regarded as more difficult than convex optimization. In this work, we propose two reliable algorithms for optimizing arbitrary functions over the space of quantum measurements by combining the so-called Gilbert's algorithm for convex optimization with the direct-gradient (DG) algorithm as well as the accelerated projected gradient (APG) algorithm. With extensive applications, we demonstrate the efficacy of our algorithms with both convex and nonconvex functions.

This work is organized as follows: In Section 2, we propose two reliable algorithms for optimizing over quantum measurement space by combining Gilbert's algorithm with the DG and APG algorithms, respectively. The universality of our method is demonstrated by several examples with both convex and nonconvex functions in Section 3. The last Section 4 provides the conclusions.

2. Function Optimization

In the quantum state space $\mathcal{Q}$, an arbitrary state ρ should satisfy the conditions

$$\rho \geq 0, \tag{1}$$
$$\mathrm{tr}(\rho) = 1. \tag{2}$$

Given a smaller convex subset $\mathcal{C} \in \mathcal{Q}$, Gilbert's algorithm can be used to approximately find the closest state $\rho^{\mathcal{C}} \in \mathcal{C}$ with respect to ρ [9]. In general, for an arbitrary matrix M in the matrix space $\mathcal{M}$, we employ Gilbert's algorithm to search for the closest quantum state $\rho^{\mathcal{Q}} \in \mathcal{Q}$ with respect to M. Throughout this work, let us denote the operation by using Gilbert's algorithm as

$$\rho^{\mathcal{Q}} \equiv \mathcal{S}(M). \tag{3}$$

Given experimental data, it is critical to identify the measurement settings that are most compatible with the data. Here, we consider the quantum measurement space Ω as all the positive operator-valued measures (POVMs). A quantum measurement device is characterized by a set of operators $\{\Pi_l\}$, which have to satisfy two constraints

$$\Pi_l \geq 0, \tag{4}$$
$$\sum_{l=1}^{L} \Pi_l = \mathbb{I}, \tag{5}$$

where L is the total number of operators in the set. Denote a function $\mathcal{F}[\{\Pi_l\}]$ defined over the quantum measurement space Ω. We assume that $\mathcal{F}[\{\Pi_l\}]$ is differentiable with the gradient $\nabla \mathcal{F}[\{\Pi_l\}] \equiv G[\{\Pi_l\}]$. The objective is to optimize $\mathcal{F}[\{\Pi_l\}]$ over the entire quantum measurement space, and we have

$$\text{optimize} \quad \mathcal{F}[\{\Pi_l\}], \tag{6a}$$
$$\text{s.t.} \quad \{\Pi_l\} \in \Omega. \tag{6b}$$

A simple gradient method is very likely to take $\{\Pi_l\}$ outside of the quantum measurement space; for this, we employ Gilbert's algorithm to guarantee the condition in Equation (4). In addition, we rewrite the POVM as $\{\Pi_l\} = \{\Pi_1, \Pi_2, \ldots, \Pi_{L-1}, \mathbb{I} - \sum_{l=1}^{L-1} \Pi_l\}$ to satisfy the condition in Equation (5). Then, the structure of optimization proceeds as follows.

Taking the to-be-minimized objective function as an example, for the $(k+1)$th iteration, first update the $(L-1)$ elements foremost of the measurement operators with the DG scheme to obtain

$$\begin{aligned}
\Pi_{l,k+1} &= \Pi_{l,k} - \epsilon G\big(\Pi_{l,k}\big) \\
&\equiv DG\Big[\Pi_{l,k}, G\big(\Pi_{l,k}\big), \epsilon\Big].
\end{aligned} \tag{7}$$

Here, ϵ represents the step size of the update which can be any positive value, and k is the number of iterations. Second, normalize the measurement operators $\Pi_{l,k+1}$ as density matrices $\rho_{l,k+1}$, such that

$$\rho_{l,k+1} = \frac{\Pi_{l,k+1}}{\mathrm{tr}(\Pi_{l,k+1})}, \tag{8}$$

which could be nonphysical. Third, use Gilbert's algorithm to project $\rho_{l,k+1}$ back to the quantum state space $\mathcal{Q}$, i.e., $\rho_{l,k+1} \to \rho_{l,k+1}^{\mathcal{Q}} = \mathcal{S}(\rho_{l,k+1})$. Finally, reconstruct the physical measurement operators as

$$\big\{\Pi_{l,k+1}^{\Omega} = \rho_{l,k+1}^{\mathcal{Q}} t_{l,k+1}\big\}_{l=1}^{L-1}, \tag{9}$$

$$\Pi_{L,k+1}^{\Omega} = \mathbb{I} - \sum_{l=1}^{L-1} \Pi_{l,k+1}^{\Omega}, \tag{10}$$

where the parameter t_l is obtained by fixing the obtained $\rho_{l,k+1}^{\mathcal{Q}}$ to obtain $\big\{t_{l,k+1}\big\}_{l=1}^{L-1} = \mathrm{argmin}\, \mathcal{F}\big[\{t_{l,k+1}\}_{l=1}^{L-1}\big]$. Here, to ensure that the first $(L-1)$ measurement operators satisfy condition Equation (4), only $t_{l,k+1} \geq 0$ is required since $\rho_{l,k+1}^{\mathcal{Q}} \geq 0$ is guaranteed by using Gilbert's algorithm. Meanwhile, in order to ensure that the last element of the new POVM satisfies the condition in Equation (4), let

$$\Pi_{L,k+1}^{\Omega} = \mathbb{I} - \sum_{l=1}^{L-1} \big(\rho_{l,k+1}^{\mathcal{Q}} t_{l,k+1}\big) \geq 0. \tag{11}$$

Hence, we obtain the new POVM $\big\{\Pi_{k+1,l}^{\Omega}\big\}$ that satisfies the condition in Equation (6b) after each iteration. Whenever the difference between the values of the adjacent iterations is less than a certain threshold, the iteration stops, and the optimal POVM is obtained. Otherwise, continue with the iteration and the step size is controlled by a step factor β. When $\mathcal{F}_k < \mathcal{F}_{k-1}$, the step size is appropriately selected. When $\mathcal{F}_k > \mathcal{F}_{k-1}$, it indicates that the step size selection is too large, and the step factor β needs to be used to adjust the step size. See the DG algorithm in Algorithm 1.

However, the DG algorithm has some disadvantages, such as slow optimization speed and low accuracy. For faster convergence, one can choose the APG algorithm [5,24]. The APG algorithm adjusts the direction of the gradient at each step, which improves the convergence speed of the algorithm. In simple terms, the APG algorithm has introduced a companion operator $E_{l,k} = \Pi_{l,k} + \frac{\theta_{k-1}-1}{\theta_k}\big(\Pi_{l,k} - \Pi_{l,k-1}\big)$, which provides the momentum of the previous step controlled by the parameter θ, in order to update the measurement operators $\Pi_{l,k} = E_{l,k-1} - \epsilon\, G\big(E_{l,k-1}\big)$. See the specific process shown in Algorithm 2.

Algorithm 1: DG algorithm

Input: $\epsilon > 0, 0 < \beta < 1$, choose any $\{\Pi_{l,0}\}_{l=1}^{L-1} \in \Omega, \mathcal{F}_0 = \mathcal{F}[\{\Pi_{l,0}\}]$.
Output: $\{\Pi_l\}$.

1 **for** $k = 1, \cdots, \textbf{do}$
2   **for** $l = 1, \cdots, L-1$ **do**
3   Update $\Pi_{l,k} = \text{DG}[\Pi_{l,k-1}, G(\Pi_{l,k-1}), \epsilon]$. Calculate $\rho_{l,k}$ and $\rho_{l,k}^{\mathcal{Q}} = \mathcal{S}(\rho_{l,k})$.
4   **end**
5   Gain $\{t_{l,k}\}_{l=1}^{L-1} = \text{argmin } \mathcal{F}_k\left[\{t_{l,k}\}_{l=1}^{L-1}\right]$. Calculate $\{\Pi_{l,k}^{\Omega}\}, \mathcal{F}_k = \mathcal{F}[\{\Pi_{l,k}^{\Omega}\}]$.
6   Termination criterion!
7   **if** $\mathcal{F}_k > \mathcal{F}_{k-1}$ **then**
8   Reset $\epsilon = \beta\epsilon$, and $\{\Pi_{l,k}\} = \{\Pi_{l,k-1}^{\Omega}\}$.
9   **end**
10 **end**

Algorithm 2: APG algorithm

Input: $\epsilon > 0, 0 < \beta < 1$, choose any $\{\Pi_{l,0}\}_{l=1}^{L-1} \in \Omega, \{E_{l,0}\} = \{\Pi_{l,0}\}, \theta_0 = 1$, and
     $\mathcal{F}_0 = \mathcal{F}[\{\Pi_{l,0}\}]$.

1 . **Output:** $\{\Pi_l\}$.
2 **for** $k = 1, \cdots, \textbf{do}$
3   **for** $l = 1, \cdots, L-1$ **do**
4   Update $\Pi_{l,k} = E_{l,k-1} - \epsilon\, G(E_{l,k-1})$. Calculate $\rho_{l,k}$ and $\rho_{l,k}^{\mathcal{Q}} = \mathcal{S}(\rho_{l,k})$.
5   **end**
6   Gain $\{t_{l,k}\}_{l=1}^{L-1} = \text{argmin } \mathcal{F}_k\left[\{t_{l,k}\}_{l=1}^{L-1}\right]$. Calculate $\{\Pi_{l,k}^{\Omega}\}, \mathcal{F}_k = \mathcal{F}[\{\Pi_{l,k}^{\Omega}\}]$.
7   Termination criterion!
8   **if** $\mathcal{F}_k > \mathcal{F}_{k-1}$ **then**
9   Reset $\epsilon = \beta\epsilon$, and $\{\Pi_{l,k}\} = \{\Pi_{l,k-1}^{\Omega}\}$. $\{E_{l,k}\} = \{\Pi_{l,k}\}$, and $\theta_k = 1$.
10   **else**
11   Set $\theta_k = \frac{1}{2}\left(1 + \sqrt{1 + 4\theta_{k-1}^2}\right)$;
12   Update $\{E_{l,k}\} = \left\{\Pi_{l,k} + \frac{\theta_{k-1}-1}{\theta_k}(\Pi_{l,k} - \Pi_{l,k-1})\right\}$.
13   **end**
14 **end**

---

3. Applications

In this section, we demonstrate the efficacy of our algorithms by optimizing arbitrary convex as well as nonconvex functions over the space of quantum measurements.

3.1. Convex Functions

In quantum measurement tomography [25–27], a set of known probe states ρ_m is measured to provide the information needed to reconstruct an unknown POVM $\{\Pi_l\}$. The probability that the device would respond to the quantum state ρ_m by producing the outcome Π_l is given by

$$p_{lm} = \text{tr}(\rho_m \Pi_l). \tag{12}$$

Typically, the linear inversion method [28] can be used to obtain the ideal POVM, but nonphysical results are likely to be obtained. Then, the maximum likelihood estimation (MLE) [29] is proposed to reconstruct the POVM that satisfies all the conditions. However, MLE fails to return any meaningful results when the target POVM is of low rank, which is quite typical, especially in higher-dimensional spaces. These problems can be avoided by using our algorithms.

To estimate the operators $\{\Pi_l\}$, we maximize the likelihood function

$$\mathcal{L}\big[\{\Pi_l\}\big] = \prod_{l=1}^{L}\prod_{m=1}^{M}\Big[\mathrm{tr}\big(\rho_m\Pi_l\big)\Big]^{f_{lm}}, \tag{13}$$

where M is the number of different input states ρ_m, and

$$f_{lm} = \frac{n_{lm}}{n}, \tag{14}$$

with n_{lm} denoting the number of lth outcome when measuring the mth state ρ_m, and n representing the total number of measured input states. One can see that $\mathcal{L}\big[\{\Pi_l\}\big]$ is not strictly concave, while the log-likelihood $\ln\mathcal{L}\big[\{\Pi_l\}\big]$ is. Here, we minimize the negative log-likelihood function $\mathcal{F}\big[\{\Pi_l\}\big] = -\ln\mathcal{L}\big[\{\Pi_l\}\big]$ with

$$\ln\mathcal{L}\big[\{\Pi_l\}\big] = \sum_{l=1}^{L}\sum_{m=1}^{M} f_{lm}\ln p_{lm}. \tag{15}$$

To satisfy the condition in Equation (5), rewrite the objective function as

$$\ln\mathcal{L}\big[\{\Pi_l\}\big] = \sum_{l=1}^{L-1}\sum_{m=1}^{M} f_{lm}\ln\big[\mathrm{tr}\big(\rho_m\Pi_l\big)\big] + \sum_{m=1}^{M} f_{Lm}\ln\left\{\mathrm{tr}\left[\rho_m\left(\mathbb{I} - \sum_{l=1}^{L-1}\Pi_l\right)\right]\right\}. \tag{16}$$

The gradient of $\ln\mathcal{L}\big[\{\Pi_l\}\big]$ with respect to Π_l is

$$\nabla\ln\mathcal{L}\big(\Pi_l\big) = \sum_{m=1}^{M}\left[\left(\frac{f_{lm}}{p_{lm}} - \frac{f_{lm}}{1 - \sum_{l=1}^{L-1} p_{lm}}\right)\rho_m\right]. \tag{17}$$

For numerical simulations, we mainly consider Pauli measurements which are the most commonly-used measurements in quantum information processing. Then, the cases of one qubit, one qutrit, two qubits, and two qutrits are used for the experimental setup, respectively. Specifically, the setups of these four scenarios are described below.

3.1.1. One Qubit

For one qubit, we take the eigenstates of σ_z and the superposition states $-\frac{1}{\sqrt{2}}\big(|0_z\rangle \pm |1_z\rangle\big)$ and $\frac{1}{\sqrt{2}}\big(|0_z\rangle \pm i|1_z\rangle\big)$ as the input states. In the measurement setup, we select the projection of the spin along the x-axis, i.e.,

$$\Pi_1 = |0_x\rangle\langle 0_x|; \quad \Pi_2 = |1_x\rangle\langle 1_x|. \tag{18}$$

3.1.2. One Qutrit

For one qutrit, we use 12 different input states: three eigenstates of σ_z, $|-1_z\rangle$, $|0_z\rangle$ and $|1_z\rangle$, and nine superposition states $\frac{1}{\sqrt{2}}\big(|-1_z\rangle + e^{i\psi_j}|0_z\rangle\big)$, $\frac{1}{\sqrt{2}}\big(|0_z\rangle + e^{i\psi_j}|1_z\rangle\big)$ and $\frac{1}{\sqrt{2}}\big(|-1_z\rangle + e^{i\Psi_j}|1_z\rangle\big)$, where $j = 1,2,3$; and $\psi_1 = 0$, $\psi_2 = \frac{\pi}{2}$, and $\psi_3 = \pi$. The device measures the projection of the spin along the x-axis, and the POVM are projectors

$$\Pi_1 = |-1_x\rangle\langle -1_x|; \quad \Pi_2 = |0_x\rangle\langle 0_x|; \quad \Pi_3 = |1_x\rangle\langle 1_x|. \tag{19}$$

3.1.3. Two Qubits

In the case of two qubits, we take the tensor products of the four eigenstates of two Pauli-Z operators $|0_z0_z\rangle, |1_z1_z\rangle, |0_z1_z\rangle, |1_z0_z\rangle$ and the superposition states $\frac{1}{\sqrt{2}}\left(|0_z0_z\rangle + e^{i\psi_j}|0_z1_z\rangle\right)$, $\frac{1}{\sqrt{2}}\left(|0_z0_z\rangle + e^{i\psi_j}|1_z0_z\rangle\right)$, $\frac{1}{\sqrt{2}}\left(|0_z0_z\rangle + e^{i\psi_j}|1_z1_z\rangle\right)$, $\frac{1}{\sqrt{2}}\left(|0_z1_z\rangle + e^{i\psi_j}|1_z0_z\rangle\right)$, $\frac{1}{\sqrt{2}}\left(|0_z1_z\rangle + e^{i\psi_j}|1_z1_z\rangle\right)$, $\frac{1}{\sqrt{2}}\left(|1_z0_z\rangle + e^{i\psi_j}|1_z1_z\rangle\right)$ as the probe states, where $j = 1, 2, 3$; $\psi_1 = 0$, $\psi_2 = \frac{\pi}{2}$, and $\psi_3 = \pi$. Then, we choose the following POVM for the experimental simulation:

$$
\begin{aligned}
\Pi_1 &= |0_x0_x\rangle\langle 0_x0_x|; \quad \Pi_2 = |0_x1_x\rangle\langle 0_x1_x|; \\
\Pi_3 &= |1_x0_x\rangle\langle 1_x0_x|; \quad \Pi_4 = |0_x0_x\rangle\langle 0_x0_x|.
\end{aligned}
\tag{20}
$$

3.1.4. Two Qutrits

Finally, for the case of two qutrits, we perform a numerical simulation of the Stern–Gerlach apparatus measuring two particles with spin-1. We assume 45 different input states: $|1_z - 1_z\rangle, |-1_z0_z\rangle, |-1_z1_z\rangle, |0_z - 1_z\rangle, |0_z0_z\rangle, |0_z1_z\rangle, |1_z0_z\rangle, |1_z1_z\rangle, |-1_z - 1_z\rangle$ and 36 superposition states. In the simulation, the device measures the projection of the spin along the x-axis, and the POVM are projectors

$$
\begin{aligned}
\Pi_1 &= |0_x1_x\rangle\langle 0_x1_x|; \quad \Pi_2 = |0_x - 1_x\rangle\langle 0_x - 1_x|; \\
\Pi_3 &= |1_x0_x\rangle\langle 1_x0_x|; \quad \Pi_4 = |1_x - 1_x\rangle\langle 1_x - 1_x|; \\
\Pi_5 &= |0_x0_x\rangle\langle 0_x0_x|; \quad \Pi_6 = |-1_x0_x\rangle\langle -1_x0_x|; \\
\Pi_7 &= |1_x1_x\rangle\langle 1_x1_x|; \quad \Pi_8 = |-1_x1_x\rangle\langle -1_x1_x|; \\
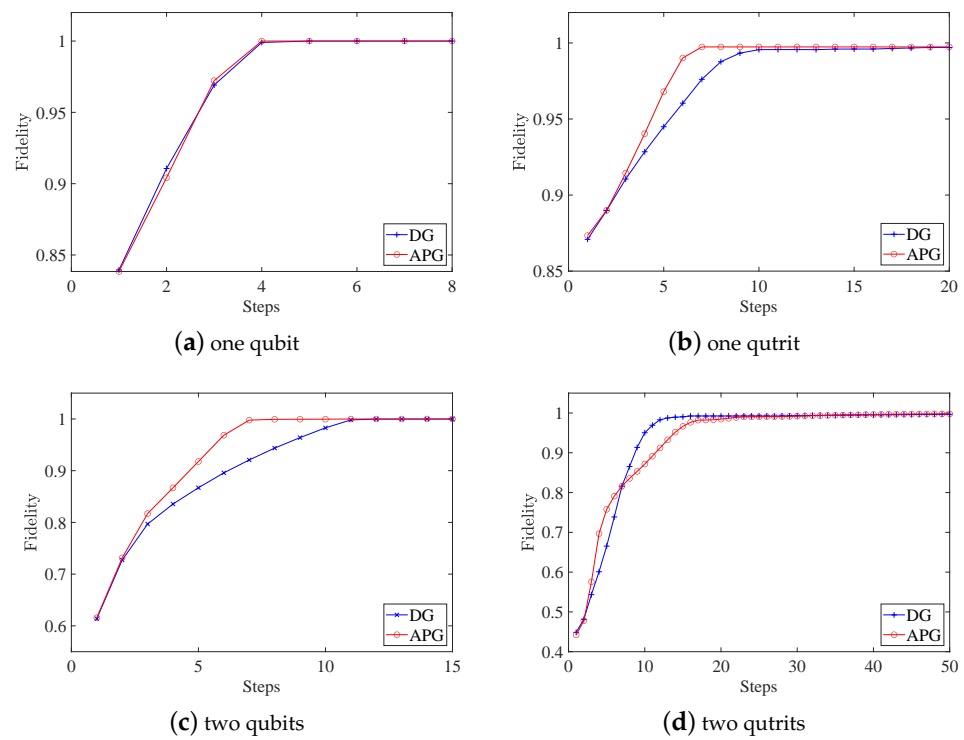\Pi_9 &= |-1_x - 1_x\rangle\langle -1_x - 1_x|.
\end{aligned}
\tag{21}
$$

For each case of simulation, the number of measurements for each probe state is 300, 10^5, 10^5, and 5×10^5, respectively. Then, according to the frequency obtained by the simulated data, we use our algorithm to reconstruct the POVM. The fidelity between different POVM elements is defined as the fidelity between the two states σ and ρ, i.e.,

$$
F(\sigma, \rho) := \left(\text{tr}\sqrt{\sqrt{\sigma}\rho\sqrt{\sigma}}\right)^2 = F\left(\frac{\Pi_l}{\text{tr}(\Pi_l)}, \frac{\Pi_j}{\text{tr}(\Pi_j)}\right).
\tag{22}
$$

In addition, the overall fidelity between two POVMs $\{\Pi_l\}_{l=1}^{L}$ and $\{\Pi_j\}_{j=1}^{L}$ on a d-dimensional Hilbert space is defined by

$$
F(\Pi_l, \Pi_j) := \left[\sum_{l=1}^{L} w_l \sqrt{F\left(\frac{\Pi_l}{\text{tr}(\Pi_l)}, \frac{\Pi_j}{\text{tr}(\Pi_j)}\right)}\right]^2,
\tag{23}
$$

with $w_l = \frac{\sqrt{\text{tr}(\Pi_l)\text{tr}(\Pi_j)}}{d}$ [30]. The overall fidelities of the reconstructed POVMs are shown in Figure 1. Figures 2 and 3 present the variations of fidelity of the POVM elements reconstructed using the DG algorithm and APG algorithm with respect to the number of iteration steps in different cases. We can see that these two algorithms are almost identical in accuracy, and the fidelities of the measurement operators are close to 1. Generally speaking, the APG algorithm converges faster than the DG algorithm. In addition, one notices that the fidelity of the last element in some of the simulations is not always increasing, which is a result of the constraint that we set in Equation (11).

Figure 1. For different cases, the two algorithms are compared to reconstruct the overall fidelity of the measurements. The number of measurements used in each simulation for each probe state is 300, 10^5, 10^5, and 5×10^5, respectively. For most cases, the APG algorithm converges faster than the DG algorithm.
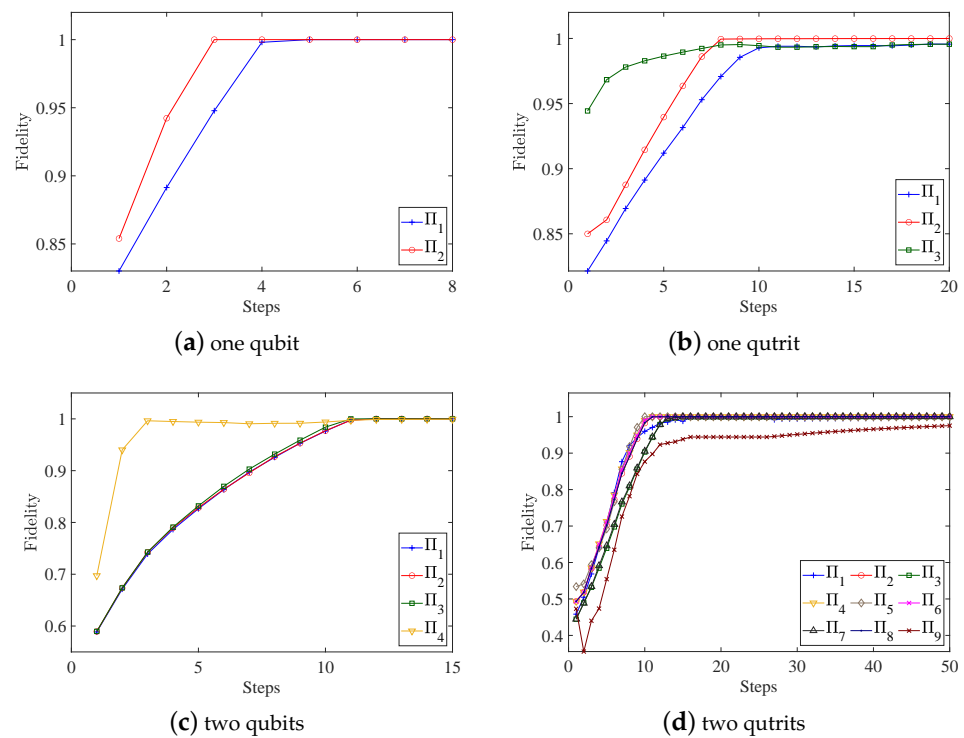
Figure 2. For different cases of the quantum measurement tomography, fidelities of the measurements obtained by the DG algorithm vary with the number of iteration steps. In general, the fidelity of each POVM element saturates to the maximum very quickly.

(**a**) one qubit

(**b**) one qutrit

(**c**) two qubits

(**d**) two qutrits
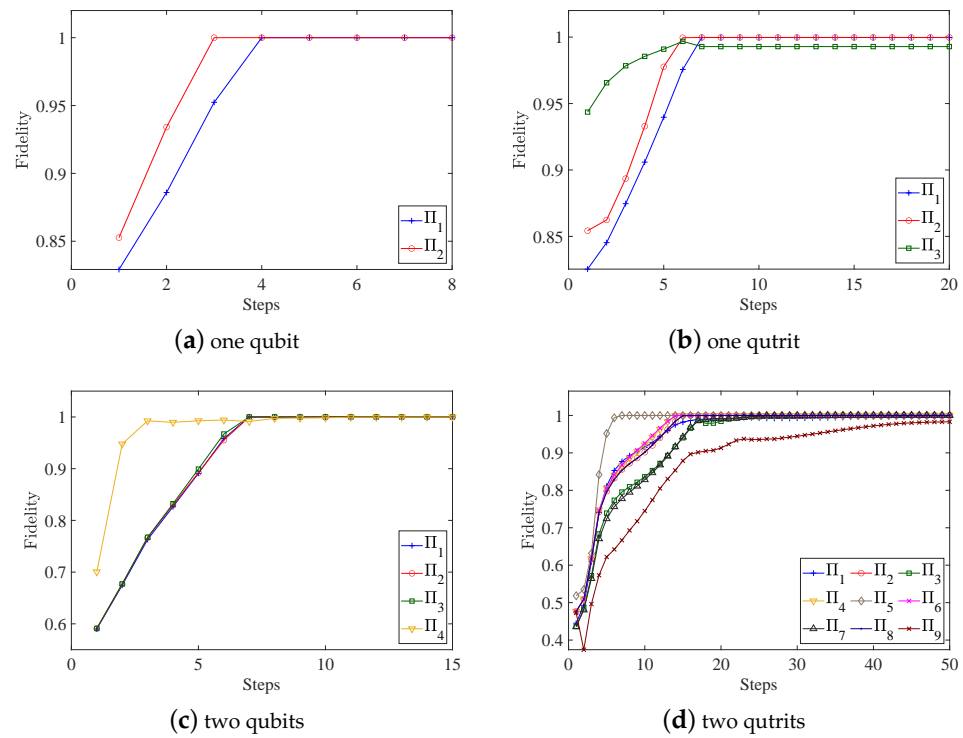
Figure 3. For different cases of the quantum measurement tomography, fidelities of the measurements obtained by the APG algorithm vary with the number of iteration steps. In general, the fidelity of each POVM element saturates to the maximum very quickly.

3.2. Nonconvex Functions

Quantum detector self-characterization (QDSC) tomography is another method for characterizing quantum measurements. Unlike quantum measurement tomography, this method does not require knowing the specific form of the input probe states, but directly optimizes the cost function based on the measurement statistic f_m to reconstruct the measurements. For POVM with L outcomes detected by m states, a data set of the measurement statistic f_{lm} is obtained. We write the distribution of the data for each state as a vector

$$\boldsymbol{f_m} = \begin{pmatrix} f_{1m} \\ f_{2m} \\ \vdots \\ f_{Lm} \end{pmatrix}. \tag{24}$$

For the one qubit case, define $N_{i,l} = \boldsymbol{b}_i^T \boldsymbol{b}_l$ and write the POVM as

$$\Pi_l = a_l I + \boldsymbol{b}_l \cdot \boldsymbol{\sigma} \tag{25}$$

under the Bloch representation, where i and l represent the number of rows i and columns l of the matrix N, $\boldsymbol{a} = (a_1 \cdots a_L)^T$, $\boldsymbol{b}_l = (b_{l,x}, b_{l,y}, b_{l,z})$, $\boldsymbol{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$, $1 \le i, l \le L$. The matrix N and vector $\boldsymbol{a}$ can be represented as

$$N_{i,l} = \boldsymbol{b}_i^T \boldsymbol{b}_l = \frac{1}{2}\mathrm{tr}(\Pi_i \Pi_l) - \frac{1}{4}\mathrm{tr}(\Pi_i)\mathrm{tr}(\Pi_l), \tag{26}$$

$$a_l = \frac{1}{2}\mathrm{tr}(\Pi_l). \tag{27}$$

Then, optimization of the cost function $\mathcal{F}(N^+, \boldsymbol{a})$ is given by [23]

$$\min \quad \sum_m \left[1 - (\boldsymbol{f_m} - \boldsymbol{a})^T N^+ (\boldsymbol{f_m} - \boldsymbol{a}) \right]^2, \tag{28a}$$

$$\text{s.t.} \quad a_l^2 - N_{l,l} >= 0, \tag{28b}$$

where N^+ stands for the Moore–Penrose pseudoinverse of N. One notices that the objective function is nonconvex. Optimization of nonconvex functions is difficult as local minima might be found. Interestingly, we find that our algorithm can also be used to optimize nonconvex functions. Since our algorithm guarantees the conditions for quantum measurements, one only needs to optimize the objective function regardless of the constraint in Equation (28b).

For numerical simulations, we choose 50 probe states:

$$\frac{1}{2}(\mathbb{I} + \sigma_z), \quad \frac{1}{2}(\mathbb{I} - \sigma_z), \quad \frac{1}{2}\left(\mathbb{I} + \sin\frac{i\pi}{4}\cos\frac{n\pi}{8}\sigma_x + \sin\frac{i\pi}{4}\sin\frac{n\pi}{8}\sigma_y + \cos\frac{i\pi}{4}\sigma_z \right), \tag{29}$$

where $i = 1, 2, \cdots, 6$; $n = 1, 2, \cdots, 8$. In addition, we use the two-dimensional SIC POVM as the measurement device, and each state is measured 200 times. The APG algorithm is used to optimize the objective function. First, select any set of POVM operators in the measurement space, and use Equations (26) and (27) to obtain the initial values N_k^+ and $\boldsymbol{a_k}$, respectively. Similarly, we calculate the gradient of the objective function in Equation (28a). The gradient of the objective function is given by

$$\delta\mathcal{F}(\boldsymbol{a}) = \sum_m 2(1 - \boldsymbol{f_m} - \boldsymbol{a})^T N^+ (\boldsymbol{f_m} - \boldsymbol{a}) \left\{ (N^+)^T \boldsymbol{f_m} + N^+ \boldsymbol{f_m} - \left[N^+ + (N^+)^T \right] \boldsymbol{a} \right\}, \tag{30}$$

$$\delta\mathcal{F}(N^+) = \sum_m -2(1 - \boldsymbol{f_m} - \boldsymbol{a})^T N^+ (\boldsymbol{f_m} - \boldsymbol{a})^2 (\boldsymbol{f_m} - \boldsymbol{a})^T. \tag{31}$$

The values of N_{k+1} and $\boldsymbol{a_{k+1}}$ are obtained by iterating over N_k and $\boldsymbol{a_k}$ using gradient descent; then, $\boldsymbol{b_{l,k+1}}$ is obtained by decomposing N_{k+1}. In the experiment, we specify that the reference frame, i.e., the vector $\boldsymbol{b_1}$ is parallel to the z-direction of the Bloch sphere, and set the xz plane of the Bloch sphere as the plane determined by the vectors $\boldsymbol{b_1}$ and $\boldsymbol{b_2}$. This is equivalent to $b_{1,x} = b_{1,y} = b_{2,y} = 0$. Then, $\{\Pi_{l,k+1}\}_{l=1}^{L-1}$ can be obtained by using Equation (25), which is the update for $\{\Pi_{l,k}\}_{l=1}^{L-1}$.

The fidelity of each POVM element can approach 1 in a very small number of iteration steps; see Figure 4. Then, the fidelities of the measurements are compared with the ones reported in [23], demonstrating that the performance of our algorithm is slightly better; see Figure 5.
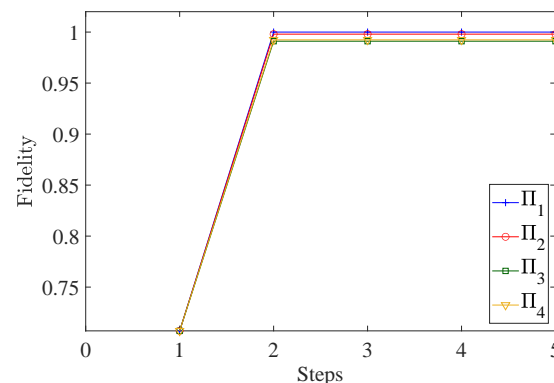
Figure 4. In the case of QDSC, the fidelity of each element of the two-dimensional SIC POVM saturates to the maximum by using only two steps.
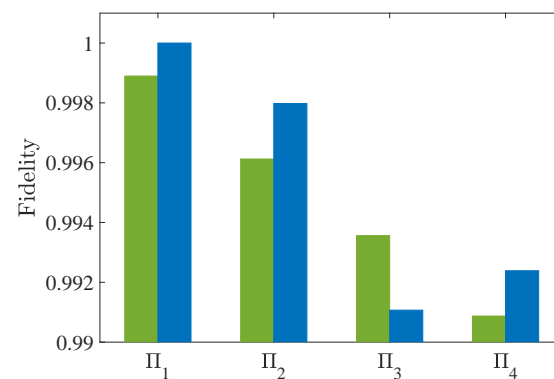
Figure 5. Comparison of the fidelities of the reconstructed quantum measurements between the APG algorithm (blue) and the method in [23] (green).

4. Conclusions

We have proposed two reliable algorithms for optimizing arbitrary functions over the quantum measurement space. For a demonstration, we have shown several examples on the convex function of quantum measurement tomography with different dimensions as well as a nonconvex function of one qubit in quantum detector self-characterization tomography. Surprisingly, our method does not encounter the problem of rank deficiency. Compared with SDP, our method can be easily applied to higher-dimensional cases as well as to optimize nonconvex functions. Moreover, our method reports better results as compared to previous approaches. For future work, we will consider the optimization over the joint space of quantum states and quantum measurements, for tasks such as calculating the capacity of quantum channels.

Author Contributions: J.L. performed the numerical calculations. All authors contributed to the interpretation of the results, preparation, and writing of the manuscript. All authors have read and agreed to the published version of the manuscript.

Institutional Review Board Statement: Not applicable.

Data Availability Statement: The data that support the findings of this study are available from the corresponding author upon reasonable request.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Gühne, O.; Tóth, G. Entanglement detection. *Phys. Rep.* **2009**, *474*, 1–75. [CrossRef]
2. Amico, L.; Fazio, R.; Osterloh, A.; Vedral, V. Entanglement in many-body systems. *Rev. Mod. Phys.* **2008**, *80*, 517. [CrossRef]
3. Grant, M.C.; Boyd, S.P. Graph implementations for nonsmooth convex programs. In *Recent Advances in Learning and Control*; Springer: London, UK, 2008; pp. 95–110.
4. Boyd, S.P.; Vandenberghe, L. Convex Optimization. *IEEE Trans. Automat. Contr.* **2006**, *51*, 1859–1859.
5. Shang, J.; Zhang, Z.; Ng, H.K. Superfast maximum-likelihood reconstruction for quantum tomography. *Phys. Rev. A* **2017**, *95*, 062336. [CrossRef]
6. Brierley, S.; Navascues, M.; Vertesi, T. Convex separation from convex optimization for large-scale problems. *arXiv* **2016**, arXiv:1609.05011.
7. Gilbert, E.G. An iterative procedure for computing the minimum of a quadratic form on a convex set. *SIAM J. Control. Optim.* **1966**, *4*, 61–80. [CrossRef]
8. Montina, A.; Wolf, S. Can non-local correlations be discriminated in polynomial time? *arXiv* **2016**, arXiv:1609.06269.
9. Shang, J.; Gühne, O. Convex optimization over classes of multiparticle entanglement. *Phys. Rev. Lett.* **2018**, *120*, 050506. [CrossRef]

10. Kyrillidis, A.; Kalev, A.; Park, D.; Bhojanapalli, S.; Caramanis, C.; Sanghavi, S. Provable compressed sensing quantum state tomography via non-convex methods. *NPJ Quantum Inf.* **2018**, *4*, 1–7. [CrossRef]
11. Smania, M.; Kleinmann, M.; Cabello, A.; Bourennane, M. Avoiding apparent signaling in Bell tests for quantitative applications. *arXiv* **2018**, arXiv:1801.05739.
12. Becker, S.R.; Candès, E.J.; Grant, M.C. Templates for convex cone problems with applications to sparse signal recovery. *Math. Program. Comput.* **2011**, *3*, 165–218. [CrossRef]
13. Kleinmann, M.; Cabello, A. Quantum correlations are stronger than all nonsignaling correlations produced by *n*-outcome measurements. *Phys. Rev. Lett.* **2016**, *117*, 150401. [CrossRef] [PubMed]
14. Kleinmann, M.; Vértesi, T.; Cabello, A. Proposed experiment to test fundamentally binary theories. *Phys. Rev. A* **2017**, *96*, 032104. [CrossRef]
15. Hu, X.M.; Liu, B.H.; Guo, Y.; Xiang, G.Y.; Huang, Y.F.; Li, C.F.; Guo, G.C.; Kleinmann, M.; Vértesi, T.; Cabello, A. Observation of stronger-than-binary correlations with entangled photonic qutrits. *Phys. Rev. Lett.* **2018**, *120*, 180402. [CrossRef]
16. Cramer, M.; Plenio, M.B.; Flammia, S.T.; Somma, R.; Gross, D.; Bartlett, S.D.; Landon-Cardinal, O.; Poulin, D.; Liu, Y.K. Efficient quantum state tomography. *Nat. Commun.* **2010**, *1*, 1–7. [CrossRef] [PubMed]
17. Torlai, G.; Mazzola, G.; Carrasquilla, J.; Troyer, M.; Melko, R.; Carleo, G. Neural-network quantum state tomography. *Nat. Phys.* **2018**, *14*, 447–450. [CrossRef]
18. Gross, D.; Liu, Y.K.; Flammia, S.T.; Becker, S.; Eisert, J. Quantum state tomography via compressed sensing. *Phys. Rev. Lett.* **2010**, *105*, 150401. [CrossRef] [PubMed]
19. Mohseni, M.; Rezakhani, A.T.; Lidar, D.A. Quantum-process tomography: Resource analysis of different strategies. *Phys. Rev. A* **2008**, *77*, 032322. [CrossRef]
20. Altepeter, J.B.; Branning, D.; Jeffrey, E.; Wei, T.; Kwiat, P.G.; Thew, R.T.; O'Brien, J.L.; Nielsen, M.A.; White, A.G. Ancilla-assisted quantum process tomography. *Phys. Rev. Lett.* **2003**, *90*, 193601. [CrossRef] [PubMed]
21. O'Brien, J.L.; Pryde, G.; Gilchrist, A.; James, D.; Langford, N.K.; Ralph, T.; White, A. Quantum process tomography of a controlled-NOT gate. *Phys. Rev. Lett.* **2004**, *93*, 080502. [CrossRef]
22. Cattaneo, M.; Borrelli, E.M.; García-Pérez, G.; Rossi, M.A.; Zimborás, Z.; Cavalcanti, D. Semidefinite programming for self-consistent quantum measurement tomography. *arXiv* **2022**, arXiv:2212.10262.
23. Zhang, A.; Xie, J.; Xu, H.; Zheng, K.; Zhang, H.; Poon, Y.T.; Vedral, V.; Zhang, L. Experimental self-characterization of quantum measurements. *Phys. Rev. Lett.* **2020**, *124*, 040402. [CrossRef] [PubMed]
24. Beck, A.; Teboulle, M. A fast iterative shrinkage-thresholding algorithm for linear inverse problems. *SIAM J. Imaging Sci.* **2009**, *2*, 183–202. [CrossRef]
25. Natarajan, C.M.; Zhang, L.; Coldenstrodt-Ronge, H.; Donati, G.; Dorenbos, S.N.; Zwiller, V.; Walmsley, I.A.; Hadfield, R.H. Quantum detector tomography of a time-multiplexed superconducting nanowire single-photon detector at telecom wavelengths. *Opt. Express* **2013**, *21*, 893–902. [CrossRef]
26. D'Ariano, G.M.; Maccone, L.; Presti, P.L. Quantum calibration of measurement instrumentation. *Phys. Rev. Lett.* **2004**, *93*, 250407. [CrossRef] [PubMed]
27. Coldenstrodt-Ronge, H.B.; Lundeen, J.S.; Pregnell, K.L.; Feito, A.; Smith, B.J.; Mauerer, W.; Silberhorn, C.; Eisert, J.; Plenio, M.B.; Walmsley, I.A. A proposed testbed for detector tomography. *J. Mod. Opt.* **2009**, *56*, 432–441. [CrossRef]
28. Fano, U. Description of states in quantum mechanics by density matrix and operator techniques. *Rev. Mod. Phys.* **1957**, *29*, 74. [CrossRef]
29. Fiurášek, J. Maximum-likelihood estimation of quantum measurement. *Phys. Rev. A* **2001**, *64*, 024102. [CrossRef]
30. Hou, Z.; Tang, J.F.; Shang, J.; Zhu, H.; Li, J.; Yuan, Y.; Wu, K.D.; Xiang, G.Y.; Li, C.F.; Guo, G.C. Deterministic realization of collective measurements via photonic quantum walks. *Nat. Commun.* **2018**, *9*, 1–7. [CrossRef]