Improved Dimension and Sample Size Scalability for Maximum-Likelihood State Tomography and Approximating PSD Matrix Permanents

Chung-En Tsai, Hao-Chung Cheng, and Yen-Huan Li (National Taiwan University)

2024 TAIPEI

We propose the currently fastest stochastic algorithm for maximum-likelihood quantum state tomography, in both theory and practice.

Optimization Problem

$$\rho^\star \in \operatorname*{arg\,min}_{\rho\in\mathcal{D}_d} \left\{ F(\rho) := \frac{1}{n}\sum_{i=1}^{n} f_i(\rho) \right\},$$

$$f_i(\rho) := -\log \operatorname{tr}(A_i \rho).$$

Applications: Maximum-likelihood quantum state tomography (MLQST), PSD matrix permanent approximation, Poisson inverse problem.

Algorithm: B-sample LB-SDA

When n and d are large, e.g., $n=\Omega(d^3)$ in MLQST, stochastic gradient-based methods are preferred.

$h(\rho) := -\log\det\rho.$
$\rho_1 = I/d.$
for all $t \in \mathbb{N}$ **do**
  Output $\bar{\rho}_t := (1/t)\rho_{1:t}$.
  Randomly pick $i_1, \ldots, i_B \in [n]$.
  $g_t = (1/B)\sum_{b=1}^{B} \nabla f_{i_b}(\rho_t)$.
  $\eta_t = O\left(\sqrt{d} / \sqrt{\sum_{\tau=1}^{t} \| g_\tau + \alpha_\tau I \|_{\rho_\tau,*}^2}\right)$.
  $\rho_{t+1} \in \operatorname*{arg\,min}_{\rho\in\mathcal{D}} \eta_t \operatorname{tr}(g_{1:t}\rho) + h(\rho)$.
end for

Non-Asymptotic Convergence

Challenge: The problem violates standard smoothness assumption.

$$\mathsf{E}\left[F(\bar{\rho}_t) - F(\rho^\star)\right] \le \tilde{O}\left(\frac{d}{t} + \frac{\sqrt{d}}{\sqrt{Bt}}\right).$$

Technical Contributions

1. ***Generalized smoothness:***

$$\|\nabla F(\rho) + \alpha I\|_{\rho,*}^2 \le 4\left(F(\rho) - F(\rho^\star)\right).$$

2. ***A local-norm-based analysis of the online-to-batch conversion.***

Time Complexity Comparison

The fastest when n is large.

Find $\hat{\rho}$ s.t. $F(\hat{\rho}) - F(\rho^\star) \le \varepsilon$. $2 \le \omega < 2.372$.

Algorithms	Time complexity ($\tilde{O}$)
d-sample LB-SDA	d^3/ε^2
SQSB, SQLBOMD	$d^{\omega+1}/\varepsilon^2$
QEM	$(nd^2 + d^\omega)/\varepsilon$
Newton's method	$(nd^\omega + d^{2\omega})\log\log(1/\varepsilon)$
EMD, Diluted iMLE	Asymptotic
GD, SGD, iMLE	May not converge

Numerical Experiments

The fastest in terms of the fidelity between the iterates and the true state.

MLQST with $d = 2^6, n = 409{,}600$.