Fast Minimization of Expected Logarithmic Loss via Stochastic Dual Averaging

Chung-En Tsai[1], Hao-Chung Cheng[2,3,4,5], and Yen-Huan Li[1,3,5]

[1]Department of Computer Science and Information Engineering, National Taiwan University
[2]Department of Electrical Engineering and Graduate Institute of Communication Engineering, National Taiwan University
[3]Department of Mathematics, National Taiwan University
[4]Hon Hai (Foxconn) Quantum Computing Centre
[5]Center for Quantum Science and Engineering, National Taiwan University

Abstract

Consider the problem of minimizing an expected logarithmic loss over either the probability simplex or the set of quantum density matrices. This problem encompasses tasks such as solving the Poisson inverse problem, computing the maximum-likelihood estimate for quantum state tomography, and approximating positive semi-definite matrix permanents with the currently tightest approximation ratio. Although the optimization problem is convex, standard iteration complexity guarantees for first-order methods do not directly apply due to the absence of Lipschitz continuity and smoothness in the loss function.

In this work, we propose a stochastic first-order algorithm named B-sample stochastic dual averaging with the logarithmic barrier. For the Poisson inverse problem, our algorithm attains an ε-optimal solution in $\tilde{O}(d^2/\varepsilon^2)$ time, matching the state of the art. When computing the maximum-likelihood estimate for quantum state tomography, our algorithm yields an ε-optimal solution in $\tilde{O}(d^3/\varepsilon^2)$ time, where d denotes the dimension. This improves on the time complexities of existing stochastic first-order methods by a factor of $d^{\omega-2}$ and those of batch methods by a factor of d^2, where ω denotes the matrix multiplication exponent. Numerical experiments demonstrate that empirically, our algorithm outperforms existing methods with explicit complexity guarantees.

1 Introduction

Denote by $\mathscr{D}_d$ the set of $d \times d$ quantum density matrices, i.e., the set of Hermitian positive semi-definite (PSD) matrices of unit traces. Let P be a distribution over the set of $d \times d$ Hermitian PSD matrices. Consider the optimization problem of minimizing an expected logarithmic loss:

$$f^\star = \min_{\rho \in \mathscr{D}_d} f(\rho), \quad f(\rho) \coloneqq \mathbb{E}_{A \sim P}[-\log \operatorname{tr}(A\rho)]. \tag{1}$$

A point $\hat{\rho} \in \mathscr{D}_d$ is said to be ε-optimal if $f(\hat{\rho}) - f^\star \leq \varepsilon$. When both A and ρ are restricted to diagonal matrices, the optimization problem (1) reduces to

$$\min_{x \in \Delta_d} f(x), \quad f(x) \coloneqq \mathbb{E}_{a \sim P'}[-\log \langle a, x \rangle] \tag{2}$$

where Δ_d denotes the probability simplex in $\mathbb{R}^d$ and P' is a distribution over $[0,\infty)^d$. We refer to the two problems (1) and (2) as the *quantum* setup and the *classical* setup, respectively. For the special cases when

$$f(\rho) \coloneqq \frac{1}{n}\sum_{i=1}^{n} -\log \operatorname{tr}(A_i\rho), \text{ or}$$
$$f(x) \coloneqq \frac{1}{n}\sum_{i=1}^{n} -\log \langle a_i, x\rangle,$$

we say the optimization problems are in the *finite-sum* setting with sample size n.

Notably, the classical setup (2) encompasses the problem of computing Kelly's criterion, an asymptotically optimal strategy in long-term investment [1, 2, 3]. It is also equivalent to solving the Poisson inverse problem, which finds applications in positron emission tomography (PET) in medical imaging and astronomical image denoising [4, 5]. Lastly, the problem is the batch counterpart of the online portfolio selection problem [6], for which designing algorithms that are optimal in both regret and computational complexity has remained unsolved for over thirty years [7, 8].

The quantum setup (1) has even broader applications. One important example is computing the maximum-likelihood (ML) estimate for quantum state tomography [9], a fundamental task for the verification of quantum devices. Another example is computing a semidefinite programming relaxation of the PSD matrix permanents [10]. The relaxation achieves the currently tightest approximation ratio and can be used to estimate the output probabilities of Boson sampling experiments [11].

Although the optimization problems (1, 2) are convex, standard convex optimization methods face two challenges. The first challenge is the lack of Lipschitz continuity and smoothness in the loss function [12]. As a result, iteration complexity guarantees of standard first-order methods, such as mirror descent and dual averaging, do not directly apply [13, 14]. While second-order methods, such as Newton's method, do possess explicit complexity guarantees [13], they still face the second challenge.

The second challenge is the scalabilities with respect to the dimension d and sample size n in the finite-sum setting. For instance, the dimension d typically reaches millions, and the sample size n can exceed hundreds of millions in PET [15, 16]. Both d and n grow exponentially with the number of qubits (quantum bits) in quantum state tomography [17]. However, the per-iteration time complexities of batch methods grow at least linearly with n, and those of second-order methods scale poorly with d as they require computing Hessian inverses.

When dealing with high dimensionality and large sample sizes, stochastic first-order algorithms, such as stochastic gradient descent, are preferred. Their per-iteration time complexities can be independent of the sample size n, and they do not require computationally demanding operations involving Hessian matrices. Nevertheless, standard stochastic first-order algorithms continue to face the challenge related to the lack of Lipschitz continuity and smoothness, as mentioned earlier.

Due to the absence of Lipschitz continuity and smoothness, mini-batch stochastic Q-Soft-Bayes (SQSB) [18, 19] and stochastic Q-LB-OMD (SQLBOMD) [20, 21] are the only two stochastic first-order methods with clear time complexity guarantees for solving the problems (1, 2). Both algorithms do not compete with batch methods in terms of the empirical convergence speed, as shown in Section 6. Stochastic mirror descent studied by D'Orazio et al. [22] is only guaranteed to solve the problems up to an arbitrarily large error[1]. Other stochastic first-order methods, such as stochastic primal-dual hybrid

[1] D'Orazio et al. [22, Theorem 3] proved that the average Bregman divergence to the minimizer asymp-

gradient (SPDHG) [23, 24], stochastic mirror-prox [25], and stochastic coordinate descent [26], are only guaranteed to converge asymptotically.

Contributions In this work, we propose a mini-batch stochastic first-order algorithm named B-sample stochastic dual averaging with the logarithmic barrier (LB-SDA, Algorithm 1) for solving the optimization problems (1, 2), where B denotes the mini-batch size. The expected optimization error of B-sample LB-SDA vanishes at a rate of $\tilde{O}(d/t + \sqrt{d/(Bt)})$. This matches the standard results of mini-batch stochastic gradient descent for minimizing smooth functions [27], regardless of the absence of smoothness in our problem.

In the classical setup, the time complexity of obtaining an ε-optimal solution via B-sample LB-SDA is $\tilde{O}(d^2/\varepsilon^2)$, matching the state of the art of stochastic first-order methods [19, 20]. In the quantum setup, the time complexity of obtaining an ε-optimal solution is $\tilde{O}(d^3/\varepsilon^2)$ when the mini-batch size is set to d. This improves the dimension dependence of existing stochastic algorithms by a factor of $d^{\omega-2}$, where $\omega \in [2, 2.372)$ denotes the matrix multiplication exponent [28]. It is worth noting that in practical implementation, such as BLAS [29], ω is effectively 3. The time complexity guarantee also improves that of the currently fastest batch algorithm by a factor of n/d. Such improvement is significant, given that $n = \Omega(d^3)$ is necessary for ML quantum state tomography [17].

Lastly, we conducted numerical experiments to demonstrate the efficiency of the proposed method. The numerical results suggest that 1-sample LB-SDA is the currently fastest method with explicit complexity guarantees for the Poisson inverse problem, and d-sample LB-SDA outperforms all methods in terms of fidelity, a standard measure of the closeness of quantum states, for ML quantum state tomography. To the best of our knowledge, this is the first empirical evidence that stochastic first-order algorithms can surpass batch ones in computing the ML estimate for quantum state tomography.

Technical Breakthroughs Our analysis consists of three key ingredients: a regret bound of Tsai et al. [30], a smoothness characterization of the logarithmic loss (Lemma 4), and a new local-norm-based analysis of the standard online-to-batch conversion [31], all of which are of independent interest.

Tsai et al. [30] proved the following regret bound for online convex optimization with the logarithmic loss on the probability simplex Δ_d (Appendix A.2):

$$R_t \leq \tilde{O}\left(\sqrt{d\sum_{\tau=1}^{t}\|\nabla f_\tau(\rho_\tau)\|_{\rho_\tau,*}^2}\right) \leq \tilde{O}\left(\sqrt{dt}\right), \tag{3}$$

where $\|\cdot\|_{\rho,*}$ is the dual local norm associated with the logarithmic barrier. Directly applying the standard online-to-batch conversion [31] with the second upper bound can only yield an optimization error bound of $\tilde{O}(\sqrt{d/t})$, which is independent of the mini-batch size. In comparison, we make use of the finer first upper bound and derive an optimization error bound of $\tilde{O}(d/t + \sqrt{d/(Bt)})$. This leads to a time complexity bound of $\tilde{O}(d^3/\varepsilon^2 + d^{\omega+1}/(B\varepsilon^2))$ for the quantum setup, which creates space for improved dimensional scalability via choosing the mini-batch size B. See Section 5.3 for a detailed discussion.

To make use of the finer first regret bound (3), we generalize the smoothness characterization of the logarithmic loss of Tsai et al. [30] for the quantum setup. The original proof of Tsai et al. [30] is challenging to generalize due to the noncommutativity in the

totically converges to σ_X^2, which equals $f^\star$ when, for example, $\|a_i\|_\infty = 1$ in the classical setup and can be arbitrarily large.

Algorithms	Iter. complexity	Per-iter. time	Time complexity
NoLips [33]	d/ε	$nd^2 + d^\omega$	$(nd^3 + d^{\omega+1})/\varepsilon$
QEM [18]	$\log d/\varepsilon$	$nd^2 + d^\omega$	$(nd^2 + d^\omega)/\varepsilon$
Frank-Wolfe [34]	n/ε	nd^ω	$n^2 d^\omega/\varepsilon$
SQSB [19]	d/ε^2	d^ω	$d^{\omega+1}/\varepsilon^2$
SQLBOMD [20]	d/ε^2	d^ω	$d^{\omega+1}/\varepsilon^2$
1-sample LB-SDA (Corollary 7)	d/ε^2	d^ω	$d^{\omega+1}/\varepsilon^2$
d-sample LB-SDA (Corollary 7)	$1/\varepsilon^2$	d^3	d^3/ε^2
B-sample LB-SDA (Corollary 7)	$d/(B\varepsilon^2)$	$Bd^2 + d^\omega$	$d^3/\varepsilon^2 + d^{\omega+1}/(B\varepsilon^2)$

Table 1: A comparison of existing first-order methods for the quantum setup (1) with explicit complexity guarantees. Iteration complexity and time complexity refer to the number of iterations and arithmetic operations required to obtain an ε-optimal solution, respectively. We assume $t \gg d^2$ and omit logarithmic factors, where t denotes the number of iterations.

quantum setup. Our generalization is based on a great simplification of their proof by utilizing self-concordance properties of the logarithmic loss (Appendix A.1).

Our analysis modifies that of the anytime online-to-batch conversion [32] to handle the local norms. It is worth noting that we also adapt the analysis of anytime online-to-batch for the standard one, as the latter has shown better empirical performance.

Notations We denote the set $\{1, 2, \ldots, n\}$ by $[\![n]\!]$ for a natural number $n \in \mathbb{N}$. We denote the ℓ_p-norm by $\|\cdot\|_p$ for $p \in [1, \infty]$. We denote the sets of $d \times d$ Hermitian matrices, Hermitian PSD matrices, and Hermitian positive definite matrices by $\mathbb{H}^d$, $\mathbb{H}^d_+$, and $\mathbb{H}^d_{++}$, respectively. We denote the relative interior of a set S by $\operatorname{ri} S$. We denote the i-th entry of a vector v by $v(i)$. We denote the conjugate transpose of a matrix U by U^*. We denote the sum of time-indexed matrices $A_1, \ldots, A_t \in \mathbb{H}^d$ by $A_{1:t}$. We define the domain of a function $f : \mathbb{H}^d \to \mathbb{R} \cup \{\infty\}$ by $\operatorname{dom} f \coloneqq \{\rho \in \mathbb{H}^d \mid f(\rho) < \infty\}$.

2 Related Work

The relationships between this work and the works of Tsai et al. [30] and Cutkosky [32] have been addressed in Section 1. This section focuses on optimization algorithms.

Although standard optimization methods are not suitable for solving the problems (1, 2), several methods with clear complexity guarantees have been proposed in the last decade. Table 1 summarizes existing results. We focus on batch methods as stochastic methods have already been discussed in Section 1.

QEM [18] is the current theoretically fastest batch method that solves the optimization problem (1) with clear complexity guarantees. Its classical counterpart EM was proposed by Shepp and Vardi [35] and Cover [36] independently. While NoLips, QEM, and Frank-Wolfe have clear complexity guarantees, their time complexities scale at least linearly with the sample size, which is undesirable when the sample size is large.

Other batch methods lack explicit complexity guarantees and, as a result, are not comparable to our algorithm. For instance, the convergence rates of proximal gradient methods [37] and several variants of the Frank-Wolfe method [38, 39, 40] involve unknown parameters. Diluted iterative MLE (iMLE) [41, 42] and entropic mirror descent (EMD)

with Armijo line search [12] are only guaranteed to converge asymptotically. Additionally, OSEM [43] for PET and iMLE [44] for ML quantum state tomography are commonly used heuristics but do not converge in general [43, 41].

3 Applications

3.1 Kelly's Criterion

Denote by Δ_d the probability simplex in $\mathbb{R}^d$. Consider long-term investment in a market with d investment alternatives. Let $\{a_t\}$ be a stochastic process taking values in $[0,\infty)^d$. On day t, the investor first selects a portfolio $x_t \in \Delta_d$ that indicates the distribution of their assets among the d investment alternatives. Then, the investor observes a_t that provides the price relatives of the investment alternatives for that day. The investor's goal is to maximize the wealth growth rate.

Kelly's criterion suggests choosing x_{t+1} by maximizing the expected logarithmic loss conditional on the past, i.e.,

$$x_{t+1} \in \operatorname*{argmin}_{x\in\Delta_d} \mathbb{E}_{a_{t+1}}[-\log\langle a_{t+1}, x\rangle \mid a_1, \dots, a_t],$$

which requires solving the classical setup (2).

3.2 Poisson Inverse Problem

In the Poisson inverse problem, our goal is to recover an unknown signal $\lambda^\natural \in [0,\infty)^d$ based on n independent measurement outcomes $\{y_i\}$. Each outcome y_i follows a Poisson distribution with mean $\langle b_i, \lambda^\natural\rangle$, where $b_i \in [0,\infty)^d$ is known and depends on the measurement setup. In positron emission tomography, $\lambda^\natural(i)$ represents the emitter density of the i-th region, and y_i represents the number of photons detected by the i-th sensor.

The ML estimate is given by [35]

$$\hat{\lambda} \in \operatorname*{argmin}_{\lambda\in[0,\infty)^d} \sum_{i=1}^{n} (\langle b_i, \lambda\rangle - y_i \log\langle b_i, \lambda\rangle). \tag{4}$$

Vardi and Lee [4] and Ben-Tal et al. [15] showed that by setting

$$Y = \sum_{i=1}^{n} y_i, \quad \hat{\lambda}(i) = \frac{Y\hat{x}(i)}{\sum_{j=1}^{n} a_j(i)}, \quad \forall i \in [\![n]\!],$$

and

$$a_i(j) = \frac{Y b_i(j)}{\sum_{k=1}^{n} b_k(j)}, \quad \forall i \in [\![n]\!],\ j \in [\![d]\!],$$

the ML estimate can be reformulated as

$$\hat{x} \in \operatorname*{argmin}_{x\in\Delta_d} \sum_{i=1}^{n} -\frac{y_i}{Y} \log\langle a_i, x\rangle,$$

which is equivalent to the classical setup (2) in the finite-sum setting with $\mathbb{P}(a = a_i) = y_i/Y$ for all $i \in [\![n]\!]$.

3.3 ML Quantum State Tomography

A quantum state is described by a density matrix $\rho \in \mathscr{D}_d$, which is a $d \times d$ Hermitian PSD matrix of unit trace. For a state consisting of q qubits, d equals 2^q. Denote by $\mathscr{D}_d$ the set of density matrices. The set $\mathscr{D}_d$ can be regarded as a quantum generalization of Δ_d, in the sense that the vector of eigenvalues of any $\rho \in \mathscr{D}_d$ lies in Δ_d.

Given n measurement outcomes from an unknown quantum state $\rho^\natural$, ML estimation is a standard and widely used approach to estimate $\rho^\natural$ [9, 45, 46, 47]. The ML estimate is given by

$$\hat{\rho}_{\mathrm{ML}} \in \operatorname*{argmin}_{\rho \in \mathscr{D}_d} \frac{1}{n} \sum_{i=1}^{n} -\log \operatorname{tr}(A_i \rho),$$

for some known $A_i \in \mathbb{H}^d_+$ related to the i-th measurement outcome. Note that solving $\hat{\rho}_{\mathrm{ML}}$ requires solving the quantum setup (1) in the finite-sum setting.

3.4 PSD Matrix Permanents

The permanent of a matrix $A \in \mathbb{C}^{d \times d}$ is defined as

$$\operatorname{per} A \coloneqq \sum_{\pi \in S_d} \sum_{i=1}^{d} A_{i,\pi(i)},$$

where S_d is the set of all permutations of $[\![d]\!]$. Let $\{v_i\}_{i=1}^d$ be the eigenvectors of A. Yuan and Parrilo [10] proposed the following approximation of $\operatorname{per} A$ when $A \in \mathbb{H}^d_+$:

$$\operatorname{rel} A \coloneqq \max_{\rho \in \mathscr{D}_d} \prod_{i=1}^{d} \operatorname{tr}((d v_i v_i^*)\rho).$$

The approximation is equivalent to the quantum setup (1) in the finite-sum setting with $A_i = d v_i v_i^*$. As noted by Meiburg [48], $\operatorname{rel} A$ achieves the currently tightest approximation ratio of 4.85^d.

4 Characterizations of Logarithmic Loss

This section aims to address the aforementioned lack of Lipschitz continuity and smoothness in the logarithmic loss. We first set up a few notations. For $\rho \in \mathbb{H}^d_{++}$, let

$$h(\rho) \coloneqq -\log\det \rho \tag{5}$$

be the *logarithmic barrier*. Let $\|\cdot\|_\rho \coloneqq (D^2 h(\rho)[\cdot,\cdot])^{1/2}$ be the *local norm* associated with h at $\rho \in \operatorname{dom} h$. The following lemma gives explicit formulae for the local norm and its dual norm. The proof is deferred to Appendix B.1.

Lemma 1. *For $\rho \in \mathbb{H}^d_{++}$ and $X \in \mathbb{H}^d$, the local norm and its dual norm associated with h are given by*

$$\begin{aligned} \|X\|_\rho &= \sqrt{\operatorname{tr}((\rho^{-1/2} X \rho^{-1/2})^2)} = \sqrt{\operatorname{tr}((\rho^{-1} X)^2)}, \\ \|X\|_{\rho,*} &= \sqrt{\operatorname{tr}((\rho^{1/2} X \rho^{1/2})^2)} = \sqrt{\operatorname{tr}((\rho X)^2)}. \end{aligned} \tag{6}$$

4.1 "Lipschitz Continuity"

A continuously differentiable function $f : \mathbb{H}^d \to \mathbb{R}$ is said to be G-Lipschitz with respect to a norm $\|\cdot\|$ if its gradients are bounded by G in the dual norm, i.e., $\|\nabla f(\rho)\|_* \leq G$.

Although the loss function is not Lipschitz, Lemma 2 below shows that ∇f is bounded in the dual local norm associated with h. This Lipschitz-type property enables us to control the distance between iterates and exploit local properties of the loss function, in particular, the local smoothness property of self-concordant functions (Theorem 11).

Lemma 2 is a simple quantum generalization of Lemma 4.3 of Tsai et al. [30]. Its proof is deferred to Appendix B.2.

Lemma 2. *Let f be defined in the quantum setup* (1)*. Then, $\|\nabla f(\rho)\|_{\rho,*} \leq 1$ for all $\rho \in \mathbb{H}^d_{++}$.*

4.2 "Smoothness"

A continuously differentiable function $f : \mathbb{H}^d \to \mathbb{R}$ is said to be L-smooth with respect to a norm $\|\cdot\|$ if its gradient is L-Lipschitz with respect to $\|\cdot\|$, i.e.,

$$\|\nabla f(\rho) - \nabla f(\rho')\|_* \leq L\|\rho - \rho'\|, \quad \forall \rho, \rho' \in \mathbb{H}^d.$$

Lemma 3, known as the self-bounding property, is a consequence of smoothness [49]. A proof of Lemma 3 can be found in Lemma 4.23 of Orabona [50].

Lemma 3. *Let $f : \mathbb{R}^d \to \mathbb{R}$ be L-smooth with respect to $\|\cdot\|$ with $\operatorname{dom} f = \mathbb{R}^d$. Then, for any $x \in \mathbb{R}^d$, it holds that*

$$\|\nabla f(x)\|_*^2 \leq 2L\left(f(x) - \inf_{x' \in \mathbb{R}^d} f(x')\right).$$

Although the loss function is not smooth, Lemma 4 below establishes a self-bounding-type property of the loss function. As discussed in Section 1, the lemma generalizes Lemma 4.7 of Tsai et al. [30] to the quantum setup, and greatly simplifies the proof therein. The proof is deferred to Appendix B.3.

For any $\rho \in \operatorname{ri} \mathscr{D}_d$ and $X \in \mathbb{H}^d$, define

$$\alpha_\rho(X) \coloneqq -\frac{\operatorname{tr}(\rho X \rho)}{\operatorname{tr}(\rho^2)} \in \operatorname*{argmin}_{\alpha \in \mathbb{R}} \|X + \alpha I\|_{\rho,*}^2. \tag{7}$$

Lemma 4. *Let f be defined in the quantum setup* (1)*. Then, for any $\rho \in \operatorname{ri} \mathscr{D}_d$, it holds that*

$$\|\nabla f(\rho) + \alpha_\rho(\nabla f(\rho))I\|_{\rho,*}^2 \leq 4\left(f(\rho) - \min_{\rho' \in \mathscr{D}_d} f(\rho')\right).$$

5 Algorithms and Convergence Guarantees

This section presents LB-SDA and its theoretical guarantee. We focus on the quantum setup (1) since it includes the classical setup (2) as a special case.

5.1 Algorithm

LB-SDA is presented in Algorithm 1, where h is the logarithmic barrier (5) and $\|\cdot\|_\rho$ and $\|\cdot\|_{\rho,*}$ are the local norm and the dual local norm associated with h at ρ (6), respectively. A *stochastic first-order oracle* is a randomized function $\mathcal{O}$ that outputs an unbiased estimate $\mathcal{O}(\rho) \in \mathbb{H}^d$ of the gradient $\nabla f(\rho)$ given an input $\rho \in \mathscr{D}_d$.

We will make the following assumptions on the stochastic first-order oracle. It is notable that the boundedness is defined in terms of the dual local norm, which deviates from existing literature.

Algorithm 1 Stochastic Dual Averaging with the Logarithmic Barrier (LB-SDA) for the quantum setup (1)

Input: A stochastic first-order oracle $\mathcal{O}$.

$h(\rho) \coloneqq -\log\det\rho$.
$\rho_1 = I/d \in \operatorname{argmin}_{\rho\in\mathscr{D}_d} h(\rho)$.
for all $t \in \mathbb{N}$ **do**
  Output $\bar{\rho}_t \coloneqq (1/t)\rho_{1:t}$.
  $g_t = \mathcal{O}(\rho_t)$.
  Compute a learning rate $\eta_t > 0$.
  $\rho_{t+1} \in \operatorname{argmin}_{\rho\in\mathscr{D}_d} \eta_t \operatorname{tr}(g_{1:t}\rho) + h(\rho)$.
end for

Assumption 1. *Conditional on the past, the stochastic gradients $\{g_t\}$ in Algorithm 1 are unbiased and bounded, and their variances are also bounded, i.e., for all $t \in \mathbb{N}$,*

- $\mathbb{E}[g_t|\mathcal{H}_t] = \nabla f(\rho_t)$,
- $\mathbb{E}\left[\|g_t\|^2_{\rho_t,*}\middle|\mathcal{H}_t\right] \le G^2$,
- $\mathbb{E}\left[\|g_t - \nabla f(\rho_t)\|^2_{\rho_t,*}\middle|\mathcal{H}_t\right] \le \sigma^2$,

where $\mathcal{H}_t = \{g_1, \dots, g_{t-1}, \rho_1, \dots, \rho_t\}$ is the past information before obtaining g_t.

The unbiasedness and bounded variance assumptions are standard in the literature. Regarding the bounded gradient assumption, by the triangle inequality and Lemma 2,

$$\begin{aligned}\|g_t\|^2_{\rho_t,*} &\le (\|g_t - \nabla f(\rho_t)\|_{\rho_t,*} + \|\nabla f(\rho_t)\|_{\rho_t,*})^2 \\ &\le \|g_t - \nabla f(\rho_t)\|^2_{\rho_t,*} + 2\|g_t - \nabla f(\rho_t)\|_{\rho_t,*} + 1.\end{aligned}$$

Taking expectations on both sides and using the inequality $\mathbb{E}X \le \sqrt{\mathbb{E}[X^2]}$, we can verify that the bounded gradient assumption always holds with $G = 1 + \sigma$. Nevertheless, since G can be smaller than $1 + \sigma$, we include the assumption for a tighter result.

An important example of the oracle is

$$\mathcal{O}_B(\rho) \coloneqq \frac{1}{B}\sum_{b=1}^{B} \nabla\ell_b(\rho) \tag{8}$$

where $B \in \mathbb{N}$, $\ell_b(\rho) \coloneqq -\log\operatorname{tr}(A_b\rho)$, and $A_1, \dots, A_B$ are independently drawn from P. The resulting algorithm is called *B-sample LB-SDA*. The following lemma justifies the use of $\mathcal{O}_B$, whose proof is deferred to Appendix B.4.

Lemma 5. *The oracle $\mathcal{O}_B$ (8) satisfies Assumption 1 with $G = 1$ and $\sigma^2 = 4/B$.*

5.2 Convergence Guarantee

The non-asymptotic convergence guarantee of Algorithm 1 is presented in Theorem 6 below. The analysis follows the online-to-batch approach, where we use the following regret bound of Tsai et al. [30] in Appendix A.2:

$$R_t \le \tilde{O}\left(\sqrt{d\sum_{\tau=1}^{t}\|g_\tau + \alpha_{\rho_\tau}(g_\tau)I\|^2_{\rho_\tau,*}}\right) \le \tilde{O}(\sqrt{dt}).$$

Note that applying the online-to-batch conversion on the right upper bound can only yield a convergence rate of $\tilde{O}(\sqrt{d/t})$, independent of the variance σ^2. Since the effect of batch

size is unclear without the variance term, the direct approach fails to improve the time complexity guarantee.

Deriving the variance term typically requires smoothness of the loss function in the literature, and this is where the self-bounding-type property (Lemma 4) comes into play. It bounds the square of the dual local norm by

$$\mathbb{E}\|g_\tau + \alpha_{\rho_\tau}(g_\tau)I\|^2_{\rho_\tau,*} \leq 4\mathbb{E}\left[f(\rho_\tau) - \min_{\rho\in\mathscr{D}_d} f(\rho)\right] + \sigma^2,$$

which results in a "self-bounding" inequality of $\mathbb{E}R_t$:

$$\mathbb{E}R_t \leq \tilde{O}\left(\sqrt{d\mathbb{E}R_t + \sigma^2 dt}\right)$$

Our analysis can be seen as a local-norm extension of that of the anytime online-to-batch conversion [32]. The proof is deferred to Appendix B.5.

Theorem 6. *Consider the quantum setup* (1)*. Under Assumption 1, let* $\{\bar{\rho}_t\}$ *be the iterates generated by Algorithm 1 with*

$$\eta_t = \frac{\sqrt{d}}{\sqrt{\sum_{\tau=1}^t \|g_\tau + \alpha_{\rho_\tau}(g_\tau)I\|^2_{\rho_\tau,*} + 4dG^2 + G^2}}.$$

Then, for all $t \in \mathbb{N}$*, it holds that*

$$\begin{aligned}
&\mathbb{E}\left[f(\bar{\rho}_t) - \min_{\rho\in\mathscr{D}_d} f(\rho)\right] \\
&\quad\leq \frac{4dC_t^3 + 2C_t\sqrt{\sigma^2 dt + 4d^2G^2 + dG^2} + 1}{t} \\
&\quad= O\left(\frac{dG(\log t)^3}{t} + \frac{\sigma\sqrt{d}\log t}{\sqrt{t}}\right),
\end{aligned}$$

where $C_t \coloneqq \log t + 3$ *and the expectation is taken with respect to* $\{g_t\}$*.*

Plugging in the estimates in Lemma 5, we obtain the following result for the mini-batch case.

Corollary 7. *Consider the quantum setup* (1)*. Let* $\{\bar{\rho}_t\}$ *be the iterates of B-sample LB-SDA with learning rates*

$$\eta_t = \frac{\sqrt{d}}{\sqrt{\sum_{\tau=1}^t \|g_\tau + \alpha_{\rho_\tau}(g_\tau)I\|^2_{\rho_\tau,*} + 4d + 1}}.$$

Then, for all $t \in \mathbb{N}$*, it holds that*

$$\mathbb{E}\left[f(\bar{\rho}_t) - \min_{\rho\in\mathscr{D}_d} f(\rho)\right] = O\left(\frac{d(\log t)^3}{t} + \frac{\sqrt{d}\log t}{\sqrt{Bt}}\right).$$

where the expectation is taken with respect to $\{g_t\}$*.*

Remark 8. *Proving high-probability guarantees for B-sample LB-SDA is challenging since the logarithmic loss violates the boundedness assumption required by standard analysis [50]. We left this extension as a future research direction.*

5.3 Time Complexity Analysis

This section discusses the time complexity of B-sample LB-SDA and compares it with existing methods. Comparisons of time complexities of existing first-order methods have been presented in Table 1 and discussed in Section 1.

First, note that the 6th line in Algorithm 1 cannot be solved exactly. Nevertheless, after an eigendecomposition, which takes $\tilde{O}(d^\omega)$ time [51], the 6th line reduces to an one-dimensional convex optimization problem, which can be efficiently solved by Newton's method on the real line in $\tilde{O}(d)$ time (see, e.g., Appendix A.2 of Nesterov [13]). As a result, the time complexity of the 6th line is $\tilde{O}(d^\omega)$. Second, the time complexity of the 5th line, which requires implementing the oracle $\mathcal{O}_B$, is $O(Bd^2)$. Lastly, since the 5th and the 6th lines are the most time-consuming parts, the per-iteration time complexity of B-sample LB-SDA is $O(Bd^2 + d^\omega)$.

By Corollary 7, the iteration complexity of B-sample LB-SDA to obtain an ε-optimal solution is $\tilde{O}(d/(B\varepsilon^2))$. Combining with the per-iteration time complexity, the overall time complexity is $\tilde{O}(d^3/\varepsilon^2 + d^{\omega+1}/(B\varepsilon^2))$. In particular, the overall time complexity is $\tilde{O}(d^3/\varepsilon^2)$ when $B = \Omega(d^{\omega-2})$. Since ω is 3 in practical implementation [29], we will often choose $B = d$.

Since the eigendecomposition is no longer needed in the classical setup, the per-iteration time complexity of B-sample LB-SDA is reduced to $\tilde{O}(Bd)$. Because the iteration complexity of obtaining an ε-optimal solution is $\tilde{O}(d/(B\varepsilon^2))$, the overall time complexity is $\tilde{O}(d^2/\varepsilon^2)$ for any $B \in \mathbb{N}$ in the classical setup.

6 Numerical Results

We have shown that LB-SDA achieves the currently best time complexity guarantees in the previous section. In this section, we show that LB-SDA also performs well empirically. We consider solving the Poisson inverse problem and computing the ML estimate for quantum state tomography. All results in this section are presented in terms of the elapsed time. Results in terms of the number of iterations can be found in Appendix C.

Both experiments were conducted on a machine with an Intel Xeon Gold 5218 CPU of 2.30GHz and 131,621,512kB memory. The elapsed time records the actual running time of the method on the machine. All methods are implemented in the Julia programming language [52] with the Intel Math Kernel Library, and the number of threads in BLAS is set to 8. It is important to note that the empirical speed is highly dependent on the specific implementations. The source code of the experiments is available at `https://github.com/chungentsai/pip` and `https://github.com/chungentsai/mlqst` for the Poisson inverse problem and ML quantum state tomography, respectively.

The approximate optimization error at an iterate is defined as the difference between its function value and the smallest one obtained in the experiments.

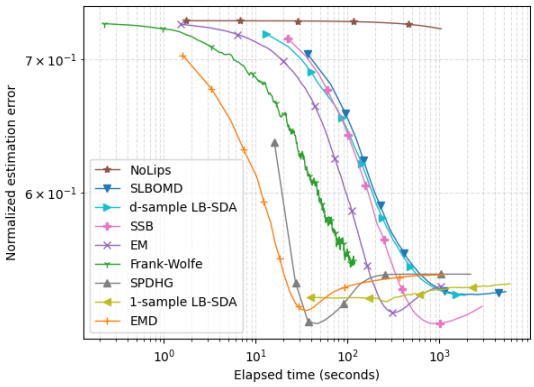
6.1 Poisson Inverse Problem

Consider the Poisson inverse problem in Section 3.2 with a synthetic dataset, where d equals 256 and n equals $1,000,000$. The unknown signal $\lambda^\natural$ is the gray intensities of the Shepp-Logan phantom image [53] of size 16×16 multiplied by $1,000$. The signal is presented in Appendix C. The vectors $\{b_i\}$ are generated following the scheme of Raginsky et al. [54]. Each entry of b_i is assigned to either 0 or $1/n$ with equal probability.

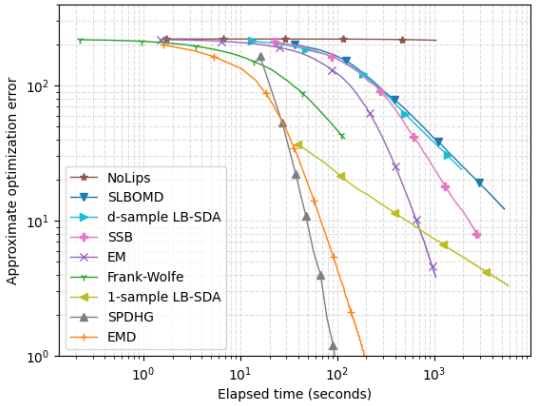
We consider all algorithms in Table 1 that have explicit complexity guarantees. EM [35], SSB [55], and SLBOMD [20] are the classical counterparts of QEM, SQSB, and SQLBOMD, respectively. Additionally, we include SPDHG [23] and EMD with Armijo line search [56] for comparison. The former is well-known in practice and the latter

Figure 1: Performances of all algorithms in Table 1, SPDHG, and EMD with line search for solving the Poisson inverse problem.

(a) Normalized estimation error versus the elapsed time.



(b) Approximate optimization error versus the elapsed time.



is known to converge fast empirically. However, they are only guaranteed to converge asymptotically. Their parameters are set according to the cited works. We do not include batch PDHG as it is slow in practice.

We solve the Poisson inverse problem based on the equivalence between it and the classical setup (2) in Section 3.2. Figure 1 presents the numerical results. For an iterate $\hat{\lambda}$, the normalized estimation error is defined as $\|\hat{\lambda} - \lambda^\natural\|_2 / \|\lambda^\natural\|_2$. Since the goal of the Poisson inverse problem is to recover the unknown signal $\lambda^\natural$, rather than minimizing the loss function, results presented in terms of the normalized estimation error is more important than results presented in terms of the optimization error.

Observe that 1-sample LB-SDA outperforms all methods with explicit complexity guarantees in terms of the normalized estimation error. Although it is slower than EMD with line search and SPDHG, the latter two methods are only guaranteed to converge asymptotically, whereas LB-SDA has an explicit non-asymptotic complexity guarantee.

LB-SDA converges faster than SLBOMD and SSB in terms of the optimization error, although they have the same theoretical time complexity of $\tilde{O}(d^2/\varepsilon^2)$. This can be explained by the use of time-varying learning rates in LB-SDA, in contrast to the fixed learning rates used by the other two methods. The time-varying learning rates are large in the beginning, which leads to a fast convergence in practice.

6.2 ML Quantum State Tomography

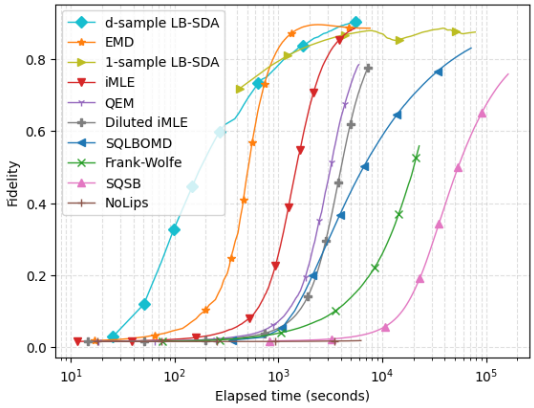
Consider the problem of ML quantum state tomography in Section 3.3. We construct a synthetic dataset, following the setup of Häffner et al. [45]. The number of qubits q is 6, the dimension d is $2^6 = 64$, and the sample size n is 409,600. The unknown quantum state is the W state, which corresponds to a rank-1 density matrix. The Hermitian matrices $\{A_i\}$ are generated following the procedure of Lin et al. [18], where each A_i is of rank $d/2$.

We compare all algorithms in Table 1, along with iMLE [44], diluted iMLE [42], and EMD with Armijo line search [56]. Their parameters are set according to the cited works. Although iMLE does not always converge [41], we include it because it is often considered as a benchmark. We do not include the accelerated projected gradient descent [57] as it is slower than iMLE in experiments [58].

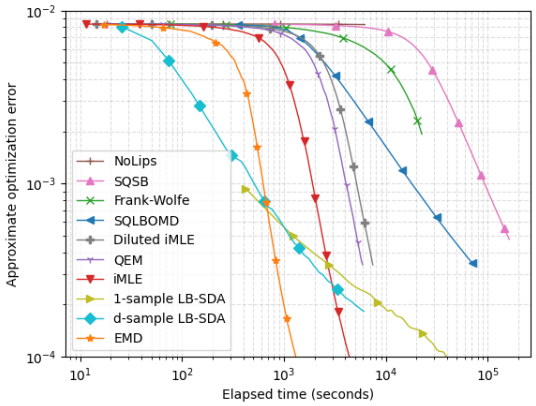
Figure 2 presents the numerical results. The fidelity between two quantum states $\rho, \rho' \in \mathscr{D}_d$ is defined as $F(\rho, \rho') := \left(\operatorname{tr} \sqrt{\sqrt{\rho}\rho'\sqrt{\rho}}\right)^2 \in [0, 1]$. It is a standard measure of the closeness of two quantum states, with $F(\rho, \rho') = 1$ if and only if $\rho = \rho'$. Similar to

Figure 2: Performances of all algorithms in Table 1, iMLE, diluted iMLE, and EMD with line search for computing the ML estimate for quantum state tomography.

(a) Fidelity between the iterates and the W state versus the elapsed time.



(b) Approximate optimization error versus the elapsed time.



the Poisson inverse problem, as the goal of quantum state tomography is to recover the unknown quantum state, results presented in terms of the fidelity is more important than results presented in terms of the optimization error.

Observe that d-sample LB-SDA outperforms all methods in terms of the fidelity. We conclude that d-sample LB-SDA achieves the currently best theoretical time complexity and the currently best empirical performance.

Note that d-sample LB-SDA performs better than SQLBOMD and SQSB in terms of the optimization error. It also outperforms QEM and diluted iMLE when the optimization error is not smaller than 10^{-3}. Recall that the time complexity of QEM has a worse dimension dependence and a better optimization error dependence than that of d-sample LB-SDA, and diluted iMLE lacks a non-asymptotic complexity guarantee. Although d-sample LB-SDA is slower than EMD with line search and iMLE in terms of the optimization error, the latter two methods lack clear complexity results.

While it is theoretically known that stochastic methods outperform batch ones when the dimension and the sample size are sufficiently large [59], empirical results presented in the literature did not confirm this phenomenon. In this work, we observed that d-sample LB-SDA outperforms all methods in terms of the fidelity. This marks the first empirical evidence that stochastic methods can be more efficient than batch methods for computing the ML estimate for quantum state tomography.

7 Concluding Remarks

We have proposed a stochastic first-order method named B-sample LB-SDA for solving the Poisson inverse problem, computing the ML estimate for quantum state tomography, and approximating PSD matrix permanents. In particular, d-sample LB-SDA takes $\tilde{O}(d^3/\varepsilon^2)$ time to obtain an ε-optimal solution in the quantum setup, improving the time complexities of existing first-order methods. The improvement is based on a new analysis for mini-batch methods, which relies on a novel self-bounding-type property of the logarithmic loss and a new local-norm based analysis of the online-to-batch conversion. Lastly, we have shown that LB-SDA performs better empirically than all methods with explicit complexity guarantees.

Several research directions arise. One direction is to design accelerated or variance-reduced methods for solving the optimization problem (1) based on the smoothness char-

acterization. Another direction is to generalize our argument to other non-smooth loss functions.

Acknowledgements

C.-E. Tsai and Y.-H. Li are supported by the Young Scholar Fellowship (Einstein Program) of the National Science and Technology Council of Taiwan under grant number NSTC 112-2636-E-002-003, by the 2030 Cross-Generation Young Scholars Program (Excellent Young Scholars) of the National Science and Technology Council of Taiwan under grant number NSTC 112-2628-E-002-019-MY3, by the research project "Pioneering Research in Forefront Quantum Computing, Learning and Engineering" of National Taiwan University under grant number NTU-CC-112L893406, and by the Academic Research-Career Development Project (Laurel Research Project) of National Taiwan University under grant number NTU- CDP-112L7786.

H.-C. Cheng is supported by Grants No. NSTC 112-2636-E-002-009, No. NSTC 112-2119-M-007-006, No. NSTC 112-2119-M-001-006, No. NSTC 112-2124-M-002-003, No. NTU-112V1904-4, No. NTU-112L900702, and by the research project "Pioneering Research in Forefront Quantum Computing, Learning and Engineering" of National Taiwan University under Grant No. NTC-CC-112L893405.

References

[1] John L. Kelly. A new interpretation of information rate. *Bell Syst. tech. j.*, 35(4): 917–926, 1956.

[2] Paul H. Algoet and Thomas M. Cover. Asymptotic optimality and asymptotic equipartition properties of log-optimum investment. *Ann. Probab.*, pages 876–898, 1988.

[3] Leonard C. MacLean, Edward O. Thorp, and William T. Ziemba. *The Kelly Capital Growth Investment Criterion: Theory and Practice*, volume 3. World Scientific, 2011.

[4] Y. Vardi and D. Lee. From image deblurring to optimal investments: Maximum likelihood solutions for positive linear inverse problems. *J. R. Stat. Soc. Series B Stat. Methodol.*, 55(3):569–598, 1998.

[5] Mario Bertero, Patrizia Boccacci, and Valeria Ruggiero. *Inverse Imaging with Poisson Data*. 2053-2563. IOP Publishing, 2018.

[6] Thomas M. Cover. Universal portfolios. *Math. Financ.*, 1(1):1–29, 1991.

[7] Tim van Erven, Dirk van der Hoeven, Wojciech Kotłowski, and Wouter M. Koolen. Open problem: Fast and optimal online portfolio selection. In *Proc. 33rd Annu. Conf. Learning Theory*, pages 3864–3869, 2020.

[8] Julian Zimmert, Naman Agarwal, and Satyen Kale. Pushing the efficiency-regret Pareto frontier for online learning of portfolios and quantum states. In *Proc. 35th Annu. Conf. Learning Theory*, pages 182–226, 2022.

[9] Z. Hradil. Quantum-state estimation. *Phys. Rev. A*, 55, 1997.

[10] Chenyang Yuan and Pablo A. Parrilo. Maximizing products of linear forms, and the permanent of positive semidefinite matrices. *Math. Program.*, 193(1):499–510, 2022.

[11] Scott Aaronson and Alex Arkhipov. The computational complexity of linear optics. In *Proc. 43rd Annu. ACM Symp. Theory of Computing*, pages 333–342, 2011.

[12] Yen-Huan Li and Volkan Cevher. Convergence of the exponentiated gradient method with Armijo line search. *J. Optim. Theory Appl.*, 181(2):588–607, 2019.

[13] Yurii Nesterov. *Lectures on Convex Optimization.* Springer, Cham, CH, second edition, 2018.

[14] Guanghui Lan. *First-order and Stochastic Optimization Methods for Machine Learning*, volume 1. Springer, 2020.

[15] Aharon Ben-Tal, Tamar Margalit, and Arkadi Nemirovski. The ordered subsets mirror descent optimization method with applications to tomography. *SIAM J. Optim.*, 12 (1):79–108, 2001.

[16] Matthias J. Ehrhardt, Pawel Markiewicz, Antonin Chambolle, Peter Richtárik, Jonathan Schott, and Carola-Bibiane Schönlieb. Faster PET reconstruction with a stochastic primal-dual hybrid gradient method. In *Wavelets and Sparsity XVII*, volume 10394, page 103941O. SPIE, 2017.

[17] Sitan Chen, Brice Huang, Jerry Li, Allen Liu, and Mark Sellke. When does adaptivity help for quantum state learning? In *IEEE 64th Annu. Symp. Foundations of Computer Science*, 2023.

[18] Chien-Ming Lin, Hao-Chung Cheng, and Yen-Huan Li. Maximum-likelihood quantum state tomography by Cover's method with non-asymptotic analysis. *arXiv preprint arXiv:2110.00747*, 2021.

[19] Chien-Ming Lin, Yu-Ming Hsu, and Yen-Huan Li. Maximum-likelihood quantum state tomography by soft-bayes. *arXiv preprint arXiv:2012.15498*, 2022.

[20] Chung-En Tsai, Hao-Chung Cheng, and Yen-Huan Li. Faster stochastic first-order method for maximum-likelihood quantum state tomography. *arXiv preprint arXiv:2211.12880*, 2022.

[21] Filip Hanzely and Peter Richtárik. Fastest rates for stochastic mirror descent methods. *Comput. Optim. Appl.*, 79(3):717–766, 2021.

[22] Ryan D'Orazio, Nicolas Loizou, Issam Laradji, and Ioannis Mitliagkas. Stochastic mirror descent: Convergence analysis and adaptive variants via the mirror stochastic polyak stepsize. *arXiv preprint arXiv:2110.15412*, 2021.

[23] Antonin Chambolle, Matthias J. Ehrhardt, Peter Richtárik, and Carola-Bibiane Schönlieb. Stochastic primal-dual hybrid gradient algorithm with arbitrary sampling and imaging applications. *SIAM J. Optim.*, 28(4):2783–2808, 2018.

[24] Ahmet Alacaoglu, Olivier Fercoq, and Volkan Cevher. On the convergence of stochastic primal-dual hybrid gradient. *SIAM J. Optim.*, 32(2):1288–1318, 2022.

[25] Niao He, Zaid Harchaoui, Yichen Wang, and Le Song. Point process estimation with mirror prox algorithms. *Appl. Math. Optim.*, 82(3):919–947, 2020.

[26] Olivier Fercoq and Pascal Bianchi. A coordinate-descent primal-dual algorithm with large step size and possibly nonseparable functions. *SIAM J. Optim.*, 29(1):100–134, 2019.

[27] Ofer Dekel, Ran Gilad-Bachrach, Ohad Shamir, and Lin Xiao. Optimal distributed online prediction using mini-batches. *J. Mach. Learn. Res.*, 13(1), 2012.

[28] Virginia Vassilevska Williams, Yinzhan Xu, Zixuan Xu, and Renfei Zhou. New bounds for matrix multiplication: from alpha to omega. In *Proc. Annu. ACM-SIAM Symp. Discrete Algorithms*, pages 3792–3835, 2024.

[29] J. J. Dongarra, Jeremy Du Croz, Sven Hammarling, and I. S. Duff. A set of level 3 basic linear algebra subprograms. *ACM Trans. Math. Softw.*, 16(1):1–17, 1990.

[30] Chung-En Tsai, Ying-Ting Lin, and Yen-Huan Li. Data-dependent bounds for online portfolio selection without Lipschitzness and smoothness. *arXiv preprint arXiv:2305.13946*, 2023.

[31] N. Cesa-Bianchi, A. Conconi, and C. Gentile. On the generalization ability of on-line learning algorithms. *IEEE Trans. Inf. Theory*, 50(9):2050–2057, 2004.

[32] Ashok Cutkosky. Anytime online-to-batch, optimism and acceleration. In *Proc. 36th Int. Conf. Machine Learning*, volume 97, pages 1446–1454, 2019.

[33] Heinz H. Bauschke, Jérôme Bolte, and Marc Teboulle. A descent lemma beyond Lipschitz gradient continuity: first-order methods revisited and applications. *Math. Oper. Res.*, 42(2):330–348, 2017.

[34] Renbo Zhao and Robert M. Freund. Analysis of the Frank–Wolfe method for convex composite optimization involving a logarithmically-homogeneous barrier. *Math. Program.*, 199(1):123–163, 2023.

[35] L. A. Shepp and Y. Vardi. Maximum likelihood reconstruction for emission tomography. *IEEE Trans. Med. Imaging*, 1(2):113–122, 1982.

[36] Thomas M. Cover. An algorithm for maximizing expected log investment return. *IEEE Trans. Inf. Theory*, 30(2):369–373, 1984.

[37] Quoc Tran-Dinh, Anastasios Kyrillidis, and Volkan Cevher. Composite self-concordant minimization. *J. Mach. Learn. Res.*, 16(12):371–416, 2015.

[38] Pavel Dvurechensky, Kamil Safin, Shimrit Shtern, and Mathias Staudigl. Generalized self-concordant analysis of Frank–Wolfe algorithms. *Math. Program.*, 198(1):255–323, 2023.

[39] Alejandro Carderera, Mathieu Besançon, and Sebastian Pokutta. Simple steps are all you need: Frank-Wolfe and generalized self-concordant functions. In *Adv. Neural Information Processing Systems*, volume 34, pages 5390–5401, 2021.

[40] Deyi Liu, Volkan Cevher, and Quoc Tran-Dinh. A Newton Frank–Wolfe method for constrained self-concordant minimization. *J. Glob. Optim.*, 83(2):273–299, 2022.

[41] Jaroslav Řeháček, Zdeněk Hradil, E. Knill, and A. I. Lvovsky. Diluted maximum-likelihood algorithm for quantum tomography. *Phys. Rev. A*, 75(4):042108, 2007.

[42] Doglas S. Gonçalves, Márcia A. Gomes-Ruggiero, and Carlile Lavor. Global convergence of diluted iterations in maximum-likelihood quantum tomography. *Quantum Info. Comput.*, 14(11–12):966–980, 2014.

[43] H. M. Hudson and R. S. Larkin. Accelerated image reconstruction using ordered subsets of projection data. *IEEE Trans. Med. Imaging*, 13(4):601–609, 1994.

[44] A. I. Lvovsky. Iterative maximum-likelihood reconstruction in quantum homodyne tomography. *J. Opt. B: Quantum and Semiclass. Opt.*, 6(6):S556, 2004.

[45] H. Häffner, W. Hänsel, C. F. Roos, J. Benhelm, D. Chek-al kar, M. Chwalla, T. Körber, U. D. Rapol, M. Riebe, P. O. Schmidt, C. Becher, O. Gühne, W. Dür, and R. Blatt. Scalable multiparticle entanglement of trapped ions. *Nature*, 438(7068): 643–646, 2005.

[46] Adriano Macarone Palmieri, Egor Kovlakov, Federico Bianchi, Dmitry Yudin, Stanislav Straupe, Jacob D. Biamonte, and Sergei Kulik. Experimental neural network enhanced quantum tomography. *npj Quantum Inf.*, 6(1):20, 2020.

[47] M. O. Brown, S. R. Muleady, W. J. Dworschack, R. J. Lewis-Swan, A. M. Rey, O. Romero-Isart, and C. A. Regal. Time-of-flight quantum tomography of an atom in an optical tweezer. *Nat. Phys.*, 19(4):569–573, 2023.

[48] Alexander Meiburg. Inapproximability of positive semidefinite permanents and quantum state tomography. In *IEEE 63rd Annu. Symp. Foundations of Computer Science*, pages 58–68, 2022.

[49] Nathan Srebro, Karthik Sridharan, and Ambuj Tewari. Smoothness, low-noise and fast rates. In *Adv. Neural Information Processing Systems*, volume 23, 2010.

[50] Francesco Orabona. A modern introduction to online learning. *arXiv preprint arXiv:1912.13213v6*, 2023.

[51] James Demmel, Ioana Dumitriu, and Olga Holtz. Fast linear algebra is stable. *Numer. Math.*, 108(1):59–91, 2007.

[52] Jeff Bezanson, Alan Edelman, Stefan Karpinski, and Viral B Shah. Julia: A fresh approach to numerical computing. *SIAM Review*, 59(1):65–98, 2017.

[53] Lawrence A Shepp and Benjamin F Logan. The Fourier reconstruction of a head section. *IEEE Trans. Nuc. Sci*, 21(3):21–43, 1974.

[54] Maxim Raginsky, Rebecca M. Willett, Zachary T. Harmany, and Roummel F. Marcia. Compressed sensing performance bounds under Poisson noise. *IEEE Trans. Signal Process.*, 58(8):3990–4002, 2010.

[55] Yen-Huan Li. Online positron emission tomography by online portfolio selection. In *IEEE Int. Conf. Acoustics, Speech and Signal Processing*, pages 1110–1114, 2020.

[56] Yen-Huan Li, Carlos A Riofrío, and Volkan Cevher. A general convergence result for mirror descent with Armijo line search. *arXiv preprint arXiv:1805.12232*, 2018.

[57] Jiangwei Shang, Zhengyun Zhang, and Hui Khoon Ng. Superfast maximum-likelihood reconstruction for quantum tomography. *Phys. Rev. A*, 95:062336, 2017.

[58] Shahnawaz Ahmed, Carlos Sánchez Muñoz, Franco Nori, and Anton Frisk Kockum. Quantum state tomography with conditional generative adversarial networks. *Phys. Rev. Lett.*, 127:140502, 2021.

[59] Léon Bottou and Olivier Bousquet. The tradeoffs of large scale learning. In *Adv. Neural Information Processing Systems*, volume 20, 2007.

[60] Yurii Nesterov and Arkadii Nemirovskii. *Interior-Point Polynomial Algorithms in Convex Programming*. SIAM, Philadelphia, PA, 1994.

[61] Haihao Lu, Robert M. Freund, and Yurii Nesterov. Relatively smooth convex optimization by first-order methods, and applications. *SIAM J. Optim.*, 28(1):333–354, 2018.

[62] Chung-En Tsai, Hao-Chung Cheng, and Yen-Huan Li. Online self-concordant and relatively smooth minimization, with applications to online portfolio selection and learning quantum states. In *Proc. 34th Int. Conf. Algorithmic Learning Theory*, pages 1481–1483, 2023. arXiv:2210.00997.

[63] Elad Hazan. Introduction to Online Convex Optimization. *Found. Trends Opt.*, 2 (3–4):157–325, 2016.

[64] Stephen P. Boyd and Lieven Vandenberghe. *Convex Optimization*. Cambridge university press, 2004.

[65] Fumio Hiai and Dénes Petz. *Introduction to Matrix Analysis and Applications*. Springer, Cham, 2014.

[66] Haipeng Luo, Chen-Yu Wei, and Kai Zheng. Efficient online portfolio with logarithmic regret. In *Adv. Neural Information Processing Systems*, volume 31, 2018.

A Preliminaries

Throughout this section, let $(\mathbb{V}, \langle \cdot, \cdot \rangle)$ be a finite-dimensional real Hilbert space, such as $\mathbb{R}^d$ with the standard inner product and $\mathbb{H}^d$ with the Hilbert-Schmidt inner product $\langle U, V \rangle \coloneqq \operatorname{tr}(U^* V)$. Let $\mathcal{X} \subseteq \mathbb{V}$ be a convex set.

A.1 Self-Concordance and Relative Smoothness

This section provides necessary background information on the notions of *self-concordance* [60, 13] and *relative smoothness* [33, 61], which form the basis of the smoothness characterization in Section 4. We begin with self-concordance. Define $\omega(t) \coloneqq t - \log(1+t)$ and its Fenchel conjugate $\omega_*(t) = -t - \log(1-t)$.

Definition 9 (Self-concordance [13])**.** *A closed convex function $\varphi : \mathbb{V} \to (-\infty, \infty]$ with an open domain $\operatorname{dom} \varphi$ is said to be M-self-concordant if it is three-times continuously differentiable on $\operatorname{dom} \varphi$ and*

$$|D^3 \varphi(x)[u,u,u]| \leq 2M (D^2 \varphi(x)[u,u])^{3/2}, \quad \forall x \in \operatorname{dom} \varphi,\ u \in \mathbb{V}.$$

Theorem 10 (Theorem 5.1.5 of Nesterov [13])**.** *Let φ be an M-self-concordant function. Let $\|\cdot\|_x \coloneqq (D^2 \varphi(x)[\cdot,\cdot])^{1/2}$ be the local norm associated with φ at x. Define the Dikin ellipsoid $W(x) \coloneqq \{y \in \mathbb{V} \mid \|y - x\|_x < 1/M\}$. Then, $W(x) \subseteq \operatorname{dom} \varphi$ for all $x \in \operatorname{dom} \varphi$.*

Theorem 11 is an important local smoothness-type property of self-concordant functions.

Theorem 11 (Theorem 5.1.9 and Lemma 5.1.5 of [13])**.** *Let φ be an M-self-concordant function. Let $\|\cdot\|_x \coloneqq (D^2 \varphi(x)[\cdot,\cdot])^{1/2}$ be the local norm associated with φ at x. Then, for $x, y \in \operatorname{dom} \varphi$ such that $\|y - x\|_x < 1/M$, it holds that*

$$\varphi(y) \leq \varphi(x) + \langle \nabla \varphi(x), y - x \rangle + \frac{1}{M^2} \omega_*(M \|y - x\|_x).$$

Moreover, if $\|y - x\|_x < 1/(2M)$, then

$$\varphi(y) \leq \varphi(x) + \langle \nabla \varphi(x), y - x \rangle + \|y - x\|_x^2.$$

Lemma 12 (Proposition 5.4.5 of Nesterov and Nemirovskii [60])**.** *The logarithmic barrier $h(\rho) = -\log\det \rho$ is 1-self-concordant.*

Now we introduce the notion of relative smoothness.

Definition 13 (Lu et al. [61])**.** *Let $f, h : \mathbb{V} \to (-\infty, \infty]$. f is said to be L-smooth relative to h on $\mathcal{X}$ for some $L > 0$ if $Lh - f$ is convex on $\mathcal{X}$.*

Lemma 14 (Proposition 7 of Tsai et al. [62])**.** *Let $f(\rho) \coloneqq \mathbb{E}[-\log \operatorname{tr}(A\rho)]$ and $h(\rho) \coloneqq -\log\det \rho$ be the logarithmic barrier. Then, f is 1-smooth relative to h on $\mathbb{H}^d_{++}$.*

A.2 FTRL with Self-Concordant Regularizer

This section presents the regret bound of follow-the-regularized-leader (FTRL) with self-concordant regularizers of Tsai et al. [30] in a slightly general form. An *online linear optimization* problem is a multi-round game between two players, say Learner and Reality. In the t-th round,

- first, Learner announces an action $x_t \in \mathcal{X}$;

- then, Reality reveals a loss function $f_t(x) \coloneqq \langle v_t, x \rangle$ for some $v_t \in \mathbb{V}$;
- lastly, Learner suffers a loss $f_t(x_t)$.

The goal of Learner is to minimize the regret $\sup_{x \in \mathcal{X}} R_t(x)$, defined as

$$R_t(x) \coloneqq \sum_{\tau=1}^{t} f_\tau(x_\tau) - \sum_{\tau=1}^{t} f_\tau(x), \quad \forall x \in \mathcal{X}.$$

We refer readers to the lecture notes of Orabona [50] and Hazan [63] for a general introduction to online convex optimization.

Algorithm 2 FTRL for online linear optimization

$x_1 \in \operatorname{argmin}_{x \in \mathcal{X}} \eta_0^{-1} \varphi(x)$.
for all $t \in \mathbb{N}$ **do**
  Announce x_t and receive $v_t \in \mathbb{V}$.
  Compute a learning rate $\eta_t > 0$.
  $x_{t+1} \leftarrow \operatorname{argmin}_{x \in \mathcal{X}} \langle v_{1:t}, x \rangle + \eta_t^{-1} \varphi(x)$.
end for

FTRL is presented in Algorithm 2. We assume that the regularizer φ is a self-concordant function.

Assumption 2. *The function φ is an M-self-concordant function such that $\mathcal{X}$ is contained in the closure of $\operatorname{dom} \varphi$ and $\min_{x \in \mathcal{X}} \varphi(x) = 0$. The Hessian $\nabla^2 \varphi(x)$ is positive definite for all $x \in \mathcal{X} \cap \operatorname{dom} \varphi$.*

Let $\|\cdot\|_x \coloneqq (D^2 \varphi(x)[\cdot, \cdot])^{1/2}$ be the local norm associated with φ at x and $\|\cdot\|_{x,*}$ be its dual norm. The theorem below bounds the regret of Algorithm 2.

Theorem 15 (Theorem 3.2 of Tsai et al. [30])**.** *Assume that Assumption 2 holds and $\eta_{t-1} \|v_t\|_{x_t,*} \leq 1/(2M)$ for all $t \in \mathbb{N}$. Then, Algorithm 2 satisfies*

$$R_t(x) \leq \frac{\varphi(x)}{\eta_t} + \sum_{\tau=1}^{t} \eta_{\tau-1} \|v_\tau\|_{x_\tau,*}^2, \quad \forall t \in \mathbb{N}.$$

Remark 16. *It is important to notice that the regret analysis of Tsai et al. [30] directly extends for the quantum setup.*

The following corollary has appeared in the proof of Theorem 6.2 of Tsai et al. [30] implicitly. We provide the statement and the proof for completeness.

Corollary 17. *Assume that Assumption 2 holds. Moreover, assume that $\|v_t\|_{x_t,*} \leq G$ for all $t \in \mathbb{N}$. Then, for any $D > 0$, Algorithm 2 with*

$$\eta_t = \frac{D}{\sqrt{\sum_{\tau=1}^{t} \|v_\tau\|_{x_\tau,*}^2 + 4M^2 G^2 D^2 + G^2}}, \quad \forall t \in \mathbb{N},$$

satisfies

$$R_t(x) \leq \left(\frac{\varphi(x)}{D} + 2D \right) \sqrt{\sum_{\tau=1}^{t} \|v_\tau\|_{x_\tau,*}^2 + 4M^2 G^2 D^2 + G^2}, \quad \forall t \in \mathbb{N}.$$

Proof. First, the learning rates satisfy $\eta_{t-1}\|v_t\|_{x_t,*} \leq 1/(2M)$ for all $t \in \mathbb{N}$ because

$$\eta_{t-1}\|v_t\|_{x_t,*} = \frac{D\|v_t\|_{x_t,*}}{\sqrt{\sum_{\tau=1}^{t-1}\|v_\tau\|_{x_\tau,*}^2 + 4M^2G^2D^2 + G^2}} \leq \frac{DG}{\sqrt{4M^2G^2D^2}} = \frac{1}{2M}.$$

By Theorem 15 and Lemma 4.13 of Orabona [50], we have

$$\begin{aligned}
R_t(x) &\leq \frac{\varphi(x)}{D}\sqrt{\sum_{\tau=1}^{t}\|v_\tau\|_{x_\tau,*}^2 + 4M^2G^2D^2 + G^2} + D\sum_{\tau=1}^{t}\frac{\|v_\tau\|_{x_\tau,*}^2}{\sqrt{\sum_{s=1}^{\tau-1}\|v_s\|_{x_s,*}^2 + 4M^2G^2D^2 + G^2}} \\
&\leq \frac{\varphi(x)}{D}\sqrt{\sum_{\tau=1}^{t}\|v_\tau\|_{x_\tau,*}^2 + 4M^2G^2D^2 + G^2} + D\sum_{\tau=1}^{t}\frac{\|v_\tau\|_{x_\tau,*}^2}{\sqrt{\sum_{s=1}^{\tau}\|v_s\|_{x_s,*}^2}} \\
&\leq \frac{\varphi(x)}{D}\sqrt{\sum_{\tau=1}^{t}\|v_\tau\|_{x_\tau,*}^2 + 4M^2G^2D^2 + G^2} + 2D\sqrt{\sum_{\tau=1}^{t}\|v_\tau\|_{x_\tau,*}^2} \\
&\leq \left(\frac{\varphi(x)}{D} + 2D\right)\sqrt{\sum_{\tau=1}^{t}\|v_\tau\|_{x_\tau,*}^2 + 4M^2G^2D^2 + G^2}.
\end{aligned}$$

This completes the proof. □

A.3 Online-to-Batch Conversion

This section recaps the online-to-batch conversion proposed by Cesa-Bianchi et al. [31]. Consider the following optimizaiton problem:

$$\min_{x\in\mathcal{X}} f(x),$$

with a stochastic first-order oracle $\mathcal{O}$ that returns an unbiased estimate $\mathcal{O}(x) \in \mathbb{V}$ of $\nabla f(x)$ given any $x \in \mathcal{X}$. Algorithm 3 presents the online-to-batch conversion and Theorem 18 presents its theoretical guarantee.

Algorithm 3 Online-to-batch conversion

Input: An online learning algorithm $\mathcal{A}$.

Get x_1 from $\mathcal{A}$.
for all $t \in \mathbb{N}$ **do**
  Output $\bar{x}_t \coloneqq (1/t)x_{1:t}$.
  $g_t = \mathcal{O}(x_t)$.
  Send $f_t(x) \coloneqq \langle g_t, x\rangle$ to $\mathcal{A}$.
  Get x_{t+1} from $\mathcal{A}$.
end for

Theorem 18. *Let $R_t(x)$ be the regret of the online algorithm $\mathcal{A}$ against $x \in \mathcal{X}$. Assume that the stochastic gradients are unbiased, i.e., $\mathbb{E}[g_t|g_1,\dots,g_{t-1},x_1,\dots,x_t] = \nabla f(x_t)$. Then, for any $x \in \mathcal{X}$, Algorithm 3 satisfies*

$$\mathbb{E}\left[f(\bar{x}_t) - f(x)\right] \leq \frac{\mathbb{E}[R_t(x)]}{t}, \quad \forall t \in \mathbb{N},$$

and

$$\mathbb{E}\left[\sum_{\tau=1}^{t}(f(\bar{x}_\tau) - f(x))\right] \leq (1+\log t)\max_{1\leq\tau\leq t}\mathbb{E}[R_\tau(x)], \quad \forall t \in \mathbb{N},$$

where $R_t(x) \coloneqq \sum_{\tau=1}^{t} \langle g_\tau, x_\tau - x \rangle$ *is the regret, and the expectation is taken with respect to the stochastic gradients* $\{g_t\}$.

In Theorem 18, the first inequality can be found in Theorem 3.1 of Orabona [50], and the second inequality follows immediately by summing the first one over t.

B Proofs

B.1 Local Norm and Dual Local Norm

Lemma 19. *The local norm* $\|\cdot\|_\rho$ *is given by* $\|X\|_\rho = \sqrt{\operatorname{tr}((\rho^{-1}X)^2)}$.

Proof. By Appendix A.4.1 of Boyd and Vandenberghe [64] and Example 3.20 of Hiai and Petz [65], we write

$$\begin{aligned}
\|X\|_\rho^2 &= D^2 h(\rho)[X, X] \\
&= \frac{\mathrm{d}^2}{\mathrm{d}t^2} - \log\det(\rho + tX)\bigg|_{t=0} \\
&= \frac{\mathrm{d}}{\mathrm{d}t} - \operatorname{tr}((\rho + tX)^{-1}X)\bigg|_{t=0} \\
&= \operatorname{tr}\left(-\frac{\mathrm{d}}{\mathrm{d}t}(\rho + tX)^{-1}\bigg|_{t=0} X\right) \\
&= \operatorname{tr}(\rho^{-1}X\rho^{-1}X).
\end{aligned}$$

The lemma follows. □

Lemma 20. *The dual norm of* $\|\cdot\|_\rho$ *is given by* $\|X\|_{\rho,*} = \sqrt{\operatorname{tr}((\rho X)^2)}$.

Proof. By the definition of dual norm,

$$\|X\|_{\rho,*} = \sup_{\|\tau\|_\rho = 1} |\operatorname{tr}(X\tau)| = \sup_{\|\tau\|_\rho = 1} |\operatorname{tr}(\rho^{1/2}X\rho^{1/2}\rho^{-1/2}\tau\rho^{-1/2})|.$$

By the Cauchy-Schwarz inequality and $\|\tau\|_\rho = 1$, we have

$$|\operatorname{tr}(\rho^{1/2}X\rho^{1/2}\rho^{-1/2}\tau\rho^{-1/2})| \le \sqrt{\operatorname{tr}((\rho^{1/2}X\rho^{1/2})^2)\operatorname{tr}((\rho^{-1/2}\tau\rho^{-1/2})^2)} = \sqrt{\operatorname{tr}((\rho^{1/2}X\rho^{1/2})^2)}.$$

Then,

$$\sup_{\|\tau\|_\rho = 1} |\operatorname{tr}(X\tau)| \le \sup_{\|\tau\|_\rho = 1} \sqrt{\operatorname{tr}((\rho^{1/2}X\rho^{1/2})^2)} = \sqrt{\operatorname{tr}((\rho^{1/2}X\rho^{1/2})^2)}.$$

The equality can be achieved by taking

$$\tau = \frac{\rho X \rho}{\sqrt{\operatorname{tr}((\rho^{1/2}X\rho^{1/2})^2)}}.$$

The lemma follows. □

B.2 Proof of Lemma 2

We write

$$\|\nabla f(\rho)\|_{\rho,*}^2 = \operatorname{tr}\left(\left(\mathbb{E}\frac{\rho^{1/2}A\rho^{1/2}}{\operatorname{tr}(\rho A)}\right)^2\right) \le \mathbb{E}\operatorname{tr}\left(\left(\frac{\rho^{1/2}A\rho^{1/2}}{\operatorname{tr}(\rho A)}\right)^2\right) = \mathbb{E}\frac{\operatorname{tr}((\rho^{1/2}A\rho^{1/2})^2)}{(\operatorname{tr}(\rho^{1/2}A\rho^{1/2}))^2} \le 1,$$

where the first inequality follows from the convexity of $\operatorname{tr}(A^2)$ and Jensen's inequality, and the second inequality follows from the inequality $0 \le \operatorname{tr}(A^2) \le (\operatorname{tr} A)^2$ for $A \in \mathbb{H}_+^d$.

B.3 Proof of Lemma 4

The first few steps follow from Lemma 4.7 of Tsai et al. [30]. The main simplification of the original proof is the use of Theorem 10, which we will see later.

Write $\alpha_\rho = \alpha_\rho(\nabla f(\rho))$ for simplicity and assume $\|\nabla f(\rho) + \alpha_\rho I\|_{\rho,*} \neq 0$. Otherwise, the lemma holds immediately. Fix $\rho \in \operatorname{ri} \mathscr{D}_d$. By relative smoothness of f (Definition 13 and Lemma 14), we have

$$f(\rho') \leq f(\rho) + \langle \nabla f(\rho), \rho' - \rho \rangle + \left[h(\rho') - h(\rho) - \langle \nabla h(\rho), \rho' - \rho \rangle \right], \quad \forall \rho' \in \operatorname{ri} \mathscr{D}_d,$$

where $\langle U, V \rangle \coloneqq \operatorname{tr}(U^* V)$ is the Hilbert-Schmidt inner product on $\mathbb{H}^d$. Then, by self-concordance of h (Theorem 11 and Lemma 12), we have

$$f(\rho') \leq f(\rho) + \langle \nabla f(\rho), \rho' - \rho \rangle + \|\rho' - \rho\|_\rho^2, \quad \forall \rho' \in \operatorname{ri} \mathscr{D}_d : \|\rho' - \rho\|_\rho \leq 1/2,$$

where $\|\cdot\|_\rho$ is the local norm associated with h. Since $\langle I, \rho - \rho' \rangle = 0$, we write

$$f(\rho') \leq f(\rho) + \langle \nabla f(\rho) + \alpha_\rho I, \rho' - \rho \rangle + \|\rho' - \rho\|_\rho^2, \quad \forall \rho' \in \operatorname{ri} \mathscr{D}_d : \|\rho' - \rho\|_\rho \leq 1/2.$$

Rearraging the terms and taking supremum over all possible ρ', we obtain

$$\sup_{\rho' \in \operatorname{ri} \mathscr{D}_d : \|\rho' - \rho\|_\rho \leq 1/2} \langle -\nabla f(\rho) - \alpha_\rho I, \rho' - \rho \rangle - \|\rho' - \rho\|_\rho^2 \leq f(\rho) - \min_{\rho' \in \mathscr{D}_d} f(\rho'). \tag{9}$$

Next, for $\beta \in [0, 1/2]$, define

$$\rho'_\beta \coloneqq \rho - \beta \frac{\rho(\nabla f(\rho) + \alpha_\rho I)\rho}{\|\nabla f(\rho) + \alpha_\rho I\|_{\rho,*}}.$$

We will plug ρ'_β into the supremum (9) and must verify that ρ'_β satisfies the constraints. Since

$$\begin{aligned}
\operatorname{tr}\left(\rho(\nabla f(\rho) + \alpha_\rho I)\rho\right) &= \operatorname{tr}\left(\rho \left(-\mathbb{E} \frac{A}{\operatorname{tr}(A\rho)} + \mathbb{E} \frac{\operatorname{tr}(A\rho^2)}{\operatorname{tr}(A\rho)\operatorname{tr}(\rho^2)} I \right) \rho \right) \\
&= \mathbb{E}\left[-\frac{\operatorname{tr}(A\rho^2)}{\operatorname{tr}(A\rho)} + \frac{\operatorname{tr}(A\rho^2)\operatorname{tr}(\rho^2)}{\operatorname{tr}(A\rho)\operatorname{tr}(\rho^2)} \right] \\
&= 0,
\end{aligned}$$

we have $\operatorname{tr} \rho'_\beta = \operatorname{tr} \rho = 1$. Second, by the definition of α_ρ (7) and Lemma 2, we have $\|\rho'_\beta - \rho\|_\rho = \beta \leq 1/2$. At last, we need to verify $\rho'_\beta > 0$. Since $\|\rho'_\beta - \rho\|_\rho \leq 1/2$, by Theorem 10, we have $\rho'_\beta \in \operatorname{dom} h = \mathbb{H}^d_{++}$ and $\rho'_\beta > 0$. The application of Theorem 10 simplifies the proof of Tsai et al. [30] because we no longer need to check $\rho'_\beta > 0$ explicitly. Plugging ρ'_β and lower bounding the supremum (9), we have

$$\sup_{0 \leq \beta \leq 1/2} (\beta - \beta^2) \|\nabla f(\rho) + \alpha_\rho I\|_{\rho,*}^2 \leq f(\rho) - \min_{\rho' \in \mathscr{D}_d} f(\rho').$$

The lemma follows by noticing that $\beta = 1/2$ achieves the supremum.

B.4 Proof of Lemma 5

First, it is clear that the unbiasedness property holds. By Lemma 2, we can take $G = 1$. For the variance, we write

$$\begin{aligned}
&\mathbb{E}\left[\|g_t - \nabla f(\rho_t)\|_{\rho_t,*}^2 \mid \mathcal{H}_t\right] \\
&\quad= \mathbb{E}\left[\left\|\frac{1}{B}\sum_{b=1}^{B}(\nabla \ell_b(\rho_t) - \nabla f(\rho_t))\right\|_{\rho_t,*}^2 \middle| \mathcal{H}_t\right] \\
&\quad= \frac{1}{B^2}\mathbb{E}\left[\sum_{1\le b,b'\le B}\operatorname{tr}((\nabla \ell_b(\rho_t) - \nabla f(\rho_t))\rho_t(\nabla \ell_{b'}(\rho_t) - \nabla f(\rho_t))\rho_t)\middle| \mathcal{H}_t\right] \\
&\quad= \frac{1}{B^2}\mathbb{E}\left[\sum_{b=1}^{B}\|\nabla \ell_b(\rho_t) - \nabla f(\rho_t)\|_{\rho_t,*}^2\middle| \mathcal{H}_t\right],
\end{aligned}$$

where the second equality follows from the explicit formula of the dual local norm (Lemma 20); the third equality follows from the independence of b, b' and unbiasedness of $\nabla \ell_b$ and $\nabla \ell_{b'}$. Finally, by the triangle inequality and Lemma 2,

$$\begin{aligned}
\mathbb{E}\left[\|\nabla \ell_b(\rho_t) - \nabla f(\rho_t)\|_{\rho_t,*}^2\middle|\mathcal{H}_t\right] &\le \mathbb{E}\left[\left(\|\nabla \ell_b(\rho_t)\|_{\rho_t,*} + \|\nabla f(\rho_t)\|_{\rho_t,*}\right)^2 \mid \mathcal{H}_t\right] \\
&\le \mathbb{E}\left[(1+1)^2 \mid \mathcal{H}_t\right] \le 4.
\end{aligned}$$

The lemma follows.

B.5 Proof of Theorem 6

Let $h(\rho) = -\log\det\rho - d\log d$ be the logarithmic barrier. Note that Algorithm 1 is derived by applying Algorithm 2 with the regularizer h to the online linear optimization problem (Appendix A.2), followed by the online-to-batch conversion (Algorithm 3).

Fix $\rho \in \mathscr{D}_d$. To deal with the unboundedness of h on $\mathscr{D}_d$, we first apply the technique used in Lemma 10 of Luo et al. [66]. Let any $\rho \in \mathscr{D}_d$, define $\tilde{\rho} = (t/(t+1))\rho + 1/(t+1)(I/d) \in \mathscr{D}_d$. Then, by convexity,

$$\begin{aligned}
f(\tilde{\rho}) - f(\rho) \le \langle \nabla f(\tilde{\rho}), \tilde{\rho} - \rho\rangle &= \mathbb{E}\left[\frac{\operatorname{tr}(A\rho) - \operatorname{tr}(A\tilde{\rho})}{\operatorname{tr}(A\tilde{\rho})}\right] \\
&= \mathbb{E}\left[\frac{(1+1/t)\operatorname{tr}(A\tilde{\rho}) - (1/dt)\operatorname{tr}A - \operatorname{tr}(A\tilde{\rho})}{\operatorname{tr}(A\tilde{\rho})}\right] \le \frac{1}{t}.
\end{aligned}$$

Therefore,

$$\mathbb{E}\left[f(\bar{\rho}_t) - f(\rho)\right] \le \mathbb{E}\left[f(\bar{\rho}_t) - f(\tilde{\rho})\right] + \frac{1}{t}.$$

Applying the first inequality in Theorem 18, we have

$$\mathbb{E}\left[f(\bar{\rho}_t) - f(\rho)\right] \le \frac{\mathbb{E}[R_t(\tilde{\rho})] + 1}{t} \le \frac{\max_{1\le\tau\le t}\mathbb{E}R_\tau(\tilde{\rho}) + 1}{t}, \tag{10}$$

where

$$R_t(\tilde{\rho}) = \sum_{\tau=1}^{t}\langle g_\tau, \rho_\tau - \tilde{\rho}\rangle = \sum_{\tau=1}^{t}\langle g_\tau + \alpha_{\rho_\tau}(g_\tau)I, \rho_\tau - \tilde{\rho}\rangle$$

is the regret of Algorithm 2 applied to the online linear optimization problem where $v_t = g_t + \alpha_{\rho_\tau}(g_\tau)I$ (see Appendix A.2).

Next, by assumption and the definition of α_ρ (7), we have $\|g_\tau + \alpha_{\rho_\tau}(g_\tau)I\|_{\rho_\tau,*} \le \|g_\tau\|_{\rho_\tau,*} \le G$. Applying Corollary 17 with $M = 1$ and $D = \sqrt{d}$, we obtain

$$R_t(\tilde{\rho}) \le \left(\frac{h(\tilde{\rho})}{\sqrt{d}} + 2\sqrt{d}\right)\sqrt{G_t + 4dG^2 + G^2} \le (\log t + 3)\sqrt{dG_t + 4d^2G^2 + dG^2},$$

where $G_t = \sum_{\tau=1}^t \|g_\tau + \alpha_{\rho_\tau}(g_\tau)I\|^2_{\rho_\tau,*}$. The last inequality follows from $h(\tilde{\rho}) \le d\log(t+1) \le d\log t + d$. By Jensen's inequality, the expected regret is bounded by

$$\mathbb{E}R_t(\tilde{\rho}) \le (\log t + 3)\sqrt{d\mathbb{E}G_t + 4d^2G^2 + dG^2}.$$

Then, we take maximum from $t = 1$ to $t = t$ on both sides. Since the above upper bound is increasing in t, we obtain

$$\max_{1\le\tau\le t} \mathbb{E}R_\tau(\tilde{\rho}) \le (\log t + 3)\sqrt{d\mathbb{E}G_t + 4d^2G^2 + dG^2}. \tag{11}$$

Now, we upper bound $\mathbb{E}G_t$ by $\max_{1\le\tau\le t}\mathbb{E}R_\tau(\tilde{\rho})$. Denote by $\mathcal{H}_t = \{g_1, \dots, g_{t-1}, \rho_1, \dots, \rho_t\}$ the past information before obtaining g_t. By the linearity of α_ρ (7), we have

$$\mathbb{E}[\alpha_{\rho_t}(g_t)|\mathcal{H}_\tau] = \alpha_{\rho_t}(\nabla f(\rho_t)).$$

Then, by the law of total expectation and the variance decomposition $\mathbb{E}[X^2] = \mathbb{E}[(X - \mathbb{E}X)^2] + (\mathbb{E}X)^2$, we have

$$\begin{aligned}
&\mathbb{E}\left[\|g_\tau + \alpha_{\rho_\tau}(g_\tau)I\|^2_{\rho_\tau,*}\right] \\
&\quad= \mathbb{E}_{\mathcal{H}_\tau}\mathbb{E}\left[\|g_\tau + \alpha_{\rho_\tau}(g_\tau)I\|^2_{\rho_\tau,*} \mid \mathcal{H}_\tau\right] \\
&\quad= \mathbb{E}_{\mathcal{H}_\tau}\mathbb{E}\left[\|g_\tau + \alpha_{\rho_\tau}(g_\tau)I - \nabla f(\rho_\tau) - \alpha_{\rho_\tau}(\nabla f(\rho_\tau))I\|^2_{\rho_\tau,*} \mid \mathcal{H}_\tau\right] \\
&\qquad\qquad + \mathbb{E}_{\mathcal{H}_\tau}\left[\|\nabla f(\rho_\tau) + \alpha_{\rho_\tau}(\nabla f(\rho_\tau))I\|^2_{\rho_\tau,*}\right] \\
&\quad= \mathbb{E}_{\mathcal{H}_\tau}\mathbb{E}\left[\|g_\tau + \alpha_{\rho_\tau}(g_\tau)I - \nabla f(\rho_\tau) - \alpha_{\rho_\tau}(\nabla f(\rho_\tau))I\|^2_{\rho_\tau,*} \mid \mathcal{H}_\tau\right] \\
&\qquad\qquad + \mathbb{E}\left[\|\nabla f(\rho_\tau) + \alpha_{\rho_\tau}(\nabla f(\rho_\tau))I\|^2_{\rho_\tau,*}\right]
\end{aligned} \tag{12}$$

We bound the two terms separately. By the definition of α_ρ (7) and the bounded-variance assumption,

$$\begin{aligned}
&\mathbb{E}_{\mathcal{H}_\tau}\mathbb{E}\left[\|g_\tau + \alpha_{\rho_\tau}(g_\tau)I - \nabla f(\rho_\tau) - \alpha_{\rho_\tau}(\nabla f(\rho_\tau))I\|^2_{\rho_\tau,*} \mid \mathcal{H}_\tau\right] \\
&\quad= \mathbb{E}_{\mathcal{H}_\tau}\mathbb{E}\left[\|g_\tau - \nabla f(\rho_\tau) + \alpha_{\rho_\tau}(g_\tau - \nabla f(\rho_\tau))I\|^2_{\rho_\tau,*} \mid \mathcal{H}_\tau\right] \\
&\quad\le \mathbb{E}_{\mathcal{H}_\tau}\mathbb{E}\left[\|g_\tau - \nabla f(\rho_\tau)\|^2_{\rho_\tau,*} \mid \mathcal{H}_\tau\right] \le \sigma^2.
\end{aligned}$$

Furthermore, let $\delta_\tau \coloneqq \mathbb{E}f(\rho_\tau) - f(\tilde{\rho})$. By the self-bounding-type property (Lemma 4), we have

$$\mathbb{E}\left[\|\nabla f(\rho_\tau) + \alpha_{\rho_\tau}(\nabla f(\rho_\tau))I\|^2_{\rho_\tau,*}\right] \le \mathbb{E}f(\rho_\tau) - \min_{\rho\in\mathscr{D}_d} f(\rho) \le \delta_\tau.$$

Therefore, we have (12) $\le \sigma^2 + 4\delta_\tau$ and

$$\mathbb{E}G_t = \sum_{\tau=1}^t \mathbb{E}\left[\|g_\tau + \alpha_{\rho_\tau}(g_\tau)I\|^2_{\rho_\tau,*}\right] \le \sigma^2 t + 4\delta_{1:t}.$$

By the second inequality in Theorem 18, we have $\delta_{1:t} \le (1 + \log t)\max_{1\le\tau\le t}\mathbb{E}R_\tau(\tilde{\rho})$. Hence,

$$\mathbb{E}G_t \le \sigma^2 t + 4(1 + \log t)\max_{1\le\tau\le t}\mathbb{E}R_\tau(\tilde{\rho}).$$

Finally, combining with the bound on the expected regret (11), we have

$$\max_{1\le\tau\le t}\mathbb{E}R_\tau(\tilde{\rho}) \le (3+\log t)\sqrt{4d(1+\log t)\max_{1\le\tau\le t}\mathbb{E}R_\tau(\tilde{\rho}) + 4d^2G^2 + dG^2 + \sigma^2 dt}.$$

By Lemma 4.24 of Orabona [50], solving $\max_{1\le\tau\le t}\mathbb{E}R_\tau(\tilde{\rho})$ gives

$$\max_{1\le\tau\le t}\mathbb{E}R_\tau(\tilde{\rho}) \le 4d(3+\log t)^3 + 2(3+\log t)\sqrt{\sigma^2 dt + 4d^2G^2 + dG^2}.$$

The theorem follows by combining the above inequality with (10).

C Additional Numerical Results

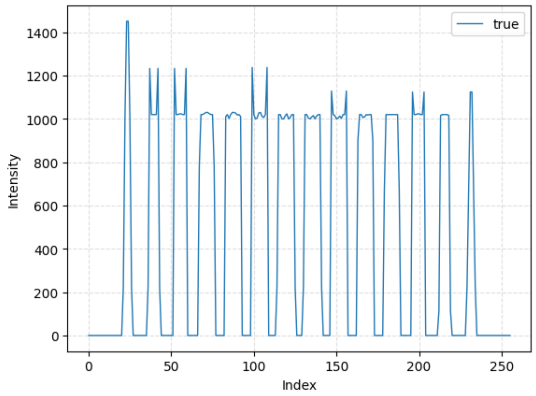
This section presents additional numerical results in terms of the number of epochs, where one *epoch* refers to one full pass of the dataset. Specifically, one epoch of B-sample LB-SDA corresponds to n/B iterations, one epoch of other stochastic methods corresponds to n iterations, and one epoch of batch methods corresponds to one iteration. The number of epochs is proportional to the number of gradient evaluated by the algorithm. Note that EMD and diluted iMLE compute function values in the Armijo line search procedure, which pass through the dataset for more than once. Therefore, the numerical results in terms of the number of epochs favor these two algorithms. Since computing the function values is much faster than computing the gradients, we still present numerical results in terms of the number of epochs.

Figure 3 presents results for the experiment of the Poisson inverse problem in Section 6.1, while Figure 4 presents results for the experiment of ML quantum state tomography in Section 6.2. For the Poisson inverse problem, we randomly generated 20 problem instances under the setup described in Section 6.1. We then reported the average performance of the algorithms over them. For ML quantum state tomography, we considered only one problem instance because the experiment already took one week.

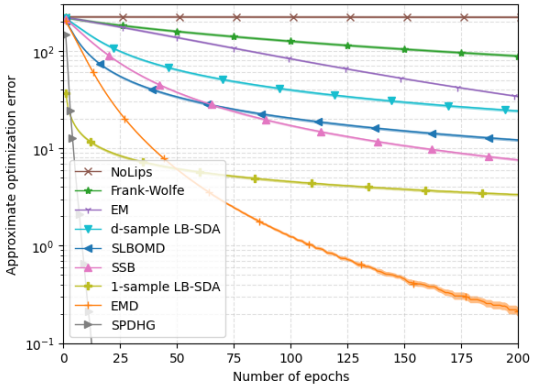
In terms of the optimization error, 1-sample LB-SDA outperforms other methods with clear complexity guarantees. Note that 1-sample LB-SDA converges faster than d-sample LB-SDA. The reason is that after the s-th epoch, 1-sample LB-SDA has performed ns iterations, whereas d-sample LB-SDA has only performed ns/d iterations. According to Corollary 7, the optimization error bound of 1-sample LB-SDA is $\tilde{O}(d/(ns) + \sqrt{d/(ns)})$, smaller than the $\tilde{O}(d^2/(ns) + \sqrt{d/(ns)})$ optimization error bound of d-sample LB-SDA. In terms of the fidelity, 1-sample LB-SDA achieves the best performance among all methods with explicit complexity guarantees for computing the ML estimate for quantum state tomography.

Figure 3: Performances of all algorithms in Table 1, SPDHG, and EMD with line search for solving 20 randomly generated Poisson inverse problem instances. For each algorithm, the solid line represents the average error, and the shaded region indicates the 95% confidence interval.

(a) The true signal.



(b) Approximate optimization error versus the number of epochs.



(c) Normalized estimation error versus the number of epochs.

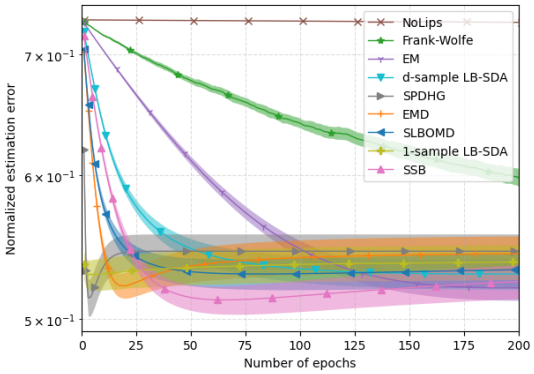
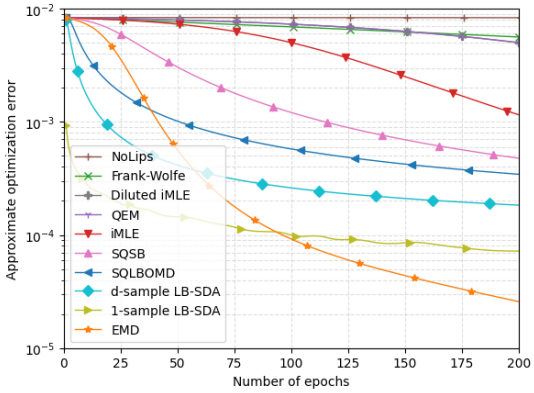


Figure 4: Performances of all algorithms in Table 1, iMLE, diluted iMLE, and EMD with line search for computing the ML estimate for quantum state tomography.

(a) Approximate optimization error versus the number of epochs.



(b) Fidelity between the iterates and the W state versus the number of epochs.