photonics

MDPI

Article

Fast Quantum State Reconstruction via Accelerated Non-Convex Programming

Junhyung Lyle Kim [1], George Kollias [2], Amir Kalev [3], Ken X. Wei [2] and Anastasios Kyrillidis [1,*]

[1] Computer Science Department, Rice University, Houston, TX 77005, USA
[2] IBM Quantum, IBM T.J. Watson Research Center, Yorktown Heights, NY 10598, USA
[3] Information Sciences Institute, University of Southern California, Marina del Rey, CA 90292, USA
* Correspondence: anastasios@rice.edu

Abstract: We propose a new quantum state reconstruction method that combines ideas from compressed sensing, non-convex optimization, and acceleration methods. The algorithm, called Momentum-Inspired Factored Gradient Descent (`MiFGD`), extends the applicability of quantum tomography for larger systems. Despite being a non-convex method, `MiFGD` converges provably close to the true density matrix at an accelerated linear rate asymptotically in the absence of experimental and statistical noise, under common assumptions. With this manuscript, we present the method, prove its convergence property and provide the Frobenius norm bound guarantees with respect to the true density matrix. From a practical point of view, we benchmark the algorithm performance with respect to other existing methods, in both synthetic and real (noisy) experiments, performed on the IBM's quantum processing unit. We find that the proposed algorithm performs orders of magnitude faster than the state-of-the-art approaches, with similar or better accuracy. In both synthetic and real experiments, we observed accurate and robust reconstruction, despite the presence of experimental and statistical noise in the tomographic data. Finally, we provide a ready-to-use code for state tomography of multi-qubit systems.

Keywords: quantum state tomography; non-convex optimization; matrix factorization; acceleration

Citation: Kim, J.L.; Kollias, G.; Kalev, A.; Wei, K.X.; Kyrillidis, A. Fast Quantum State Reconstruction via Accelerated Non-Convex Programming. *Photonics* **2023**, *10*, 116. https://doi.org/10.3390/photonics10020116

Received: 30 December 2022
Revised: 17 January 2023
Accepted: 19 January 2023
Published: 22 January 2023

1. Introduction

Quantum tomography is one of the main procedures to identify the nature of imperfections and deviations in quantum processing unit (QPU) implementations [1,2]. Generally, quantum tomography is composed of two main parts: (i) measuring the quantum system, and (ii) analyzing the measurement data to obtain an estimate of the density matrix (in the case of state tomography [1]), or of the quantum process (in the case of process tomography [3]). In this manuscript, we focus on the state tomography.

Quantum tomography is generally considered as a non-scalable protocol [4], as the number of free parameters that define quantum states and processes scale exponentially with the number of subsystems. In particular, quantum state tomography (QST) suffers from two bottlenecks. The first concerns about the large amount of data one needs to collect to perform tomography; the second concerns about numerically searching in an exponentially increasing space for a density matrix that is consistent with the data.

There have been various approaches to improve the scalability of QST, in terms of the amount of data required [5–7]. To address the data collection bottleneck, prior information about the unknown quantum state is often assumed. For example, in compressed sensing QST [4,8], the density matrix of the quantum system is assumed to be low-rank. In neural network QST [9–11], the wave-functions are often assumed to be real and positive, confining the landscape of quantum states (To handle more complex wave-functions, neural network approaches require a proper re-parameterization of the Restricted Boltzmann machines [9]). The prior information considered in these cases is that they are characterized by structured

quantum states [9] (Such assumptions are often the reason behind accurate solutions of neural network QST in high-dimensional spaces.) Ref. [9] considers also the case of a completely unstructured case and test the limitation of this technique, which does not perform as expected, due to lack of any exploitable structure). Similarly, in matrix-product-state tomography [12,13], one assumes that the quantum state can be represented with low bond-dimension matrix-product state.

To address the computational bottleneck, several works introduce sophisticated numerical methods to improve the efficiency of QST. In particular, variants of gradient-based convex solvers—e.g., [14–17]—have been tested on synthetic scenarios [17]. The problem is that, achieving these results often requires utilizing special-purpose hardwares, such as Graphics Processing Units (GPUs), on top of carefully designing a proper distributed system [18]. Thus, going beyond current capabilities requires novel methods that can efficiently search in the space of density matrices under more realistic scenarios. Importantly, such numerical methods should come with rigorous guarantees on their convergence and performance.

The setup we consider here is the estimation of an n-qubit state, under the prior assumption that the state is close to a pure state, and thus its density matrix is of low-rank. This assumption is justified by the state-of-the-art experiments, where the aim is to manipulate the pure states with unitary maps. From a theoretical perspective, the low-rank assumption implies that we can use compressed sensing techniques [19], which allow the recovery of the density matrix from relatively fewer measurement data [20,21].

Indeed, compressed sensing QST is widely used for estimating highly-pure quantum states; e.g., [4,22–24]. However, compressed sensing QST usually relies on convex optimization for the estimation [8], which limits the applicability to relatively small system sizes [4] (In particular, convex solvers over low-rank structures utilize the nuclear norm over $2^n \times 2^n$ matrices. This assumes calculating all 2^n eigenvalues of such matrices per iteration, which has cubic complexity $O((2^n)^3)$). On the other hand, non-convex optimization approaches could preform much faster than their convex counterparts [25]. Although non-convex optimization typically lacks convergence guarantees, it was recently shown that one can formulate the compressed sensing QST as a non-convex problem, and solve it with rigorous convergence guarantees (under certain but generic conditions), allowing the state estimation of larger system sizes [26].

Following the non-convex path, we introduce a new algorithm to the toolbox of QST—the Momentum-Inspired Factored Gradient Descent (`MiFGD`). Our approach combines the ideas from compressed sensing, non-convex optimization, and acceleration/momentum techniques to scale QST beyond the current capabilities. `MiFGD` includes acceleration motions per iteration, meaning that it uses two previous iterates to update the next estimate; see Section 2 for details. The intuition is that if the k-th and $(k-1)$-th estimates were pointing to the correct direction, then both information should be useful to determine the $(k+1)$-th estimate. Of course such approach requires an additional estimate to be stored—yet, we show both theoretically and experimentally that momentum results in faster estimation. We emphasize that the analysis becomes non-trivially challenging due to the inclusion of two previous iterates.

The contributions of the paper are summarized as follows:

(i) We prove that the non-convex `MiFGD` algorithm asymptotically enjoys an *accelerated linear* convergence rate in terms of the iterate distance, in the noiseless measurement data case and under common assumptions.

(ii) We provide QST results using the real measurement data from IBM's quantum computers up to 8-qubits, contributing to recent efforts on testing QST algorithms in real quantum data [22]. Our synthetic examples scale up to 12-qubits effortlessly, leaving the space for an efficient and hardware-aware implementation open for future work.

(iii) We show through extensive empirical evaluations that `MiFGD` allows faster estimation of quantum states compared to the state-of-the-art convex and non-convex algorithms,

including recent deep learning approaches [9–11,27], even in the presence of statistical noise in the measurement data.

(iv) We further increase the efficiency of `MiFGD` by extending its implementation to utilize parallel execution over the shared and distributed memory systems. We experimentally showcase the scalability of our approach, which is particularly critical for the estimation of larger quantum system.

(v) We provide the implementation of our approach at https://github.com/gidiko/MiFGD (accessed on 20 January 2023), which is compatible with the open-source software Qiskit [28].

The rest of this manuscript is organized as follows. In Section 2, we set up the problem in detail, and present our proposed method: `MiFGD`. Then, we detail the experimental set up in Section 3, followed by the results in Section 4. Finally, we discuss related and future works with concluding remarks in Section 5.

2. Methods

2.1. Problem Setup

We consider the estimation of a low-rank density matrix $\rho^\star \in \mathbb{C}^{d\times d}$ on an n-qubit Hilbert space with dimension $d = 2^n$, through the following ℓ_2-norm reconstruction objective:

$$\begin{aligned} \min_{\rho\in\mathbb{C}^{d\times d}} \quad & f(\rho) := \tfrac{1}{2}\|\mathcal{A}(\rho) - y\|_2^2 \\ \text{subject to} \quad & \rho \succeq 0,\ \texttt{rank}(\rho) \leq r. \end{aligned} \tag{1}$$

Here, $y \in \mathbb{R}^m$ is the measurement data (observations) (Specific description on how y is generated and what it represents will follow), and $\mathcal{A}(\cdot) : \mathbb{C}^{d\times d} \to \mathbb{R}^m$ is the linear sensing map, where $m \ll d^2$. The sensing map relates the density matrix $\rho^\star$ to (expected, noiseless) observations through the Born rule: $(\mathcal{A}(\rho))_i = \texttt{Tr}(A_i\rho)$, where $\{A_i\}_{i=1}^m \in \mathbb{C}^{d\times d}$ are matrices closely related to the measured observable or the POVM elements of appropriate dimensions.

The objective function in Equation (1) has two constraints: the positive semi-definite constraint: $\rho \succeq 0$, and the low-rank constraint: $\texttt{rank}(\rho) \leq r$. The former is a convex constraint, whereas the latter is a non-convex one, rendering Equation (1) to be a non-convex optimization problem (Convex optimization problem requires both the objective function as well as the constraints to be convex). Following compressed sensing QST results [8], the unit trace constraint $\texttt{Tr}(\rho) = 1$ (which should be satisfied by any density matrix by definition) can be disregarded, without affecting the precision of the final estimate.

A pivotal assumption to apply compressed sensing results is that the linear sensing map $\mathcal{A}$ should satisfy the *restricted isometry property*, which we recall below.

Definition 1 (Restricted Isometry Property (RIP) [29]). *A linear operator $\mathcal{A} : \mathbb{C}^{d\times d} \to \mathbb{R}^m$ satisfies the RIP on rank-r matrices with the RIP constant $\delta_r \in (0, 1)$, if the following holds with high probability for any rank-r matrix $X \in \mathbb{C}^{d\times d}$:*

$$(1 - \delta_r) \cdot \|X\|_F^2 \leq \|\mathcal{A}(X)\|_2^2 \leq (1 + \delta_r) \cdot \|X\|_F^2. \tag{2}$$

Such maps (almost) preserve the Frobenius norm of low-rank matrices, and, as an extension, of low-rank Hermitian matrices. The intuition behind RIP is that the operator $\mathcal{A}(\cdot)$ behaves almost as a bijection between the subspaces $\mathbb{C}^{d\times d}$ and $\mathbb{R}^m$ for low-rank matrices.

Following recent works [26], instead of solving Equation (1), we propose to solve a factorized version of it:

$$\min_{U\in\mathbb{C}^{d\times r}} \ \tfrac{1}{2}\|\mathcal{A}(UU^\dagger) - y\|_2^2, \tag{3}$$

where $U^\dagger \in \mathbb{C}^{r\times d}$ denotes the adjoint of U, and ρ is re-parametrized with $\rho = UU^\dagger$. The motivation for this reformulation is two-folds. First, by representing the $d \times d$ dimensional density matrix ρ with (the outer product of) its $d \times r$ dimensional low-rank factors U,

the search space for the density matrix (that is consistent with the measurement data) significantly reduces, given that $r \ll d$. Second, via the reformulation $\rho = UU^\dagger$, both the PSD constraint and the low-rank constraint are automatically satisfied, transforming the constrained optimization problem in Equation (1) to the *unconstrained* optimization problem in Equation (3). An important implication is that, to solve Equation (3), one can bypass the projection step onto the PSD and low-rank subspace, which requires a sigular value decomposition (SVD) of the estimate of the density matrix ρ on every iteration. This is prohibitively expensive when the dimension $d = 2^n$ is large, which is the case for even moderate number of qubits n. As such, working in the factored space was shown to improve time and space complexities [26,30–34].

A common approach to solve a factored objective as in Equation (3) is to use gradient descent [35] on the parameter U, with iterates as follows (We assume the case where $\nabla f(\cdot) = \nabla f(\cdot)^\dagger$. If this does not hold, the theory still holds by carrying around $\nabla f(\cdot) + \nabla f(\cdot)^\dagger$ instead of just $\nabla f(\cdot)$, after proper scaling):

$$U_{k+1} = U_k - \eta \nabla f(U_k U_k^\dagger) \cdot U_k \tag{4}$$

$$= U_k - \eta \mathcal{A}^\dagger\Big(\mathcal{A}(U_k U_k^\dagger) - y\Big) \cdot U_k. \tag{5}$$

Here, $U_k \in \mathbb{C}^{d\times r}$ is the k-th iterate, and the operator $\mathcal{A}^\dagger : \mathbb{R}^m \to \mathbb{C}^{d\times d}$ is the adjoint of $\mathcal{A}$, defined as $\mathcal{A}^\dagger(x) = \sum_{i=1}^m x_i A_i$, for $x \in \mathbb{R}^m$. The hyperparameter $\eta > 0$ is the step size. This algorithm has been studied in [25,32,34,36–38]. We will refer to the above iteration as the *factored gradient descent* (FGD) algorithm, as in [30]. In what follows, we will introduce our proposed method, the `MiFGD` algorithm: momentum-inspired factored gradient descent.

2.2. The `MiFGD` Algorithm

The `MiFGD` algorithm is a two-step variant of FGD, which iterates as follows:

$$U_{k+1} = Z_k - \eta \mathcal{A}^\dagger\Big(\mathcal{A}(Z_k Z_k^\dagger) - y\Big) \cdot Z_k, \tag{6}$$

$$Z_{k+1} = U_{k+1} + \mu(U_{k+1} - U_k). \tag{7}$$

Here, $Z_k \in \mathbb{C}^{d\times r}$ is a rectangular matrix (with the same dimension as U_k) that accumulates the "momentum" of the iterates U_k [39]. μ is the momentum parameter that weighs the amount of mixture of the previous estimate U_k and the current U_{k+1} to generate Z_{k+1}. The above iteration is an adaptation of Nesterov's accelerated first-order method for *convex* problems [40]. We borrow this momentum formulation, and study how the choice of the momentum parameter μ affects the overall performance in *non-convex* problem formulations, such as Equation (3). *We note that the theory and algorithmic configurations in [40] do not generalize to non-convex problems*, which is one of the contributions of this work. Albeit being a non-convex problem, we show that `MiFGD` asymptotically converges at an accelerated linear rate around a neighborhood of the optimal value, akin to convex optimization results [40].

An important observation is that the factorization $\rho = UU^\dagger$ is not unique. For instance, suppose that $U^\star$ is an optimal solution for Equation (3); then, for any rotation matrix $R \in \mathbb{C}^{r\times r}$ satisfying $RR^\dagger = I$, the matrix $\widehat{U} = U^\star R$ is also optimal for Equation (3) (To see this, observe that $\rho^\star = U^\star U^{\star\dagger} = U^\star I U^{\star\dagger} = U^\star RR^\dagger U^{\star\dagger} = \widehat{U}\widehat{U}^\dagger$). To resolve this ambiguity, we use the distance between a pair of matrices as the minimum distance $\min_{R\in\mathcal{O}} \|U - U^\star R\|_F$ up to rotations, where $\mathcal{O} = \{R \in \mathbb{C}^{r\times r} \mid RR^\dagger = I\}$. In words, we want to track how close an estimate U is to $U^\star$, up to the minimizing rotation matrix.

Algorithm 1 contains the details of the `MiFGD`. As Problem (3) is non-convex, the initialization plays an important role in achieving global convergence. The initial point U_0 is either randomly initialized [36,41,42], or set according to Lemma 4 in [26]:

$$\rho_0 = U_0 U_0^\dagger = \Pi_{\mathcal{C}}\left(\tfrac{-1}{1+\delta_{2r}} \cdot \nabla f(0)\right) = \tfrac{1}{1+\delta_{2r}} \Pi_{\mathcal{C}}\left(\sum_{i=1}^{m} y_i A_i\right), \tag{8}$$

where $\Pi_{\mathcal{C}}(\cdot)$ is the projection onto the set of PSD matrices, $\delta_{2r} \in (0,1)$ is the RIP constant from Definition 1, and $\nabla f(0)$ is the gradient of $f(\cdot)$ evaluated at all-zero matrix. Since computing the RIP constant is NP-hard, in practice we compute U_0 through $\rho_0 = \frac{-1}{\widehat{L}} \Pi_{\mathcal{C}}\left(\sum_{i=1}^{m} y_i A_i\right)$, where $\widehat{L} \in (1, 11/10]$; see Theorem 1 below for details.

Algorithm 1 Momentum-Inspired Factored Gradient Descent (`MiFGD`)

Input: $\mathcal{A}$ (sensing map), y (measurement data), r (rank), and μ (momentum parameter).
- Set U_0 randomly or as in Equation (8).
- Set $Z_0 = U_0$.
- Set η as in Equation (9).

for $k = 0, 1, 2, \ldots$ **do**
$U_{k+1} = Z_k - \eta \mathcal{A}^\dagger\left(\mathcal{A}(Z_k Z_k^\dagger) - y\right) \cdot Z_k$
$Z_{k+1} = U_{k+1} + \mu(U_{k+1} - U_k)$
end for
Output: $\rho = U_{k+1} U_{k+1}^\dagger$

Compared to randomly selecting U_0, the initialization scheme in Equation (8) involves a gradient and a top-r eigenvalue computations. Yet, Equation (8) provides a more informed initial point, as it is based on the data $\{y_i, A_i\}_{i=1}^m$, which could lead to convergence in fewer iterations in practice, and satisfies the initialization condition of Theorem 1 for small enough κ (Based on our experiments, in practice, both initializations are applicable and useful).

For the step size η in Algorithm 1, it is set to the following based on our theoretical analysis (c.f., Lemma A6):

$$\eta = \frac{1}{4\left((1+\delta_{2r})\|Z_0 Z_0^\dagger\|_2 + \|\mathcal{A}^\dagger\left(\mathcal{A}(Z_0 Z_0^\dagger) - y\right)\|_2\right)}, \tag{9}$$

where $Z_0 = U_0$. Similarly to the above, in practice we replace the RIP constant δ_{2r} with $\widehat{L}$. The step size η remains constant at every iteration, and requires only two top-eigenvalue computations to obtain the spectral norms $\|Z_0 Z_0^\dagger\|_2$ and $\|\mathcal{A}^\dagger\left(\mathcal{A}(Z_0 Z_0^\dagger - y\right)\|_2$. These computations can be efficiently implemented by any off-the-shelf eigenvalue solver, such as the Power Method or the Lanczos method [43].

2.3. Theoretical Guarantees of the `MiFGD` Algorithm

We now present the formal convergence theorem, where under certain conditions, `MiFGD` asymptotically achieves an accelerated linear rate.

Theorem 1 (Accelerated asymptotic convergence rate)**.** *Assume that $\mathcal{A}(\cdot)$ satisfies the RIP in Definition 1 with the constant $\delta_{2r} \leq 1/10$. Initialize $U_0 = U_{-1}$ such that*

$$\min_{R \in \mathcal{O}} \|U_0 - U^\star R\|_F = \min_{R \in \mathcal{O}} \|U_{-1} - U^\star R\|_F \leq \frac{\sqrt{\sigma_r(\rho^\star)}}{10^3 \sqrt{\kappa \tau(\rho^\star)}},$$

where $\kappa := \frac{1+\delta_{2r}}{1-\delta_{2r}}$ is the (inverse) condition number of $\mathcal{A}(\cdot)$, $\tau(\rho) := \frac{\sigma_1(\rho)}{\sigma_r(\rho)}$ is the condition number of ρ with $\mathtt{rank}(\rho) = r$, and $\sigma_i(\rho)$ is the i-th singular value of ρ. Set the step size η such that

$$\left[1 - \left(\frac{\sqrt{1+\delta_{2r}} - \sqrt{1-\delta_{2r}}}{(\sqrt{2}+1)\sqrt{1+\delta_{2r}}}\right)^4\right] \cdot \frac{10}{4\sigma_r(\rho^\star)(1-\delta_{2r})} \leq \eta \leq \frac{10}{4\sigma_r(\rho^\star)(1-\delta_{2r})},$$

and the momentum parameter $\mu = \frac{\varepsilon_\mu}{2\cdot 10^3 r\tau(\rho^\star)\sqrt{\kappa}}$*, for user-defined* $\varepsilon_\mu \in (0,1]$*. Then, for the (noiseless) measurement data* $y = \mathcal{A}(\rho^\star)$ *with* $\mathtt{rank}(\rho^\star) = r$*, the output of the* `MiFGD` *in Algorithm 1 satisfies the following: for any* $\epsilon > 0$*, there exist constants* C_ϵ *and* $\tilde{C}_\epsilon$ *such that, for all k,*

$$\begin{aligned}\left(\min_{R\in\mathcal{O}} \|U_{k+1} - U^\star R\|_F^2 + \min_{R\in\mathcal{O}} \|U_k - U^\star R\|_F^2\right)^{1/2} \\ \leq C_\epsilon\left(1 - \sqrt{\tfrac{1-\delta_{2r}}{1+\delta_{2r}}} + \epsilon\right)^{k+1} \min_{R\in\mathcal{O}} \|U_0 - U^\star R\|_F + \xi\cdot\mu\cdot\sigma_1(\rho^\star)^{1/2}\cdot r\cdot\tilde{C}_\epsilon \\ \approx C_\epsilon\left(1 - \sqrt{\tfrac{1-\delta_{2r}}{1+\delta_{2r}}} + \epsilon\right)^{k+1} \min_{R\in\mathcal{O}} \|U_0 - U^\star R\|_F + O(\mu).\end{aligned} \tag{10}$$

where $\xi = \sqrt{1 - \frac{4\eta\sigma_r(\rho^\star)(1-\delta_{2r})}{10}}$*. That is, the* `MiFGD` *asymptotically enjoys an accelerated linear convergence rate in iterate distances up to a constant proportional to the momentum parameter* μ*.*

Theorem 1 can be interpreted as follows. The right hand side of Equation (10) depends on the initial distance $\min_{R\in\mathcal{O}} \|U_0 - U^\star R\|_F$ akin to convex optimization results, where asymptotically $O\left(1 - \sqrt{\frac{1-\delta_{2r}}{1+\delta_{2r}}}\right)$ appear as the contraction factor. In contrast, the contraction factor of vanilla FGD [26] is of the order $O\left(1 - \frac{1-\delta_{2r}}{1+\delta_{2r}}\right)$.

The main assumption is that the sensing map $\mathcal{A}(\cdot)$ satisfies RIP. This assumption implies that the condition number of f depends on the RIP constants δ_{2r} such that $\frac{L}{\mu} \propto \frac{1+\delta_{2r}}{1-\delta_{2r}}$, since the eigenvalues of the Hessian of f, i.e., $\mathcal{A}^\dagger A(\cdot)$, lie between $1-\delta_{2r}$ and $1+\delta_{2r}$ (when restricted to low-rank matrices) (In this sense, the RIP assumption plays the similar role to assuming f is μ-strongly convex and L-smooth (when restricted to low-rank matrices)). Such assumption has become standard in the optimization and the signal processing community [19,25,29]. Hence, `MiFGD` has better dependency on the (inverse) condition number of f compared to FGD. Such improvement of the dependency on the condition number is referred to as "acceleration" in the convex optimization literature [44,45]. Thus, assuming that the initial points U_0 and U_{-1} are close enough to the optimum as stated in the theorem, `MiFGD` decreases its distance to $U^\star$ at an accelerated linear rate, up to an "error" level that depends on the momentum parameter μ, which is bounded by $\frac{1}{2\cdot 10^3 r\tau(\rho^\star)\sqrt{\kappa}}$.

Theorem 1 requires a strong assumption on the momentum parameter μ, which depends on quantities that might not be known *a priori* for general problems. However, we note that for the special case of QST, we know these quantities exactly: r is the rank of density matrix—thus, for pure states $r = 1$; $\tau(\rho^\star)$ is the (rank-restricted) condition number of the density matrix ρ—for pure states, $\tau(\rho^\star) = \frac{\sigma_1(\rho)}{\sigma_r(\rho)} = \frac{\sigma_1(\rho)}{\sigma_1(\rho)} = 1$; and finally, κ is the condition number of the sensing map, and satisfies: $\kappa \leq 11/9$ given the constraint $\delta_{2r} \leq 1/10$. This analysis leads to a momentum value $\mu \approx \epsilon_\mu/2211$ (This is the numerical value of $\mu^\star$ we use in experiments in Section 4). However, as we show in both real and synthetic experiments in Section 4 (and further in Appendix A), the theory is conservative; much larger values of μ lead to fast, stable, and improved performance. Finally, the bound on the condition number in Theorem 1 is not strict, and comes out of the analysis we follow; we point the reader to similar assumptions made where $\tau(\rho^\star)$ is assumed to be constant: $O(1)$ [46].

The detailed proof of Theorem 1 is provided in Appendix B. The proof differs from state of the art proofs for non-accelerated factored gradient descent: due to the inclusion of the memory term, three different terms—U_{k+1}, U_k, U_{k-1}—need to be handled simultaneously. Further, the proof differs from other recent proofs on non-convex, but non-factored, gradient descent methods, as in [47]: the distance metric over rotations $\min_{R\in\mathcal{O}} \|Z_k - U^\star R\|_F$, where Z_k includes estimates from two steps in history, is not amenable to simple triangle inequality bounds. As a result, a careful analysis is required, including the design of two-dimensional dynamical systems, where we characterize and bound the eigenvalues of a 2×2 contraction matrix.

3. Experimental Setup

3.1. $\rho^\star$ Density Matrices and Quantum Circuits

In our numerical and real experiments, we have considered (different subsets of) the following n-qubit pure quantum states (The content in this subsection is implemented in the `states.py` component of our complementary software package: https://github.com/gidiko/MiFGD (accessed on 20 January 2023)):

1. The (generalized) GHZ state:

$$|\texttt{GHZ}(n)\rangle = \frac{|0\rangle^{\otimes n} + |1\rangle^{\otimes n}}{\sqrt{2}}, \ n > 2.$$

2. The (generalized) GHZ-minus state:

$$|\texttt{GHZ}_-(n)\rangle = \frac{|0\rangle^{\otimes n} - |1\rangle^{\otimes n}}{\sqrt{2}}, \ n > 2.$$

3. The Hadamard state:

$$|\texttt{Hadamard}(n)\rangle = \left(\frac{|0\rangle + |1\rangle}{\sqrt{2}}\right)^{\otimes n}.$$

4. A random state $|\texttt{Random}(n)\rangle$.

We have implemented these states (on the IBM quantum simulator and/or the IBM's QPU) using the following circuits. The GHZ state $|\texttt{GHZ}(n)\rangle$ is generated by applying the Hadamard gate to one of the qubits, and then applying $n-1$ CNOT gates between this qubit (as a control) and the remaining $n-1$ qubits (as targets). The GHZ-minus state $|\texttt{GHZ}_-(n)\rangle$ is generated by applying the X gate to one of the qubits (e.g., the first qubit) and the Hadamard gate to the remaining $n-1$ qubits, followed by applying $n-1$ CNOT gates between the first qubit (as a target) and the other $n-1$ qubits (as controls). Finally, we apply the Hadamard gate to all of the qubits. The Hadamard state $|\texttt{Hadamard}(n)\rangle$ is a separable state, and is generated by applying the Hadamard gate to all of the qubits. The random state $|\texttt{Random}(n)\rangle$ is generated by a random quantum gate selection: In particular, for a given circuit depth, we uniformly select among generic single-qubit rotation gates with 3 Euler angles, and controlled-X gates, for every step in the circuit sequence. For the rotation gates, the qubits involved are selected uniformly at random, as well as the angles from the range $[0, 1]$. For the controlled-X gates, the source and target qubits are also selected uniformly at random.

We generically denote the density matrix that correspond to pure state $|\psi\rangle$ as $\rho^\star = |\psi\rangle\langle\psi|$. For clarity, we will drop the bra-ket notation when we refer to $|\texttt{GHZ}(n)\rangle$, $|\texttt{GHZ}_-(n)\rangle$, $|\texttt{Hadamard}(n)\rangle$ and $|\texttt{Random}(n)\rangle$. While the density matrices of the $\texttt{GHZ}(n)$ and $\texttt{GHZ}_-(n)$ are sparse in the $\{|0\rangle, |1\rangle\}^n$ basis, the density matrix of $\texttt{Hadamard}(n)$ state is fully-dense in this basis, and the sparsity of the density matrix that of $\texttt{Random}(n)$ may be different form one state to another.

3.2. Measuring Quantum States

The quantum measurement model. In our experiments (both synthetic and real), we measure the qubits in the Pauli basis (This is the non-commutative analogue of the Fourier basis, for the case of sparse vectors [19,48]). A Pauli basis measurement on an n-qubit system has $d = 2^n$ possible outcomes. The Pauli basis measurement is uniquely defined by the *measurement setting*. A Pauli measurement is a string of n letters $\alpha := (\alpha_1, \alpha_2, \ldots, \alpha_n)$ such that $\alpha_k \in \{x, y, z\}$ for all $k \in [n]$. Note that there are at most 3^n distinct Pauli strings. To define the Pauli basis measurement associated with a given measurement string α, we first define the the following three bases on $\mathbb{C}^{2\times 2}$:

$$\mathcal{B}_x = \left\{ |x,0\rangle := \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle),\ |x,1\rangle := \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle) \right\},$$
$$\mathcal{B}_y = \left\{ |y,0\rangle := \frac{1}{\sqrt{2}}(|0\rangle + i|1\rangle),\ |y,1\rangle := \frac{1}{\sqrt{2}}(|0\rangle - i|1\rangle) \right\},$$
$$\mathcal{B}_z = \{ |z,0\rangle := |0\rangle,\ |z,1\rangle := |1\rangle \}.$$

These are the eigenbases of the single-qubit Pauli operators, σ_x, σ_y, and σ_z, whose 2×2 matrix representations are given by:

$$\sigma_x = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \sigma_y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \quad \text{and} \quad \sigma_z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}.$$

Given a Pauli setting α, the Pauli basis measurement Π_α is defined by the 2^n projectors:

$$\Pi_\alpha = \left\{ \left|v_\ell^{(\alpha)}\right\rangle\left\langle v_\ell^{(\alpha)}\right| = \bigotimes_{k=1}^{n} |\alpha_k, \ell_k\rangle\langle\alpha_k, \ell_k| : \ell_k \in \{0,1\}\ \forall k \in [1,n] \right\},$$

where ℓ denotes the bit string $(\ell_{k_1}, \ell_{k_2}, \ldots, \ell_{k_n})$. Since there are 3^n distinct Pauli measurement settings, there are the same number of possible Pauli basis measurements.

Technically, this set forms a positive operator-valued measure (POVM). The projectors that form Π_α are the measurement outcomes (or POVM elements) and the probability to obtain an outcome $|v_\ell^{(\alpha)}\rangle\langle v_\ell^{(\alpha)}|$ –when the state of the system is $\rho^\star$– is given by the Born rule: $\langle v_\ell^{(\alpha)}|\rho^\star|v_\ell^{(\alpha)}\rangle = \mathtt{Tr}\left(|v_\ell^{(\alpha)}\rangle\langle v_\ell^{(\alpha)}| \cdot \rho^\star\right)$.

The RIP and expectation values of Pauli observables. Starting with the requirements of our algorithm, the sensing map $\mathcal{A} : \mathbb{C}^{d\times d} \to \mathbb{R}^m$ we consider is comprised of a collection of matrices $\{A_i \in \mathbb{C}^{d\times d}\}_{i=1}^m$, such that $y_i = \mathtt{Tr}(A_i\rho^\star)$. We denote the vector $(y_1, \ldots, y_m)$ by y.

When no prior information about the quantum state is assumed, to ensure its (robust) recovery, one must choose a set m sensing matrices A_i, so that d^2 of them are linearly independent. One example of such choice is the POVM elements of the 3^n Pauli basis measurements.

Yet, when it is known that the state-to-be-reconstructed is of low-rank, theory on low-rank recovery problems suggests that A_i could just be "incoherent" enough with respect to $\rho^\star$ [20], so that recovery is possible from a limited set of measurements, i.e., with $m \ll d^2$. In particular, it is known [4,20,21] that if the sensing matrices correspond to random *Pauli monomials*, then $m = O(r \cdot d \cdot \mathrm{poly}(\log d))$ A_i's are sufficient for a successful recovery of $\rho^\star$, using convex solvers for Equation (1) (The main difference between [4,20] and [21] is that the former guarantees recovery for almost all choices of $m = O(r \cdot d \cdot \mathrm{poly}(\log d))$ random Pauli monomials, while the latter proves that there exists a *universal* set of $m = O(r \cdot d \cdot \mathrm{poly}(\log d))$ Pauli monomials A_i that guarantees successful recovery).

A Pauli monomial P_i is an operator in the set $P_i \in \{\mathbb{1}, \sigma_x, \sigma_y, \sigma_z\}^{\otimes n}$, that is, an n-fold tensor product of single-qubit Pauli operators (including the identity operator). For convenience we re-label the single-qubit Pauli operators as $\sigma_0 := \mathbb{1}, \sigma_1 := \sigma_x, \sigma_2 := \sigma_y$, and $\sigma_3 := \sigma_z$, so that we can also write $P_i = \bigotimes_{k=1}^n \sigma_{i_k}$ with $i_k \in \{0, \ldots, 3\}$ for all $k \in [n]$. These results [4,20,21] are feasible since the Pauli-monomial-based sensing map $\mathcal{A}(\cdot)$ obeys the RIP property, as in Definition 1 (In particular, the RIP is satisfied for the sensing mechanisms that obeys $(\mathcal{A}(\rho^\star))_i = \frac{d}{\sqrt{m}}\mathtt{Tr}(A_i^*\rho^\star)$, $i = 1, \ldots, m$. Further, the case considered in [21] holds for a slightly larger set than the set of rank-r density matrices: for all $\rho \in \mathbb{C}^{d\times d}$ such that $\|\rho\|_* \leq \sqrt{r}\|\rho\|_F$). For the remainder of this manuscript, we will use the term "*Pauli expectation value*" to denote $\mathtt{Tr}(A_i\rho^\star) = \mathtt{Tr}(P_i\rho^\star)$.

From Pauli basis measurements to Pauli expectation values. While the theory for compressed sensing was proven for Pauli expectation values, in real QPUs, experimental data

is obtained from Pauli basis measurements. Therefore, to make sure we are respecting the compressed sensing requirements on the sensing map, we follow this protocol:

(i) We sample $m = O(r \cdot d \cdot \text{poly}(\log d))$ or $m = \texttt{measpc} \cdot d^2$ Pauli monomials uniformly over $\{\sigma_i\}^{\otimes n}$ with $i \in \{0, \ldots, 3\}$, where $\texttt{measpc} \in [0, 1]$ represents the percentage of measurements out of full tomography.

(ii) For every monomial, P_i, in the generated set, we identify an experimental setting $\alpha(i)$ that corresponds to the monomial. There, qubits, for which their Pauli operator in P_i is the identity operator, are measured, without loss of generality, in the σ_3 basis. For example, for $n = 3$ and $P_i = \sigma_0 \otimes \sigma_1 \otimes \sigma_1$, we identify the measurement setting $\alpha(i) = (z, x, x)$.

(iii) We measure the quantum state in the Pauli basis that corresponds to $\alpha(i)$, and record the outcomes.

To connect the measurement outcomes to the expectation value of the Pauli monomial, we use the relation:

$$\texttt{Tr}(P_i \rho^\star) = \sum_{\ell \in \{0,1\}^n} (-1)^{\chi_{f(\ell)}} \cdot \texttt{Tr}\Big(|v_\ell^{(\alpha(i))}\rangle\langle v_\ell^{(\alpha(i))}| \cdot \rho^\star \Big), \tag{11}$$

where $f(\ell) : \{0,1\}^n \to \{0,1\}^n$ is a mapping that takes a bit string ℓ and returns a new bit string $\tilde{\ell}$ (of the same size) such that $\tilde{\ell}_k = 0$ for all k's for which $i_k = 0$ (that is, the locations of the identity operators in P_i), and $\chi_{\tilde{\ell}}$ is the parity of the bit string $\tilde{\ell}$.

3.3. Algorithmic Setup

In our implementation, we explore a number of control parameters, including the maximum number of iterations `maxiters`, the step size η, the relative error from successive state iterates `reltol`, the momentum parameter μ, the percentage of the complete set of measurements (i.e., over all possible Pauli monomials) `measpc`, and the `seed`. In the sequel experiments we set $\texttt{maxiters} = 1000$, $\eta = 10^{-3}$, and $\texttt{reltol} = 5 \times 10^{-4}$, unless stated differently. Regarding acceleration, $\mu = 0$ when acceleration is muted; we experiment over the range of values $\mu \in \{\frac{1}{8}, \frac{1}{4}, \frac{1}{3}, \frac{3}{4}\}$ when investigating the acceleration effect, beyond the theoretically suggested $\mu^\star$. In order to explore the dependence of our approach on the number of measurements available, `measpc` varies over the set of $\{5\%, 10\%, 15\%, 20\%, 40\%, 60\%\}$; `seed` is used for differentiating repeating runs with all other parameters kept fixed (`maxiters` is `num_iterations` in the code; also `reltol` is `relative_error_tolerance`, `measpc` is `complete_measurements_percentage`).

Denoting $\widehat{\rho}$ the estimate of $\rho^\star$ by `MiFGD`, we report on outputs including:

- The evolution with respect to the distance between $\widehat{\rho}$ and $\rho^\star$: $\|\widehat{\rho} - \rho^\star\|_F$, for various μ's.
- The number of iterations to reach `reltol` to $\rho^\star$ for various μ's.
- The fidelity of $\widehat{\rho}$, defined as $\text{Tr}(\rho^\star \widehat{\rho})$ (for rank-1 $\rho^\star$), as a function of the acceleration parameter μ in the default set.

In our plots, we sweep over our default set of `measpc` values, repeat 5 times for each individual setup, varying supplied seed, and depict their 25-, 50- and 75-percentiles.

3.4. Experimental Setup on Quantum Processing Unit (QPU)

We show empirical results on 6- and 8-qubit real data, obtained on the 20-qubit IBM QPU `ibmq_boeblingen`. The layout/connectivity of the device is shown in Figure 1. The 6-qubit data was from qubits $[0, 1, 2, 3, 8, 9]$, and the 8-qubit data was from $[0, 1, 2, 3, 8, 9, 6, 4]$. The T_1 coherence times are $[39.1, 75.7, 66.7, 100.0, 120.3, 39.2, 70.7, 132.3]$ μs, and T_2 coherence times are $[86.8, 94.8, 106.8, 63.6, 156.5, 66.7, 104.5, 134.8]$ μs. The circuit for generating 6-qubit and 8-qubit GHZ states are shown in Figure 1. The typical two qubit gate errors measured from randomized benchmarking (RB) for relevant qubits are summarized in Table 1.

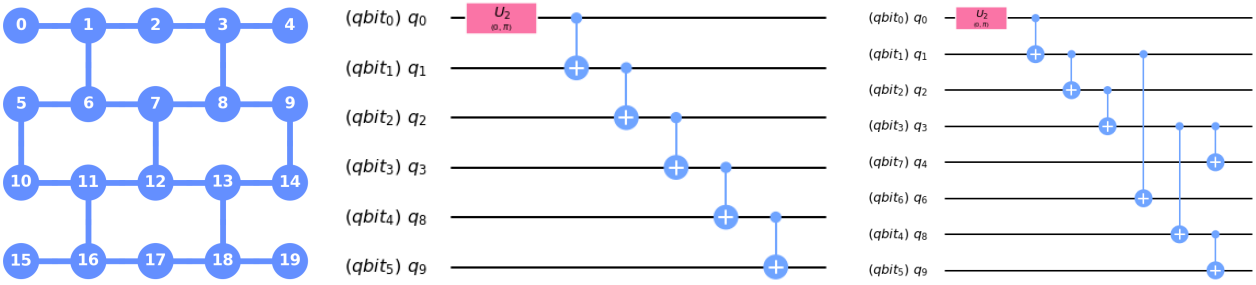

Figure 1. Left panel: Layout connectivity of IBM backend `ibmq_boeblingen`; **Middle and right panels:** Circuits used to generate 6-qubit state (**left**) and 8-qubit GHZ state (**right**). *qbit* refers to the quantum registers used in qiskit, and *q* corresponds to qubits on the real device.

Table 1. Two qubit error rates for the relevant gates used in generating 6-qubit and 8-qubit GHZ states on `ibmq_boeblingen`.

C_0X_1	C_1X_2	C_2X_3	C_3X_8	C_8X_9	C_3X_4	C_1X_6
0.0072	0.0062	0.0087	0.0077	0.0152	0.0167	0.0133

The QST circuits were generated using the tomography module in `qiskit-ignis` (https://github.com/Qiskit/qiskit-ignis (accessed on 20 January 2023)). For complete QST of a n-qubits state 3^n circuits are needed. The result of each circuit is averaged over 8192, 4096 or 2048, for different n-qubit scenarios. To mitigate for readout errors, we prepare and measure all of the 2^n computational basis states in the computation basis to construct a calibration matrix C. C has dimension 2^n by 2^n, where each column vector corresponds to the measured outcome of a prepared basis state. In the ideal case of no readout error, C is an identity matrix. We use C to correct for the measured outcomes of the experiment by minimizing the function:

$$\min_{v^{\text{cal}} \in \mathbb{R}^d} \quad \|Cv^{\text{cal}} - v^{\text{meas}}\|^2 \quad \text{subject to} \quad \sum_i v_i^{\text{cal}} = 1,\ v_i^{\text{cal}} \geq 0, \forall i = 1, \ldots, d \tag{12}$$

Here v^{meas} and v^{cal} are the measured and calibrated outputs, respectively. The minimization problem is then formulated as a convex optimization problem and solved by quadratic programming using the package `cvxopt` [49].

4. Results

4.1. `MiFGD` on 6- and 8-Qubit Real Quantum Data

We realize two types of quantum states on IBM QPUs, parameterized by the number of qubits n for each case: the `GHZ_`(n) and the `Hadamard`(n) circuits. We collected measurements over all possible Pauli settings by repeating the experiment for each setting a number of times: these are the number of `shots` for each setting. The (circuit, number of `shots`) measurement configurations from IBM Quantum devices are summarized in Table 2.

Table 2. QPU settings.

Circuit	`GHZ_`(6)	`GHZ_`(6)	`GHZ_`(8)	`GHZ_`(8)	`Hadamard`(6)	`Hadamard`(8)
# `shots`	2048	8192	2048	4096	8192	4096

In Appendix A, we provide target error list plots for the evolution of $\|\widehat{\rho} - \rho^\star\|_F^2$ for reconstructing all the settings in Table 2, both for real data and for simulated scenarios. Further, we provide plots that relate the effect of momentum acceleration on the final fidelity observed for these cases. For clarity, in Figure 2, we summarize the efficiency of momentum acceleration, by showing the reconstruction error only for the following

settings: `maxiters` $= 1000$, $\eta = 10^{-3}$, `reltol` $= 5 \times 10^{-4}$, and `measpc` $= 20\%$. In the plots, $\mu = 0$ corresponds to the FGD algorithm in [30], $\mu^\star$ corresponds to the value obtained through our theory, while we use $\mu \in \left\{\frac{1}{8}, \frac{1}{4}, \frac{1}{3}, \frac{3}{4}\right\}$ to study the acceleration effect. For $\mu^\star$, per our theory, we follow the rule $\mu^\star \approx \varepsilon_\mu / 2211$ for $\varepsilon_\mu \in (0, 1]$; see also Section 2 for details (For this application, $\sigma_r(\rho^\star) = 1$, $\tau(\rho^\star) = 1$, and $r = 1$ by construction; we also approximated $\kappa = 1.223$, which, for user-defined $\varepsilon_\mu = 1$, results in $\mu^\star = 4.5 \cdot 10^{-4}$. Note that smaller ε_μ values result into a smaller radius of the convergence region; however, more pessimistic ε_μ values result into small μ, with no practical effect in accelerating the algorithm). Note that, in most of the cases, the curve corresponding to $\mu = 0$ is hidden behind the curve corresponding to $\mu \approx \mu^\star$. We run each QST experiment for 5 times for random initializations. We record the evolution of the $\|\widehat{\rho} - \rho^\star\|_F^2$ error at each step, and stop when the relative error of successive iterates gets smaller than `reltol` or the number of iterations exceeds `maxiters` (whichever happens first). To implement `measpc` $= 20\%$, we follow the description given in Equation (11) with $m =$ `measpc` $\cdot\, d^2$.

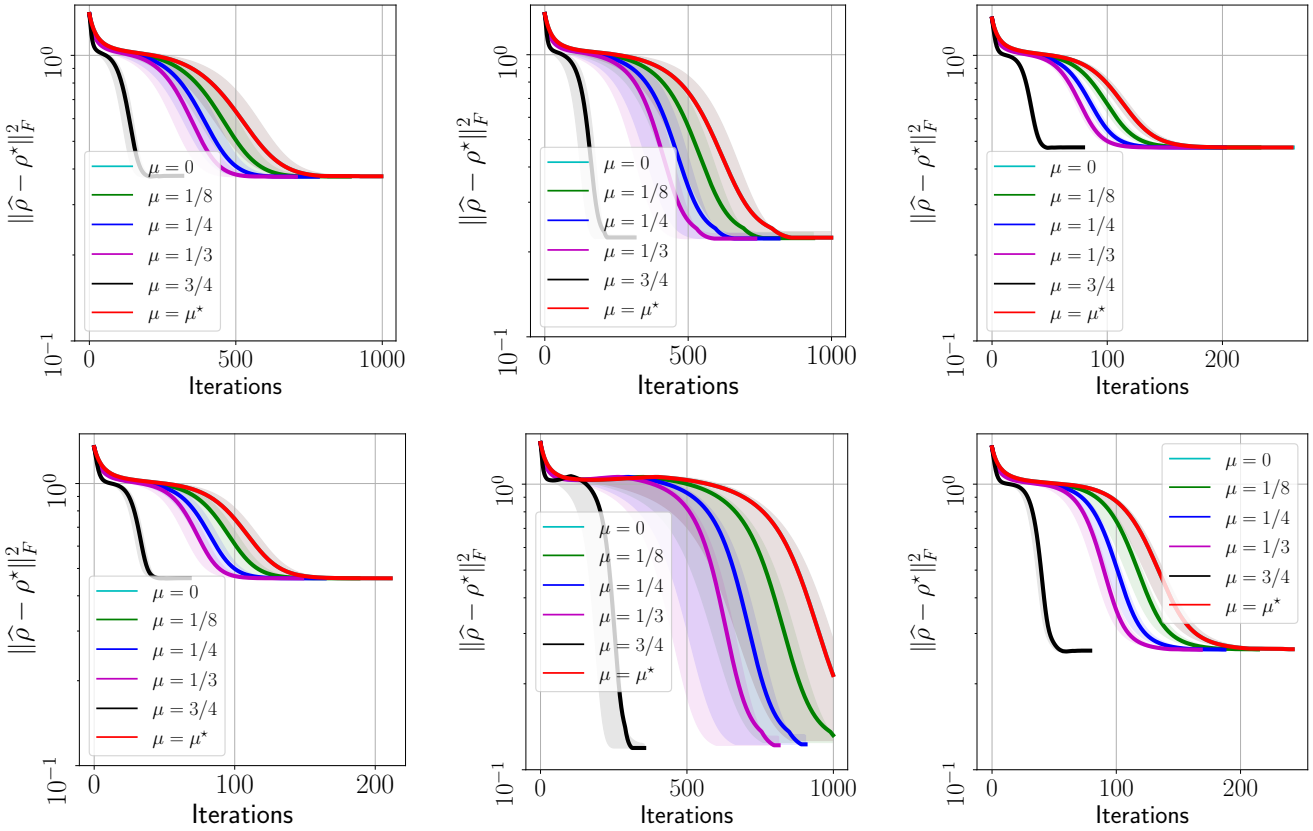

Figure 2. Target error list plots $\|\widehat{\rho} - \rho^\star\|_F^2$ versus method iterations using real IBM QPU data. **Top-left**: `GHZ_`(6) with 2048 shots; **Top-middle**: `GHZ_`(6) with 8192 shots; **Top-right**: `GHZ_`(8) with 2048 shots; **Bottom-left**: `GHZ_`(8) with 4096 shots/copies of $\rho^\star$; **Bottom-middle**: `Hadamard`(6) with 8192 shots; **Bottom-right**: `Hadamard`(8) with 4096 shots. All cases have `measpc` $= 20\%$. Shaded area denotes standard deviation around the mean over repeated runs in all cases.

To highlight the level of noise existing in real quantum data, in Figure 3, we repeat the same setting using the QASM simulator in `qiskit-aer`. This is a parallel, high performance quantum circuit simulator written in C++ that can support a variety of realistic circuit level noise models.

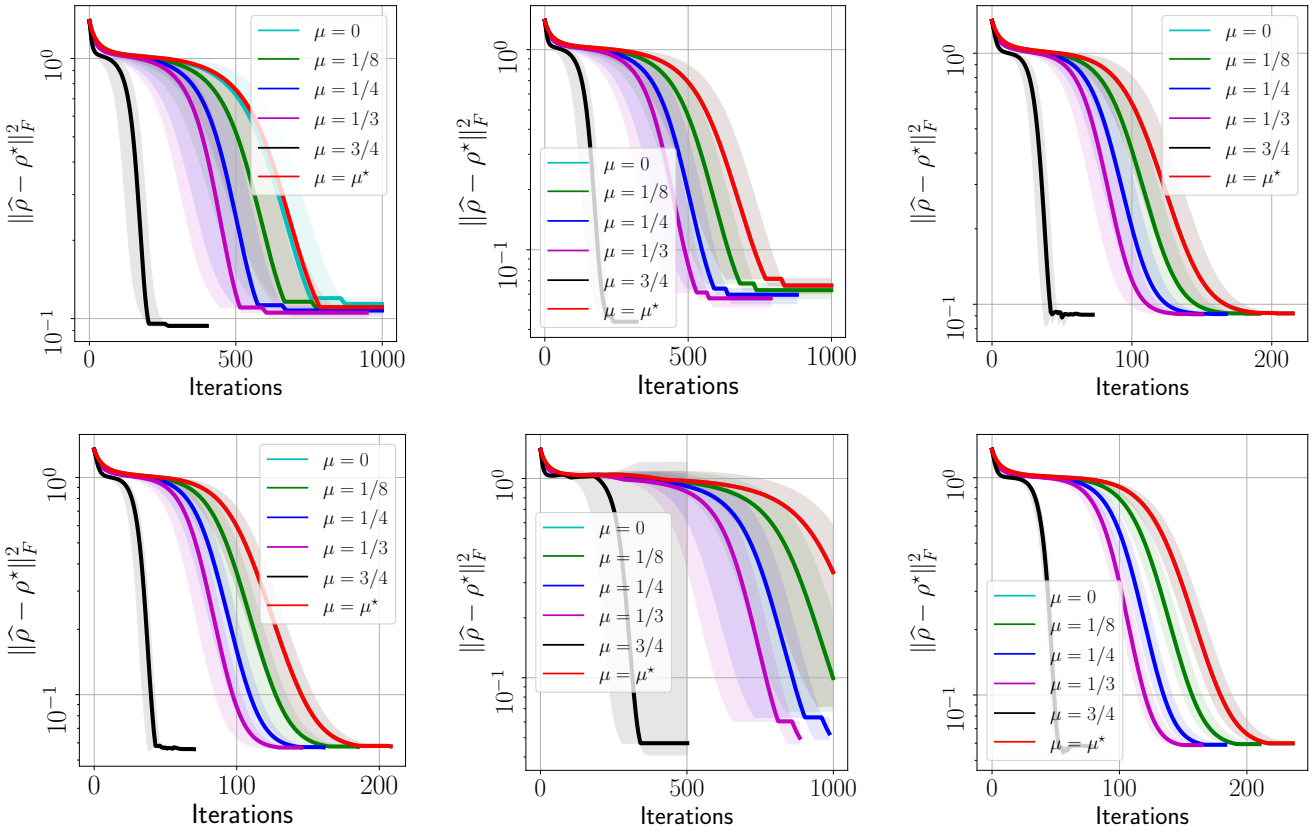

Figure 3. Target error list plots $\|\widehat{\rho} - \rho^\star\|_F^2$ versus method iteration using synthetic IBM's quantum simulator data. **Top-left**: `GHZ_`(6) with 2048 shots; **Top-middle**: `GHZ_`(6) with 8192 shots; **Top-right**: `GHZ_`(8) with 2048 shots; **Bottom-left**: `GHZ_`(8) with 4096 shots; **Bottom-middle**: `Hadamard`(6) with 8192 shots; **Bottom-right**: `Hadamard`(8) with 4096 shots. All cases have `measpc` $= 20\%$. Shaded area denotes standard deviation around the mean over repeated runs in all cases.

Figure 2 summarizes the performance of our proposal on different $\rho^\star$, and for different μ values on real IBM QPU data. All plots show the evolution of $\|\widehat{\rho} - \rho^\star\|_F$ across iterations, featuring a steep dive to convergence for the largest value of μ we tested: we report that we also tested $\mu = 0$, which shows only slight worse performances than $\mu^\star$. Figure 2 highlights the universality of our approach: its performance is oblivious to the quantum state reconstructed, as long as it satisfies purity or it is close to a pure state. Our method does not require any additional structure assumptions in the quantum state.

To highlight the effect of real noise on the performance of `MiFGD`, we further plot its performance on the same settings but using measurements coming from an idealized quantum simulator. Figure 3 considers the exact same settings as in Figure 2. It is obvious that `MiFGD` can achieve better reconstruction performance when data are less erroneous. This also highlights that, in real noisy scenarios, the radius of the convergence region of `MiFGD` around $\rho^\star$ is controlled mostly by the the noise level, rather than by the inclusion of momentum acceleration.

Finally, in Figure 4, we depict the fidelity of $\widehat{\rho}$ achieved using `MiFGD`, defined as $\mathrm{Tr}\left(\rho^\star\widehat{\rho}\right)$, versus various μ values and for different circuits $(\rho^\star)$. Shaded area denotes standard deviation around the mean over repeated runs in all cases. The plots show the significant gap in performance when using real quantum data versus using synthetic simulated data within a controlled environment.

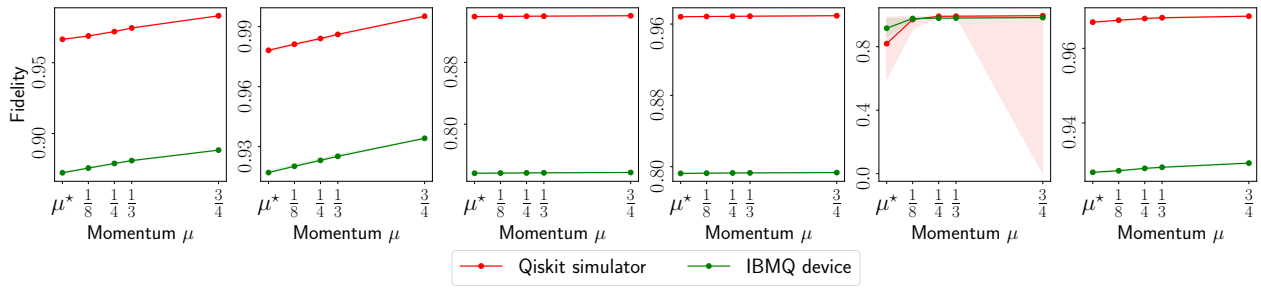

Figure 4. Fidelity list plots where we depict the fidelity of $\widehat{\rho}$ to $\rho^\star$. From **left** to **right**: (i) $\texttt{GHZ\_}(6)$ with 2048 shots; (ii) $\texttt{GHZ\_}(6)$ with 8192 shots; (iii) $\texttt{GHZ\_}(8)$ with 2048 shots; (iv) $\texttt{GHZ\_}(8)$ with 4096 shots; (v) $\texttt{Hadamard}(6)$ with 8192 shots; (vi) $\texttt{Hadamard}(8)$ with 4096 shots. All cases have $\texttt{measpc} = 20\%$. Shaded area denotes standard deviation around the mean over repeated runs in all cases.

4.2. Performance Comparison with Full Tomography Methods in Qiskit

We compare the `MiFGD` with publicly available implementations for QST reconstruction. Two common techniques for QST, included in the `qiskit-ignis` distribution [28], are: (i) the `CVXPY` fitter method, that uses the `CVXPY` convex optimization package [50,51]; and (ii) the `lstsq` method, that uses least-squares fitting [52]. Both methods solve *the full tomography problem* (In [8], it was sown that the minimization program (13) yields a robust estimation of low-rank states in the compressed sensing. Thus, one can use `CVXPY` fitter method to solve Equation (13) with $m \ll d^2$ Pauli expectation value to obtain a robust reconstruction of $\rho^\star$) according to the following expression:

$$\begin{aligned} \min_{\rho \in \mathbb{C}^{d\times d}} \quad & f(\rho) := \tfrac{1}{2}\|\mathcal{A}(\rho) - y\|_2^2 \\ \text{subject to} \quad & \rho \succeq 0, \texttt{Tr}(\rho) = 1. \end{aligned} \tag{13}$$

We note that `MiFGD` is not restricted to "tall" U scenarios to encode PSD and rank constraints: even without rank constraints, one could still exploit the matrix decomposition $\rho = UU^\dagger$ to avoid the PSD projection, $\rho \succeq 0$, where $U \in \mathbb{C}^{d\times d}$. For the `lstsq` fitter method, the putative estimate $\widehat{\rho}$ is rescaled using the method proposed in [52]. For `CVXPY`, the convex constraint makes the optimization problem a semidefinite programming (SDP) instance. By default, `CVXPY` calls the `SCS` solver that can handle all problems (including SDPs) [53,54]. Further comparison results with matrix factorization techniques from the machine learning community is provided in the Appendix for $n = 12$.

The settings we consider for full tomography are the following: $\texttt{GHZ}(n)$, $\texttt{Hadamard}(n)$ and $\texttt{Random}(n)$ quantum states (for $n = 3, \ldots, 8$). We focus on fidelity of reconstruction and computation timings performance between `CVXPY`, `lstsq` and `MiFGD`. We use 100% of the measurements. We experimented with states simulated in QASM and measured taking 2048 shots. For `MiFGD`, we set $\eta = 0.001$, $\mu = \frac{3}{4}$, and stopping criterion/tolerance $\texttt{reltol} = 10^{-5}$. All experiments are run on a Macbook Pro with 2.3 GHz Quad-Core Intel Core i7CPU and 32GB RAM.

The results are shown in Figure 5; higher-dimensional cases are provided in Table 3. Some notable remarks: (i) For small-scale scenarios ($n = 3, 4$), `CVXPY` and `lstsq` attain almost perfect fidelity, while being comparable or faster than `MiFGD`. (ii) The difference in performance becomes apparent from $n = 6$ and on: while `MiFGD` attains 98% fidelity in <5 s, `CVXPY` and `lstsq` require up to hundreds of seconds to find a good solution. (iii) Finally, while `MiFGD` gets to high-fidelity solutions in seconds for $n = 7, 8$, `CVXPY` and `lstsq` methods could not finish tomography as their memory usage exceeded the system's available memory.

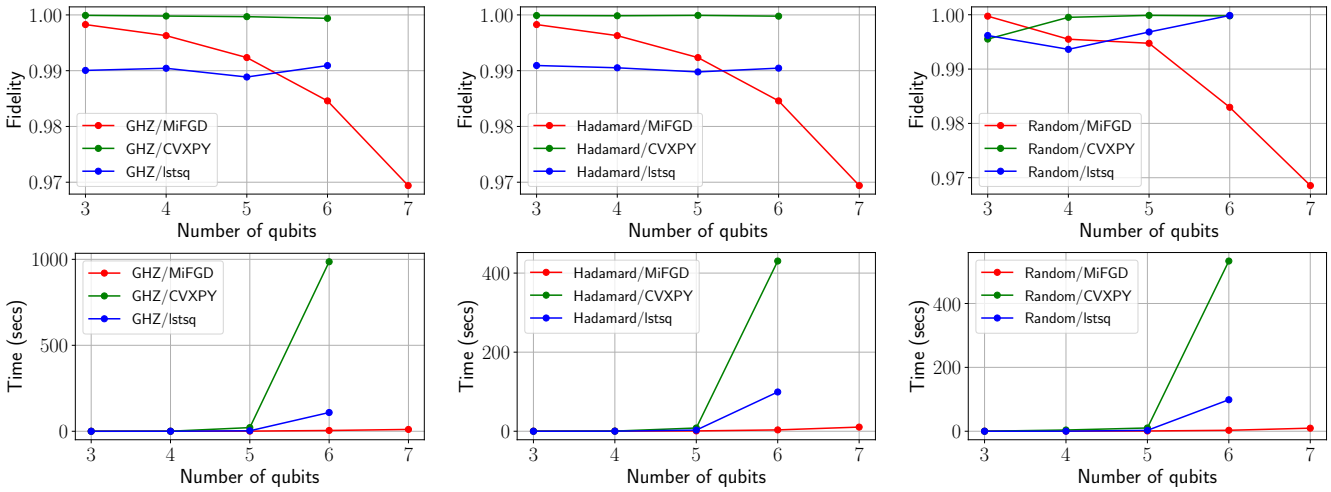

Figure 5. Fidelity versus time plots using synthetic IBM's quantum simulator data. **Left panel**: `GHZ_`(n) for $n = 3, 4$; **Middle panel**: `Hadamard_`(n) for $n = 3, 4$; **Right panel**: `Random_`(n) for $n = 3, 4$.

Table 3. Fidelity of reconstruction and computation timings using 100% of the complete measurements. Rows correspond to combinations of number of qubits (7∼8), synthetic circuit, and tomographic method (`MiFGD`, Qiskit's `lstsq` and `CVXPY` fitters. 2048 shots per measurement circuit. For `MiFGD`, $\eta = 0.001$, $\mu = \frac{3}{4}$, `reltol` $= 10^{-5}$. All experiments are run on a 13" Macbook Pro with 2.3 GHz Quad-Core Intel Core i7 CPU and 32 GB RAM.

Circuit	Method	Fidelity	Time (s)
`GHZ`(7)	`MiFGD`	0.969397	10.6709
`Hadamard`(7)	`MiFGD`	0.969397	10.4926
`Random`(7)	`MiFGD`	0.968553	9.59607
All above	`lstsq`, `CVXPY`	Memory limit exceeded	
`GHZ`(8)	`MiFGD`	0.940389	35.0666
`Hadamard`(8)	`MiFGD`	0.940390	37.5331
`Random`(8)	`MiFGD`	0.942815	36.3251
All above	`lstsq`, `CVXPY`	Memory limit exceeded	

It is noteworthy that the reported fidelities for `MiFGD` are the fidelities at the last iteration, before the stopping criterion is activated, or the maximum number of iterations is exceeded. However, the reported fidelity is not necessarily the best one during the whole execution: for all cases, we observe that `MiFGD` finds intermediate solutions with fidelity >99%. Though, it is not realistic to assume that the iteration with the best fidelity is known a priori, and this is the reason we report only the final iteration fidelity.

4.3. Performance Comparison of `MiFGD` with Neural-Network Quantum State Tomography

We compare the performance of `MiFGD` with neural network approaches. Per [9–11,27], we model a quantum state with a two-layer Restricted Boltzmann Machine (RBM). RBMs are stochastic neural networks, where each layer contains a number of binary stochastic variables: the size of the visible layer corresponds to the number of input qubits, while the size of the hidden layer is a hyperparameter controlling the representation error. We experiment with three types of RBMs for reconstructing either the positive-real wave function, the complex wave function, or the density matrix of the quantum state. In the first two cases the state is assumed pure while in the last, general mixed quantum states can be represented. We leverage the implementation in QuCumber [10], `PositiveRealWaveFunction` (`PRWF`), `ComplexWaveFunction` (`CWF`), and `DensityMatrix` (`DM`), respectively.

We reconstruct GHZ(n), Hadamard(n) and Random(n) quantum states (for $n = 3, \ldots, 8$), by training PRWF, CWF, and DM neural networks (We utilize GPU (NVidia GeForce GTX 1080 TI,11GB RAM) for faster training of the neural networks) with measurements collected by the QASM Simulator.

For our setting, we consider measpc = 50% and shots = 2048. The set of measurements is presented to the RBM implementation, along with the target positive-real wave function (for PRWF), complex wavefunction (for CWF) or the target density matrix (for DM) in a suitable format for training. We train Hadamard and Random states with 20 epochs, and GHZ state with 100 epochs (We experimented higher number of epochs (up to 500) for all cases, but after the reported number of epochs, Qucumber methods did not improve, if not worsened). We set the number of hidden variables (and also of additional auxilliary variables for DM) to be equal to the number of input variables n and we use 100 data points for both the positive and the negative phase of the gradient (as per the recommendation for the defaults). We choose $k = 10$ contrastive divergence steps and fixed the learning rate to 10 (per hyperparameter tuning). Lastly, we limit the fitting time of Qucumber methods (excluding data preparation time) to be three hours. To compare to the RBM results, we run MiFGD with $\eta = 0.001$, $\mu = \frac{3}{4}$, reltol $= 10^{-5}$ and using measpc = 50%, keeping previously chosen values for all other hyperparameters.

We report the fidelity of the reconstruction as a function of elapsed training time for $n = 3, 4$ in Figure 6 for PRWF, CWF, and DM. We observe that for all cases, Qucumber methods are orders of magnitude slower than MiFGD. E.g., for $n = 8$, for all three states, CWF and DM did not finish a single epoch in 3 h, while MiFG achieves high fidelity in less than 30 s. For the Hadamard(n) and Random(n), reaching reasonable fidelities is significantly slower for both CWF and DM, while PRWF hardly improves its performance throughout the training. For the GHZ case, CWF and DM also shows *non-monotonic* behaviors: even after a few *thousands of seconds*, fidelities have not "stabilized", while PRWF stabilizes in very low fidelities. In comparison MiFGD is several orders of magnitude faster than both CWF and DM and fidelity smoothly increases to comparable or higher values. Further, in Table 4, we report *final* fidelities (within the 3 h time window), and reported times.

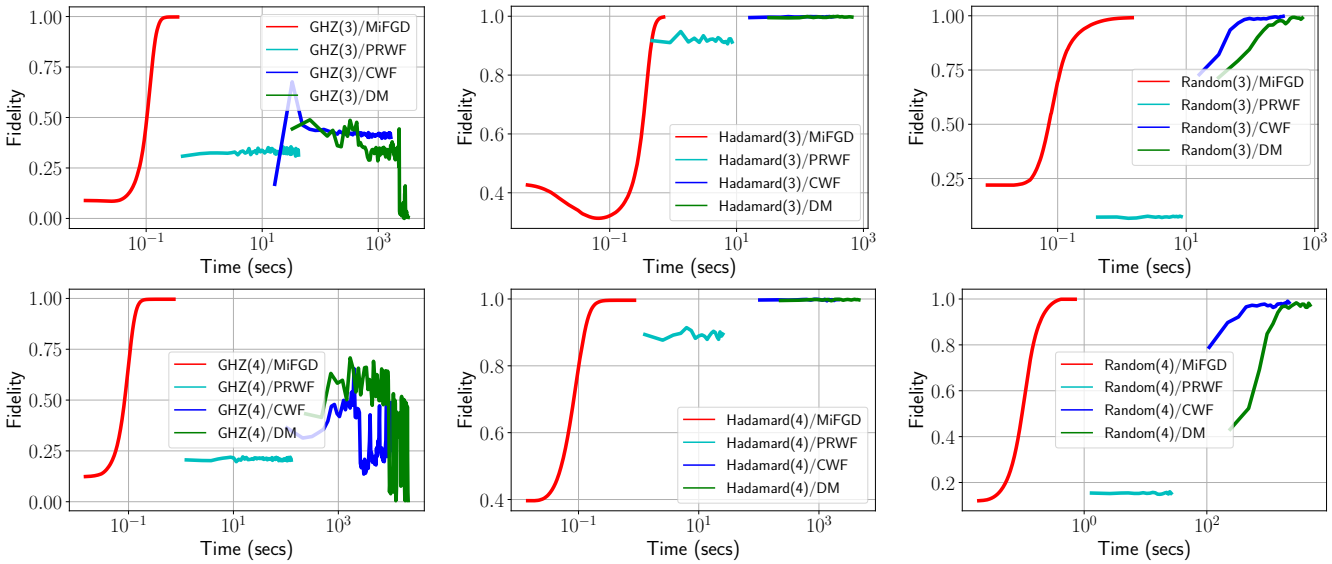

Figure 6. Fidelity versus time plots on MiFGD, PRWF, CWF, and DM, using synthetic IBM's quantum simulator data. **Left panel**: GHZ(n) for $n = 3, 4$; **Middle panel**: Hadamard(n) for $n = 3, 4$; **Right panel**: Random(n) for $n = 3, 4$.

Table 4. Fidelity of reconstruction and computation timings using `measpc` $= 50\%$ and `shots` $= 2048$. Rows correspond to combinations of number of qubits (3∼8), final fidelity within the 3h time limit, and computation time. For `MiFGD`, $\eta = 0.001$, $\mu = \frac{3}{4}$, `tol` $= 10^{-5}$. For `FGD`, $\eta = 0.001$, `tol` $= 10^{-5}$. "N/A" indicates that the method could not complete a single epoch in 3 h training time limit, and thus could not provide any fidelity result. All experiments are run on a NVidia GeForce GTX 1080 TI, 11 GB RAM.

Circuit		Method				
		MiFGD	**FGD**	**PRWF**	**CWF**	**DM**
GHZ(3)	Fidelity	0.997922	0.997857	0.314167	0.401737	0.005389
	Time (s)	0.348652	1.061421	42.27607	1649.224	3279.118
Hadamard(3)	Fidelity	0.997229	0.994191	0.912268	0.997914	0.997222
	Time (s)	0.706872	2.399405	8.492405	325.7040	656.6696
Random(3)	Fidelity	0.991063	0.988746	0.074774	0.997493	0.989754
	Time (s)	1.447057	3.431218	8.345135	322.4730	640.8185
GHZ(4)	Fidelity	0.996029	0.996041	0.204313	0.276491	0.138459
	Time (s)	0.733128	2.081035	126.2749	10756.87	>3 h
Hadamard(4)	Fidelity	0.996078	0.996083	0.894883	0.998071	0.997389
	Time (s)	0.852895	2.368223	25.15520	2087.540	4613.964
Random(4)	Fidelity	0.998850	0.998876	0.152971	0.984164	0.972877
	Time (s)	0.713302	2.380326	26.18863	2185.091	4802.495
GHZ(5)	Fidelity	0.992105	0.992106	0.132725	0.274665	0.005138
	Time (s)	0.946350	3.287358	395.3379	>3 h	>3 h
Hadamard(5)	Fidelity	0.992102	0.992100	0.869603	0.998246	0.996516
	Time (s)	1.183290	3.895312	79.39444	9319.140	>3 h
Random(5)	Fidelity	0.995126	0.995109	0.015913	0.623273	0.086777
	Time (s)	0.988173	3.407487	79.22450	9275.836	>3 h
GHZ(6)	Fidelity	0.984352	0.984340	0.089355	0.437323	0.310067
	Time (s)	3.829866	13.306954	1167.985	>3 h	>3 h
Hadamard(6)	Fidelity	0.984384	0.984377	0.842515	0.990849	0.998077
	Time (s)	2.500354	8.661999	246.0011	>3 h	>3 h
Random(6)	Fidelity	0.989543	0.989536	0.143145	0.784873	0.302534
	Time (s)	1.991154	7.604232	237.7037	>3 h	>3 h
GHZ(7)	Fidelity	0.969174	0.969168	0.058387	0.080648	N/A
	Time (s)	6.174129	15.895504	3633.082	>3 h	>3 h
Hadamard(7)	Fidelity	0.969156	0.969156	0.818174	0.996586	N/A
	Time (s)	6.324469	16.283301	713.9404	>3 h	>3 h
Random(7)	Fidelity	0.967640	0.967619	0.141745	0.06568	N/A
	Time (s)	6.802577	16.594162	746.2630	>3 h	>3 h
GHZ(8)	Fidelity	0.940601	0.940600	0.0400391	N/A	N/A
	Time (s)	21.16011	36.892739	>3 h	>3 h	>3 h
Hadamard(8)	Fidelity	0.940638	0.940638	0.794892	N/A	N/A
	Time (s)	22.30246	41.472961	2344.796	>3 h	>3 h
Random(8)	Fidelity	0.939418	0.939416	0.050521	N/A	N/A
	Time (s)	22.81059	41.193810	2196.259	>3 h	>3 h

4.4. The Effect of Parallelization

We study the effect of parallelization in running `MiFGD`. We parallelize the iteration step across a number of processes, that can be either distributed and network connected, or sharing memory in a multicore environment. Our approach is based on Message Passing Interface (MPI) specification [55], which is the lingua franca for interprocess communication

in high performance parallel and supercomputing applications. A MPI implementation provides facilities for launching processes organized in a virtual topology and highly tuned primitives for point-to-point and collective communication between them.

We assign to each process a subset of the measurement labels consumed by the parallel computation. At each step, a process first computes the local gradient-based corrections due only to its assigned Pauli monomials and corresponding measurements. These local gradient-based corrections will then (i) need to be communicated, so that they can be added, and (ii) finally, their sum will be shared across all processes to produce a global update for U for next step. We accomplish this structure in MPI using `MPI_Allreduce` collective communication primitive with `MPI_SUM` as its reduction operator: the underlying implementation will ensure minimum communication complexity for the operation (e.g., $\log p$ steps for p processes organized in a communication ring) and thus maximum performance (This communication pattern can alternatively be realized in two stages, as naturally suggested in its structure: (i) first invoke MPI's `MPI_Reduce` primitive, with `MPI_SUM` as its reduction operator, which results in the element-wise accumulation of local corrections (vector sum) at a single, designated *root* process, and (ii) finally, send a "copy" of this sum from *root* process to each process participating in the parallel computation (broadcasting); `MPI_Bcast` primitive can be utilized for this latter stage. However, `MPI_Allreduce` is typically faster, since its actual implementation is not constrained by the requirement to have the sum available at a specific, *root* process, at an intermediate time point - as the two-stage approach implies). We leverage `mpi4py` [56] bindings to issue MPI calls in our parallel Python code.

We conducted our parallel experiments on a server equipped with 4 × E7-4850 v2 CPUs @ 2.30GHz (48/96 physical/virtual cores), 256 GB RAM, using shared memory multiprocessing over multiple cores. We experimented with states simulated in QASM and measured taking 8192 `shots`; parallel `MiFGD` runs with default parameters and using all measurements (`measpc` = 100%). Reported times are wall-clock computation time. These exclude initialization time for all processes to load Pauli monomials and measurements: we here target parallelizing computation proper in `MiFGD`.

In our first round of experiments, we investigate the scalability of our approach. We vary the number p of parallel processes ($p = 1, 2, 4, 8, 16, 32, 48, 64, 80, 96$), collect timings for reconstructing $\texttt{GHZ}(4)$, $\texttt{Random}(6)$ and $\texttt{GHZ\_}(8)$ states and report speedups T_p/T_1 we gain from `MiFGD` in Figure 7 Left . We observe that the benefits of parallelization are pronounced for bigger problems (here: $n = 8$ qubits) and maximum scalability results when we use all physical cores (48 in our platform).

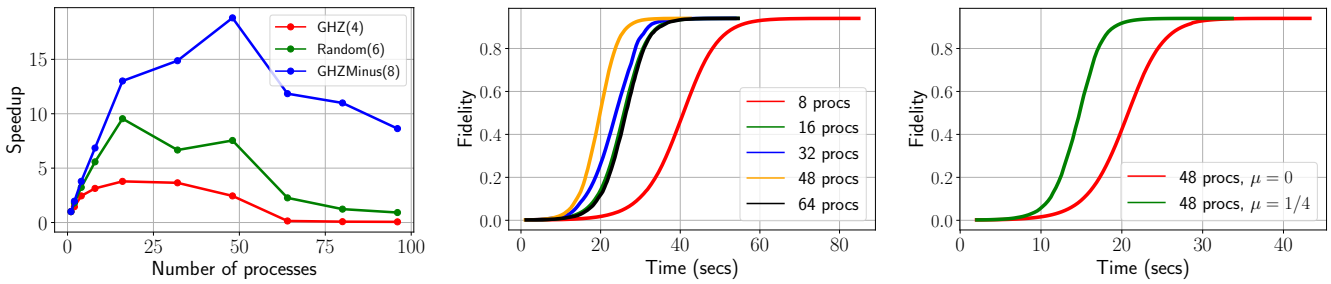

Figure 7. Left panel: Scalability of our approach as we vary the number p of parallel processes. **Middle panel:** Fidelity function versus time consumed for different number of processes p. **Right panel:** The effect of momentum for a fixed scenario with $\texttt{Hadamard}(10)$ state, $p = 48$, and varying momentum from $\mu = 0$ to $\mu = \frac{1}{4}$.

Further, we move to larger problems ($n = 10$ qubits: reporting on reconstructing $\texttt{Hadamard}(10)$ state) and focus on the effect parallelization to achieving a given level of fidelity in reconstruction. In Figure 7 Middle, we illustrate the fidelity as a function of the time spent in the iteration loop of `MiFGD` for ($p = 8, 16, 32, 48, 64$): we observe the smooth path to convergence in all p counts which again minimizes compute time for $p = 48$. Note that in this case we use `measpc` = 10% and $\mu = \frac{1}{4}$.

Finally, in Figure 7 Right, we fix the number of processes to $p = 48$, in order to minimize compute time and increase the percentage of used measurements to 20% of the total available for `Hadamard`(10). We vary the acceleration parameter, $\mu = 0$ (no acceleration) to $\mu = \frac{1}{4}$ and confirm that we indeed get faster convergence times in the latter case while the fidelity value remains the same (i.e., coinciding upper plateau value in the plots). We can also compare with the previous fidelity versus time plot, where the same μ but half the measurements are consumed: more measurements translate to faster convergence times (plateau is reached roughly 25% faster; compare the green line with the yellow line in the previous plot).

5. Conclusions and Discussions

We have introduced the `MiFGD` algorithm for the factorized form of the low-rank QST problems. We proved that, under certain assumptions on the problem parameters, `MiFGD` converges linearly to a neighborhood of the optimal solution, whose size depends on the momentum parameter μ, while using acceleration motions in a non-convex setting. We demonstrate empirically, using both simulated and real data, that `MiFGD` outperforms non-accelerated methods on both the original problem domain and the factorized space, contributing to recent efforts on testing QST algorithms in real quantum data [22]. These results expand on existing work in the literature illustrating the promise of factorized methods for certain low-rank matrix problems. Finally, we provide a publicly available implementation of our approach, compatible to the open-source software Qiskit [28], where we further exploit parallel computations in `MiFGD` by extending its implementation to enable efficient, parallel execution over shared and distributed memory systems.

Despite our theory does not apply to the Pauli basis measurement directly (i.e., *using randomly selected Pauli bases* Π_α*, does not lead to the* ℓ_2*-norm RIP*), using the data from random Pauli basis measurements directly could provide excellent tomographic reconstruction with `MiFGD`. Preliminary results suggest that only $O(r \cdot \log d)$ random Pauli bases should be taken for a reconstruction, with the same level of accuracy as with $O(r \cdot d \cdot \log d)$ expectation values of random Pauli matrices. We leave the analysis of our algorithm in this case for future work, along with detailed experiments.

5.1. Related Work

Matrix sensing. The problem of low-rank matrix reconstruction from few samples was first studied within the paradigm of convex optimization, using the nuclear norm minimization [29,57,58]. The use of non-convex approaches for low-rank matrix recovery—by imposing rank-constraints—has been proposed in [59–61]. In all these works, the convex and non-convex algorithms involve a full, or at least a truncated, singular value decomposition (SVD) per algorithm iteration. Since SVD can be prohibitive, these methods are limited to relatively small system sizes.

Momentum acceleration methods are used regularly in the convex setting, as well as in machine learning practical scenarios [62–67]. While momentum acceleration was previously studied in non-convex programming setups, it mostly involve non-convex constraints with a convex objective function [47,61,68,69]; and generic non-convex settings but only considering with the question of whether momentum acceleration leads to fast convergence to a saddle point or to a local minimum, rather than to a global optimum [45,70–72].

The factorized version for semi-definite programming was popularized in [73]. Effectively the factorization of a the set of PSD matrices to a product of rectangular matrices results in a non-convex setting. This approach have been heavily studied recently, due to computational and space complexity advantages [25,26,30–34,36–38,41,74–76]. None of the works above consider the inclusion and analysis of momentum. Moreover, the *Procrustes Flow* approach [32,34] uses certain initializations techniques, and thus relies on multiple SVD computations. Our approach on the other hand uses a single, unique, top-r SVD computation. Comparison results beyond QST are provided in the appendix.

Compressed sensing QST using non-convex optimization. There are only few works that study non-convex optimization in the context of compressed sensing QST. The authors of [16] propose a hybrid algorithm that (i) starts with a conjugate-gradient (CG) algorithm in the factored space, in order to get initial rapid descent, and (ii) switch over to accelerated first-order methods in the original ρ space, provided one can determine the switch-over point cheaply. Using the multinomial maximum likelihood objective, in the initial CG phase, the Hessian of the objective is computed per iteration (i.e., a $4^n \times 4^n$ matrix), along with its eigenvalue decomposition. Such an operation is costly, even for moderate values of qubit number n, and heuristics are proposed for its completion. From a theoretical perspective, [16] provide no convergence or convergence rate guarantees.

From a different perspective, [77] relies on spectrum estimation techniques [78,79] and the Empirical Young Diagram algorithm [80,81] to prove that $O(rd/\varepsilon)$ copies suffice to obtain an estimate $\hat{\rho}$ that satisfies $\|\hat{\rho} - \rho^\star\|_F^2 \leq \varepsilon$; however, to the best of our knowledge, there is no concrete implementation of this technique to compare with respect to scalability.

Ref. [82] proposes an efficient quantum tomography protocol by determining the permutationally invariant part of the quantum state. The authors determine the minimal number of local measurement settings, which scales quadratically with the number of qubits. The paper determines which quantities have to be measured in order to get the smallest uncertainty possible. See [83] for a more recent work on permutationally invariant tomography. The method has been tested in a six-qubit experiment in [84].

Ref. [22] presented an experimental implementation of compressed sensing QST of a $n = 7$ qubit system, where only 127 Pauli basis measurements are available. To achieve recovery in practice, the authors proposed a computationally efficient estimator, based on gradient descent method in the factorized space. The authors of [22] focus on the experimental efficiency of the method, and provide no specific results on the optimization efficiency, neither convergence guarantees of the algorithm. Further, there is no available implementation publicly available.

Similar to [22], Ref. [26] also proposes a non-convex projected gradient decent algorithm that works on the factorized space in the QST setting. The authors prove a rigorous convergence analysis and show that, under proper initialization and step-size, the algorithm is guaranteed to converge to the global minimum of the problem, thus ensuring a provable tomography procedure. *Our results extend these results by including acceleration techniques in the factorized space*. The key contribution of our work is proving convergence of the proposed algorithm in a *linear* rate to the global minimum of the problem, under common assumptions. Proving our results required developing a whole set of new techniques, which are not based on a mere extension of existing results.

Compressed sensing QST using convex optimization. The original formulation of compressed sensing QST [4] is based on convex optimization methods, solving the trace-norm minimization, to obtain an estimation of the low-rank state. It was later shown [8] that essentially any convex optimization program can be used to robustly estimate the state. In general, there are two drawbacks in using convex optimization optimization in QST. Firstly, as the dimension of density matrices grow exponentially in the number of qubits, the search space in convex optimization grows exponentially in the number of qubits. Secondly, the optimization requires projection onto the PSD cone at every iteration, which becomes exponentially hard in the number of qubits. *We avoid these two drawbacks by working in the factorized space*. Using this factorization results in a search space that is substantially smaller than its convex counterpart, and moreover, in a single use of top-r SVD during the entire execution algorithm. Bypassing these drawbacks, together with accelerating motions, allows us to estimate quantum states of larger qubit systems than state-of-the-art algorithms.

Full QST using non-convex optimization. The use of non-convex algorithms in QST was studied in the context of full tomography as well. By "full tomography" we refer to the situation where an informationally complete measurement is performed, so that the input data to the algorithm is of size 4^n. The exponential scaling of the data size restrict the applicability of full tomography to relatively small system sizes. In this setting non-convex

algorithms which work in the factored space were studied [85–89]. Except of the work [88], we are not aware of theoretical results on the convergence of the proposed algorithm due to the presence of spurious local minima. The authors of [88] characterize the local vs. the global behavior of the objective function under the factorization $\rho = UU^\dagger$ and discuss how existing methods fail due to improper stopping criteria or due to the lack of algorithmic convergence results. Their work highlights the lack of rigorous convergence results of non-convex algorithms used in full quantum state tomography. There is no available implementation publicly available for these methods as well.

Full QST using convex optimization. Despite the non-scalability of full QST, and the limitation of convex optimization, a lot of research was devoted to this topic. Here, we mention only a few notable results that extend the applicability of full QST using specific techniques in convex optimization. Ref [52] shows that for given measurement schemes the solution for the maximum likelihood is given by a linear inversion scheme, followed by a projection onto the set of density matrices. More recently, the authors of [18] used a combination of the techniques of [52] with the sparsity of the Pauli matrices and the use of GPUs to perform a full QST of 14 qubits. While pushing the limit of full QST using convex optimization, obtaining full tomographic *experimental* data for more than a dozen qubits is significantly time-intensive. Moreover, this approach is highly centralized, in comparison to our approach that can be distributed. Using the sparsity pattern property of the Pauli matrices and GPUs is an excellent candidate approach to further enhance the performance of non-convex compressed sensing QST.

QST using neural networks. Deep neural networks are ubiquitous, with many applications to science and industry. Recently, [9–11,27] show how machine learning and neural networks can be used to perform QST, driven by experimental data. The neural network architecture used is based on restricted Boltzmann machines (RBMs) [90], which feature a visible and a hidden layer of stochastic binary neurons, fully connected with weighted edges. Test cases considered include reconstruction of W state, magnetic observables of local Hamiltonians, the unitary dynamics induced by Hamiltonian evolution. Comparison results are provided in the Main Results section. Alternative approaches include conditional generative adversarial networks (CGANs) [91,92]: in this case, two dueling neural networks, a generator and a discriminator, learn to generate and identify multi-modal models from data.

QST for Matrix Product States (MPS). This is the case of highly structured quantum states where the state is well-approximated by a MPS of low bond dimension [12,13]. The idea behind this approach is, in order to overcome exponential bottlenecks in the general QST case, we require highly structured subsets of states, similar to the assumptions made in compressed sensing QST. MPS QST is considered an alternative approach to reduce the computational and storage complexity of QST.

Direct fidelity estimation. Rather than focusing on entrywise estimation of density matrices, the direct fidelity estimation procedure focuses on checking how close is the state of the system to a target state, where closeness is quantified by the fidelity metric. Classic techniques require up to $2^n/\epsilon^4$ number of samples, where ϵ denotes the accuracy of the fidelity term, when considering a general quantum state [93,94], but can be reduced to almost dimensionality-free $1/\epsilon^2$ number of samples for specific cases, such as stabilizer states [95–97]. Shadow tomography is considered as an alternative and generalization of this technique [98,99]; however, as noted in [94], the procedure in [98,99] requires exponentially long quantum circuits that act collectively on all the copies of the unknown state stored in a quantum memory, and thus has not been implemented fully on real quantum machines. A recent neural network-based implementation of such indirect QST learning methods is provided here [100].

The work in [93,94], goes beyond simple fidelity estimation, and utilizes random single qubit rotations to learn a minimal sketch of the unknown quantum state by which one that can predict arbitrary linear function of the state. Such methods constitute a favorable alternative to QST as they do not require number of samples that scale polynomially with

the dimension; however, this, in turn, implies that these methods cannot be used in general to estimate the density matrix itself.

Author Contributions: Conceptualization, A.K. (Amir Kalev) and A.K. (Anastasios Kyrillidis); methodology, J.L.K. and A.K. (Anastasios Kyrillidis); software, J.L.K. and G.K.; formal analysis, J.L.K. and A.K. (Anastasios Kyrillidis); investigation, J.L.K., G.K., A.K. (Amir Kalev) and A.K. (Anastasios Kyrillidis); data curation, K.X.W. and G.K.; writing—original draft preparation, J.L.K. and A.K. (Anastasios Kyrillidis); writing—review and editing, J.L.K., G.K., A.K. (Amir Kalev) and A.K. (Anastasios Kyrillidis); visualization, J.L.K. and G.K.; supervision, A.K. (Amir Kalev) and A.K. (Anastasios Kyrillidis); project administration, A.K. (Anastasios Kyrillidis); funding acquisition, A.K. (Amir Kalev) and A.K. (Anastasios Kyrillidis). All authors have read and agreed to the published version of the manuscript.

Funding: Anastasios Kyrillidis and Amir Kalev acknowledge funding by the NSF (CCF-1907936).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement:
Not applicable.

Data Availability Statement: The empirical results were obtained via synthetic and real experiments; the algorithm's implementation is available at https://github.com/gidiko/MiFGD (accessed on 20 January 2023).

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

Appendix A. Additional Experiments

Appendix A.1. IBM Quantum System Experiments: `GHZ_`(6) Circuit, 2048 `Shots`

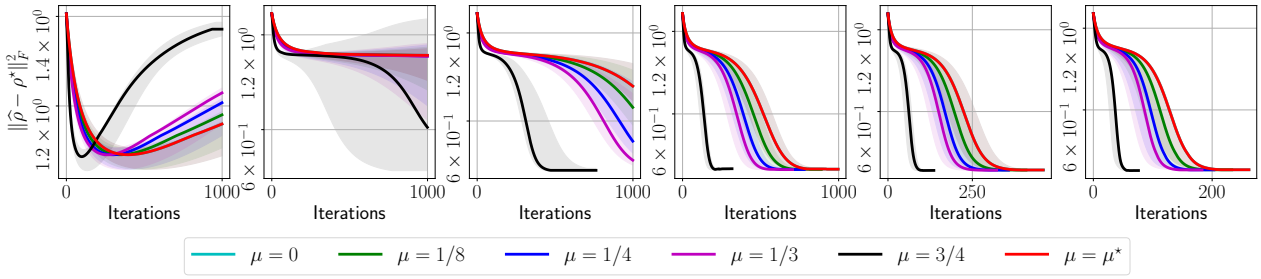

Figure A1. Target error list plots for reconstructing `GHZ_`(6) circuit using real measurements from IBM Quantum system experiments.

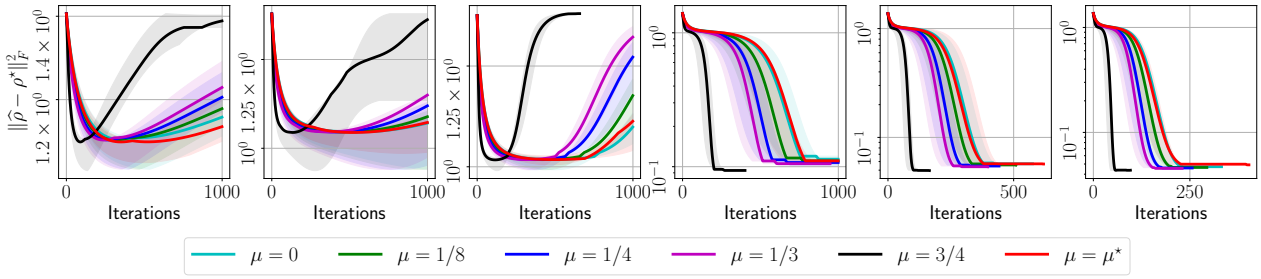

Figure A2. Target error list plots for reconstructing `GHZ_`(6) circuit using synthetic measurements from IBM's quantum simulator.

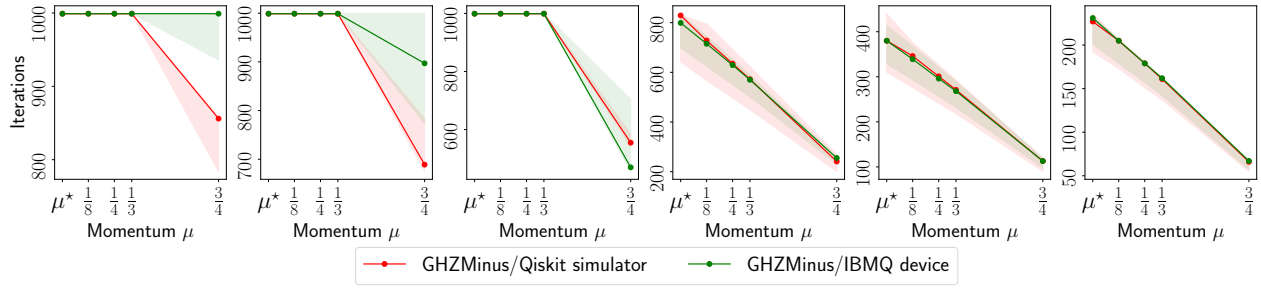

Figure A3. Convergence iteration plots for reconstructing GHZ_(6) circuit using using real measurements from IBM Quantum system experiments and synthetic measurements from Qiskit simulation experiments.

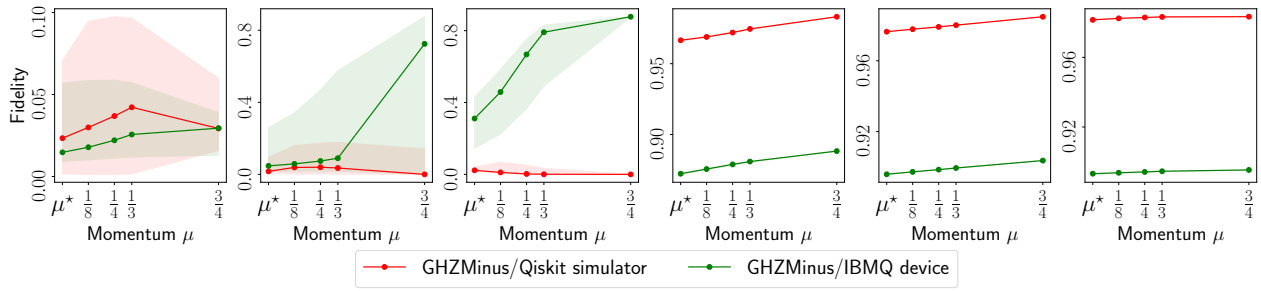

Figure A4. Fidelity list plots for reconstructing GHZ_(6) circuit using using real measurements from IBM Quantum system experiments and synthetic measurements from Qiskit simulation experiments.

Appendix A.2. IBM Quantum System Experiments: GHZ_(8) Circuit, 2048 Shots

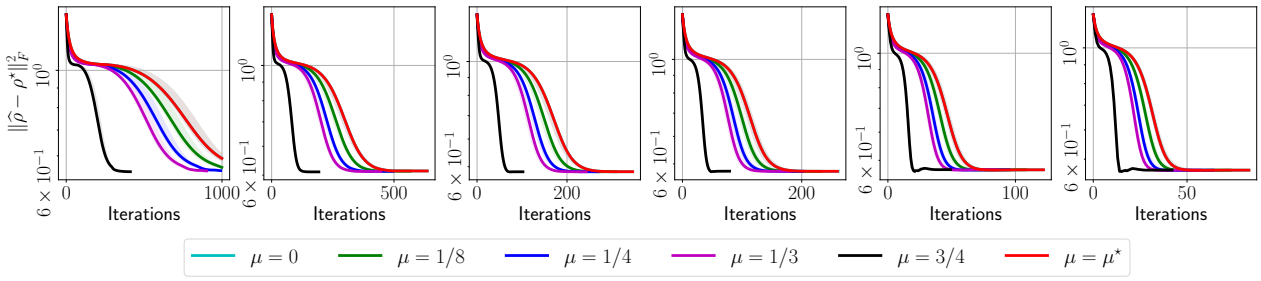

Figure A5. Target error list plots for reconstructing GHZ_(8) circuit using real measurements from IBM Quantum system experiments.

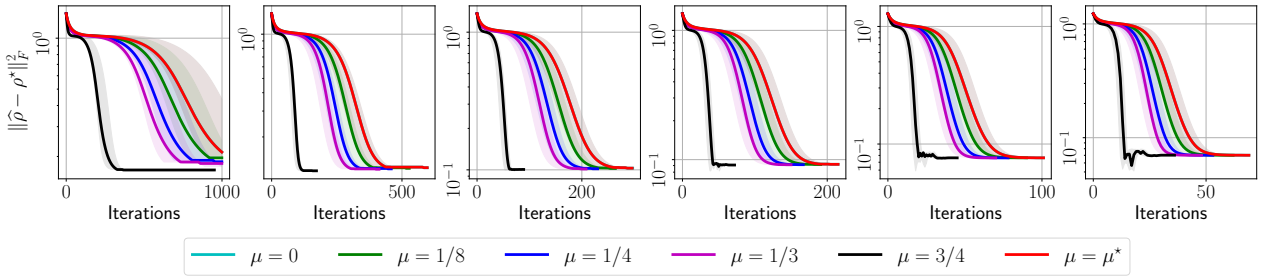

Figure A6. Target error list plots for reconstructing GHZ_(8) circuit using synthetic measurements from IBM's quantum simulator.

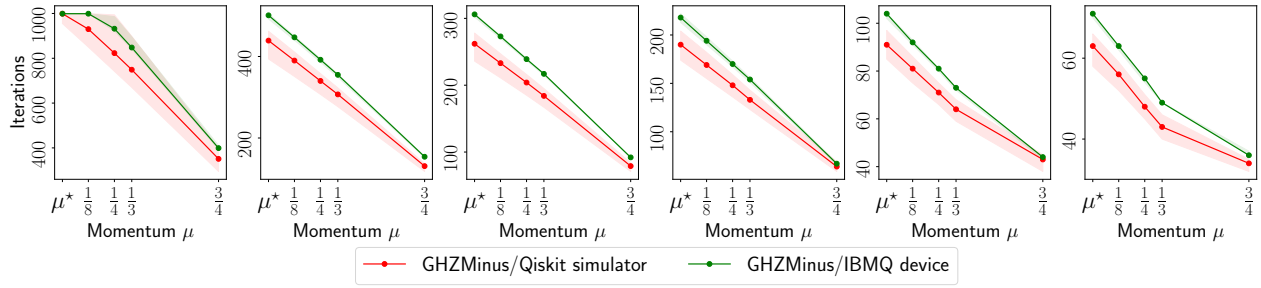

Figure A7. Convergence iteration plots for reconstructing GHZ_(8) circuit using using real measurements from IBM Quantum system experiments and synthetic measurements from Qiskit simulation experiments.

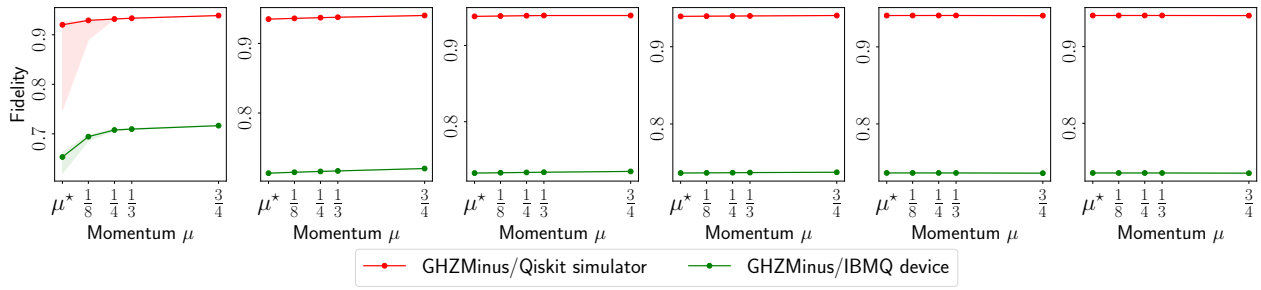

Figure A8. Fidelity list plots for reconstructing GHZ_(8) circuit using using real measurements from IBM Quantum system experiments and synthetic measurements from Qiskit simulation experiments.

Appendix A.3. IBM Quantum System Experiments: GHZ_(8) Circuit, 4096 Shots

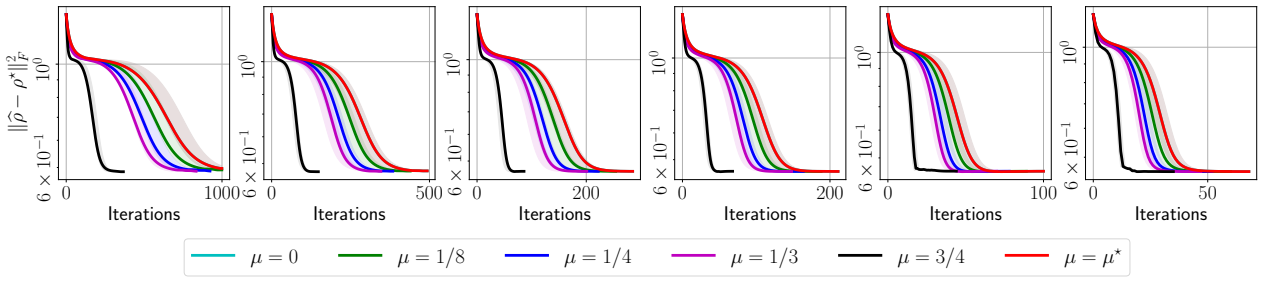

Figure A9. Target error list plots for reconstructing GHZ_(8) circuit using real measurements from IBM Quantum system experiments.

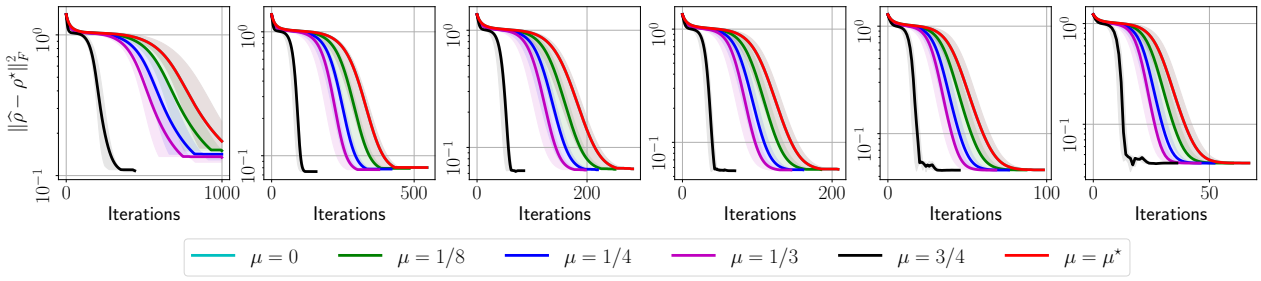

Figure A10. Target error list plots for reconstructing GHZ_(8) circuit using synthetic measurements from IBM's quantum simulator.

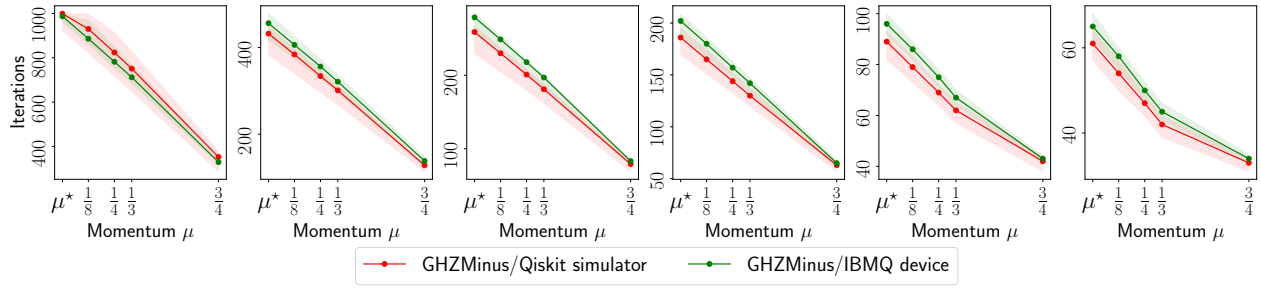

Figure A11. Convergence iteration plots for reconstructing `GHZ_`(8) circuit using using real measurements from IBM Quantum system experiments and synthetic measurements from Qiskit simulation experiments.

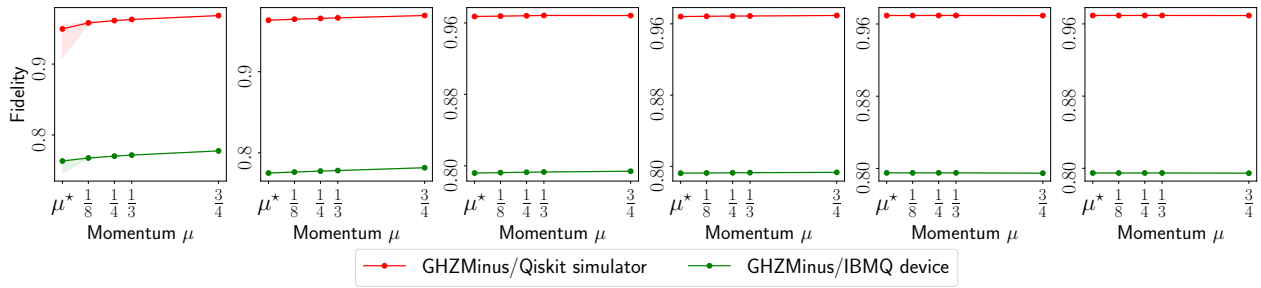

Figure A12. Fidelity list plots for reconstructing `GHZ_`(8) circuit using using real measurements from IBM Quantum system experiments and synthetic measurements from Qiskit simulation experiments.

Appendix A.4. IBM Quantum System Experiments: `Hadamard`(6) *Circuit, 8192* `Shots`

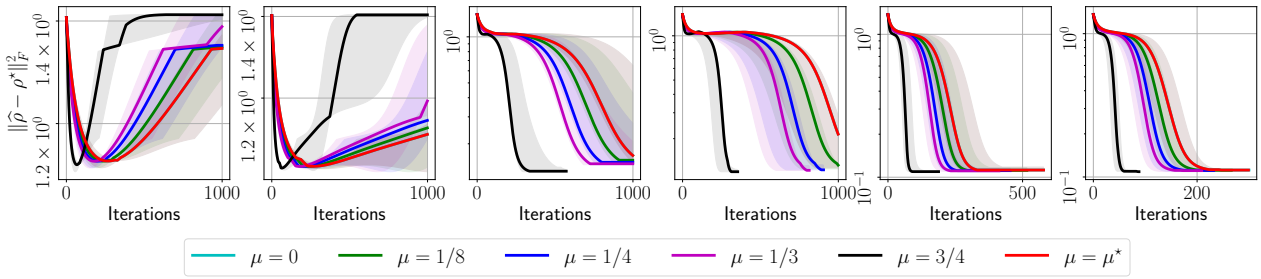

Figure A13. Target error list plots for reconstructing `Hadamard`(6) circuit using real measurements from IBM Quantum system experiments.

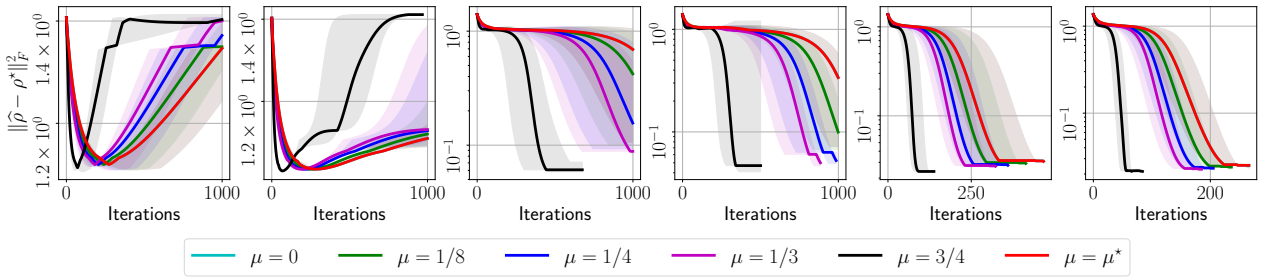

Figure A14. Target error list plots for reconstructing `Hadamard`(6) circuit using synthetic measurements from IBM's quantum simulator.

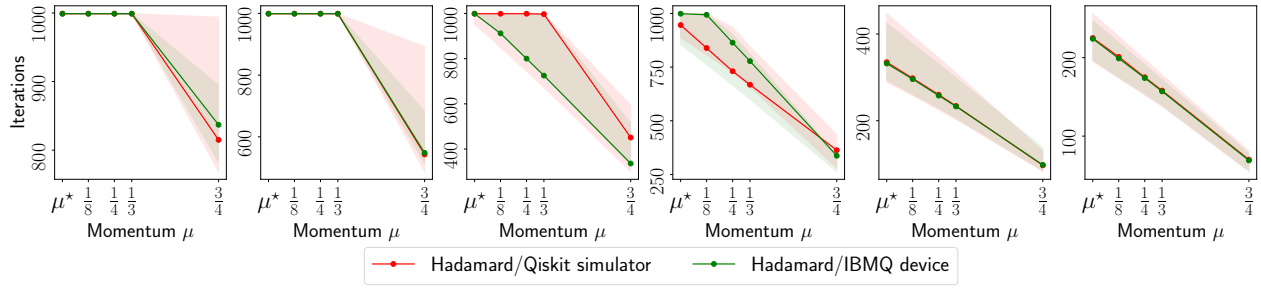

Figure A15. Convergence iteration plots for reconstructing `Hadamard`(6) circuit using using real measurements from IBM Quantum system experiments and synthetic measurements from Qiskit simulation.

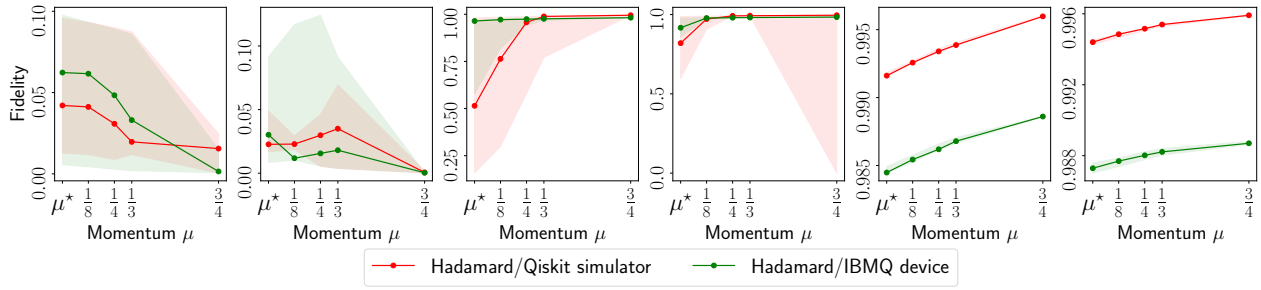

Figure A16. Fidelity list plots for reconstructing `Hadamard`(6) circuit using using real measurements from IBM Quantum system experiments and synthetic measurements from Qiskit simulation experiments.

Appendix A.5. IBM Quantum System Experiments: `Hadamard`(8) *Circuit, 4096* `Shots`

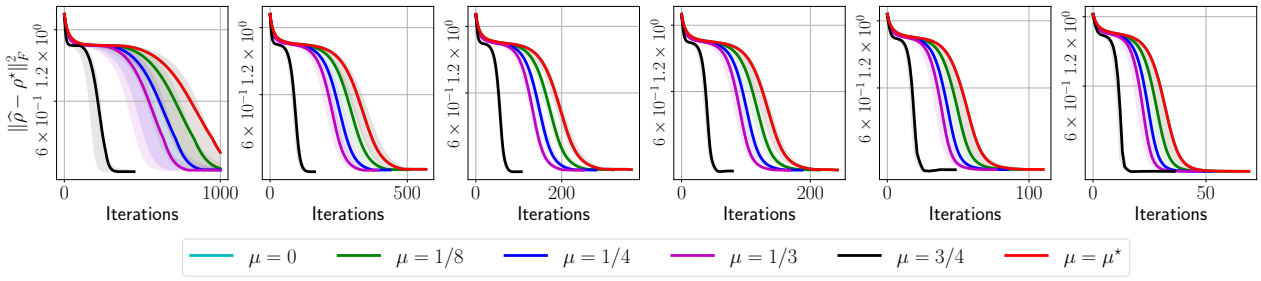

Figure A17. Target error list plots for reconstructing `Hadamard`(8) circuit using real measurements from IBM Quantum system experiments.

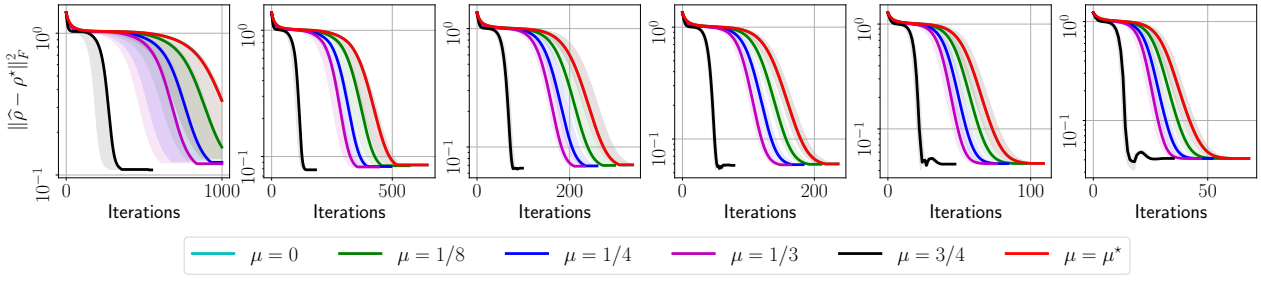

Figure A18. Target error list plots for reconstructing `Hadamard`(8) circuit using synthetic measurements from IBM's quantum simulator.

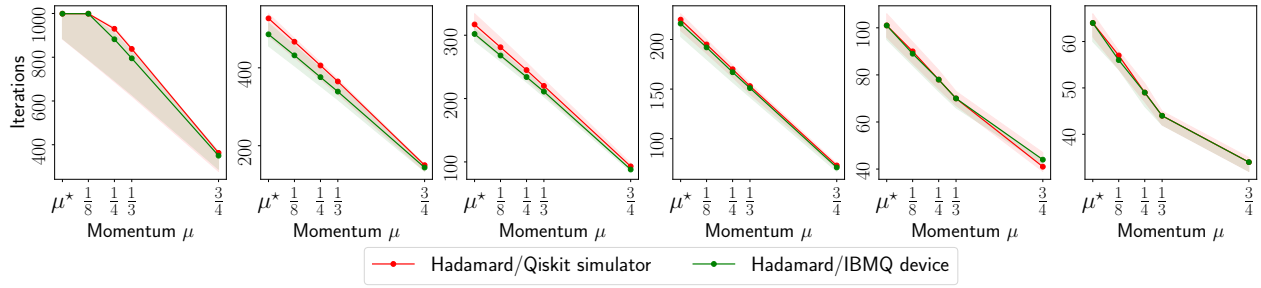

Figure A19. Convergence iteration plots for reconstructing `Hadamard`(8) circuit using using real measurements from IBM Quantum system experiments and synthetic measurements from Qiskit simulation.

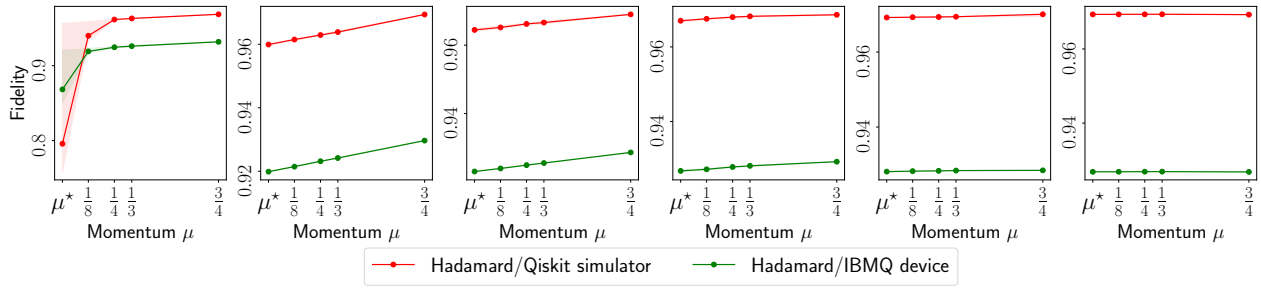

Figure A20. Fidelity list plots for reconstructing `Hadamard`(8) circuit using using real measurements from IBM Quantum system experiments and synthetic measurements from Qiskit simulation experiments.

Appendix A.6. Synthetic Experiments for $n = 12$

We compare `MiFGD` with (i) the `Matrix ALPS` framework [61], *a state of the art projected gradient descent algorithm, and an optimized version of matrix iterative hard thresholding*, operating on the full matrix variable ρ, with adaptive step size η (we note that this algorithm has outperformed most of the schemes that work on the original space ρ; see [61]); (ii) the plain Procrustes Flow/`FGD` algorithm [25,26,32], where we use the step size as reported in [25], since the later has reported better performance than vanilla Procrustes Flow. *We note that the Procrustes Flow/`FGD` algorithm is similar to our algorithm without acceleration.* Further, the original Procrustes Flow/`FGD` algorithm relies on performing many iterations in the original space ρ as an initialization scheme, which is often prohibitive as the problem dimensions grow. Both for our algorithm and the plain Procrustes Flow/`FGD` scheme, we use random initialization.

To properly compare the algorithms in the above list, we pre-select a common set of problem parameters. We fix the dimension $d = 4096$ (equivalent to $n = 12$ qubits), and the rank of the optimal matrix $\rho^\star \in \mathbb{R}^{d\times d}$ to be $r = 10$ (equivalent to a mixed quantum state reconstruction). Similar behavior has been observed for other values of r, and are omitted. We fix the number of observables m to be $m = c \cdot d \cdot r$, where $c \in \{3, 5\}$. In all algorithms, we fix the maximum number of iterations to 4000, and we use the same stopping criterion: $\|\rho_{i+1} - \rho_i\|_F / \|\rho_i\|_F \leq \texttt{tol} = 10^{-3}$. For the implementation of `MiFGD`, we have used the momentum parameter $\mu = \frac{2}{3}$, as well as the theoretical μ value.

The procedure to generate synthetically the data is as follows: The observations y are set to $y = \mathcal{A}(\rho^\star) + w$ for some noise vector w; while the theory holds for the noiseless case, we show empirically that noisy cases are robustly handled by the same algorithm. We use permuted and subsampled noiselets for the linear operator $\mathcal{A}$ [101]. The optimal matrix $\rho^\star$ is generated as the multiplication of a tall matrix $U^\star \in \mathbb{R}^{d\times r}$ such that $\rho^\star = U^\star U^{\star\top}$, and $\|\rho^\star\|_F = 1$, without loss of generality. The entries of $U^\star$ are drawn i.i.d. from a Gaussian distribution with zero mean and unit variance. In the noisy case, w has the same dimensions with y, its entries are drawn from a zero mean Gaussian distribution with

norm $\|w\|_2 = 0.01$. The random initialization is defined as U_0 drawn i.i.d. from a Gaussian distribution with zero mean and unit variance.

The results are shown in Figure A21. Some notable remarks: (i) While factorization techniques might take more iterations to converge compared to non-factorized algorithms, the per iteration time complexity is much less, such that overall, factorized gradient descent converges more quickly in terms of total execution time. (ii) Our proposed algorithm, *even under the restrictive assumptions on acceleration parameter* μ, performs better than the non-accelerated factored gradient descent algorithms, such as Procrustes Flow. (iii) Our theory is conservative: using a much larger μ we obtain a faster convergence; the proof for less strict assumptions for μ is an interesting future direction. In all cases, our findings illustrate the effectiveness of the proposed schemes on different problem configurations.

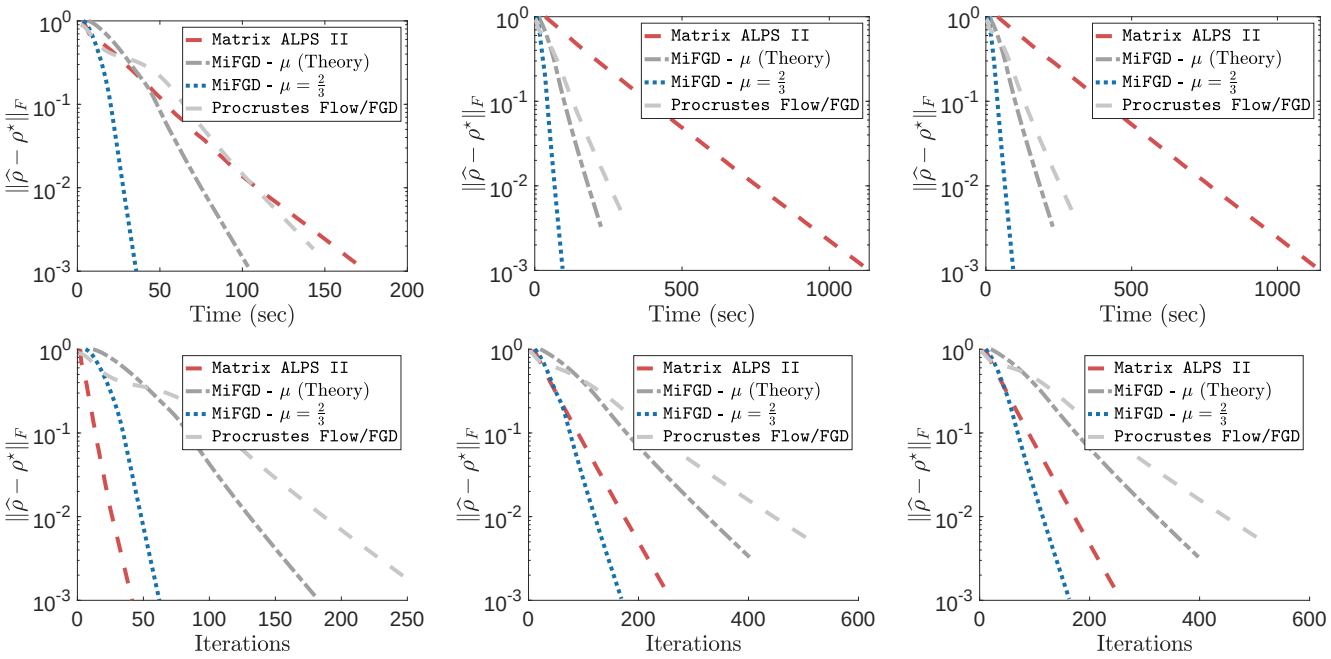

Figure A21. Synthetic example results on low-rank matrix sensing in higher dimensions (equivalent to $n = 12$ qubits). **Top row**: Convergence behavior vs. time elapsed. **Bottom row:** Convergence behavior vs. number of iterations. **Left panel**: $c = 5$, noiseless case; **Center panel:** $c = 3$, noiseless case; **Right panel:** $c = 5$, noisy case, $\|w\|_2 = 0.01$.

Appendix A.7. Asymptotic Complexity Comparison of `lstsq`*,* `CVXPY`*, and* `MiFGD`

We first note that `lstsq` can be only applied to the case we have a full tomographic set of measurements; this makes `lstsq` algorithm inapplicable in the compressed sensing scenario, where the number of measurements can be significantly reduced. Yet, we make the comparison by providing information-theoretically complete set of measurements to `lstsq` and `CVXPY`, as well as to `MiFGD`, to highlight the efficiency of our proposed method, even in the scenario that is not exactly intended in our work. Given this, we compare in detail the asymptotic scailing of `MiFGD` with `lstsq` and `CVXPY` below:

- `lstsq` is based on the computation of eigenvalues/eigenvector pairs (among other steps) of a matrix of size equal to the density matrix we want to reconstruct. Based on our notation, the density matrices are denoted as ρ with dimensions $2^n \times 2^n$. Here, n is the number of qubits in the quantum system. Standard libraries for eigenvalue/eigenvector calculations, like LAPACK, reduce a Hermitian matrix to tridiagonal form using the Householder method, which takes overall a $O\big((2^n)^3\big)$ computational complexity. The other steps in the `lstsq` procedure either take constant time, or $O(2^n)$ complexity. Thus, the actual run-time of an implementation depends on the eigensystem solver that is being used.

- `CVXPY` is distributed with the open source solvers; for the case of SDP instances, `CVXPY` utilizes the Splitting Conic Solver (SCS) (https://github.com/cvxgrp/scs (accessed on 20 January 2023)), a general numerical optimization package for solving large-scale convex cone problems. SCS applies Douglas-Rachford splitting to a homogeneous embedding of the quadratic cone program. Based on the PSD constraint, this again involves the computation of eigenvalues/eigenvector pairs (among other steps) of a matrix of size equal to the density matrix we want to reconstruct. This takes overall a $O((2^n)^3)$ computational complexity, *not including the other steps performed within the SCS solver*. This is an iterative algorithm that requires such complexity per iteration. Douglas-Rachford splitting methods enjoy $O(\frac{1}{\varepsilon})$ convergence rate in general [53,102,103]. This leads to a rough $O((2^n)^3 \cdot \frac{1}{\varepsilon})$ overall iteration complexity (This is an optimistic complexity bound since we have skipped several details within the Douglas-Rachford implementation of `CVXPY`).
- For `MiFGD`, and for sufficiently small momentum value, we require $O(\sqrt{\kappa} \cdot \log(\frac{1}{\varepsilon}))$ iterations to get close to the optimal value. Per iteration, `MiFGD` does not involve any expensive eigensystem solvers, but relies only on matrix-matrix and matrix-vector multiplications. In particular, the main computational complexity per iteration origins from the iteration:

$$U_{k+1} = Z_k - \eta \mathcal{A}^\dagger\Big(\mathcal{A}(Z_k Z_k^\dagger) - y\Big) \cdot Z_k,$$
$$Z_{k+1} = U_{k+1} + \mu(U_{k+1} - U_k).$$

  Here, $U_k, Z_k \in \mathbb{R}^{2^n \times r}$ for all k. Observe that $\mathcal{A}(Z_k Z_k^\dagger) \in \mathbb{R}^m$ where each element is computed independently. For an index $j \in [m]$, $(\mathcal{A}(Z_k Z_k^\dagger))_j = \mathtt{Tr}(A_j Z_k Z_k^\dagger)$ requires $O((2^n)^2 \cdot r)$ complexity, and thus computing $\mathcal{A}(Z_k Z_k^\dagger) - y$ requires $O((2^n)^2 \cdot r)$ complexity, overall. By definition the adjoing operation $\mathcal{A}^\dagger : \mathbb{R}^m \to \mathbb{C}^{2^n \times 2^n}$ satisfies: $\mathcal{A}^\dagger(x) = \sum_{i=1}^m x_i A_i$; thus, the operation $\mathcal{A}^\dagger\big(\mathcal{A}(Z_k Z_k^\dagger) - y\big)$ is still dominated by $O((2^n)^2 \cdot r)$ complexity. Finally, we perform one more matrix-matrix multiplication with Z_i, which results into an additional $O((2^n)^2 \cdot r)$ complexity. The rest of the operations involve adding $2^n \times r$ matrices, which does not dominate the overall complexity. Combining the iteration complexity with the per-iteration computational complexity, `MiFGD` has a $O((2^n)^2 \cdot r \cdot \sqrt{\kappa} \cdot \log(\frac{1}{\varepsilon}))$ complexity.

  Combining the above, we summarize the following complexities:

$$\underbrace{O((2^n)^3)}_{\texttt{lstsq}} \quad \text{vs} \quad \underbrace{O((2^n)^3 \cdot \tfrac{1}{\varepsilon})}_{\texttt{CVXPY}} \quad \text{vs} \quad \underbrace{O((2^n)^2 \cdot r \cdot \sqrt{\kappa} \cdot \log(\tfrac{1}{\varepsilon}))}_{\texttt{MiFGD}}$$

Observe that (i) `MiFGD` has the best dependence on the number of qubits and the ambient dimension of the problem, 2^n; (ii) `MiFGD` applies to cases that `lstsq` is inapplicable; (iii) `MiFGD` has a better iteration complexity than other iterative algorithms, while has a better polynomial dependency on 2^n.

Appendix B. Detailed Proof of Theorem 1

For notational brevity, we first denote $U_+ \equiv U_{k+1}$, $U \equiv U_k$, $U_- \equiv U_{k-1}$ and $Z \equiv Z_k$. Let us start with the following equality. For $R_Z \in \mathcal{O}$ as the minimizer of $\min_{R \in \mathcal{O}} \|Z - U^\star R\|_F$, we have:

$$\|U_+ - U^\star R_Z\|_F^2 = \|U_+ - Z + Z - U^\star R_Z\|_F^2 \tag{A1}$$
$$= \|U_+ - Z\|_F^2 + \|Z - U^\star R_Z\|_F^2 - 2\langle U_+ - Z, U^\star R_Z - Z\rangle. \tag{A2}$$

The proof focuses on how to bound the last part on the right-hand side. By the definition of U_+, we get:

$$\begin{aligned}\langle U_+ - Z, U^\star R_Z - Z\rangle &= \Big\langle Z - \eta\mathcal{A}^\dagger\Big(\mathcal{A}(ZZ^\dagger) - y\Big)Z - Z, U^\star R_Z - Z\Big\rangle \\ &= \eta\Big\langle \mathcal{A}^\dagger\Big(\mathcal{A}(ZZ^\dagger) - y\Big)Z, Z - U^\star R_Z\Big\rangle\end{aligned}$$

Observe the following:

$$\begin{aligned}\Big\langle \mathcal{A}^\dagger\Big(\mathcal{A}(ZZ^\dagger) - y\Big)Z, Z - U^\star R_Z\Big\rangle &= \Big\langle \mathcal{A}^\dagger\Big(\mathcal{A}(ZZ^\dagger) - y\Big), ZZ^\dagger - U^\star R_Z Z^\dagger\Big\rangle \\ &= \Big\langle \mathcal{A}^\dagger\Big(\mathcal{A}(ZZ^\dagger) - y\Big), ZZ^\dagger - \tfrac{1}{2}U^\star U^{\star\dagger} + \tfrac{1}{2}U^\star U^{\star\dagger} - U^\star R_Z Z^\dagger\Big\rangle \\ &= \tfrac{1}{2}\Big\langle \mathcal{A}^\dagger\Big(\mathcal{A}(ZZ^\dagger) - y\Big), ZZ^\dagger - U^\star U^{\star\dagger}\Big\rangle \\ &\quad + \Big\langle \mathcal{A}^\dagger\Big(\mathcal{A}(ZZ^\dagger) - y\Big), \tfrac{1}{2}(ZZ^\dagger + U^\star U^{\star\dagger}) - U^\star R_Z Z^\dagger\Big\rangle \\ &= \tfrac{1}{2}\Big\langle \mathcal{A}^\dagger\Big(\mathcal{A}(ZZ^\dagger) - y\Big), ZZ^\dagger - U^\star U^{\star\dagger}\Big\rangle \\ &\quad + \tfrac{1}{2}\Big\langle \mathcal{A}^\dagger\Big(\mathcal{A}(ZZ^\dagger) - y\Big), (Z - U^\star R_Z)(Z - U^\star R_Z)^\dagger\Big\rangle.\end{aligned}$$

By Lemmata A7 and A8, we have:

$$\begin{aligned}\|U_+ - U^\star R_Z\|_F^2 &= \|U_+ - Z\|_F^2 + \|Z - U^\star R_Z\|_F^2 - 2\langle U_+ - Z, U^\star R_Z - Z\rangle \\ &= \eta^2\|\mathcal{A}^\dagger\Big(\mathcal{A}(ZZ^\dagger) - y\Big)Z\|_F^2 + \|Z - U^\star R_Z\|_F^2 \\ &\quad - \eta\Big\langle \mathcal{A}^\dagger\Big(\mathcal{A}(ZZ^\dagger) - y\Big), ZZ^\dagger - U^\star U^{\star\dagger}\Big\rangle \\ &\quad - \eta\Big\langle \mathcal{A}^\dagger\Big(\mathcal{A}(ZZ^\dagger) - y\Big), (Z - U^\star R_Z)(Z - U^\star R_Z)^\dagger\Big\rangle \\ &\le \eta^2\|\mathcal{A}^\dagger\Big(\mathcal{A}(ZZ^\dagger) - y\Big)Z\|_F^2 + \|Z - U^\star R_Z\|_F^2 \\ &\quad - 1.0656\eta^2\Big\|\mathcal{A}^\dagger(\mathcal{A}(ZZ^\dagger) - y)Z\Big\|_F^2 - \eta\tfrac{1-\delta_{2r}}{2}\|U^\star U^{\star\dagger} - ZZ^\dagger\|_F^2 \\ &\quad + \eta\Bigg(\theta\sigma_r(\rho^\star)\cdot\|Z - U^\star R_Z\|_F^2 \\ &\quad + \tfrac{1}{200}\beta^2\cdot\widehat{\eta}\cdot\frac{\left(\frac{3}{2}+2|\mu|\right)^2}{\left(1-\left(\frac{3}{2}+2|\mu|\right)\frac{1}{10^3}\right)^2}\cdot\|\mathcal{A}^\dagger(\mathcal{A}(ZZ^\dagger) - y)\cdot Z\|_F^2\Bigg)\end{aligned}$$

Next, we use the following lemma:

Lemma A1 ([32] (Lemma 5.4))**.** *For any $W, V \in \mathbb{C}^{d\times r}$, the following holds:*

$$\|WW^\dagger - VV^\dagger\|_F^2 \ge 2(\sqrt{2}-1)\cdot\sigma_r(VV^\dagger)\cdot\min_{R\in\mathcal{O}}\|W - VR\|_F^2.$$

From Lemma A1, the quantity $\|U^\star U^{\star\dagger} - ZZ^\dagger\|_F^2$ satisfies:

$$\|U^\star U^{\star\dagger} - ZZ^\dagger\|_F^2 \ge 2(\sqrt{2}-1)\cdot\sigma_r(\rho^\star)\cdot\min_{R\in\mathcal{O}}\|Z - U^\star R\|_F^2 = 2(\sqrt{2}-1)\cdot\sigma_r(\rho^\star)\cdot\|Z - U^\star R_Z\|_F^2,$$

which, in our main recursion, results in:

$$
\begin{aligned}
\|U_+ - U^\star R_Z\|_F^2 &\le \eta^2 \|\mathcal{A}^\dagger\Big(\mathcal{A}(ZZ^\dagger) - y\Big)Z\|_F^2 + \|Z - U^\star R_Z\|_F^2 \\
&\quad - 1.0656\eta^2 \Big\|\mathcal{A}^\dagger(\mathcal{A}(ZZ^\dagger) - y)Z\Big\|_F^2 - \eta(\sqrt{2}-1)(1-\delta_{2r})\sigma_r(\rho^\star)\|Z - U^\star R_Z\|_F^2 \\
&\quad + \eta\Bigg(\theta\sigma_r(\rho^\star)\cdot\|Z - U^\star R_Z\|_F^2 \\
&\quad\quad + \tfrac{1}{200}\beta^2\cdot\widehat{\eta}\cdot\frac{\left(\frac{3}{2}+2|\mu|\right)^2}{\left(1-\left(\frac{3}{2}+2|\mu|\right)\frac{1}{10^3}\right)^2}\cdot\|\mathcal{A}^\dagger(\mathcal{A}(ZZ^\dagger)-y)\cdot Z\|_F^2\Bigg) \\
&\overset{(i)}{\le} \eta^2\|\mathcal{A}^\dagger\Big(\mathcal{A}(ZZ^\dagger) - y\Big)Z\|_F^2 + \|Z - U^\star R_Z\|_F^2 \\
&\quad - 1.0656\eta^2 \Big\|\mathcal{A}^\dagger(\mathcal{A}(ZZ^\dagger) - y)Z\Big\|_F^2 - \eta(\sqrt{2}-1)(1-\delta_{2r})\sigma_r(\rho^\star)\|Z - U^\star R_Z\|_F^2 \\
&\quad + \eta\Bigg(\theta\sigma_r(\rho^\star)\cdot\|Z - U^\star R_Z\|_F^2 \\
&\quad\quad + \tfrac{1}{200}\beta^2\cdot\tfrac{10}{9}\eta\cdot\frac{\left(\frac{3}{2}+2|\mu|\right)^2}{\left(1-\left(\frac{3}{2}+2|\mu|\right)\frac{1}{10^3}\right)^2}\cdot\|\mathcal{A}^\dagger(\mathcal{A}(ZZ^\dagger)-y)\cdot Z\|_F^2\Bigg) \\
&\overset{(ii)}{=} \left(1 + \tfrac{1}{200}\beta^2\cdot\tfrac{10}{9}\cdot\frac{\left(\frac{3}{2}+2|\mu|\right)^2}{\left(1-\left(\frac{3}{2}+2|\mu|\right)\frac{1}{10^3}\right)^2} - 1.0656\right)\eta^2\|\mathcal{A}^\dagger(\mathcal{A}(ZZ^\dagger)-y)\cdot Z\|_F^2 \\
&\quad + \Big(1 + \eta\theta\sigma_r(\rho^\star) - \eta(\sqrt{2}-1)(1-\delta_{2r})\sigma_r(\rho^\star)\Big)\|Z - U^\star R_Z\|_F^2
\end{aligned}
$$

where (i) is due to Lemma A6, and (ii) is due to the definition of U_+.

Under the assumptions that $\mu = \frac{\sigma_r(\rho^\star)^{1/2}}{10^3\sqrt{\kappa\tau(\rho^\star)}}\cdot\frac{\varepsilon}{4\cdot\sigma_1(\rho^\star)^{1/2}\cdot r}$, for $\varepsilon\in(0,1)$ user-defined, and $\delta_{2r}\le\frac{1}{10}$, the main constants in our proof so far simplify to:

$$
\beta = \frac{1+\left(\frac{3}{2}+2|\mu|\right)\cdot\frac{1}{10^3}}{1-\left(\frac{3}{2}+2|\mu|\right)\cdot\frac{1}{10^3}} = 1.003, \quad \text{and} \quad \beta^2 = 1.006,
$$

by Corollary A3. Thus:

$$
1 + \tfrac{1}{200}\beta^2\cdot\tfrac{10}{9}\cdot\frac{\left(\frac{3}{2}+2|\mu|\right)^2}{\left(1-\left(\frac{3}{2}+2|\mu|\right)\frac{1}{10^3}\right)^2} - 1.0656 \le -0.0516,
$$

and our recursion becomes:

$$
\begin{aligned}
\|U_+ - U^\star R_Z\|_F^2 &\le -0.0516\cdot\eta^2\cdot\|\mathcal{A}^\dagger(\mathcal{A}(ZZ^\dagger)-y)\cdot Z\|_F^2 \\
&\quad + \Big(1 + \eta\theta\sigma_r(\rho^\star) - \eta(\sqrt{2}-1)(1-\delta_{2r})\sigma_r(\rho^\star)\Big)\|Z - U^\star R_Z\|_F^2.
\end{aligned}
$$

Finally, we have

$$
\begin{aligned}
\theta &= \frac{(1-\delta_{2r})\left(1+\left(\frac{3}{2}+2|\mu|\right)\frac{1}{10^3}\right)^2}{10^3} + (1+\delta_{2r})\left(2+\left(\tfrac{3}{2}+2|\mu|\right)\cdot\tfrac{1}{10^3}\right)\left(\tfrac{3}{2}+2|\mu|\right)\cdot\tfrac{1}{10^3} \\
&\overset{(i)}{=} (1-\delta_{2r})\cdot\left(\frac{\left(1+(\frac{3}{2}+2|\mu|)\frac{1}{10^3}\right)^2}{10^3} + \kappa\left(2+\left(\tfrac{3}{2}+2|\mu|\right)\cdot\tfrac{1}{10^3}\right)\left(\tfrac{3}{2}+2|\mu|\right)\cdot\tfrac{1}{10^3}\right) \\
&\le 0.0047\cdot(1-\delta_{2r}).
\end{aligned}
$$

where (i) is by the definition of $\kappa := \frac{1+\delta_{2r}}{1-\delta_{2r}} \leq 1.223$ for $\delta_{2r} \leq \frac{1}{10}$, which is one of our assumptions as explained above. Combining the above in our main inequality, we obtain:

$$\begin{aligned}\|U_+ - U^\star R_Z\|_F^2 &\leq -0.0516 \cdot \eta^2 \cdot \|\mathcal{A}^\dagger(\mathcal{A}(ZZ^\dagger) - y) \cdot Z\|_F^2 \\ &\quad + \Big(1 + \eta\sigma_r(\rho^\star)(1-\delta_{2r}) \cdot (0.0047 - \sqrt{2} + 1)\Big)\|Z - U^\star R_Z\|_F^2 \\ &\leq \Big(1 - \tfrac{4\eta\sigma_r(\rho^\star)(1-\delta_{2r})}{10}\Big)\|Z - U^\star R_Z\|_F^2. \end{aligned} \tag{A3}$$

Taking square root on both sides, we obtain:

$$\|U_+ - U^\star R_Z\|_F \leq \sqrt{1 - \tfrac{4\eta\sigma_r(\rho^\star)(1-\delta_{2r})}{10}} \cdot \|Z - U^\star R_Z\|_F$$

Let us define $\xi = \sqrt{1 - \frac{4\eta\sigma_r(\rho^\star)(1-\delta_{2r})}{10}}$. Using the definitions $Z = U + \mu(U - U_-)$ and $R_Z \in \arg\min_{R\in\mathcal{O}} \|Z - U^\star R\|_F$, we get

$$\begin{aligned}\|U_+ - U^\star R_Z\|_F &\leq \xi \cdot \min_{R\in\mathcal{O}} \|Z - U^\star R\|_F = \xi \cdot \min_{R\in\mathcal{O}} \|U + \mu(U - U_-) - U^\star R\|_F \\ &= \xi \cdot \min_{R\in\mathcal{O}} \|U + \mu(U - U_-) - (1 - \mu + \mu)U^\star R\|_F \\ &\overset{(i)}{\leq} \xi \cdot |1+\mu| \cdot \min_{R\in\mathcal{O}} \|U - U^\star R\|_F + \xi \cdot |\mu| \cdot \min_{R\in\mathcal{O}} \|U_- - U^\star R\|_F + \xi \cdot |\mu| \cdot r\sigma_1(\rho^\star)^{1/2}\end{aligned}$$

where (i) follows from steps similar to those in Lemma A5. Further observe that

$$\min_{R\in\mathcal{O}} \|U_+ - U^\star R\|_F \leq \|U_+ - U^\star R_Z\|_F,$$

which leads to:

$$\begin{aligned}&\min_{R\in\mathcal{O}} \|U_+ - U^\star R\|_F \\ &\quad \leq \xi \cdot |1+\mu| \cdot \min_{R\in\mathcal{O}} \|U - U^\star R\|_F + \xi \cdot |\mu| \cdot \min_{R\in\mathcal{O}} \|U_- - U^\star R\|_F + \xi \cdot |\mu| \cdot r\sigma_1(\rho^\star)^{1/2}.\end{aligned} \tag{A4}$$

Including two subsequent iterations in a single two-dimensional first-order system, we get the following characterization:

$$\begin{aligned}\begin{bmatrix}\min_{R\in\mathcal{O}} \|U_{k+1} - U^\star R\|_F \\ \min_{R\in\mathcal{O}} \|U_k - U^\star R\|_F\end{bmatrix} &\leq \begin{bmatrix}\xi \cdot |1+\mu| & \xi \cdot |\mu| \\ 1 & 0\end{bmatrix} \cdot \begin{bmatrix}\min_{R\in\mathcal{O}} \|U_k - U^\star R\|_F \\ \min_{R\in\mathcal{O}} \|U_{k-1} - U^\star R\|_F\end{bmatrix} \\ &\quad + \begin{bmatrix}1 \\ 0\end{bmatrix} \cdot \xi \cdot |\mu| \cdot \sigma_1(\rho^\star)^{1/2} \cdot r.\end{aligned}$$

Now, let $x_j := \min_{R\in\mathcal{O}} \|U_j - U^\star R\|_F$. Then, we can write the above relation as

$$\begin{bmatrix}x_{k+1} \\ x_k\end{bmatrix} \leq \underbrace{\begin{bmatrix}\xi \cdot |1+\mu| & \xi \cdot |\mu| \\ 1 & 0\end{bmatrix}}_{:=A} \cdot \begin{bmatrix}x_k \\ x_{k-1}\end{bmatrix} + \begin{bmatrix}1 \\ 0\end{bmatrix} \cdot \xi \cdot |\mu| \cdot \sigma_1(\rho^\star)^{1/2} \cdot r,$$

where we denote the "contraction matrix" by A. Observe that A has non-negative values. Unfolding the above recursion for k iterations, we obtain:

$$\begin{bmatrix}x_{k+1} \\ x_k\end{bmatrix} \leq A^{k+1} \cdot \begin{bmatrix}x_0 \\ x_{-1}\end{bmatrix} + \left(\sum_{t=0}^{k} A^t\right) \cdot \begin{bmatrix}1 \\ 0\end{bmatrix} \cdot \xi \cdot |\mu| \cdot \sigma_1(\rho^\star)^{1/2} \cdot r \tag{A5}$$

Taking norms on both sides, we get

$$\begin{aligned}\left\|\begin{bmatrix}x_{k+1}\\ x_k\end{bmatrix}\right\| &\leq \left\|A^{k+1}\cdot\begin{bmatrix}x_0\\ x_{-1}\end{bmatrix}+\left(\sum_{t=0}^{k}A^t\right)\cdot\begin{bmatrix}1\\ 0\end{bmatrix}\cdot\xi\cdot|\mu|\cdot\sigma_1(\rho^\star)^{1/2}\cdot r\right\| \\ &\overset{(i)}{\leq}\left\|A^{k+1}\cdot\begin{bmatrix}x_0\\ x_{-1}\end{bmatrix}\right\|+\left\|\left(\sum_{t=0}^{k}A^t\right)\cdot\begin{bmatrix}1\\ 0\end{bmatrix}\cdot\xi\cdot|\mu|\cdot\sigma_1(\rho^\star)^{1/2}\cdot r\right\| \\ &\overset{(ii)}{\leq}\left\|A^{k+1}\right\|\cdot\left\|\begin{bmatrix}x_0\\ x_{-1}\end{bmatrix}\right\|+\left\|\left(\sum_{t=0}^{k}A^t\right)\cdot\begin{bmatrix}1\\ 0\end{bmatrix}\cdot\xi\cdot|\mu|\cdot\sigma_1(\rho^\star)^{1/2}\cdot r\right\|, \end{aligned} \tag{A6}$$

where (i) is by triangle inequality, and (ii) is by submultiplicativity of matrix norms.

To bound Equation (A6), we will use spectral analysis. We first recall the definition of the spectral radius of a matrix M:

$$\rho(M) := \max\{|\lambda| : \lambda \in \mathrm{Sp}(M)\},$$

where $\mathrm{Sp}(M)$ is the set of all eigenvalues of M. We then use the following lemma:

Lemma A2 (Lemma 11 in [104]). *Given a matrix M and $\epsilon > 0$, there exists a matrix norm $\|\cdot\|$ such that*

$$\|M\| \leq \rho(M) + \epsilon.$$

We further use the Gelfand's formula:

Lemma A3 (Theorem 12 in [104]). *Given any matrix norm $\|\cdot\|$, the following holds:*

$$\rho(M) = \lim_{k\to\infty}\|M^k\|^{1/k}$$

The proofs for the above lemmata can be found in [104]. Using Lemmas A2 and A3, we have that for any $\epsilon > 0$, there exists $K_\epsilon \in \mathbb{N}$ such that

$$\|A^k\|^{1/k} \leq (\rho(A)+\epsilon) \quad \text{for all} \quad k.$$

Further, let $C_\epsilon := \max_{k<K_\epsilon}\max\left\{1, \frac{\|A^k\|}{(\rho(A)+\epsilon)^k}\right\}$. Then, we have

$$\|A^k\| \leq C_\epsilon(\rho(A)+\epsilon)^k \quad \text{for all} \quad k \geq K_\epsilon. \tag{A7}$$

Hence, using Equation (A7) in Equation (A6), we have

$$\begin{aligned}\left\|\begin{bmatrix}x_{k+1}\\ x_k\end{bmatrix}\right\| &\leq \left\|A^{k+1}\right\|\cdot\left\|\begin{bmatrix}x_0\\ x_{-1}\end{bmatrix}\right\|+\left\|\left(\sum_{t=0}^{k}A^t\right)\cdot\begin{bmatrix}1\\ 0\end{bmatrix}\cdot\xi\cdot|\mu|\cdot\sigma_1(\rho^\star)^{1/2}\cdot r\right\| \\ &\leq C_\epsilon(\rho(A)+\epsilon)^{k+1}\left\|\begin{bmatrix}x_0\\ x_{-1}\end{bmatrix}\right\| \\ &\quad+\left\|\begin{bmatrix}1\\ 0\end{bmatrix}\cdot\xi\cdot|\mu|\cdot\sigma_1(\rho^\star)^{1/2}\cdot r\cdot C_\epsilon\cdot\sum_{t=0}^{k}(\rho(A)+\epsilon)^t\right\|. \end{aligned} \tag{A8}$$

Therefore, asymptotically, the convergence rate is $O(\rho(A)^{k+1})$, where $\rho(A)$ is the spectral radius of the contraction matrix A. We thus compute $\rho(A)$ below. Since A is a 2×2 matrix, its eigenvalues are:

$$\lambda_{1,2} = \frac{\xi \cdot |1+\mu|}{2} \pm \sqrt{\frac{\xi^2(1+\mu)^2}{4} + \xi \cdot |\mu|}$$
$$\overset{(i)}{\Longrightarrow} \rho(A) := \max\{\lambda_1, \lambda_2\} = \lambda_1 = \frac{\xi \cdot |1+\mu|}{2} + \sqrt{\frac{\xi^2(1+\mu)^2}{4} + \xi \cdot |\mu|},$$

where (i) follows since every term in $\lambda_{1,2}$ is positive.

To show an accelerated convergence rate, we want the above eigenvalue (which determines the convergence rate) to be bounded by $1 - \sqrt{\frac{1-\delta_{2r}}{1+\delta_{2r}}}$ (This is akin to the notion of acceleration and optimal method (for certain function classes) from convex optimization literature, where the condition number appears inside the $\sqrt{\cdot}$ term. For details, see [39,44]). To show this, first note that this term is bounded above as follows:

$$\begin{aligned}\lambda_1 = \frac{\xi \cdot |1+\mu|}{2} + \sqrt{\frac{\xi^2(1+\mu)^2}{4} + \xi \cdot |\mu|} &\overset{(i)}{\leq} \xi + \sqrt{\xi^2 + \xi} \\ &\overset{(ii)}{\leq} \xi + \sqrt{2\xi} \\ &\overset{(ii)}{\leq} (\sqrt{2}+1)\sqrt{\xi},\end{aligned}$$

where (i) is by the conventional bound on momentum: $0 < \mu < 1$, and (ii) is by the relation $\xi^2 \leq \xi \leq \sqrt{\xi}$ for $0 \leq \xi \leq 1$. Therefore, to show an accelerated rate of convergence, we want the following relation to hold:

$$(\sqrt{2}+1)\sqrt{\xi} \leq 1 - \sqrt{\tfrac{1-\delta_{2r}}{1+\delta_{2r}}} \iff \sqrt{\xi} \leq \frac{\sqrt{1+\delta_{2r}} - \sqrt{1-\delta_{2r}}}{(\sqrt{2}+1)\sqrt{1+\delta_{2r}}}. \tag{A9}$$

Recalling our definition of $\xi = \sqrt{1 - \frac{4\eta\sigma_r(\rho^\star)(1-\delta_{2r})}{10}}$, the problem boils down to choosing the right step size η such that the above inequality on ξ in Equation (A9) is satisfied. With simple algebra, we can show the following lower bound on η:

$$\left[1 - \left(\frac{\sqrt{1+\delta_{2r}} - \sqrt{1-\delta_{2r}}}{(\sqrt{2}+1)\sqrt{1+\delta_{2r}}}\right)^4\right] \cdot \frac{10}{4\sigma_r(\rho^\star)(1-\delta_{2r})} \leq \eta$$

Finally, the argument inside the $\sqrt{\cdot}$ term of $\xi = \sqrt{1 - \frac{4\eta\sigma_r(\rho^\star)(1-\delta_{2r})}{10}} > 0$ has to be non-negative, yielding the following upper bound on η:

$$\eta \leq \frac{10}{4\sigma_r(\rho^\star)(1-\delta_{2r})}.$$

Combining two inequalities, and noting that the term $\left[1 - \left(\frac{\sqrt{1+\delta_{2r}} - \sqrt{1-\delta_{2r}}}{(\sqrt{2}+1)\sqrt{1+\delta_{2r}}}\right)^4\right]$ is bounded above by 1, we arrive at the following bound on η:

$$\left[1 - \left(\frac{\sqrt{1+\delta_{2r}} - \sqrt{1-\delta_{2r}}}{(\sqrt{2}+1)\sqrt{1+\delta_{2r}}}\right)^4\right] \cdot \frac{10}{4\sigma_r(\rho^\star)(1-\delta_{2r})} \leq \eta \leq \frac{10}{4\sigma_r(\rho^\star)(1-\delta_{2r})}. \tag{A10}$$

In sum, for the specific η satisfying Equation (A10), we have shown that

$$\rho(A) = \lambda_1 = \frac{\xi \cdot |1+\mu|}{2} + \sqrt{\frac{\xi^2(1+\mu)^2}{4} + \xi \cdot |\mu|} \leq 1 - \sqrt{\frac{1-\delta_{2r}}{1+\delta_{2r}}}$$

Above bound translates Equation (A8) into:

$$\begin{aligned}\left\|\begin{bmatrix}x_{k+1}\\ x_k\end{bmatrix}\right\| &\leq C_\epsilon(\rho(A)+\epsilon)^{k+1}\left\|\begin{bmatrix}x_0\\ x_{-1}\end{bmatrix}\right\| + \left\|\begin{bmatrix}1\\0\end{bmatrix}\cdot\xi\cdot|\mu|\cdot\sigma_1(\rho^\star)^{1/2}\cdot r\cdot C_\epsilon\cdot\sum_{t=0}^{k}(\rho(A)+\epsilon)^t\right\| \\ &\leq C_\epsilon\left(1-\sqrt{\tfrac{1-\delta_{2r}}{1+\delta_{2r}}}+\epsilon\right)^{k+1}\left\|\begin{bmatrix}x_0\\ x_{-1}\end{bmatrix}\right\| + \left\|\begin{bmatrix}1\\0\end{bmatrix}\cdot\xi\cdot|\mu|\cdot\sigma_1(\rho^\star)^{1/2}\cdot r\cdot C_\epsilon\cdot\sum_{t=0}^{k}(\rho(A)+\epsilon)^t\right\|\end{aligned}$$

Without loss of generality, we can take the Euclidean norm from the above (By the equivalence of norms [105], C_ϵ can absorb additional constants), which yields:

$$\begin{aligned}\left\|\begin{bmatrix}x_{k+1}\\ x_k\end{bmatrix}\right\|_2 &\leq C_\epsilon\left(1-\sqrt{\tfrac{1-\delta_{2r}}{1+\delta_{2r}}}+\epsilon\right)^{k+1}\left\|\begin{bmatrix}x_0\\ x_{-1}\end{bmatrix}\right\|_2 + \left\|\begin{bmatrix}1\\0\end{bmatrix}\cdot\xi\cdot|\mu|\cdot\sigma_1(\rho^\star)^{1/2}\cdot r\cdot C_\epsilon\cdot\sum_{t=0}^{k}(\rho(A)+\epsilon)^t\right\|_2 \\ &= C_\epsilon\left(1-\sqrt{\tfrac{1-\delta_{2r}}{1+\delta_{2r}}}+\epsilon\right)^{k+1}\sqrt{x_0^2+x_{-1}^2} + \xi\cdot|\mu|\cdot\sigma_1(\rho^\star)^{1/2}\cdot r\cdot C_\epsilon\cdot\sum_{t=0}^{k}(\rho(A)+\epsilon)^t \\ &= C_\epsilon\left(1-\sqrt{\tfrac{1-\delta_{2r}}{1+\delta_{2r}}}+\epsilon\right)^{k+1}\sqrt{x_0^2+x_{-1}^2} + \xi\cdot|\mu|\cdot\sigma_1(\rho^\star)^{1/2}\cdot r\cdot C_\epsilon\cdot\tfrac{1-(\rho(A)+\epsilon)^{k+1}}{1-(\rho(A)+\epsilon)}.\end{aligned}$$

Re-substituting $x_j = \min_{R\in\mathbb{O}}\|U_j - U^\star R\|_F$, and using the same initialization for U_0 and U_{-1}, we get:

$$\begin{aligned}&\left(\min_{R\in\mathcal{O}}\|U_{k+1}-U^\star R\|_F^2 + \min_{R\in\mathcal{O}}\|U_k - U^\star R\|_F^2\right)^{1/2} \\ &\quad\leq C_\epsilon\left(1-\sqrt{\tfrac{1-\delta_{2r}}{1+\delta_{2r}}}+\epsilon\right)^{k+1}\left(2\min_{R\in\mathcal{O}}\|U_0-U^\star R\|_F^2\right)^{1/2} + \xi\cdot|\mu|\cdot\sigma_1(\rho^\star)^{1/2}\cdot r\cdot C_\epsilon\cdot\tfrac{1-(\rho(A)+\epsilon)^{k+1}}{1-(\rho(A)+\epsilon)} \\ &\quad= C_\epsilon\left(1-\sqrt{\tfrac{1-\delta_{2r}}{1+\delta_{2r}}}+\epsilon\right)^{k+1}\min_{R\in\mathcal{O}}\|U_0-U^\star R\|_F + \xi\cdot|\mu|\cdot\sigma_1(\rho^\star)^{1/2}\cdot r\cdot\tilde{C}_\epsilon \\ &\quad\approx C_\epsilon\left(1-\sqrt{\tfrac{1-\delta_{2r}}{1+\delta_{2r}}}+\epsilon\right)^{k+1}\min_{R\in\mathcal{O}}\|U_0-U^\star R\|_F + O(\mu),\end{aligned}$$

where in the equality, with slight abuse of notation, we absorbed $\sqrt{2}$ factor to C_ϵ, and similarly absorbed the last term such that $\tilde{C}_\epsilon = C_\epsilon\cdot\frac{1-(\rho(A)+\epsilon)^{k+1}}{1-(\rho(A)+\epsilon)}$. This concludes the proof for Theorem 1.

Supporting Lemmata

In this subsection, we present a series of lemmata, used for the proof of Theorem 1.

Lemma A4. *Let $U\in\mathbb{C}^{d\times r}$ and $U^\star\in\mathbb{C}^{d\times r}$, such that $\|U-U^\star R\|_F \leq \frac{\sigma_r(\rho^\star)^{1/2}}{10^3\sqrt{\kappa\tau(\rho^\star)}}$ for some $R\in\mathcal{O}$, where $\rho^\star = U^\star U^{\star\dagger}$, $\kappa := \frac{1+\delta_{2r}}{1-\delta_{2r}} > 1$, for $\delta_{2r}\leq\frac{1}{10}$, and $\tau(\rho^\star) := \frac{\sigma_1(\rho^\star)}{\sigma_r(\rho^\star)} > 1$. Then:*

$$\begin{aligned}\sigma_1(\rho^\star)^{1/2}\left(1-\tfrac{1}{10^3}\right) \leq \sigma_1(U) \leq \sigma_1(\rho^\star)^{1/2}\left(1+\tfrac{1}{10^3}\right) \\ \sigma_r(\rho^\star)^{1/2}\left(1-\tfrac{1}{10^3}\right) \leq \sigma_r(U) \leq \sigma_r(\rho^\star)^{1/2}\left(1+\tfrac{1}{10^3}\right)\end{aligned}$$

Proof. By the fact $\|\cdot\|_2\leq\|\cdot\|_F$ and using Weyl's inequality for perturbation of singular values [106] (Theorem 3.3.16), we have:

$$|\sigma_i(U)-\sigma_i(U^\star)| \leq \frac{\sigma_r(\rho^\star)^{1/2}}{10^3\sqrt{\kappa\tau(\rho^\star)}} \leq \frac{\sigma_r(\rho^\star)^{1/2}}{10^3}, \quad 1\leq i\leq r.$$

Then,

$$\begin{aligned}
-\tfrac{\sigma_r(\rho^\star)^{1/2}}{10^3} \leq \sigma_1(U) - \sigma_1(U^\star) \leq \tfrac{\sigma_r(\rho^\star)^{1/2}}{10^3} \Rightarrow \\
\sigma_1(\rho^\star)^{1/2} - \tfrac{\sigma_r(\rho^\star)^{1/2}}{10^3} \leq \sigma_1(U) \leq \sigma_1(\rho^\star)^{1/2} + \tfrac{\sigma_r(\rho^\star)^{1/2}}{10^3} \Rightarrow \\
\sigma_1(\rho^\star)^{1/2}\Big(1 - \tfrac{1}{10^3}\Big) \leq \sigma_1(U) \leq \sigma_1(\rho^\star)^{1/2}\Big(1 + \tfrac{1}{10^3}\Big).
\end{aligned}$$

Similarly:

$$\begin{aligned}
-\tfrac{\sigma_r(\rho^\star)^{1/2}}{10^3} \leq \sigma_r(U) - \sigma_r(U^\star) \leq \tfrac{\sigma_r(\rho^\star)^{1/2}}{10^3} \Rightarrow \\
\sigma_r(\rho^\star)^{1/2} - \tfrac{\sigma_r(\rho^\star)^{1/2}}{10^3} \leq \sigma_r(U) \leq \sigma_r(\rho^\star)^{1/2} + \tfrac{\sigma_r(\rho^\star)^{1/2}}{10^3} \Rightarrow \\
\sigma_r(\rho^\star)^{1/2}\Big(1 - \tfrac{1}{10^3}\Big) \leq \sigma_r(U) \leq \sigma_r(\rho^\star)^{1/2}\Big(1 + \tfrac{1}{10^3}\Big).
\end{aligned}$$

In the above, we used the fact that $\sigma_i(U^\star) = \sigma_i(\rho^\star)^{1/2}$, for all i, and the fact that $\sigma_i(\rho^\star)^{1/2} \geq \sigma_j(\rho^\star)^{1/2}$, for $i \leq j$. □

Lemma A5. *Let $U \in \mathbb{C}^{d\times r}, U_- \in \mathbb{C}^{d\times r}$, and $U^\star \in \mathbb{C}^{d\times r}$, such that $\min_{R\in\mathcal{O}} \|U - U^\star R\|_F \leq \frac{\sigma_r(\rho^\star)^{1/2}}{10^3\sqrt{\kappa\tau(\rho^\star)}}$ and $\min_{R\in\mathcal{O}} \|U_- - U^\star R\|_F \leq \frac{\sigma_r(\rho^\star)^{1/2}}{10^3\sqrt{\kappa\tau(\rho^\star)}}$, where $\rho^\star = U^\star U^{\star\dagger}$, and $\kappa := \frac{1+\delta_{2r}}{1-\delta_{2r}} > 1$, for $\delta_{2r} \leq \frac{1}{10}$, and $\tau(\rho^\star) := \frac{\sigma_1(\rho^\star)}{\sigma_r(\rho^\star)} > 1$. Set the momentum parameter as $\mu = \frac{\sigma_r(\rho^\star)^{1/2}}{10^3\sqrt{\kappa\tau(\rho^\star)}} \cdot \frac{\varepsilon}{4\cdot\sigma_1(\rho^\star)^{1/2}\cdot r}$, for $\varepsilon \in (0,1)$ user-defined. Then,*

$$\|Z - U^\star R_Z\|_F \leq \big(\tfrac{3}{2} + 2|\mu|\big) \cdot \tfrac{\sigma_r(\rho^\star)^{1/2}}{10^3\sqrt{\kappa\tau(\rho^\star)}}.$$

Proof. Let $R_U \in \arg\min_{R\in\mathcal{O}} \|U - U^\star\|_F$ and $R_{U_-} \in \arg\min_{R\in\mathcal{O}} \|U_- - U^\star R\|_F$. By the definition of the distance function:

$$\begin{aligned}
\|Z - U^\star R_Z\|_F &= \min_{R\in\mathcal{O}} \|Z - U^\star R\|_F = \min_{R\in\mathcal{O}} \|U + \mu(U - U_-) - U^\star R\|_F \\
&= \min_{R\in\mathcal{O}} \|U + \mu(U - U_-) - (1 - \mu + \mu)U^\star R\|_F \\
&\leq |1+\mu| \cdot \|U - U^* R_U\|_F + |\mu| \cdot ||U_- - U^* R_{U_-}||_F \\
&= |1+\mu| \cdot \|U - U^* R_U\|_F + |\mu| \cdot \|U_- - U^* R_U - U^* R_{U_-} + U^* R_{U_-}\|_F \\
&= |1+\mu| \cdot \|U - U^* R_U\|_F + |\mu| \cdot \|(U_- - U^* R_{U_-}) + U^*(R_{U_-} - R_U)\|_F \\
&\leq |1+\mu| \cdot \min_{R\in\mathcal{O}} \|U - U^\star R\|_F + |\mu| \cdot \min_{R\in\mathcal{O}} \|U_- - U^\star R\|_F \\
&\quad + |\mu| \cdot \|U^\star(R_U - R_{U_-})\|_F \\
&\leq |1+\mu| \cdot \min_{R\in\mathcal{O}} \|U - U^\star R\|_F + |\mu| \cdot \min_{R\in\mathcal{O}} \|U_- - U^\star R\|_F + 2|\mu| \cdot \sigma_1(\rho^\star)^{1/2} r \\
&\overset{(i)}{\leq} \big(\tfrac{3}{2} + 2|\mu|\big) \cdot \tfrac{\sigma_r(\rho^\star)^{1/2}}{10^3\sqrt{\kappa\tau(\rho^\star)}}
\end{aligned}$$

where (i) is due to the fact that $\mu \leq \frac{\sigma_r(\rho^\star)^{1/2}}{10^3\sqrt{\kappa\tau(\rho^\star)}} \cdot \frac{1}{4\cdot\sigma_1(\rho^\star)^{1/2}\cdot r}$. We keep μ in the expression, but we use it for clarity for the rest of the proof. □

Corollary A1. *Let $Z \in \mathbb{C}^{d\times r}$ and $U^\star \in \mathbb{C}^{d\times r}$, such that $\|Z - U^\star R\|_F \leq \big(\frac{3}{2} + 2|\mu|\big) \cdot \frac{\sigma_r(\rho^\star)^{1/2}}{10^3\sqrt{\kappa\tau(\rho^\star)}}$ for some $R \in \mathcal{O}$, and $\rho^\star = U^\star U^{\star\dagger}$. Then:*

$$\begin{aligned}
\sigma_1(\rho^\star)^{1/2}\Big(1 - \big(\tfrac{3}{2} + 2|\mu|\big)\tfrac{1}{10^3}\Big) \leq \sigma_1(Z) \leq \sigma_1(\rho^\star)^{1/2}\Big(1 + \big(\tfrac{3}{2} + 2|\mu|\big)\tfrac{1}{10^3}\Big) \\
\sigma_r(\rho^\star)^{1/2}\Big(1 - \big(\tfrac{3}{2} + 2|\mu|\big)\tfrac{1}{10^3}\Big) \leq \sigma_r(Z) \leq \sigma_r(\rho^\star)^{1/2}\Big(1 + \big(\tfrac{3}{2} + 2|\mu|\big)\tfrac{1}{10^3}\Big).
\end{aligned}$$

Given that $\mu = \frac{\sigma_r(\rho^\star)^{1/2}}{10^3\sqrt{\kappa\tau(\rho^\star)}} \cdot \frac{\varepsilon}{4\cdot\sigma_1(\rho^\star)^{1/2}\cdot r} \leq \frac{1}{10^3}$ *, we get:*

$$0.998 \cdot \sigma_1(\rho^\star)^{1/2} \leq \sigma_1(Z) \leq 1.0015 \cdot \sigma_1(\rho^\star)^{1/2}$$
$$0.998 \cdot \sigma_r(\rho^\star)^{1/2} \leq \sigma_r(Z) \leq 1.0015 \cdot \sigma_r(\rho^\star)^{1/2}.$$

Proof. The proof follows similar motions as in Lemma A4. □

Corollary A2. *Under the same assumptions of Lemma A4 and Corollary A1, and given the assumptions on* μ*, we have:*

$$\begin{aligned} \tfrac{99}{100} \cdot \|\rho^\star\|_2 \leq \|ZZ^\dagger\|_2 \leq \tfrac{101}{100} \cdot \|\rho^\star\|_2 \\ \tfrac{99}{100} \cdot \|\rho^\star\|_2 \leq \|Z_0Z_0^\dagger\|_2 \leq \tfrac{101}{100} \cdot \|\rho^\star\|_2 \quad \textit{and} \\ \tfrac{99}{101} \cdot \|Z_0Z_0^\dagger\|_2 \leq \|ZZ^\dagger\|_2 \leq \tfrac{101}{99} \cdot \|Z_0Z_0^\dagger\|_2 \end{aligned}$$

Proof. The proof is easily derived based on the quantities from Lemma A4 and Corollary A1. □

Corollary A3. *Let* $Z \in \mathbb{C}^{d\times r}$ *and* $U^\star \in \mathbb{C}^{d\times r}$*, such that* $\|Z - U^\star R\|_F \leq \left(\frac{3}{2} + 2|\mu|\right) \cdot \frac{\sigma_r(\rho^\star)^{1/2}}{10^3\sqrt{\kappa\tau(\rho^\star)}}$ *for some* $R \in \mathcal{O}$*, and* $\rho^\star = U^\star U^{\star\dagger}$*. Define* $\tau(W) = \frac{\sigma_1(W)}{\sigma_r(W)}$*. Then:*

$$\tau(ZZ^\dagger) \leq \beta^2\tau(\rho^\star), \tag{A11}$$

where $\beta := \frac{1+\left(\frac{3}{2}+2|\mu|\right)\cdot\frac{1}{10^3}}{1-\left(\frac{3}{2}+2|\mu|\right)\cdot\frac{1}{10^3}} > 1$*. for* $\mu \leq \frac{\sigma_r(\rho^\star)^{1/2}}{10^3\sqrt{\kappa\tau(\rho^\star)}} \cdot \frac{1}{4\cdot\sigma_1(\rho^\star)^{1/2}\cdot r}$*.*

Proof. The proof uses the definition of the condition number $\tau(\cdot)$ and the results from Lemma A4 and and Corollary A1. □

Lemma A6. *Consider the following three step sizes:*

$$\eta = \frac{1}{4\big((1+\delta_{2r})\|Z_0Z_0^\dagger\|_2 + \|\mathcal{A}^\dagger\big(\mathcal{A}(Z_0Z_0^\dagger) - y\big)\|_2\big)}$$
$$\widehat{\eta} = \frac{1}{4\big((1+\delta_{2r})\|ZZ^\dagger\|_2 + \|\mathcal{A}^\dagger(\mathcal{A}(ZZ^\dagger) - y)Q_ZQ_Z^\dagger\|_2\big)}$$
$$\eta^\star = \frac{1}{4\big((1+\delta_{2r})\|\rho^{\star\top}\|_2 + \|\mathcal{A}^\dagger(\mathcal{A}(\rho^\star) - y)\|_2\big)}.$$

Here, $Z_0 \in \mathbb{C}^{d\times r}$ *is the initial point,* $Z \in \mathbb{C}^{d\times r}$ *is the current point,* $\rho^\star \in \mathbb{C}^{d\times d}$ *is the optimal solution, and* Q_Z *denotes a basis of the column space of* Z*. Then, under the assumptions that* $\min_{R\in\mathcal{O}} \|U - U^\star R\|_F \leq \frac{\sigma_r(\rho^\star)^{1/2}}{10^3\sqrt{\kappa\tau(\rho^\star)}}$*, and* $\min_{R\in\mathcal{O}} \|Z - U^\star R\|_F \leq \left(\frac{3}{2} + 2|\mu|\right) \cdot \frac{\sigma_r(\rho^\star)^{1/2}}{10^3\sqrt{\kappa\tau(\rho^\star)}}$*, and assuming* $\mu = \frac{\sigma_r(\rho^\star)^{1/2}}{10^3\sqrt{\kappa\tau(\rho^\star)}} \cdot \frac{\varepsilon}{4\cdot\sigma_1(\rho^\star)^{1/2}\cdot r}$*, for the user-defined parameter* $\varepsilon \in (0,1)$*, we have:*

$$\tfrac{10}{9}\eta \geq \widehat{\eta} \geq \tfrac{10}{10.5}\eta, \quad \textit{and} \quad \tfrac{100}{102}\eta^\star \leq \eta \leq \tfrac{102}{100}\eta^\star$$

Proof. The assumptions of the lemma are identical to that of Corollary A2. Thus, we have: $\frac{99}{100} \cdot \|U^\star\|_2^2 \leq \|Z\|_2^2 \leq \frac{101}{100} \cdot \|U^\star\|_2^2$, $\frac{99}{100} \cdot \|U^\star\|_2^2 \leq \|Z_0\|_2^2 \leq \frac{101}{100} \cdot \|U^\star\|_2^2$, and $\frac{99}{101} \cdot \|Z_0\|_2^2 \leq \|Z\|_2^2 \leq \frac{101}{99} \cdot \|Z_0\|_2^2$. We focus on the inequality $\widehat{\eta} \geq \frac{10}{10.5}\eta$. Observe that:

$$\begin{aligned}\left\|\mathcal{A}^\dagger\left(\mathcal{A}(ZZ^\dagger)-y\right)Q_ZQ_Z^\dagger\right\|_2 &\le \left\|\mathcal{A}^\dagger\left(\mathcal{A}(ZZ^\dagger)-y\right)\right\|_2\\ &= \left\|\mathcal{A}^\dagger\left(\mathcal{A}(ZZ^\dagger)-y\right)-\mathcal{A}^\dagger\left(\mathcal{A}(Z_0Z_0^\dagger)-y\right)+\mathcal{A}^\dagger\left(\mathcal{A}(Z_0Z_0^\dagger)-y\right)\right\|_2\\ &\overset{(i)}{\le} (1+\delta_{2r})\left\|ZZ^\dagger-Z_0Z_0^\dagger\right\|_F+\left\|\mathcal{A}^\dagger\left(\mathcal{A}(Z_0Z_0^\dagger)-y\right)\right\|_2\\ &\le (1+\delta_{2r})\left\|ZZ^\dagger-U^\star U^{\star\dagger}\right\|_F+(1+\delta_{2r})\left\|Z_0Z_0^\dagger-U^\star U^{\star\dagger}\right\|_F\\ &\quad+\left\|\mathcal{A}^\dagger\left(\mathcal{A}(Z_0Z_0^\dagger)-y\right)\right\|_2\end{aligned}$$

where (i) is due to smoothness via RIP constants of the objective and the fact $\|\cdot\|_2 \le \|\cdot\|_F$. For the first two terms on the right-hand side, where R_Z is the minimizing rotation matrix for Z, we obtain:

$$\begin{aligned}\|ZZ^\dagger-U^\star U^{\star\dagger}\|_F &= \|ZZ^\dagger-U^\star R_Z Z^\dagger+U^\star R_Z Z^\dagger-U^\star U^{\star\dagger}\|_F\\ &= \|(Z-U^\star R_Z)Z^\dagger+U^\star R_Z(Z-U^\star R_Z)^\dagger\|_F\\ &\le \|Z\|_2\cdot\|Z-U^\star R_Z\|_F+\|U^\star\|_2\cdot\|Z-U^\star R_Z\|_F\\ &\le (\|Z\|_2+\|U^\star\|_2)\cdot\|Z-U^\star R_Z\|_F\\ &\overset{(i)}{\le} \left(\sqrt{\tfrac{101}{99}}+\sqrt{\tfrac{100}{99}}\right)\|Z_0\|_2\cdot\|Z-U^\star R_Z\|_F\\ &\overset{(ii)}{\le} \left(\sqrt{\tfrac{101}{99}}+\sqrt{\tfrac{100}{99}}\right)\|Z_0\|_2\cdot 0.001\sigma_r(\rho^\star)^{1/2}\\ &\le \left(\sqrt{\tfrac{101}{99}}+\sqrt{\tfrac{100}{99}}\right)\cdot 0.001\cdot\sqrt{\tfrac{100}{99}}\cdot\|Z_0\|_2^2\end{aligned}$$

where (i) is due to the relation of $\|Z\|_2$ and $\|U^\star\|_2$ derived above, (ii) is due to Lemma A5. Similarly:

$$\|Z_0Z_0^\dagger-U^\star U^{\star\dagger}\|_F \le \left(\sqrt{\tfrac{101}{99}}+\sqrt{\tfrac{100}{99}}\right)\cdot 0.001\cdot\sqrt{\tfrac{100}{99}}\cdot\|Z_0\|_2^2$$

Using these above, we obtain:

$$\left\|\mathcal{A}^\dagger\left(\mathcal{A}(ZZ^\dagger)-y\right)Q_ZQ_Z^\dagger\right\|_2 \le \tfrac{4.1(1+\delta_{2r})}{10^3}\|Z_0Z_0^\dagger\|_2+\left\|\mathcal{A}^\dagger\left(\mathcal{A}(Z_0Z_0^\dagger)-y\right)\right\|_2$$

Thus:

$$\begin{aligned}\widehat{\eta} &= \frac{1}{4\left((1+\delta_{2r})\|ZZ^\dagger\|_2+\|\mathcal{A}^\dagger(\mathcal{A}(ZZ^\dagger)-y)Q_ZQ_Z^\dagger\|_2\right)}\\ &\ge \frac{1}{4\left((1+\delta_{2r})\frac{101}{99}\|Z_0Z_0\|_2+\right)+\frac{4.1(1+\delta_{2r})}{10^3}\|Z_0Z_0^\dagger\|_2+\left\|\mathcal{A}^\dagger\left(\mathcal{A}(Z_0Z_0^\dagger)-y\right)\right\|_2}\\ &\ge \frac{1}{4\left(\frac{10.5}{10}\cdot(1+\delta_{2r})\|Z_0Z_0^\dagger\|_2+\|\mathcal{A}^\dagger\left(\mathcal{A}(Z_0Z_0^\dagger)-y\right)\|_2\right)}\\ &\ge \tfrac{10}{10.5}\eta\end{aligned}$$

Similarly, one gets $\widehat{\eta} \le \frac{10}{9}\eta$.

For the relation between η and $\eta^\star$, we will prove here the lower bound; similar motions lead to the upper bound also. By definition, and using the relations in Corollary A2, we get:

$$\begin{aligned}\eta &= \frac{1}{4\big((1+\delta_{2r})\|Z_0Z_0^\dagger\|_2 + \|\mathcal{A}^\dagger\big(\mathcal{A}(Z_0Z_0^\dagger)-y\big)\|_2\big)}\\ &\geq \frac{1}{4\Big((1+\delta_{2r})\frac{101}{100}\|\rho^{\star\top}\|_2 + \|\mathcal{A}^\dagger\big(\mathcal{A}(Z_0Z_0^\dagger)-y\big)\|_2\Big)}\end{aligned}$$

For the gradient term, we observe:

$$\begin{aligned}\Big\|\mathcal{A}^\dagger\Big(\mathcal{A}(Z_0Z_0^\dagger)-y\Big)\Big\|_2 &\leq \Big\|\mathcal{A}^\dagger\Big(\mathcal{A}(Z_0Z_0^\dagger)-y\Big) - \mathcal{A}^\dagger(\mathcal{A}(\rho^\star)-y)\Big\|_2 + \Big\|\mathcal{A}^\dagger(\mathcal{A}(\rho^\star)-y)\Big\|_2\\ &\overset{(i)}{=} \Big\|\mathcal{A}^\dagger\Big(\mathcal{A}(Z_0Z_0^\dagger)-y\Big) - \mathcal{A}^\dagger(\mathcal{A}(\rho^\star)-y)\Big\|_2\\ &\overset{(ii)}{\leq} (1+\delta_{2r})\Big\|Z_0Z_0^\dagger - U^\star U^{\star\dagger}\Big\|_F\\ &\overset{(iii)}{\leq} (1+\delta_{2r})(\|Z_0\|_2 + \|U^\star\|_2)\cdot\|Z-U^\star R_Z\|_F\\ &\overset{(iv)}{\leq} (1+\delta_{2r})\Big(\sqrt{\tfrac{101}{100}}+1\Big)\|U^\star\|_2\cdot 0.001\cdot\|U^\star\|_2^2\\ &\leq 0.002\cdot(1+\delta_{2r})\|\rho^\star\|_2\end{aligned}$$

where (i) is due to $\big\|\mathcal{A}^\dagger(\mathcal{A}(\rho^\star)-y)\big\|_2 = 0$, (ii) is due to the restricted smoothness assumption and the RIP, (iii) is due to the bounds above on $\big\|Z_0Z_0^\dagger - U^\star U^{\star\dagger}\big\|_F$, (iv) is due to the bounds on $\|Z_0\|_2$, w.r.t. $\|U^\star\|_2$, as well as the bound on $\|Z-U^\star R\|_F$.

Thus, in the inequality above, we get:

$$\begin{aligned}\eta &\geq \frac{1}{4\Big((1+\delta_{2r})\frac{101}{100}\|\rho^{\star\top}\|_2 + \|\mathcal{A}^\dagger\big(\mathcal{A}(Z_0Z_0^\dagger)-y\big)\|_2\Big)}\\ &\geq \frac{1}{4\Big((1+\delta_{2r})\frac{101}{100}\|\rho^{\star\top}\|_2 + 0.001\cdot(1+\delta_{2r})\|\rho^\star\|_2 + \|\mathcal{A}^\dagger(\mathcal{A}(\rho^\star)-y)\|_2\Big)}\\ &\geq \frac{1}{4\Big((1+\delta_{2r})\frac{102}{100}\|\rho^{\star\top}\|_2 + \|\mathcal{A}^\dagger(\mathcal{A}(\rho^\star)-y)\|_2\Big)} \geq \tfrac{100}{102}\eta^\star\end{aligned}$$

Similarly, one can show that $\frac{102}{100}\eta^\star \geq \eta$. □

Lemma A7. *Let $U \in \mathbb{C}^{d\times r}, U_- \in \mathbb{C}^{d\times r}$, and $U^\star \in \mathbb{C}^{d\times r}$, such that $\min_{R\in\mathcal{O}}\|U-U^\star R\|_F \leq \frac{\sigma_r(\rho^\star)^{1/2}}{10^3\sqrt{\kappa\tau(\rho^\star)}}$ and $\min_{R\in\mathcal{O}}\|U_- - U^\star R\|_F \leq \frac{\sigma_r(\rho^\star)^{1/2}}{10^3\sqrt{\kappa\tau(\rho^\star)}}$, where $\rho^\star = U^\star U^{\star\dagger}$, and $\kappa := \frac{1+\delta_{2r}}{1-\delta_{2r}} > 1$, for $\delta_{2r} \leq \frac{1}{10}$, and $\tau(\rho^\star) := \frac{\sigma_1(\rho^\star)}{\sigma_r(\rho^\star)} > 1$. By Lemma A5, the above imply also that: $\|Z-U^\star R_Z\|_F \leq \Big(\frac{3}{2}+2|\mu|\Big)\cdot\frac{\sigma_r(\rho^\star)^{1/2}}{10^3\sqrt{\kappa\tau(\rho^\star)}}$. Then, under RIP assumptions of the mapping $\mathcal{A}$, we have:*

$$\begin{aligned}&\Big\langle\mathcal{A}^\dagger(\mathcal{A}(ZZ^\dagger)-y),(Z-U^\star R_Z)(Z-U^\star R_Z)^\dagger\Big\rangle\\ &\geq -\left(\theta\sigma_r(\rho^\star)\cdot\|Z-U^\star R_Z\|_F^2 + \tfrac{10.1}{100}\beta^2\cdot\widehat{\eta}\cdot\frac{(1+2|\mu|)^2}{\Big(1-(1+2|\mu|)\frac{1}{200}\Big)^2}\cdot\|\mathcal{A}^\dagger(\mathcal{A}(ZZ^\dagger)-y)\cdot Z\|_F^2\right)\end{aligned}$$

where

$$\theta = \frac{(1-\delta_{2r})\Big(1+(1+2|\mu|)\frac{1}{200}\Big)^2}{10^3} + (1+\delta_{2r})\Big(2+(1+2|\mu|)\cdot\tfrac{1}{200}\Big)(1+2|\mu|)\cdot\tfrac{1}{200},$$

and $\widehat{\eta} = \frac{1}{4((1+\delta_r)\|ZZ^\dagger\|_2 + \|\mathcal{A}^\dagger\left(\mathcal{A}(ZZ^\dagger) - y\right)Q_Z Q_Z^\dagger\|_2)}$.

Proof. First, denote $\Delta := Z - U^\star R_Z$. Then:

$$\begin{aligned}
&\left\langle \mathcal{A}^\dagger(\mathcal{A}(ZZ^\dagger) - y), (Z - U^\star R_Z)(Z - U^\star R_Z)^\dagger \right\rangle \\
&\overset{(i)}{=} \left\langle \mathcal{A}^\dagger(\mathcal{A}(ZZ^\dagger) - y) \cdot Q_\Delta Q_\Delta^\dagger, \Delta_Z \Delta_Z^\dagger \right\rangle \\
&\geq -\left| \mathrm{Tr}\left(\mathcal{A}^\dagger(\mathcal{A}(ZZ^\dagger) - y) \cdot Q_\Delta Q_\Delta^\dagger \cdot \Delta_Z \Delta_Z^\dagger \right) \right| \\
&\overset{(ii)}{\geq} -\|\mathcal{A}^\dagger(\mathcal{A}(ZZ^\dagger) - y) \cdot Q_\Delta Q_\Delta^\dagger\|_2 \cdot \mathrm{Tr}(\Delta_Z \Delta_Z^\dagger) \\
&\overset{(iii)}{\geq} -\left(\|\mathcal{A}^\dagger(\mathcal{A}(ZZ^\dagger) - y) \cdot Q_Z Q_Z^\dagger\|_2 + \|\mathcal{A}^\dagger(\mathcal{A}(ZZ^\dagger) - y) \cdot Q_{U^\star} Q_{U^\star}^\dagger\|_2 \right) \|Z - U^\star R_Z\|_F^2
\end{aligned} \tag{A12}$$

Note that (i) follows from the fact $\Delta_Z = \Delta_Z Q_\Delta Q_\Delta^\dagger$, for a matrix Q that spans the row space of Δ_Z, and (ii) follows from $|\mathtt{Tr}(AB)| \leq \|A\|_2 \mathtt{Tr}(B)$, for PSD matrix B (Von Neumann's trace inequality [107]). For the transformation in (iii), we use that fact that the row space of Δ_Z, $\textsc{Span}(\Delta_Z)$, is a subset of $\textsc{Span}(Z \cup U^\star)$, as Δ_Z is a linear combination of U and $U^\star$.

To bound the first term in equation (A12), we observe:

$$\begin{aligned}
&\|\mathcal{A}^\dagger(\mathcal{A}(ZZ^\dagger) - y) \cdot Q_Z Q_Z^\dagger\|_2 \cdot \|Z - U^\star R_Z\|_F^2 \\
&\quad \overset{(i)}{=} \widehat{\eta} \cdot 4\Big((1 + \delta_{2r})\|ZZ^\dagger\|_2 \\
&\qquad + \|\mathcal{A}^\dagger(\mathcal{A}(ZZ^\dagger) - y) \cdot Q_Z Q_Z^\dagger\|_2 \Big) \cdot \|\mathcal{A}^\dagger(\mathcal{A}(ZZ^\dagger) - y) \cdot Q_Z Q_Z^\dagger\|_2 \cdot \|Z - U^\star R_Z\|_F^2 \\
&\quad = \underbrace{4\widehat{\eta}(1 + \delta_{2r})\|ZZ^\dagger\|_2 \|\mathcal{A}^\dagger(\mathcal{A}(ZZ^\dagger) - y) \cdot Q_Z Q_Z^\dagger\|_2 \cdot \|Z - U^\star R_Z\|_F^2}_{:=A} \\
&\qquad + 4\widehat{\eta}\|\mathcal{A}^\dagger(\mathcal{A}(ZZ^\dagger) - y) \cdot Q_Z Q_Z^\dagger\|_2^2 \cdot \|Z - U^\star R_Z\|_F^2
\end{aligned}$$

where (i) is due to the definition of $\widehat{\eta}$.

To bound term A, we observe that $\|\mathcal{A}^\dagger(\mathcal{A}(ZZ^\dagger) - y) \cdot Q_Z Q_Z^\dagger\|_2 \leq \frac{(1-\delta_{2r})\sigma_r(ZZ^\dagger)}{10^3}$ or $\|\mathcal{A}^\dagger(\mathcal{A}(ZZ^\dagger) - y) \cdot Q_Z Q_Z^\dagger\|_2 \geq \frac{(1-\delta_{2r})\sigma_r(ZZ^\dagger)}{10^3}$. This results into bounding A as follows:

$$\begin{aligned}
&4\widehat{\eta}(1 + \delta_{2r})\|ZZ^\dagger\|_2 \|\mathcal{A}^\dagger(\mathcal{A}(ZZ^\dagger) - y) \cdot Q_Z Q_Z^\dagger\|_2 \cdot \|Z - U^\star R_Z\|_F^2 \\
&\quad \leq \max\Big\{ \tfrac{4 \cdot \widehat{\eta} \cdot (1+\delta_{2r})\|ZZ^\dagger\|_2 \cdot (1-\delta_{2r})\sigma_r(ZZ^\dagger)}{10^3} \cdot \|Z - U^\star R_Z\|_F^2, \\
&\qquad\qquad \widehat{\eta} \cdot 4 \cdot 10^3 \kappa \tau(ZZ^\dagger) \|\mathcal{A}^\dagger(\mathcal{A}(ZZ^\dagger) - y) \cdot Q_Z Q_Z^\dagger\|_2^2 \cdot \|Z - U^\star R_Z\|_F^2 \Big\} \\
&\quad \leq \tfrac{4 \cdot \widehat{\eta} \cdot (1-\delta_{2r}^2)\|ZZ^\dagger\|_2 \cdot \sigma_r(ZZ^\dagger)}{10^3} \cdot \|Z - U^\star R_Z\|_F^2 \\
&\qquad\qquad + \widehat{\eta} \cdot 4 \cdot 10^3 \kappa \tau(ZZ^\dagger) \|\mathcal{A}^\dagger(\mathcal{A}(ZZ^\dagger) - y) \cdot Q_Z Q_Z^\dagger\|_2^2 \cdot \|Z - U^\star R_Z\|_F^2.
\end{aligned}$$

Combining the above inequalities, we obtain:

$$
\begin{aligned}
&\|\mathcal{A}^\dagger(\mathcal{A}(ZZ^\dagger)-y)\cdot Q_Z Q_Z^\dagger\|_2\cdot\|Z-U^\star R_Z\|_F^2\\
&\overset{(i)}{\leq}\tfrac{(1-\delta_{2r})\sigma_r(ZZ^\dagger)}{10^3}\cdot\|Z-U^\star R_Z\|_F^2\\
&\qquad+(10^3\kappa\tau(ZZ^\dagger)+1)\cdot 4\cdot\widehat{\eta}\|\mathcal{A}^\dagger(\mathcal{A}(ZZ^\dagger)-y)\cdot Q_Z Q_Z^\dagger\|_2^2\cdot\|Z-U^\star R_Z\|_F^2\\
&\overset{(ii)}{\leq}\tfrac{(1-\delta_{2r})\sigma_r(ZZ^\dagger)}{10^3}\cdot\|Z-U^\star R_Z\|_F^2\\
&\qquad+(10^3\beta^2\kappa\tau(\rho^\star)+1)\cdot 4\cdot\widehat{\eta}\|\mathcal{A}^\dagger(\mathcal{A}(ZZ^\dagger)-y)\cdot Q_Z Q_Z^\dagger\|_2^2\cdot\tfrac{\left(\frac{3}{2}+2|\mu|\right)^2}{\kappa\tau(\rho^\star)}\tfrac{1}{10^6}\sigma_r(\rho^\star)\\
&\overset{(iii)}{\leq}\tfrac{(1-\delta_{2r})\sigma_r(ZZ^\dagger)}{10^3}\cdot\|Z-U^\star R_Z\|_F^2\\
&\qquad+4\cdot 1001\beta^2\cdot\widehat{\eta}\cdot\|\mathcal{A}^\dagger(\mathcal{A}(ZZ^\dagger)-y)\cdot Q_Z Q_Z^\dagger\|_2^2\cdot\frac{\left(\frac{3}{2}+2|\mu|\right)^2}{10^6\left(1-\left(\frac{3}{2}+2|\mu|\right)\frac{1}{10^3}\right)^2}\sigma_r(ZZ^\dagger)\\
&\overset{(iv)}{\leq}\tfrac{(1-\delta_{2r})\sigma_r(ZZ^\dagger)}{10^3}\cdot\|Z-U^\star R_Z\|_F^2\\
&\qquad+4\cdot 1001\beta^2\cdot\widehat{\eta}\cdot\frac{\left(\frac{3}{2}+2|\mu|\right)^2}{10^6\left(1-\left(\frac{3}{2}+2|\mu|\right)\frac{1}{10^3}\right)^2}\cdot\|\mathcal{A}^\dagger(\mathcal{A}(ZZ^\dagger)-y)\cdot Z\|_F^2\\
&\overset{(v)}{\leq}\frac{(1-\delta_{2r})\left(1+\left(\frac{3}{2}+2|\mu|\right)\frac{1}{10^3}\right)^2\sigma_r(\rho^\star)}{10^3}\cdot\|Z-U^\star R_Z\|_F^2\\
&\qquad+\tfrac{1}{200}\beta^2\cdot\widehat{\eta}\cdot\frac{\left(\frac{3}{2}+2|\mu|\right)^2}{\left(1-\left(\frac{3}{2}+2|\mu|\right)\frac{1}{10^3}\right)^2}\cdot\|\mathcal{A}^\dagger(\mathcal{A}(ZZ^\dagger)-y)\cdot Z\|_F^2
\end{aligned}
$$

where (i) follows from $\widehat{\eta}\leq\frac{1}{4(1+\delta_{2r})\|ZZ^\dagger\|_2}$, (ii) is due to Corollary A3, bounding $\|Z-U^\star R_Z\|_F\leq\rho\sigma_r(\rho^\star)^{1/2}$, where $\rho:=\left(\frac{3}{2}+2|\mu|\right)\frac{1}{10^3\sqrt{\kappa\tau(\rho^\star)}}$ by Lemma A5, (iii) is due to $(10^3\beta^2\kappa\tau(\rho^\star)+1)\leq 1001\beta^2\kappa\tau(\rho^\star)$, and by Corollary A1, (iv) is due to the fact

$$\sigma_r(ZZ^\dagger)\|\mathcal{A}^\dagger(\mathcal{A}(ZZ^\dagger)-y)\cdot Q_Z Q_Z^\dagger\|_2^2\leq\|\mathcal{A}^\dagger(\mathcal{A}(ZZ^\dagger)-y)Z\|_F^2,$$

and (v) is due to Corollary A1.

Next, we bound the second term in equation (A12):

$$
\begin{aligned}
&\|\mathcal{A}^\dagger(\mathcal{A}(ZZ^\dagger)-y)\cdot Q_{U^\star}Q_{U^\star}^\dagger\|_2\cdot\|Z-U^\star R_Z\|_F^2\\
&\qquad\overset{(i)}{\leq}\|\mathcal{A}^\dagger(\mathcal{A}(ZZ^\dagger)-y)-\mathcal{A}^\dagger(\mathcal{A}(\rho^\star)-y)\|_2\cdot\|Z-U^\star R_Z\|_F^2\\
&\qquad\overset{(ii)}{\leq}(1+\delta_{2r})\cdot\|ZZ^\dagger-U^\star U^{\star\dagger}\|_F\cdot\|Z-U^\star R_Z\|_F^2\\
&\qquad\overset{(iii)}{\leq}(1+\delta_{2r})(2+\rho)\cdot\rho\cdot\sigma_1(U^\star)\cdot\sigma_r(U^\star)\cdot\|Z-U^\star R_Z\|_F^2\\
&\qquad\overset{(iv)}{\leq}(1+\delta_{2r})(2+\rho)\left(\tfrac{3}{2}+2|\mu|\right)\cdot\tfrac{1}{10^3}\sigma_r(\rho^\star)\cdot\|Z-U^\star R_Z\|_F^2\\
&\qquad\leq(1+\delta_{2r})\left(2+\left(\tfrac{3}{2}+2|\mu|\right)\cdot\tfrac{1}{10^3}\right)\left(\tfrac{3}{2}+2|\mu|\right)\cdot\tfrac{1}{10^3}\sigma_r(\rho^\star)\cdot\|Z-U^\star R_Z\|_F^2,
\end{aligned}
$$

where (i) follows from $\|\mathcal{A}^\dagger(\mathcal{A}(ZZ^\dagger)-y)\cdot Q_{U^\star}Q_{U^\star}^\dagger\|_2\leq\|\mathcal{A}^\dagger(\mathcal{A}(ZZ^\dagger)-y)\|_2$ and $\mathcal{A}^\dagger(\mathcal{A}(\rho^\star)-y)=0$, (ii) is due to smoothness of f and the RIP constants, (iii) follows from [25] (Lemma 18), for $\rho=\left(\frac{3}{2}+2|\mu|\right)\cdot\frac{1}{10^3\sqrt{\kappa\tau(\rho^\star)}}$, (iv) follows from substituting ρ above, and observing that $\tau(\rho^\star)=\sigma_1(U^\star)^2/\sigma_r(U^\star)^2>1$ and $\kappa=(1+\delta_{2r})/(1-\delta_{2r})>1$.

Combining the above we get:

$$\begin{aligned}&\Big\langle \mathcal{A}^\dagger(\mathcal{A}(ZZ^\dagger)-y),(Z-U^\star R_Z)(Z-U^\star R_Z)^\dagger\Big\rangle\\&\geq -\left(\theta\sigma_r(\rho^\star)\cdot\|Z-U^\star R_Z\|_F^2+\tfrac{1}{200}\beta^2\cdot\widehat{\eta}\cdot\frac{\left(\frac{3}{2}+2|\mu|\right)^2}{\left(1-\left(\frac{3}{2}+2|\mu|\right)\frac{1}{10^3}\right)^2}\cdot\|\mathcal{A}^\dagger(\mathcal{A}(ZZ^\dagger)-y)\cdot Z\|_F^2\right)\end{aligned}\tag{A13}$$

where $\theta=\frac{(1-\delta_{2r})\left(1+\left(\frac{3}{2}+2|\mu|\right)\frac{1}{10^3}\right)^2}{10^3}+(1+\delta_{2r})\left(2+\left(\frac{3}{2}+2|\mu|\right)\cdot\frac{1}{10^3}\right)\left(\frac{3}{2}+2|\mu|\right)\cdot\frac{1}{10^3}$. □

Lemma A8. *Under identical assumptions with Lemma A7, the following inequality holds:*

$$\Big\langle \mathcal{A}^\dagger(\mathcal{A}(ZZ^\dagger)-y),ZZ^\dagger-U^\star U^{\star\dagger}\Big\rangle\geq 1.1172\eta\Big\|\mathcal{A}^\dagger(\mathcal{A}(ZZ^\dagger)-y)Z\Big\|_F^2+\tfrac{1-\delta_{2r}}{2}\|U^\star U^{\star\dagger}-ZZ^\dagger\|_F^2$$

Proof. By smoothness assumption of the objective, based on the RIP assumption, we have:

$$\begin{aligned}\tfrac{1}{2}\|\mathcal{A}(ZZ^\dagger)-y\|_2^2\geq\tfrac{1}{2}\|\mathcal{A}(U_+U_+^\dagger)-y\|_2^2\\-\Big\langle\mathcal{A}^\dagger(\mathcal{A}(ZZ^\dagger)-y),U_+U_+^\dagger-ZZ^\dagger\Big\rangle-\tfrac{1+\delta_{2r}}{2}\|U_+U_+^\dagger-ZZ^\dagger\|_F^2\Rightarrow\\\tfrac{1}{2}\|\mathcal{A}(ZZ^\dagger)-y\|_2^2\geq\tfrac{1}{2}\|\mathcal{A}(U^\star U^{\star\dagger})-y\|_2^2\\-\Big\langle\mathcal{A}^\dagger(\mathcal{A}(ZZ^\dagger)-y),U_+U_+^\dagger-ZZ^\dagger\Big\rangle-\tfrac{1+\delta_{2r}}{2}\|U_+U_+^\dagger-ZZ^\dagger\|_F^2\end{aligned}$$

due to the optimality $\|\mathcal{A}(U^\star U^{\star\dagger})-y\|_2^2=0\leq\|\mathcal{A}(VV^\dagger)-y\|_2^2$, for any $V\in\mathbb{C}^{d\times r}$. Also, by the restricted strong convexity with RIP, we get:

$$\begin{aligned}\tfrac{1}{2}\|\mathcal{A}(U^\star U^{\star\dagger})-y\|_2^2\geq\tfrac{1}{2}\|\mathcal{A}(ZZ^\dagger)-y\|_2^2\\+\Big\langle\mathcal{A}^\dagger(\mathcal{A}(ZZ^\dagger)-y),U^\star U^{\star\dagger}-ZZ^\dagger\Big\rangle+\tfrac{1-\delta_{2r}}{2}\|U^\star U^{\star\dagger}-ZZ^\dagger\|_F^2\end{aligned}$$

Adding the two inequalities, we obtain:

$$\begin{aligned}\Big\langle\mathcal{A}^\dagger(\mathcal{A}(ZZ^\dagger)-y),ZZ^\dagger-U^\star U^{\star\dagger}\Big\rangle\geq\Big\langle\mathcal{A}^\dagger(\mathcal{A}(ZZ^\dagger)-y),ZZ^\dagger-U_+U_+^\dagger\Big\rangle\\-\tfrac{1+\delta_{2r}}{2}\|U_+U_+^\dagger-ZZ^\dagger\|_F^2+\tfrac{1-\delta_{2r}}{2}\|U^\star U^{\star\dagger}-ZZ^\dagger\|_F^2\end{aligned}$$

To proceed we observe:

$$\begin{aligned}U_+U_+^\dagger&=\Big(Z-\eta\mathcal{A}^\dagger\Big(\mathcal{A}(ZZ^\dagger)-y\Big)Z\Big)\cdot\Big(Z-\eta\mathcal{A}^\dagger\Big(\mathcal{A}(ZZ^\dagger)-y\Big)Z\Big)^\dagger\\&=ZZ^\dagger-\eta ZZ^\dagger\cdot\mathcal{A}^\dagger\Big(\mathcal{A}(ZZ^\dagger)-y\Big)-\eta\mathcal{A}^\dagger\Big(\mathcal{A}(ZZ^\dagger)-y\Big)\cdot ZZ^\dagger\\&\quad+\eta^2\mathcal{A}^\dagger\Big(\mathcal{A}(ZZ^\dagger)-y\Big)\cdot ZZ^\dagger\cdot\mathcal{A}^\dagger\Big(\mathcal{A}(ZZ^\dagger)-y\Big)\\&\overset{(i)}{=}ZZ^\dagger-\Big(I-\tfrac{\eta}{2}Q_ZQ_Z^\dagger\mathcal{A}^\dagger\Big(\mathcal{A}(ZZ^\dagger)-y\Big)\Big)\cdot\eta ZZ^\dagger\cdot\mathcal{A}^\dagger\Big(\mathcal{A}(ZZ^\dagger)-y\Big)\\&\quad-\eta\mathcal{A}^\dagger\Big(\mathcal{A}(ZZ^\dagger)-y\Big)\cdot ZZ^\dagger\cdot\Big(I-\tfrac{\eta}{2}Q_ZQ_Z^\dagger\mathcal{A}^\dagger\Big(\mathcal{A}(ZZ^\dagger)-y\Big)\Big)\end{aligned}$$

where (i) is due to the fact $\mathcal{A}^\dagger\big(\mathcal{A}(ZZ^\dagger)-y\big)\cdot ZZ^\dagger\cdot\mathcal{A}^\dagger\big(\mathcal{A}(ZZ^\dagger)-y\big)=\mathcal{A}^\dagger\big(\mathcal{A}(ZZ^\dagger)-y\big)\cdot Q_ZQ_Z^\dagger\cdot ZZ^\dagger\cdot Q_ZQ_Z^\dagger\cdot\mathcal{A}^\dagger\big(\mathcal{A}(ZZ^\dagger)-y\big)$, for Q_Z a basis matrix whose columns span the column space of Z; also, I is the identity matrix whose dimension is apparent from the context. Thus:

$$\tfrac{\eta}{2}Q_ZQ_Z^\dagger\mathcal{A}^\dagger\Big(\mathcal{A}(ZZ^\dagger)-y\Big)\preceq\tfrac{10.5}{10}\tfrac{\widehat{\eta}}{2}Q_ZQ_Z^\dagger\mathcal{A}^\dagger\Big(\mathcal{A}(ZZ^\dagger)-y\Big),$$

and, hence,

$$I - \tfrac{\eta}{2} Q_Z Q_Z^\dagger \mathcal{A}^\dagger\Big(\mathcal{A}(ZZ^\dagger) - y\Big) \succeq I - \tfrac{10.5}{10}\tfrac{\widehat{\eta}}{2} Q_Z Q_Z^\dagger \mathcal{A}^\dagger\Big(\mathcal{A}(ZZ^\dagger) - y\Big).$$

Define $\Psi = I - \frac{\eta}{2} Q_Z Q_Z^\dagger \mathcal{A}^\dagger\big(\mathcal{A}(ZZ^\dagger) - y\big)$. Then, using the definition of $\widehat{\eta}$, we know that $\widehat{\eta} \le \frac{1}{4\|Q_Z Q_Z^\dagger \mathcal{A}^\dagger\left(\mathcal{A}(ZZ^\dagger) - y\right)\|_2}$, and thus:

$$\Psi \succ 0, \quad \sigma_1(\Psi) \le 1 + \tfrac{21}{160}, \quad \text{and} \quad \sigma_n(\Psi) \ge 1 - \tfrac{21}{160}.$$

Going back to the main recursion and using the above expression for $UU_+^\dagger$, we have:

$$\begin{aligned}
&\Big\langle \mathcal{A}^\dagger(\mathcal{A}(ZZ^\dagger) - y), ZZ^\dagger - U^\star U^{\star\dagger} \Big\rangle - \tfrac{1-\delta_{2r}}{2}\|U^\star U^{\star\dagger} - ZZ^\dagger\|_F^2 \\
&\quad \ge \Big\langle \mathcal{A}^\dagger(\mathcal{A}(ZZ^\dagger) - y), ZZ^\dagger - U_+U_+^\dagger \Big\rangle - \tfrac{1+\delta_{2r}}{2}\|U_+U_+^\dagger - ZZ^\dagger\|_F^2 \\
&\quad \overset{(i)}{\ge} 2\eta \Big\langle \mathcal{A}^\dagger(\mathcal{A}(ZZ^\dagger) - y), \mathcal{A}^\dagger(\mathcal{A}(ZZ^\dagger) - y) \cdot ZZ^\dagger \cdot \Psi \Big\rangle \\
&\qquad - \tfrac{1+\delta_{2r}}{2}\|2\eta \mathcal{A}^\dagger(\mathcal{A}(ZZ^\dagger) - y) \cdot ZZ^\dagger \cdot \Psi\|_F^2 \\
&\quad \overset{(ii)}{\ge} 2\Big(1 - \tfrac{21}{160}\Big)\eta \Big\|\mathcal{A}^\dagger(\mathcal{A}(ZZ^\dagger) - y)Z\Big\|_F^2 \\
&\qquad - 2(1+\delta_{2r})\eta^2 \Big\|\mathcal{A}^\dagger(\mathcal{A}(ZZ^\dagger) - y)Z\Big\|_F^2 \cdot \|Z\|_2^2 \cdot \|\Psi\|_2^2 \\
&\quad \overset{(iii)}{\ge} 2\Big(1 - \tfrac{21}{160}\Big)\eta \Big\|\mathcal{A}^\dagger(\mathcal{A}(ZZ^\dagger) - y)Z\Big\|_F^2 \\
&\qquad - 2(1+\delta_{2r})\eta^2 \Big\|\mathcal{A}^\dagger(\mathcal{A}(ZZ^\dagger) - y)Z\Big\|_F^2 \cdot \|Z\|_2^2 \cdot \Big(1 + \tfrac{21}{160}\Big)^2 \\
&\quad = 2\Big(1 - \tfrac{21}{160}\Big)\eta \Big\|\mathcal{A}^\dagger(\mathcal{A}(ZZ^\dagger) - y)Z\Big\|_F^2 \cdot \left(1 - 2(1+\delta_{2r})\eta \cdot \|Z\|_2^2 \cdot \Big(1 + \tfrac{21}{160}\Big)^2 \cdot \tfrac{1}{2(1-\frac{21}{160})}\right) \\
&\quad \overset{(iv)}{\ge} 2\Big(1 - \tfrac{21}{160}\Big)\eta \Big\|\mathcal{A}^\dagger(\mathcal{A}(ZZ^\dagger) - y)Z\Big\|_F^2 \cdot \left(1 - 2(1+\delta_{2r})\tfrac{10.5}{10}\widehat{\eta} \cdot \|Z\|_2^2 \cdot \Big(1 + \tfrac{21}{160}\Big)^2 \cdot \tfrac{1}{2(1-\frac{21}{160})}\right) \\
&\quad \overset{(v)}{\ge} 2\Big(1 - \tfrac{21}{160}\Big)\eta \Big\|\mathcal{A}^\dagger(\mathcal{A}(ZZ^\dagger) - y)Z\Big\|_F^2 \cdot \left(1 - \tfrac{10.5}{10}\tfrac{\left(1+\frac{21}{160}\right)^2}{4(1-\frac{21}{160})}\right) \\
&\quad = 1.0656\eta \Big\|\mathcal{A}^\dagger(\mathcal{A}(ZZ^\dagger) - y)Z\Big\|_F^2
\end{aligned}$$

where (i) is due to the symmetry of the objective; (ii) is due to Cauchy-Schwarz inequality and the fact:

$$\begin{aligned}
&\Big\langle \mathcal{A}^\dagger(\mathcal{A}(ZZ^\dagger) - y),\ \mathcal{A}^\dagger(\mathcal{A}(ZZ^\dagger) - y) \cdot ZZ^\dagger \cdot \Psi \Big\rangle \\
&\quad = \Big\langle \mathcal{A}^\dagger(\mathcal{A}(ZZ^\dagger) - y), \mathcal{A}^\dagger(\mathcal{A}(ZZ^\dagger) - y) \cdot ZZ^\dagger \Big\rangle \\
&\qquad - \tfrac{\eta}{2}\Big\langle \mathcal{A}^\dagger(\mathcal{A}(ZZ^\dagger) - y), \mathcal{A}^\dagger(\mathcal{A}(ZZ^\dagger) - y) \cdot ZZ^\dagger \cdot \mathcal{A}^\dagger(\mathcal{A}(ZZ^\dagger) - y) \Big\rangle \\
&\quad \overset{(i)}{\ge} \Big\langle \mathcal{A}^\dagger(\mathcal{A}(ZZ^\dagger) - y), \mathcal{A}^\dagger(\mathcal{A}(ZZ^\dagger) - y) \cdot ZZ^\dagger \Big\rangle \\
&\qquad - \tfrac{10.5}{10}\tfrac{\widehat{\eta}}{2}\Big\langle \mathcal{A}^\dagger(\mathcal{A}(ZZ^\dagger) - y), \mathcal{A}^\dagger(\mathcal{A}(ZZ^\dagger) - y) \cdot ZZ^\dagger \cdot \mathcal{A}^\dagger(\mathcal{A}(ZZ^\dagger) - y) \Big\rangle \\
&\quad \ge \Big(1 - \tfrac{10.5}{10}\tfrac{\widehat{\eta}}{2}\|Q_Z Q_Z^\dagger \mathcal{A}^\dagger(\mathcal{A}(ZZ^\dagger) - y)\|_2^2\Big) \cdot \Big\|\mathcal{A}^\dagger(\mathcal{A}(ZZ^\dagger) - y)Z\Big\|_F^2 \\
&\quad \ge \Big(1 - \tfrac{21}{160}\Big)\Big\|\mathcal{A}^\dagger(\mathcal{A}(ZZ^\dagger) - y)Z\Big\|_F^2
\end{aligned}$$

where (i) is due to $\eta \leq \frac{10.5}{10}\widehat{\eta}$, and the last inequality comes from the definition of the $\widehat{\eta}$ and its upper bound; (iii) is due to the upper bound on $\|\Psi\|_2$ above; (iv) is due to $\eta \leq \frac{10.5}{10}\widehat{\eta}$; (v) is due to $\widehat{\eta} \leq \frac{1}{4(1+\delta_{2r})\|ZZ^\dagger\|_2}$. The above lead to the desiderata:

$$\left\langle \mathcal{A}^\dagger(\mathcal{A}(ZZ^\dagger)-y), ZZ^\dagger - U^\star U^{\star\dagger} \right\rangle \geq 1.0656\eta \left\| \mathcal{A}^\dagger(\mathcal{A}(ZZ^\dagger)-y)Z \right\|_F^2 + \tfrac{1-\delta_{2r}}{2}\|U^\star U^{\star\dagger} - ZZ^\dagger\|_F^2$$

□

References

1. Altepeter, J.B.; James, D.F.; Kwiat, P.G. 4 qubit quantum state tomography. In *Quantum State Estimation*; Springer: Berlin/Heidelberg, Germany, 2004; pp. 113–145.
2. Eisert, J.; Hangleiter, D.; Walk, N.; Roth, I.; Markham, D.; Parekh, R.; Chabaud, U.; Kashefi, E. Quantum certification and benchmarking. *arXiv* **2019**, arXiv:1910.06343.
3. Mohseni, M.; Rezakhani, A.; Lidar, D. Quantum-process tomography: Resource analysis of different strategies. *Phys. Rev. A* **2008**, *77*, 032322.
4. Gross, D.; Liu, Y.K.; Flammia, S.; Becker, S.; Eisert, J. Quantum state tomography via compressed sensing. *Phys. Rev. Lett.* **2010**, *105*, 150401.
5. Vogel, K.; Risken, H. Determination of quasiprobability distributions in terms of probability distributions for the rotated quadrature phase. *Phys. Rev. A* **1989**, *40*, 2847.
6. Ježek, M.; Fiurášek, J.; Hradil, Z. Quantum inference of states and processes. *Phys. Rev. A* **2003**, *68*, 012305.
7. Banaszek, K.; Cramer, M.; Gross, D. Focus on quantum tomography. *New J. Phys.* **2013**, *15*, 125020.
8. Kalev, A.; Kosut, R.; Deutsch, I. Quantum tomography protocols with positivity are compressed sensing protocols. *NPJ Quantum Inf.* **2015**, *1*, 15018.
9. Torlai, G.; Mazzola, G.; Carrasquilla, J.; Troyer, M.; Melko, R.; Carleo, G. Neural-network quantum state tomography. *Nat. Phys.* **2018**, *14*, 447–450. https://doi.org/10.1038/s41567-018-0048-5.
10. Beach, M.J.; De Vlugt, I.; Golubeva, A.; Huembeli, P.; Kulchytskyy, B.; Luo, X.; Melko, R.G.; Merali, E.; Torlai, G. QuCumber: Wavefunction reconstruction with neural networks. *SciPost Phys.* **2019**, *7*, 009.
11. Torlai, G.; Melko, R. Machine-Learning Quantum States in the NISQ Era. *Annu. Rev. Condens. Matter Phys.* **2020**, *11*, 325-344.
12. 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. Comm.* **2010**, *1*, 149. https://doi.org/10.1038/ncomms1147.
13. Lanyon, B.; Maier, C.; Holzäpfel, M.; Baumgratz, T.; Hempel, C.; Jurcevic, P.; Dhand, I.; Buyskikh, A.; Daley, A.; Cramer, M.; et al. Efficient tomography of a quantum many-body system. *Nat. Phys.* **2017**, *13*, 1158–1162.
14. Gonçalves, D.; Gomes-Ruggiero, M.; Lavor, C. A projected gradient method for optimization over density matrices. *Optim. Methods Softw.* **2016**, *31*, 328–341.
15. Bolduc, E.; Knee, G.; Gauger, E.; Leach, J. Projected gradient descent algorithms for quantum state tomography. *NPJ Quantum Inf.* **2017**, *3*, 44.
16. Shang, J.; Zhang, Z.; Ng, H.K. Superfast maximum-likelihood reconstruction for quantum tomography. *Phys. Rev. A* **2017**, *95*, 062336. https://doi.org/10.1103/PhysRevA.95.062336.
17. Hu, Z.; Li, K.; Cong, S.; Tang, Y. Reconstructing Pure 14-Qubit Quantum States in Three Hours Using Compressive Sensing. *IFAC-PapersOnLine* **2019**, *52*, 188– 193. https://doi.org/https://doi.org/10.1016/j.ifacol.2019.09.139.
18. Hou, Z.; Zhong, H.S.; Tian, Y.; Dong, D.; Qi, B.; Li, L.; Wang, Y.; Nori, F.; Xiang, G.Y.; Li, C.F.; et al. Full reconstruction of a 14-qubit state within four hours. *New J. Phys.* **2016**, *18*, 083036.
19. Candes, E.; Tao, T. Near-optimal signal recovery from random projections: Universal encoding strategies? *IEEE Trans. Inf. Theory* **2006**, *52*, 5406–5425.
20. Gross, D. Recovering low-rank matrices from few coefficients in any basis. *IEEE Trans. Inf. Theory* **2011**, *57*, 1548–1566.
21. Liu, Y.K. Universal low-rank matrix recovery from Pauli measurements. In Proceedings of the Advances in Neural Information Processing Systems, Granada, Spain, 12–15 December 2011, pp. 1638–1646.
22. Riofrío, C.; Gross, D.; Flammia, S.; Monz, T.; Nigg, D.; Blatt, R.; Eisert, J. Experimental quantum compressed sensing for a seven-qubit system. *Nat. Commun.* **2017**, *8*, 15305, 1–8.
23. Kliesch, M.; Kueng, R.; Eisert, J.; Gross, D. Guaranteed recovery of quantum processes from few measurements. *Quantum* **2019**, *3*, 171.
24. Flammia, S.T.; Gross, D.; Liu, Y.K.; Eisert, J. Quantum tomography via compressed sensing: Error bounds, sample complexity and efficient estimators. *New J. Phys.* **2012**, *14*, 095022.
25. Bhojanapalli, S.; Kyrillidis, A.; Sanghavi, S. Dropping convexity for faster semi-definite optimization. In Proceedings of the Conference on Learning Theory, New York, NY, USA, 23–26 June 2016; pp. 530–582.
26. 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*, 36.

27. Gao, X.; Duan, L.M. Efficient representation of quantum many-body states with deep neural networks. *Nat. Commun.* **2017**, *8*, 1–6.
28. tA v, A. ; Anis, M.S.; Mitchell, A.; Abraham, H.; AduOffei.; Agarwal, R.; Agliardi, G.; Aharoni, M.; Ajith, V.; Akhalwaya, I.Y.; et al. Qiskit: An Open-Source Framework for Quantum Computing. **2021**. https://doi.org/10.5281/zenodo.2573505.
29. Recht, B.; Fazel, M.; Parrilo, P.A. Guaranteed minimum-rank solutions of linear matrix equations via nuclear norm minimization. *SIAM Rev.* **2010**, *52*, 471–501.
30. Park, D.; Kyrillidis, A.; Caramanis, C.; Sanghavi, S. Finding low-rank solutions to matrix problems, efficiently and provably. *arXiv* **2016**, arXiv:1606.03168.
31. Park, D.; Kyrillidis, A.; Bhojanapalli, S.; Caramanis, C.; Sanghavi, S. Provable Burer-Monteiro factorization for a class of norm-constrained matrix problems. *arXiv* **2016**, arXiv:1606.01316.
32. Tu, S.; Boczar, R.; Simchowitz, M.; Soltanolkotabi, M.; Recht, B. Low-rank solutions of linear matrix equations via Procrustes flow. In Proceedings of the 33rd International Conference on International Conference on Machine Learning-Volume 48, New York, NY, USA, 19–24 June 2016; pp. 964–973.
33. Zhao, T.; Wang, Z.; Liu, H. A nonconvex optimization framework for low rank matrix estimation. In Proceedings of the Advances in Neural Information Processing Systems, Montreal, QC, Canada, 7–12 December 2015; pp. 559–567.
34. Zheng, Q.; Lafferty, J. A convergent gradient descent algorithm for rank minimization and semidefinite programming from random linear measurements. In Proceedings of the Advances in Neural Information Processing Systems, Montreal, QC, Canada, 7–12 December 2015; pp. 109–117.
35. Cauchy, A. Méthode générale pour la résolution des systemes d'équations simultanées. *Comp. Rend. Sci. Paris* **1847**, *25*, 536–538.
36. Park, D.; Kyrillidis, A.; Caramanis, C.; Sanghavi, S. Non-square matrix sensing without spurious local minima via the Burer-Monteiro approach. In Proceedings of the International Conference on Artificial Intelligence and Statistics, Ft. Lauderdale, FL, USA, 20–22 April 2017; pp. 65–74.
37. Ge, R.; Jin, C.; Zheng, Y. No spurious local minima in nonconvex low rank problems: A unified geometric analysis. *arXiv* **2017**, arXiv:1704.00708.
38. Hsieh, Y.P.; Kao, Y.C.; Karimi Mahabadi, R.; Alp, Y.; Kyrillidis, A.; Cevher, V. A Non-Euclidean Gradient Descent Framework for Non-Convex Matrix Factorization. *IEEE Transactions on Signal Processing.* **2018**, *66*, pp. 5917–5926.
39. Polyak, B. Some methods of speeding up the convergence of iteration methods. *USSR Comput. Math. Math. Phys.* **1964**, *4*, 1–17.
40. Nesterov, Y. A method of solving a convex programming problem with convergence rate $O(\frac{1}{k^2})$. In Proceedings of the Soviet Mathematics Doklady, 1983; Volume 27, pp. 372–376.
41. Bhojanapalli, S.; Neyshabur, B.; Srebro, N. Global optimality of local search for low rank matrix recovery. In Proceedings of the Advances in Neural Information Processing Systems, Barcelona, Spain, 5–10 December 2016; pp. 3873–3881.
42. Stöger, D.; Soltanolkotabi, M. Small random initialization is akin to spectral learning: Optimization and generalization guarantees for overparameterized low-rank matrix reconstruction. *Adv. Neural Inf. Process. Syst.* **2021**, *34*, 23831–23843.
43. Lanczos, C. An Iteration Method for the Solution of the Eigenvalue Problem of Linear Differential and Integral Operators1. *J. Res. Natl. Bur. Stand.* **1950**, *45*.
44. Nesterov, Y. *Introductory Lectures on Convex Optimization: A Basic Course*; Springer Science & Business Media: Berlin, Germany, 2013; Volume 87.
45. Carmon, Y.; Duchi, J.; Hinder, O.; Sidford, A. Accelerated methods for non-convex optimization. *arXiv* **2016**, arXiv:1611.00756.
46. Li, Y.; Ma, C.; Chen, Y.; Chi, Y. Nonconvex Matrix Factorization from Rank-One Measurements. In Proceedings of the International Conference on Artificial Intelligence and Statistics, Naha, Japan, 16–18 April 2019.
47. Khanna, R.; Kyrillidis, A. IHT dies hard: Provable accelerated Iterative Hard Thresholding. In Proceedings of the International Conference on Artificial Intelligence and Statistics, Playa Blanca, Lanzarote, Canary Islands, 9–11 April 2018; pp. 188–198.
48. Rudelson, M.; Vershynin, R. On sparse reconstruction from Fourier and Gaussian measurements. *Commun. Pure Appl. Math. A J. Issued Courant Inst. Math. Sci.* **2008**, *61*, 1025–1045.
49. Vandenberghe, L. The CVXOPT Linear and Quadratic Cone Program Solvers. 2010. Available online: https://www.seas.ucla.edu/~vandenbe/publications/coneprog.pdf (accessed on 20 January 2023) .
50. Diamond, S.; Boyd, S. CVXPY: A Python-embedded modeling language for convex optimization. *J. Mach. Learn. Res.* **2016**, *17*, 1–5.
51. Agrawal, A.; Verschueren, R.; Diamond, S.; Boyd, S. A rewriting system for convex optimization problems. *J. Control. Decis.* **2018**, *5*, 42–60.
52. Smolin, J.A.; Gambetta, J.M.; Smith, G. Efficient method for computing the maximum-likelihood quantum state from measurements with additive gaussian noise. *Phys. Rev. Lett.* **2012**, *108*, 070502.
53. O'Donoghue, B.; Chu, E.; Parikh, N.; Boyd, S. Conic Optimization via Operator Splitting and Homogeneous Self-Dual Embedding. *J. Optim. Theory Appl.* **2016**, *169*, 1042–1068.
54. O'Donoghue, B.; Chu, E.; Parikh, N.; Boyd, S. SCS: Splitting Conic Solver, version 2.1.2. 2019. Available online: https://github.com/cvxgrp/scs (accessed on 20 January 2023) .
55. Forum, T.M. MPI: A Message Passing Interface. In Proceedings of the 1993 ACM/IEEE Conference on Supercomputing, Portland, OR, USA, 19 November 1993; Association for Computing Machinery: New York, NY, USA, 1993; pp. 878–883. https://doi.org/10.1145/169627.169855.

56. Dalcin, L.D.; Paz, R.R.; Kler, P.A.; Cosimo, A. Parallel distributed computing using Python. *Adv. Water Resour.* **2011**, *34*, 1124–1139.
57. Lee, K.; Bresler, Y. Guaranteed minimum rank approximation from linear observations by nuclear norm minimization with an ellipsoidal constraint. *arXiv* **2009**, arXiv:0903.4742.
58. Liu, Z.; Vandenberghe, L. Interior-point method for nuclear norm approximation with application to system identification. *SIAM J. Matrix Anal. Appl.* **2009**, *31*, 1235–1256.
59. Jain, P.; Meka, R.; Dhillon, I.S. Guaranteed rank minimization via singular value projection. In Proceedings of the Advances in Neural Information Processing Systems, Vancouver, BC, Canada, 6–9 December 2010; pp. 937–945.
60. Lee, K.; Bresler, Y. Admira: Atomic decomposition for minimum rank approximation. *IEEE Trans. Inf. Theory* **2010**, *56*, 4402–4416.
61. Kyrillidis, A.; Cevher, V. Matrix recipes for hard thresholding methods. *J. Math. Imaging Vis.* **2014**, *48*, 235–265.
62. Kingma, D.; Ba, J. Adam: A method for stochastic optimization. *arXiv* **2014**, arXiv:1412.6980.
63. Tieleman, T.; Hinton, G. Lecture 6.5-RMSPro: Divide the gradient by a running average of its recent magnitude. *COURSERA Neural Netw. Mach. Learn.* **2012**, *4*, 26–31.
64. Beck, A.; Teboulle, M. A fast iterative shrinkage-thresholding algorithm for linear inverse problems. *SIAM J. Imaging Sci.* **2009**, *2*, 183–202.
65. O'Donoghue, B.; Candes, E. Adaptive restart for accelerated gradient schemes. *Found. Comput. Math.* **2015**, *15*, 715–732.
66. Bubeck, S.; Lee, Y.T.; Singh, M. A geometric alternative to Nesterov's accelerated gradient descent. *arXiv* **2015**, arXiv:1506.08187.
67. Goh, G. Why Momentum Really Works. *Distill* **2017**, *2* e6. https://doi.org/10.23915/distill.00006.
68. Kyrillidis, A.; Cevher, V. Recipes on hard thresholding methods. In Proceedings of the Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP), San Juan, PR, USA, 13–16 December 2011; pp. 353–356.
69. Xu, P.; He, B.; De Sa, C.; Mitliagkas, I.; Re, C. Accelerated stochastic power iteration. In Proceedings of the International Conference on Artificial Intelligence and Statistics, Playa Blanca, Lanzarote, Canary Islands, 9–11 April 2018; pp. 58–67.
70. Ghadimi, S.; Lan, G. Stochastic first-and zeroth-order methods for nonconvex stochastic programming. *SIAM J. Optim.* **2013**, *23*, 2341–2368.
71. Lee, J.; Simchowitz, M.; Jordan, M.; Recht, B. Gradient descent only converges to minimizers. In Proceedings of the Conference on Learning Theory, New York, NY, USA, 23–26 June 2016; pp. 1246–1257.
72. Agarwal, N.; Allen-Zhu, Z.; Bullins, B.; Hazan, E.; Ma, T. Finding approximate local minima for nonconvex optimization in linear time. *arXiv* **2016**, arXiv:1611.01146.
73. Burer, S.; Monteiro, R.D. A nonlinear programming algorithm for solving semidefinite programs via low-rank factorization. *Math. Program.* **2003**, *95*, 329–357.
74. Jain, P.; Dhillon, I.S. Provable inductive matrix completion. *arXiv* **2013**, arXiv:1306.0626.
75. Chen, Y.; Wainwright, M.J. Fast low-rank estimation by projected gradient descent: General statistical and algorithmic guarantees. *arXiv* **2015**, arXiv:1509.03025.
76. Sun, R.; Luo, Z.Q. Guaranteed matrix completion via non-convex factorization. *IEEE Trans. Inf. Theory* **2016**, *62*, 6535–6579.
77. O'Donnell, R.; Wright, J. Efficient quantum tomography. In Proceedings of the Forty-Eighth Annual ACM Symposium on Theory of Computing, Cambridge, MA, USA, 19–21 June 2016; pp. 899–912.
78. Hayashi, M.; Matsumoto, K. Quantum universal variable-length source coding. *Phys. Rev. A* **2002**, *66*, 022311.
79. Christandl, M.; Mitchison, G. The spectra of quantum states and the Kronecker coefficients of the symmetric group. *Commun. Math. Phys.* **2006**, *261*, 789–797.
80. Alicki, R.; Rudnicki, S.; Sadowski, S. Symmetry properties of product states for the system of N n-level atoms. *J. Math. Phys.* **1988**, *29*, 1158–1162.
81. Keyl, M.; Werner, R.F. Estimating the spectrum of a density operator. In *Asymptotic Theory of Quantum Statistical Inference: Selected Papers*; World Scientific: Singapore, 2005; pp. 458–467.
82. Tóth, G.; Wieczorek, W.; Gross, D.; Krischek, R.; Schwemmer, C.; Weinfurter, H. Permutationally invariant quantum tomography. *Phys. Rev. Lett.* **2010**, *105*, 250403.
83. Moroder, T.; Hyllus, P.; Tóth, G.; Schwemmer, C.; Niggebaum, A.; Gaile, S.; Gühne, O.; Weinfurter, H. Permutationally invariant state reconstruction. *New J. Phys.* **2012**, *14*, 105001.
84. Schwemmer, C.; Tóth, G.; Niggebaum, A.; Moroder, T.; Gross, D.; Gühne, O.; Weinfurter, H. Experimental comparison of efficient tomography schemes for a six-qubit state. *Phys. Rev. Lett.* **2014**, *113*, 040503.
85. Banaszek, K.; D'Ariano, G.M.; Paris, M.G.A.; Sacchi, M.F. Maximum-likelihood estimation of the density matrix. *Phys. Rev. A* **1999**, *61*, 010304.
86. Paris, M.; D'Ariano, G.; Sacchi, M. Maximum-likelihood method in quantum estimation. In Proceedings of the AIP Conference Proceedings, 2001; Volume 568, pp. 456–467.
87. Řeháček, J.; Hradil, Z.; Knill, E.; Lvovsky, A.I. Diluted maximum-likelihood algorithm for quantum tomography. *Phys. Rev. A* **2007**, *75*, 042108. https://doi.org/10.1103/PhysRevA.75.042108.
88. Gonçalves, D.; Gomes-Ruggiero, M.; Lavor, C.; Farias, O.J.; Ribeiro, P. Local solutions of maximum likelihood estimation in quantum state tomography. *Quantum Inf. Comput.* **2012**, *12*, 775–790.
89. Teo, Y.S.; Řeháček, J.; Hradil, Z. Informationally incomplete quantum tomography. *Quantum Meas. Quantum Metrol.* **2013**, *1*, 57–83. https://doi.org/10.2478/qmetro-2013-0006.

90. Sutskever, I.; Hinton, G.E.; Taylor, G.W. The recurrent temporal restricted boltzmann machine. In Proceedings of the Advances in Neural Information Processing Systems, Vancouver, BC, Canada, 7–10 December 2009; pp. 1601–1608.
91. Ahmed, S.; Muñoz, C.S.; Nori, F.; Kockum, A. Quantum state tomography with conditional generative adversarial networks. *arXiv* **2020**, arXiv:2008.03240.
92. Ahmed, S.; Muñoz, C.; Nori, F.; Kockum, A. Classification and reconstruction of optical quantum states with deep neural networks. *arXiv* **2020**, arXiv:2012.02185.
93. Paini, M.; Kalev, A.; Padilha, D.; Ruck, B. Estimating expectation values using approximate quantum states. *Quantum* **2021**, *5*, 413. https://doi.org/10.22331/q-2021-03-16-413.
94. Huang, H.Y.; Kueng, R.; Preskill, J. Predicting Many Properties of a Quantum System from Very Few Measurements. *arXiv* **2020**, arXiv:2002.08953.
95. Flammia, S.T.; Liu, Y.K. Direct fidelity estimation from few Pauli measurements. *Phys. Rev. Lett.* **2011**, *106*, 230501.
96. da Silva, M.P.; Landon-Cardinal, O.; Poulin, D. Practical characterization of quantum devices without tomography. *Phys. Rev. Lett.* **2011**, *107*, 210404.
97. Kalev, A.; Kyrillidis, A.; Linke, N.M. Validating and certifying stabilizer states. *Phys. Rev. A* **2019**, *99*, 042337.
98. Aaronson, S. Shadow tomography of quantum states. In Proceedings of the 50th Annual ACM SIGACT Symposium on Theory of Computing, Los Angeles, CA, USA, 25–29 June 2018; pp. 325–338.
99. Aaronson, S.; Rothblum, G.N. Gentle measurement of quantum states and differential privacy. In Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing, Phoenix, AZ, USA, 23–26 June 2019; pp. 322–333.
100. Smith, A.; Gray, J.; Kim, M. Efficient Approximate Quantum State Tomography with Basis Dependent Neural-Networks. *arXiv* **2020**, arXiv:2009.07601.
101. Waters, A.E.; Sankaranarayanan, A.C.; Baraniuk, R. SpaRCS: Recovering low-rank and sparse matrices from compressive measurements. In Proceedings of the Advances in Neural Information Processing Systems, Granada, Spain, 12–14 December 2011; pp. 1089–1097.
102. He, B.; Yuan, X. On the convergence rate of Douglas–Rachford operator splitting method. *Math. Program.* **2015**, *153*, 715–722.
103. O'Donoghue, B. Operator Splitting for a Homogeneous Embedding of the Linear Complementarity Problem. *SIAM J. Optim.* **2021**, *31*, 1999–2023.
104. Foucart, S. Matrix Norms and Spectral Radii. 2012. Available online: https://www.math.drexel.edu/~foucart/TeachingFiles/F12/M504Lect6.pdf (accessed on 20 January 2023).
105. Johnson, S.G. Notes on the Equivalence of Norms. 2012. Available online: https://math.mit.edu/~stevenj/18.335/norm-equivalence.pdf (accessed on 20 January 2023).
106. Horn, R.; Johnson, C. *Matrix Analysis*; Cambridge University Press: Cambridge, UK, 1990.
107. Mirsky, L. A trace inequality of John von Neumann. *Monatshefte Math.* **1975**, *79*, 303–306.