Regularizing least squares quantum state tomography with classical shadows

Zhihui Zhu,[1,*] Joseph M. Lukens,[2,3,†] and Brian T. Kirby[4,5,§]

[1]*Department of Computer Science and Engineering, The Ohio State University, Columbus, Ohio 43210, USA*
[2]*Research Technology Office and Quantum Collaborative, Arizona State University, Tempe, Arizona 85287, USA*
[3]*Quantum Information Science Section, Oak Ridge National Laboratory, Oak Ridge, Tennessee 37831, USA*
[4]*DEVCOM Army Research Laboratory, Adelphi, MD 20783, USA*
[5]*Tulane University, New Orleans, LA 70118, USA*

[*]*zhu.3440@osu.edu* [†]*joseph.lukens@asu.edu* [§]*brian.t.kirby4.civ@army.mil*

Abstract: Classical shadows herald remarkable opportunities for resource-efficient quantum estimation. Although superficially disconnected from traditional inference methods, we show how classical shadows fit under a larger umbrella of least squares regularization, revealing tradeoffs with related methods.

A fixture of quantum information science, quantum state tomography represents the archetypal approach for recovering a density matrix ρ [1]. However, the exponential scaling in the required number of measurements and the classical computational cost of state estimation have inspired a variety of techniques for resource-efficient tomography, typically based on assumptions about the structure of the underlying state [2,3]. In contrast, the method of classical shadows (CS) [4] places no such restrictions; instead, by performing randomly chosen measurements, calculating from each a "shadow" $\hat{\rho}_m$, and then averaging them for the final estimate $\hat{\rho} = \frac{1}{M}\sum_m \hat{\rho}_m$, quantum observables in exceptionally large Hilbert spaces can be estimated with few measurements and no structural assumptions on the ground truth ρ. The core of CS lies in its shadow calculation, which involves the inversion of a quantum channel $\mathscr{M}$ defined as the average over a distribution of possible measurements, a superficially enigmatic choice due to its violation of positive semidefiniteness (required for any physical ρ) and reliance on a fictitious quantum channel (i.e., distinct from the actual measurements executed).

Here we summarize recent work [5] shedding light on the power of CS through formal and practical connections with traditional least squares (LS) and regularized least squares (RLS) estimation. Deriving the LS and RLS solutions to a generic quantum measurement scenario, we find that both rely on their own "shadows"—i.e., linear transformations of individual measurement results—with CS and RLS stabilizing the LS shadow in the underdetermined regime by replacing the pseudoinverse with invertible and well-conditioned operators. Collectively, our results contribute to the fundamental understanding of CS while simultaneously offering practical guidance for quantum estimation.

Figure 1 overviews the formal shadow picture of quantum tomography. M measurements described by positive operator-valued measures (POVMs) $\mathscr{A}_m$ are performed L times, each with K outcomes occurring with ground truth probabilities $\mathscr{A}_m(\rho) = [\mathrm{tr}(A_{m,1}\rho)\cdots\mathrm{tr}(A_{m,K}\rho)]^\top$. For each POVM, one obtains the empirical frequencies $\hat{\mathbf{p}}_m = [f_{m,1}\cdots f_{m,K}]^\top/L$ (a one-hot vector for $L=1$). After application of the POVM transpose, the shadow estimate follows as $\hat{\rho}_m = S(\mathscr{A}_m^\dagger(\hat{\mathbf{p}}_m))$, and the individual shadows are averaged into the final estimate $\hat{\rho}$. This explicit form highlights the remarkable similarities between the three highlighted methods, differing only in the shadow

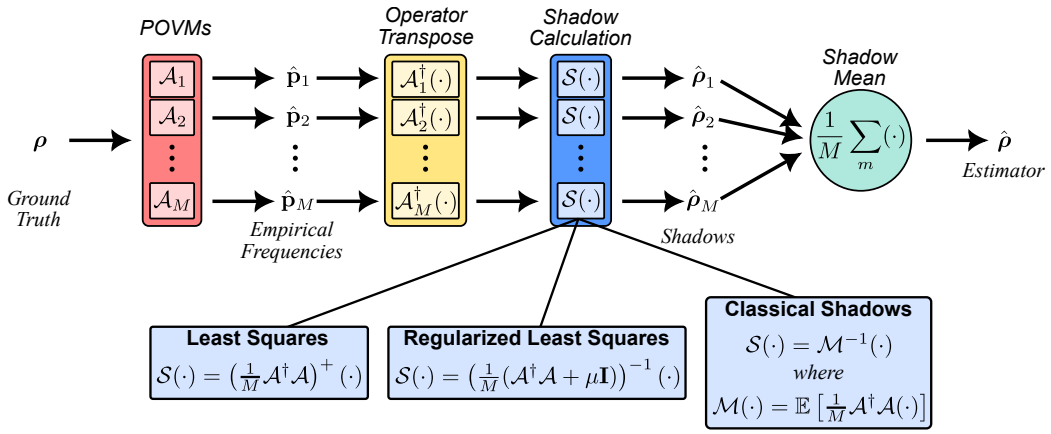

Fig. 1. Shadow picture of quantum state tomography. Empirical frequencies from quantum measurements (POVMs) are converted into shadows and then averaged. Least squares (LS), regularized least squares (RLS), and classical shadows (CS) differ only in the details of their shadow inversion.

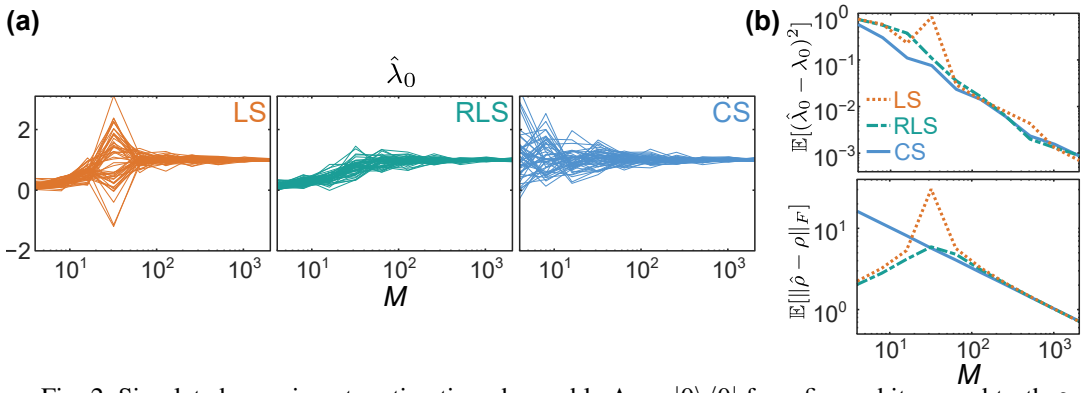

Fig. 2. Simulated experiments estimating observable $\Lambda_0 = |0\rangle\langle 0|$ for a five-qubit ground truth $\rho = |0\rangle\langle 0|$ as a function of measurements M. (a) LS, RLS, and CS results for 50 trials. (b) Average errors for the observable λ_0 (top) and state ρ (bottom).

operation S: LS employs the pseudoinverse of the entire set $\left(\frac{1}{M}\mathscr{A}^\dagger\mathscr{A}\right)^+$, where $\mathscr{A} = [\mathscr{A}_1 \cdots \mathscr{A}_M]^\top$; RLS ensures invertibility through a regularization parameter μ; and CS enlists the inverse of a quantum channel defined as the expectation over all possible measurements $\mathscr{M}(\cdot) = \mathbb{E}\left[\frac{1}{M}\mathscr{A}^\dagger\mathscr{A}(\cdot)\right]$ (treating $\mathscr{A}$ as a random variable).

As one example of how these methods handle inference, consider estimation of the observable $\Lambda_0 = |0\rangle\langle 0|$ for a $(D = 32)$-dimensional five-qubit state $\rho = |0\rangle\langle 0|$—a combination chosen in [6] due to the observable's ground truth value $\lambda_0 = \mathrm{tr}(\Lambda_0\rho) = 1$ representing an "uncommon" case (from the perspective of a uniform prior) in which CS techniques particularly shine. Considering up to $M = 2^{11}$ Haar-random unitaries as measurements with $L = 1$ shot each, the estimates $\hat{\lambda}_0 = \mathrm{tr}(\Lambda_0\hat{\rho})$ for 50 simulated trials appear in Fig. 2(a). Since no methods enforce positivity, all produce results outside of the physically allowed range $\lambda_0 \in [0, 1]$, but their specific behavior differs significantly. For LS, the estimate initially approaches the ground truth in the underdetermined regime $(M < D)$, but then varies wildly in the interpolating regime $(M \approx D)$ before converging in the overdetermined regime $(M > D)$. This behavior resembles the "double descent" phenomenon observed in deep neural networks, where performance initially improves, worsens, and then improves again with increasing model or data size [7].

One method to address double descent, RLS adopts a modified shadow in which the identity I is added before inversion; the second panel of Fig. 2(a) shows results for the same simulated data with regularization parameter $\mu = 0.1$. The fluctuations around the interpolating regime disappear completely, with the estimates approaching the ground truth monotonically as M increases. Finally, CS [third panel in Fig. 2(a)] begins with large variance that reduces steadily with M. Since the channel inverse is well conditions for any M, CS avoids double descent like RLS; however, unlike RLS, CS is unbiased at the cost of higher variance. Figure 2(b) summarizes these features on the mean squared error of the observable $\mathbb{E}[(\hat{\lambda}_0 - \lambda_0)^2]$ and the mean Frobenius error of the state $\mathbb{E}[\|\hat{\rho} - \rho\|_F]$. Compared the monotonic error reduction of CS and the interpolating spike of LS, RLS attains the advantages of the former on the observable front and the latter in lower state error in the underdetermined regime $(M < D)$.

Our results reveal a fundamental relationship between CS and RLS techniques as regularizers of the LS shadow: CS with a regularizer based on the target measurement distribution, and RLS with a regularizer based on the actual measurement set $\mathscr{A}$. As detailed further in [5], such differences result in distinct tradeoffs: (i) CS is unbiased, while RLS has lower variance; (ii) CS has a computationally simple inverse, but RLS is robust to mismatch between the designed and actual measurement distribution (e.g., measurements not being truly random); and (iii) CS shows similar scaling independent of shots L, while RLS is much more sensitive to L. Integration of LS, RLS, and CS under a unifying framework advances the overall picture of CS techniques, while also illuminating the circumstances under which either RLS or CS would be preferred in quantum tomographic contexts, such as unverified randomness for the former or unbiased estimation for the latter.

We thank S. Becker, Z. Gong, Z. Qin, M. Wakin, O. Gühne, and N. Wyderka for valuable discussions. Funding was provided by the National Science Foundation (CCF-2241298, ECCS-2409701), CQISE at the Ohio State University, and the U.S. Department of Energy (ERKJ432, ERKJ353, DE-SC0024257).

References

1. M. A. Nielsen and I. I. Chuang, *Quantum Computation and Quantum Information* (Cambridge, 2000).
2. D. Gross, Y.-K. Liu, S. T. Flammia, S. Becker, and J. Eisert, Phys. Rev. Lett. **105**, 150401 (2010).
3. M. Cramer, M. B. Plenio, S. T. Flammia, R. Somma, D. Gross, S. D. Bartlett, O. Landon-Cardinal, D. Poulin and Y.-K. Liu, Nat. Commun. **1**, 149 (2010).
4. H.-Y. Huang, R. Kueng, and J. Preskill, Nat. Phys. **16**, 1050 (2020).
5. Z. Zhu, J. M. Lukens, and B. T. Kirby, Quantum **8**, 1455 (2024).
6. J. M. Lukens, K. J. H. Law, and R. S. Bennink, npj Quantum Inf. **7**, 113 (2021).
7. M. Belkin, D. Hsu, S. Ma, and S. Mandal, Proc. Natl. Acad. Sci. U.S.A. **116**, 15849 (2019).