RANK-1 FACTORED GRADIENT DESCENT WITHOUT RESTRICTED ISOMETRY: APPLICATION TO QUANTUM PURE STATE TOMOGRAPHY ON MULTI-QUBITS

Anonymous authors
Paper under double-blind review

ABSTRACT

Factored gradient descent (FGD) is a scalable nonconvex method for low-rank positive semidefinite matrix recovery, in which the rank and positivity constraints are satisfied automatically by the factorization. Most of the existing theory that certifies benignity of its loss landscape, however, relies on a restricted isometry property (RIP) of the measurement map. In quantum state tomography, one of its applications, the measurements that can be acquired with stable statistics are limited by detection efficiency and drift to Pauli observables of low weight; we first show that under such partial-weight Pauli measurements RIP fails in the worst possible way. We then bundle the mutually commuting observables setting by setting, pushing the state forward to the probability distribution of each setting, and introduce a max-sliced total-variation loss. On the stratum where the setting attaining the maximum is unique there is no spurious local minimum, and provided only that the measurement map is injective on generic pure states, the global minimum is the truth up to a phase. Finally we verify these claims numerically on four qubits and compare against FGD with the ordinary loss.

1 INTRODUCTION

Write $\mathbb{H}_d$ for the space of Hermitian matrices. The basic problem of low-rank matrix recovery is to recover an unknown positive semidefinite matrix $M_\star \in \mathbb{H}_d$ of rank $r \ll d$ from the observation $y = \mathcal{A}(M_\star)$, where the linear measurement map $\mathcal{A} : \mathbb{H}_d \to \mathbb{R}^{m_0}$, $\mathcal{A}(M) = \big(\mathrm{tr}(A_1 M), \ldots, \mathrm{tr}(A_{m_0} M)\big)$, is determined by known measurement matrices $A_1, \ldots, A_{m_0} \in \mathbb{H}_d$. Burer–Monteiro type methods factor the optimization variable as $M = YY^\dagger$ and solve the nonconvex least-squares problem

$$\min_{Y \in \mathbb{C}^{d\times r}} \; f_r(Y) = \tfrac{1}{2}\big\|\mathcal{A}(YY^\dagger) - y\big\|_2^2 \tag{1}$$

by gradient methods (factored gradient descent, FGD) (Burer and Monteiro, 2003). The factorization makes the positivity and the rank constraint on M automatic and reduces the number of parameters from d^2 to dr. To overcome the nonconvexity that this introduces, it is standard to impose a restricted isometry property (RIP) on the measurement map: $\mathcal{A}$ satisfies rank-r RIP if there is $\delta_r \in [0, 1)$ such that

$$(1 - \delta_r)\,\|X\|_F^2 \;\le\; \|\mathcal{A}(X)\|_2^2 \;\le\; (1 + \delta_r)\,\|X\|_F^2 \tag{2}$$

for every $X \in \mathbb{H}_d$ of rank at most r; the smallest such value is the RIP constant δ_r. If δ_{2r} is small enough, the landscape of (1) is benign: there is no spurious local minimum and every saddle point is strict (Bhojanapalli et al., 2016; Ge et al., 2017; Zhu et al., 2018; Chi et al., 2019). For rank $r = 1$ even the location of the threshold is known. Zhang et al. (2019) proved that $\delta_2 < 1/2$ excludes spurious local minima, and at the same time constructed a measurement map and a ground truth with $\delta_2 = 1/2$ that does have one. By monotonicity of the RIP constant such a pair exists for every $\delta_2 \ge 1/2$, so $1/2$ is sharp in the worst-case sense over the class of measurement maps.

Quantum state tomography (QST) is a special case of this framework. The state of an n-qubit system is a positive semidefinite matrix $M \in \mathbb{H}_d$ ($d = 2^n$) with $\mathrm{tr}\, M = 1$ on the n-fold tensor product $(\mathbb{C}^2)^{\otimes n}$, and measuring an observable $A_i \in \mathbb{H}_d$ on M means obtaining $\mathrm{tr}(A_i M)$. Choosing the

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

observables $A_1, \ldots, A_m$ is therefore nothing but choosing the measurement map $\mathcal{A}$. In the noiseless idealization, QST can thus be written as the feasibility problem of finding a state consistent with the measured values $y = \mathcal{A}(M_\star)$:

$$\begin{aligned} &\text{find} && M \in \mathbb{H}_d \\ &\text{subject to} && M \succeq 0, \qquad \operatorname{tr} M = 1, \qquad \operatorname{tr}(A_i M) = y_i \quad (i = 1, \ldots, m). \end{aligned} \tag{3}$$

When the measurement determines a rank-r state uniquely among all states, the feasible set collapses to the single point $\{M_\star\}$ and solving this problem is recovery itself (Kalev et al., 2015; Baldwin et al., 2016). Moreover, what a quantum experiment aims to prepare is in most cases a pure state—a density matrix of rank 1 (Eisert et al., 2020). The convex formulation (3) cannot exploit this prior knowledge and does not guarantee that its solution has rank 1. If one wants a pure state as the output, the natural choice is the Burer–Monteiro factorization, which builds positivity and rank into the search space from the start (Kyrillidis et al., 2018; Kim et al., 2023). Indeed, replacing the equality constraints of (3) by a residual sum of squares, restricting the variable to rank-one matrices and imposing the normalization gives the spherical version of (1),

$$\min_{Y \in \mathbb{S}_d} \ \tfrac{1}{2} \big\| \mathcal{A}(YY^\dagger) - y \big\|_2^2, \qquad \mathbb{S}_d = \{\, u \in \mathbb{C}^d \mid \|u\| = 1 \,\}, \tag{4}$$

Analyzing the landscape of this nonconvex formulation (4) of pure-state QST is the main purpose of this paper.

Nevertheless, FGD has not become the mainstream tool in QST. One major reason is that most theorems certifying benignity of its landscape or escape from saddle points rest on RIP of the measurement map (Bhojanapalli et al., 2016; Kyrillidis et al., 2018). Realizing a RIP measurement system in an actual quantum experiment is difficult, and the assumption should not be used uncritically. The random Pauli measurements for which RIP is known to hold (Liu, 2011), for instance, require building the measurement pool uniformly at random under the premise that all n qubits can be read out simultaneously; in practice the success rate of such a joint readout decays exponentially in n. This is not a quirk of a particular device but a statistical consequence of incorporating imperfect detection efficiency into the model. And if one lengthens the acquisition time in order to collect such full-qubit coincidence data, drift of every device from state preparation to readout becomes unavoidable, so in principle only partial joint-readout events can be used as data (Figure 1). The situation is different for the tomography of pure states. It is known that almost every pure state is determined uniquely, among all states, by the Pauli expectations of subsystems of roughly half the size alone (Linden et al., 2002; Linden and Wootters, 2002; Jones and Linden, 2005; Chen et al., 2013). Moreover, since the search space of rank-1 FGD is the manifold of pure states itself, for the global solution to be confined to the truth it suffices that the measurement map be generically injective on the pure-state manifold. That is, as long as pure states are the target, high-weight Pauli measurements spanning all qubits are unnecessary for identifiability. In this paper we adopt the explicit construction of Zhang et al. (2024)—the Pauli expectations of weight at most N on a $2N$-qubit system determine a generic pure state uniquely—and work with this partial measurement map. What is obtained in this way, however, is identifiability and not RIP. Uniqueness is a statement about "almost every" pure state and allows an exceptional set. Indeed, the following GHZ-type family is not determined by partial measurements even among pure states (Walck and Lyons, 2008). RIP, on the other hand, is a uniform condition over all rank-2 matrices that does not depend on the truth, so a single such exceptional family breaks it in the worst possible way, $\delta_2 \geq 1$, however generic the truth may be. The following counterexample is exactly this.

Counterexample 1.1 (Counterexample to RIP with $\delta_2 \geq 1$)**.** *Let e_0, e_1 be the standard basis of $\mathbb{C}^2$ and put $\psi_\varphi = \frac{1}{\sqrt{2}}\big(e_0^{\otimes n} + e^{\mathrm{i}\varphi}\, e_1^{\otimes n}\big)$ for $\varphi \in [0, 2\pi)$. For $\varphi \neq 0$ this is a pure state different from ψ_0. Yet on any proper subset $S \subsetneq \{1, \ldots, n\}$ the reduced density matrix*

$$\rho^S := \operatorname{tr}_{S^c}\big(\psi_\varphi \psi_\varphi^\dagger\big) \ = \ \tfrac{1}{2}\Big((e_0 e_0^\dagger)^{\otimes S} \ + \ (e_1 e_1^\dagger)^{\otimes S} \Big)$$

takes a value independent of φ. Hence for every measurement of the form $A = A^S \otimes I^{S^c}$ with $|S| = N < n$ the expectation is $\operatorname{tr}(A\, \psi_\varphi \psi_\varphi^\dagger) = \operatorname{tr}\big(A^S \rho^S\big)$, so ψ_0 and ψ_φ cannot be distinguished. Consequently the rank-two Hermitian matrix $X_\varphi = \psi_0 \psi_0^\dagger - \psi_\varphi \psi_\varphi^\dagger$ determined by the difference of the two pure states lies in the kernel of any measurement map $\mathcal{A}$ built solely from the partial measurements $\{A^S \otimes I^{S^c} \mid A^S \in \mathbb{H}_d\}_{S \subsetneq \{1,\ldots,n\}}$. The lower bound in (2) therefore drops to 0 and $\delta_2 \geq 1$, i.e. RIP does not hold.

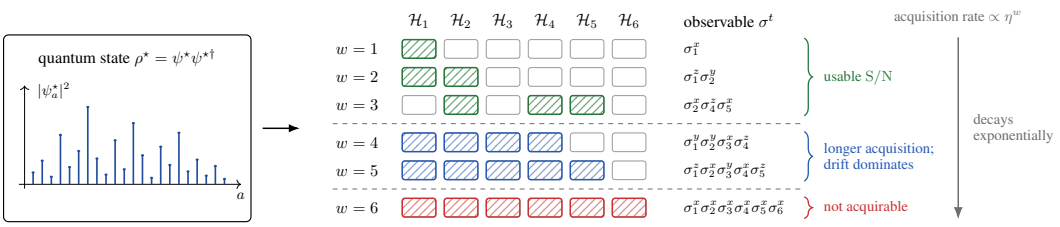


Figure 1: Partial-weight Pauli measurements. Each row is one observable σ^t; the shaded boxes mark the sites on which the string acts nontrivially, and the weight $w(t)$ is the number of such sites. Measuring an observable of weight w requires reading out w sites jointly, so the acquisition rate decays exponentially as η^w, where $\eta \in (0,1)$ is the per-site detection efficiency. Accumulating high-weight statistics therefore takes a long time, during which the apparatus drifts and the state itself changes—the samples cease to be i.i.d. and a systematic bias enters. The coherence that distinguishes the family ψ_φ of Counterexample 1.1 appears only in the $w = n$ string, which falls outside the usable regime.

The starting point of this work is to look squarely at this structural mismatch between the design of the measurement and the design of the loss. The ℓ^2 loss demands, as a regularity condition on the landscape, information about high-order coherences that a partial measurement family cannot touch even in principle. We therefore rebuild the ℓ^2 geometry so that the measurement design and the regularity condition are reconciled. The key is that within one measurement setting the observables commute. Fixing a setting τ, the observables that can be read out simultaneously are diagonalized by a common family of eigenprojectors $\{\Pi^\tau_k\}_k$ with $\sum_k \Pi^\tau_k = I$, and every one of them can be written as $A = \sum_k a_k \Pi^\tau_k$. The measurement map then reads

$$\mathrm{tr}\big(A\, uu^\dagger\big) \;=\; \sum_k a_k\, p^\tau_k(u), \quad p^\tau_k(u) \;=\; u^\dagger \Pi^\tau_k u,$$

so it depends on the state only through the probability vector $p^\tau(u) = (p^\tau_k(u))_k$. In other words, one setting is a map $u \mapsto p^\tau(u)$ that pushes the state forward to a finitely supported probability distribution. If we increase the number of settings τ and bundle the observables setting by setting into slices, noncommutativity is dissolved inside each slice and what remains is the classical problem of comparing two probability vectors. Writing $u_\star u_\star^\dagger$ for the pure state to be recovered and $q^\tau := p^\tau(u_\star)$ for its pushforward, we take as the loss the worst-case total variation between the distributions $p^\tau(u)$ and q^τ over all slices,

$$L(u) \;=\; \max_\tau \big\|p^\tau(u) - q^\tau\big\|_1.$$

This is an analogue of the max-sliced Wasserstein metric, except that the variable selecting the slice is not a free angle but the experimental setting itself, and the distance on each slice is total variation. The price is nonsmoothness; it is a price worth paying, and it is precisely that nonsmoothness which makes the analysis possible. Our main theorem takes the following form.

Theorem 1.2 (Main theorem, informal)**.** *If $L(u) > 0$ and the maximizer of L is unique, then every stationary point is a saddle point and an escape direction is constructed explicitly. Moreover, if the measurement map $\mathcal{A}$ determines u uniquely, then $L(u) = 0 \iff u = u_\star$.*

Our numerical experiments use 4 qubits ($N = 2$, pair blocks). Comparing on identical data from identical random initial points, in the noiseless case the subgradient method (with momentum) on the max-sliced loss recovers the truth in $100/100$ trials, whereas the ordinary ℓ^2 loss is trapped at a spurious local minimum in 25 of them. With finite shots, a two-stage scheme—descend on the max-sliced loss, then polish with the ℓ^2 loss—is best at every shot budget we tried.

Remark 1.3 (Related work: matrix completion)**.** *Even for ordinary FGD there are situations in which benignity is recovered without RIP, the prime example being matrix completion. There the measurement map extracts entries, and it annihilates the matrix $e_ie_j^\top + e_je_i^\top$ supported on a single unobserved entry (i,j), so RIP fails with $\delta_2 \geq 1$. Even so, Ge et al. (2016; 2017) assumed that the rows of the truth have uniformly small norms (incoherence) and added a penalty on factors with large row norms, and proved the absence of spurious local minima. The mechanism is that the kernel consists of matrices whose mass is concentrated on a few entries, so one keeps the search away from*

such matrices and recovers an isometry there. The state of Counterexample 1.1 is likewise localized on a few components, but we regard even a neighbourhood of it as an object to be recovered, so such a keep-away prescription is unavailable here.

Remark 1.4 (Related work: convex relaxations for low-rank tomography)**.** *Low-rank state tomography was originally formulated as compressed sensing: $O(rd\,\mathrm{polylog}\,d)$ random Pauli expectations determine a rank-r state, and under RIP-type conditions convex programs recover it stably (Gross et al., 2010; Liu, 2011; Flammia et al., 2012; Kueng et al., 2017; Kalev et al., 2015; Baldwin et al., 2016). Under noise, however, the rank of the convex solution is a by-product of the noise and the active constraints and generally differs from r (Blume-Kohout, 2010; Scholten and Blume-Kohout, 2018), so a rank-r description requires an a posteriori truncation, after which the estimate is no longer the solution of the convex program. Factorization-based methods avoid this by optimizing inside the model class (Kyrillidis et al., 2018; Hsu et al., 2024). Moreover, our partial measurements do not provide the uniqueness among all states that the convex guarantees presuppose—only generic injectivity on the pure-state manifold—so recovering pure states here is effectively limited to nonconvex formulations.*

2 Preliminaries

2.1 Quantum state tomography on $2N$ qubits

The state space of a qubit, the minimal unit of quantum information, is $\mathbb{C}^2$, and a quantum state on it—a positive semidefinite matrix $M \in \mathbb{H}_2$ with $\mathrm{tr}\, M = 1$—can be written uniquely as

$$M \;=\; \tfrac{1}{2}\big(I + r_x\sigma^x + r_y\sigma^y + r_z\sigma^z\big), \qquad r_a = \mathrm{tr}(\sigma^a M) \in \mathbb{R},$$

$$\sigma^x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \sigma^y = \begin{pmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{pmatrix}, \qquad \sigma^z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \qquad \sigma^0 = I. \tag{5}$$

This is the Bloch vector expansion of M. Here $M \succeq 0$ is equivalent to $\|r\| \le 1$, and purity $\mathrm{rank}\, M = 1$ to $\|r\| = 1$. The three matrices $\sigma^x, \sigma^y, \sigma^z$ appearing as coefficients are the Pauli matrices, and each component of the Bloch vector is nothing but their expectation: knowing the state of a qubit is the same as knowing the expectations of the three Pauli matrices. Expectations, however, are not directly readable quantities, and this is where the quantum measurement procedure enters. Each σ^a ($a \in \{x, y, z\}$) has eigenvalues ± 1; writing v_k^a for the unit eigenvector belonging to the eigenvalue $(-1)^k$ ($k \in \{0, 1\}$) and

$$\Pi_k^a := v_k^a (v_k^a)^\dagger = \tfrac{1}{2}\big(I + (-1)^k \sigma^a\big)$$

for the corresponding projector, the condition $\sum_{k=0,1} \Pi_k^a = I$ of a projective measurement holds. Looking at the expectation of σ^a then means reading the state out repeatedly with detectors corresponding to the two projectors $\{\Pi_0^a, \Pi_1^a\}$. One readout returns a single $k \in \{0, 1\}$, with probability given by the Born rule

$$\mathbb{P}[k] \;=\; \mathrm{tr}\big(\Pi_k^a M\big) \;=\; \tfrac{1}{2}\big(1 + (-1)^k r_a\big)$$

(and $\sum_k \mathbb{P}[k] = 1$ follows from $\sum_k \Pi_k^a = I$). Hence R readouts are i.i.d. Bernoulli trials and the number R_0 of outcomes $k = 0$ follows the binomial distribution

$$\mathbb{P}[R_0 = s] \;=\; \binom{R}{s} \Big(\frac{1 + r_a}{2}\Big)^s \Big(\frac{1 - r_a}{2}\Big)^{R-s}.$$

The estimator built from the relative frequency is $\hat{r}_a = 2R_0/R - 1$, with $\mathbb{E}[\hat{r}_a] = r_a$ and $\mathrm{Var}[\hat{r}_a] = (1 - r_a^2)/R$. The primary data of an experiment are thus counts rather than expectations, and the estimate $\hat{r}$ built from them can easily violate, at finite sample size, not only the purity condition $\|\hat{r}\| = 1$ but even the condition $\|\hat{r}\| \le 1$ of being a quantum state at all. This shows how useful it is to impose prior knowledge as a hard constraint on the optimization variable.

An n-qubit system is described by the tensor product: the state space is $(\mathbb{C}^2)^{\otimes n} \cong \mathbb{C}^d$ with $d = 2^n$, and a state is a positive semidefinite $M \in \mathbb{H}_d$ with $\mathrm{tr}\, M = 1$. We write $M \in \mathbb{S}_d$ for a pure state, i.e. an $M \in \mathbb{H}_d$ with $\mathrm{rank}\, M = 1$. The Pauli strings

$$P_b \;=\; \sigma^{b_1} \otimes \cdots \otimes \sigma^{b_n}, \qquad b = (b_1, \ldots, b_n) \in \{x, y, z, 0\}^n,$$

satisfy $\mathrm{tr}(P_b P_{b'}) = d\,\delta_{bb'}$ and form an orthogonal basis of $\mathbb{H}_d$, and the Bloch expansion generalizes to

$$M \;=\; \frac{1}{d} \sum_{b\in\{x,y,z,0\}^n} [\,\mathrm{tr}(P_b M)\,]\; P_b. \tag{6}$$

Since $\mathrm{tr}(IM) = 1$ is the normalization, a state is determined by the remaining $4^n - 1$ Pauli expectations. The number of qubits on which P_b acts nontrivially, i.e. the number of indices i with $b_i \neq I$, is called the weight of P_b and written $w(b)$. The measurement procedure is inherited from the single-qubit case. Choosing a readout axis $b_i \in \{x, y, z\}$ at each site, the product basis $e_{b,k} = \bigotimes_i v_{k_i}^{b_i}$ ($k \in \{0,1\}^n$) is an orthonormal basis of $\mathbb{C}^d$, and the projectors

$$\Pi_k^b := e_{b,k} e_{b,k}^\dagger, \qquad \sum_{k\in\{0,1\}^n} \Pi_k^b = I$$

define a projective measurement. One readout returns a single bit string k, with probability $\mathbb{P}[k] = \mathrm{tr}(\Pi_k^b M)$ by the axioms, and the counts over R readouts follow, instead of a binomial, the multinomial distribution $\mathrm{Mult}\big(R, (\mathrm{tr}(\Pi_k^b M))_k\big)$ over the 2^n cells. What differs from the single-qubit case is that one readout yields several observables at once. Expanding the tensor product of $\Pi_k^a = \frac{1}{2}\big(I + (-1)^k \sigma^a\big)$ gives

$$\Pi_k^b \;=\; 2^{-n} \sum_{T\subseteq\{1,\ldots,n\}} (-1)^{\sum_{i\in T} k_i}\, P_{b,T}, \qquad P_{b,T} = \bigotimes_{i\in T} \sigma^{b_i}, \tag{7}$$

and conversely, by orthogonality of the Walsh functions $\sum_k (-1)^{\sum_{i\in T} k_i + \sum_{i\in T'} k_i} = 2^n \delta_{T,T'}$, multiplying both sides of (7) by $(-1)^{\sum_{i\in T'} k_i}$ and summing over k leaves only the term $T = T'$, so that

$$P_{b,T} = \sum_k (-1)^{\sum_{i\in T} k_i} \Pi_k^b. \tag{8}$$

Thus choosing one axis assignment b yields simultaneously the expectations of all 2^n Pauli strings $\{P_{b,T}\}_T$ along b, as signed sums of the same counts; conversely the expectation of a string not along b cannot be obtained from that readout. Designing a measurement thus amounts to choosing which axis assignments b to prepare, and how many. Measuring a string of weight w, however, requires reading out w sites jointly, so with per-site efficiency $\eta \in (0, 1)$ the acquisition rate decays exponentially as η^w. Moreover, for large n the device drift sets in before high-weight statistics can be accumulated well above the noise, so what is used in the laboratory are measurements restricted to low weight (Figure 1). We therefore adopt the following standing assumption.

Assumption 2.1 (Measurement model)**.** *In an $n = 2N$ qubit system, the object of measurement is a pure state $M_\star = u_\star u_\star^\dagger$, and the expectations of Pauli strings up to weight N are assumed to be statistically acquirable.*

On the other hand, it is known that almost every pure state of a $2N$-qubit system is uniquely determined by its N-body reduced density matrices, i.e. by the Pauli expectations of weight at most N. Partial-weight Pauli measurements are thus a design forced by experimental constraints and justified theoretically at the same time.

Proposition 2.2 ((Zhang et al., 2024))**.** *Under the measurements of Assumption 2.1, almost every pure state is uniquely determined within the set of all pure states.*

2.2 MAX-SLICED LOSS FUNCTION

Based on Assumption 2.1 we now define the loss used in this paper, after fixing some terminology.

Definition 2.3 (Setting and slice)**.** Write v_k^a for the unit eigenvector of σ^a ($a \in \{x, y, z\}$) belonging to the eigenvalue $(-1)^k$, $k \in \{0, 1\}$. A *measurement setting* is a pair $\tau = (S, b)$ of a subset of qubits $S \subset \{1, 2, \ldots, 2N\}$ with $|S| = N$ and a label string $b \in \{x, y, z\}^S$. The measurement in the setting τ is described, through the orthonormal basis $\{e_{b,k} = \bigotimes_{i\in S} v_{k_i}^{b_i} : k \in \{0, 1\}^S\}$ of the S-qubits, by

$$\Pi_k^\tau \;=\; e_{b,k} e_{b,k}^\dagger \otimes I, \qquad \sum_k \Pi_k^\tau = I. \tag{9}$$

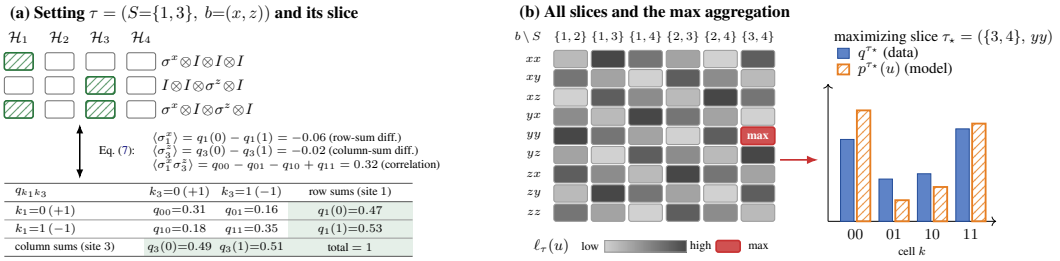


Figure 2: Structure of a slice. (a) The commuting family (slice) read out at once in the setting $\tau = (\{1,3\},(x,z))$ and the data of that setting. Data accumulate as an empirical distribution q^τ over cells (a cell k is the tuple of readout bits of the sites in the block); the interior of the contingency table is the weight-2 record and the margins are its marginalization, i.e. the weight-1 data. Each string expectation is a signed sum of counts, expectations and counts being interconvertible by (7). (b) The settings form a (basis b) $\times$ (block S) grid, of size $6 \times 9 = 54$ for $n = 4$, $N = 2$; the max-sliced loss picks the slice $\tau_\star$ (red) at which the total variation $\ell_\tau(u)$ (shading) between the Born prediction $p^\tau(u)$ and the data is worst, and its cell distribution is shown on the right. Weight-1 strings are diagonal in several settings, so the slices cover the measurement set with overlaps.

For fixed τ the $\{\Pi_k^\tau\}_k$ are mutually orthogonal projectors forming a resolution of the identity $\sum_k \Pi_k^\tau = I$. What can be read out simultaneously in the setting τ is the family of all Pauli substrings diagonal in the readout basis, hence mutually commuting. we call them a *slice* (Figure 2(a)).

Definition 2.4 (Cell distribution)**.** When the measured object is a pure state $M = uu^\dagger \in \mathbb{H}_d$, the probability that one measurement returns the bit string k is

$$p_k^\tau(u) \;=\; u^\dagger \,\Pi_k^\tau\, u \tag{10}$$

by the axioms of quantum mechanics. We call the realized value k a *cell*, $p^\tau(u) = (p_k^\tau(u))_k$ the *cell distribution* of the setting τ, and, writing q_k^τ for the relative frequency of the cell k, $q^\tau = (q_k^\tau)_k$ the *empirical cell distribution*.

Definition 2.5 (Max-sliced loss)**.** For a model $M = uu^\dagger$ with $u \in \mathbb{S}_d$, let $p^\tau(u) = (p_k^\tau(u))_k$ be the cell probabilities in the setting τ, let $q^\tau = (q_k^\tau)_k$ be the empirical cell distribution obtained from the truth $M_\star = u_\star u_\star^\dagger$, and write $\delta_k^\tau(u) = p_k^\tau(u) - q_k^\tau$ for the residual inside the setting τ. The total-variation loss of the slice τ and the max-sliced loss L are defined by

$$\ell_\tau(u) := \;\big\|p^\tau(u) - q^\tau\big\|_1 \;=\; \sum_k \big|\delta_k^\tau(u)\big|, \tag{11}$$

$$L(u) := \max_\tau\; \ell_\tau(u). \tag{12}$$

Directly from the definition, mass is conserved in every slice:

$$\sum_k \delta_k^\tau(u) \;=\; 0. \tag{13}$$

3 Main result and sketch of its proof

The purpose of this section is to analyze the structure of the stationary points of the max-sliced loss L. The outcome is that, when the maximizer is unique, every stationary point is a saddle point.

Definition 3.1 (Stationary point)**.** Let $u \in \mathbb{S}$ and let v be a unit tangent direction orthogonal to u. Write $L'(u;v) = \lim_{t\downarrow 0}\big(L(\gamma_v(t)) - L(u)\big)/t$ for the one-sided directional derivative along the great circle $\gamma_v(t) = \cos t\, u + \sin t\, v$ (existence is shown in the appendix). If $L'(u;v) \geq 0$ in every tangent direction, u is called a *stationary point* of L.

Lemma 3.2. *L is 2-Lipschitz on $\mathbb{S}_d$, and moreover L is Clarke regular.*

Remark 3.3. *As seen in Lemma 3.2, L is nonsmooth but globally Lipschitz, hence Clarke subdifferentiable and Clarke regular at every point (Clarke, 1990). The stationary points of Definition 3.1 therefore coincide with Clarke stationary points, the standard notion in nonsmooth analysis.*

Definition 3.4 (Support condition)**.** Let $L(u) > 0$ and let the maximizing slice $\tau_\star = \tau(u)$ be unique. Split the cells by the sign of the residual $\delta_k = \delta_k^{\tau_\star}(u)$ into $\mathcal{P} = \{k : \delta_k > 0\}$, $\mathcal{Z} = \{k : \delta_k = 0\}$ and $\mathcal{D} = \{k : \delta_k < 0\}$ (mass conservation (13) implies $\mathcal{P}, \mathcal{D} \neq \emptyset$). The *support condition* is that every deficient cell be disjoint from the support of u, i.e. that

$$\Pi_k^{\tau_\star} u = 0 \qquad (\forall k \in \mathcal{D}). \tag{14}$$

The main result is a classification of stationary points on the stratum where the active slice is unique.

Theorem 3.5 (Landscape theorem)**.** *Let $L(u) > 0$ and let the maximizing slice be unique. Then u is a stationary point if and only if the support condition holds. Moreover, when the support condition holds, u is not a local minimum: for any deficient cell $k_0 \in \mathcal{D}$ and any unit vector $v \in \operatorname{ran} \Pi_{k_0}^{\tau_\star}$, along the great circle γ_v and for all sufficiently small $t > 0$,*

$$L(\gamma_v(t)) \;=\; L(u) \;-\; 2P_{\mathcal{P}}(u)\,\sin^2 t, \qquad P_{\mathcal{P}}(u) = \textstyle\sum_{k\in\mathcal{P}} p_k^{\tau_\star}(u) \;>\; 0.$$

Corollary 3.6. *Every spurious local minimum of L (a local minimum with positive value) lies in the* tie set*, the set of points at which several slices share the maximum. Moreover, under Assumption 2.1, the global minimum of L is the truth up to a global phase.*

3.1 Sketch of the proof of Theorem 3.5

The complete proof is in Appendix A; the mechanism splits into four steps.

Step 1: localization. Let $\Delta = \ell_{\tau_\star}(u) - \max_{\tau\neq\tau_\star} \ell_\tau(u) > 0$ be the gap between the maximizing slice $\tau_\star$ and the runner-up. Since every ℓ_τ is 2-Lipschitz (Lemma 3.2), none of them moves by more than $\Delta/2$ inside the ball of radius $\Delta/4$, so $\tau_\star$ remains the unique maximizer there. Hence $L = \ell_{\tau_\star}$ in that neighbourhood, and stationarity, local minimality and one-sided directional derivatives of L all agree with those of the single slice $\ell_{\tau_\star}$.

Step 2: first order mass flows. Putting $dp_k^{\tau_\star}(v) = 2\,\mathrm{Re}\langle \Pi_k^{\tau_\star} u, v\rangle$, the v-directional derivative is

$$\ell'_{\tau_\star}(u; v) = \sum_{k\in\mathcal{P}} dp_k^{\tau_\star}(v) - \sum_{k\in\mathcal{D}} dp_k^{\tau_\star}(v) + \sum_{k\in\mathcal{Z}} |dp_k^{\tau_\star}(v)|,$$

where only the absolute values on $\mathcal{Z}$ carry the nonsmoothness. The first-order mass transfer is subject to two restrictions: (i) the resolution $\sum_k \Pi_k^{\tau_\star} = I$ and the tangency condition give $\sum_k dp_k^{\tau_\star}(v) = 0$, so only zero-sum flows can be realized at first order; (ii) since $p_k^{\tau_\star}(u) = \|\Pi_k^{\tau_\star} u\|^2$ is quadratic, a cell with $p_k^{\tau_\star}(u) = 0$ has $\Pi_k^{\tau_\star} u = 0$ and hence $dp_k^{\tau_\star}(v) = 0$ for every v, so no first-order inflow into an empty cell is possible. Conversely, every flow satisfying these two restrictions is realizable: for any zero-sum vector (α_k) on the support $\{k : p_k^{\tau_\star}(u) > 0\}$, the direction $v = \sum_k \frac{\alpha_k}{2p_k^{\tau_\star}(u)} \Pi_k^{\tau_\star} u$ realizes $dp_k^{\tau_\star}(v) = \alpha_k$. In short, the flows available at first order are exactly the zero-sum flows on the support.

Step 3: when the support condition fails. Suppose a deficient cell $k_- \in \mathcal{D}$ carries mass. Pick a surplus cell $k_+ \in \mathcal{P}$ (mass conservation gives $\mathcal{P} \neq \emptyset$, and on $\mathcal{P}$ we have $p_k^{\tau_\star} > q_k^{\tau_\star} \geq 0$, hence $p_k^{\tau_\star} > 0$). Both $k_\pm$ lie in the support, so by the construction of Step 2 the flow $\alpha_{k_+} = -\varepsilon$, $\alpha_{k_-} = +\varepsilon$ and 0 elsewhere is realizable. For this v every other cell has $dp_j^{\tau_\star}(v) = 0$, so only two terms survive in the expression above and

$$\ell'_{\tau_\star}(u; v) = dp_{k_+}^{\tau_\star}(v) - dp_{k_-}^{\tau_\star}(v) = (-\varepsilon) - (+\varepsilon) = -2\varepsilon < 0.$$

The ε removed from the surplus cell decreases $\delta_{k_+}^{\tau_\star}$ and the ε poured into the deficient cell decreases $|\delta_{k_-}^{\tau_\star}|$. The descent comes from both sides. Hence u is neither stationary nor a local minimum.

Step 4: when the support condition holds. Now all cells of $\mathcal{D}$ are empty, so their $dp_k^{\tau_\star}$ vanish, and the same holds for $\mathcal{Z} \cap \{p_k^{\tau_\star} = 0\}$. Rewriting the first-order mass conservation over the surviving cells gives $\sum_{k\in\mathcal{P}} dp_k^{\tau_\star}(v) = -\sum_{k\in\mathcal{Z},\, p_k^{\tau_\star}>0} dp_k^{\tau_\star}(v)$, and substituting,

$$\ell'_{\tau_\star}(u; v) = \sum_{k\in\mathcal{Z},\, p_k^{\tau_\star}>0} \big(|dp_k^{\tau_\star}(v)| - dp_k^{\tau_\star}(v)\big) \;\geq\; 0$$

for every v, i.e. u is stationary: lowering the loss would require pouring mass into a deficient cell, but those are empty and inflow into an empty cell does not occur at first order. At second order, however,

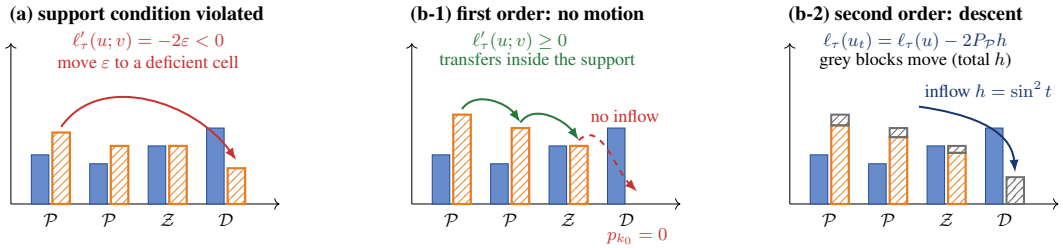


Figure 3: The support condition and the order of the mass flow. Blue is the data q^τ and orange hatching the model $p^\tau(u)$. Cells are split by the sign of the residual $\delta_k = p_k - q_k$ into surplus $\mathcal{P}$, zero $\mathcal{Z}$ and deficient $\mathcal{D}$, and what can be moved at first order is exactly a zero-sum flow on the support ($p_k > 0$) of the model. (a) When the support condition fails, a deficient cell carries mass, so the flow moving ε from a surplus cell to it is realizable and the loss decreases strictly: the point is not stationary. (b-1) Under the support condition the deficient cell is empty, no first-order inflow into it is possible, and only transfers inside the support remain, none of which decreases the loss in any direction: u is stationary. (b-2) Along the great circle $u_t = \cos t\, u + \sin t\, v$ with $v \in \operatorname{ran} \Pi^\tau_{k_0}$, however, the inflow into the empty cell, $p_{k_0}(u_t) = \sin^2 t =: h$, is achieved at second order in t while every support cell is scaled by $(1-h)$; the grey hatched blocks are the moving mass. As long as no residual changes sign this strictly decreases the loss, so u is not a local minimum.

the loss can be lowered. Take an empty deficient cell $k_0 \in \mathcal{D}$ and a unit vector $v \in \operatorname{ran} \Pi^{\tau_\star}_{k_0}$. The support condition gives $\Pi^{\tau_\star}_{k_0} u = 0$, hence $\langle u, v\rangle = 0$ and $\Pi^{\tau_\star}_k v = 0$ for $k \neq k_0$. By this orthogonality the projectors act diagonally along the great circle $u_t = \cos t\, u + \sin t\, v$, so with $h = \sin^2 t$,

$$p^{\tau_\star}_k(u_t) = (1-h)\, p^{\tau_\star}_k(u) \;\; (k \neq k_0), \qquad p^{\tau_\star}_{k_0}(u_t) = h.$$

It remains to count the change of the absolute residuals cell by cell. In the range $h \leq h_0 := \min\{\min_{k\in\mathcal{P}} \delta^{\tau_\star}_k / p^{\tau_\star}_k,\, q^{\tau_\star}_{k_0}\}$ no residual changes sign, so cells of $\mathcal{P}$ decrease by $h\, p^{\tau_\star}_k$, cells of $\mathcal{Z} \cap \{p^{\tau_\star}_k > 0\}$ increase by $h\, p^{\tau_\star}_k$, the cell k_0 decreases by h, and the rest are unchanged. Using that the whole mass of p sits on $\mathcal{P}$ and $\mathcal{Z} \cap \{p^{\tau_\star}_k > 0\}$, the total is

$$\ell_{\tau_\star}(u_t) - \ell_{\tau_\star}(u) = -h\, P_\mathcal{P} + h\, P_{\mathcal{Z}_+} - h = -2h\, P_\mathcal{P},$$

and $P_\mathcal{P} \geq \ell_{\tau_\star}(u)/2 > 0$ guarantees a strict descent. By the localization of Step 1 this identity holds for L as well for all sufficiently small t.

4 Algorithms and Experiments

Data generation. All experiments use $n = 4$ ($N = 2$, pair blocks, $d = 16$). As a target that empirically tends to create traps we start from a state with equal amplitudes on all computational basis vectors, differing only in sign. Writing e_a ($a \in \{0,1\}^n$) for the computational basis and $f(a) = \sum_{(i,j)\in E} a_i a_j \pmod 2$ for the quadratic form determined by an edge set E, put

$$\psi_0 \;=\; \frac{1}{\sqrt{d}} \sum_{a\in\{0,1\}^n} (-1)^{f(a)}\, e_a. \tag{15}$$

In quantum information this family is called a graph state, and taking E to be a cycle on n vertices gives the ring cluster state. For $n = 4$ and $E = \{(1,2),(2,3),(3,4),(4,1)\}$ only $a \in \{0011, 0110, 1001, 1100\}$ carry a negative sign. This family is highly symmetric and loses unique determination under weight-restricted measurements. Into it we mix a phased W-type state $W = (e_{0001} + \mathrm{i}\, e_{0010} - e_{0100} - \mathrm{i}\, e_{1000})/2$ orthogonal to ψ_0, forming the reference state

$$\psi^\star \;\propto\; 4\,\psi_0 + W, \qquad |\psi_0^\dagger \psi^\star|^2 = \tfrac{16}{17},$$

whose W component decisively breaks the symmetry of ψ_0. In addition, for each trial we draw an independent isotropic complex Gaussian unit direction g orthogonal to $\psi^\star$ and take as the target

$$\psi_\varepsilon \;=\; \frac{\psi^\star + \varepsilon\, g}{\|\psi^\star + \varepsilon\, g\|}, \qquad \varepsilon = 0.05 \tag{16}$$

Loss (aggregation)	Definition	Success rate ($F > 0.999$)
ℓ^2 sum of Pauli residuals (FGD)	$\|\mathcal{A}_{\mathcal{P}}(uu^\dagger) - y\|_2^2$	$75/100$
ℓ^2 sum of cell residuals	$\sum_\tau \|p^\tau(u) - q^\tau\|_2^2$	$94/100$
Sum of slice total variations	$\sum_\tau \|p^\tau(u) - q^\tau\|_1$	$74/100$
Max-sliced total variation (12)	$\max_\tau \|p^\tau(u) - q^\tau\|_1$	$\mathbf{100/100}$

Table 1: Identical noiseless data and identical initial points. The first three rows all sum over τ, only the last takes a max. The max-sliced loss reaches $F = 1.0000$ in every trial.

	ℓ^2 (FGD)		max-sliced		two-stage	
R	success	median F	success	median F	success	median F
10^5	0.75	0.9999	1.00	0.9997	**1.00**	**0.9999**
10^4	0.72	0.9990	1.00	0.9971	**1.00**	**0.9993**
10^3	0.35	0.9717	0.00	0.9722	**0.65**	**0.9937**

Table 2: Finite-shot comparison (40 trials each, $K = 54$; success means $F > 0.99$). Two-stage = ℓ^2 polishing initialized at the endpoint of the max-sliced descent.

(with $F(\psi_\varepsilon, \psi^\star) = 1/(1+\varepsilon^2) \approx 0.998$). The perturbed target lands at a generic point and falls into the region of unique determination on the pure-state manifold in the sense of Corollary 3.6, whereas the ℓ^2 landscape is not benign. The target is uniquely determined and yet least squares gets trapped, exactly the situation envisaged by Theorem 3.5. Physically it corresponds to a state prepared on hardware that deviates slightly from the ideal symmetric state. Fidelity is measured by $F(\varphi, \psi) = |\psi^\dagger \varphi|^2$, which is invariant under the global phase.

Algorithm. The slice family has $K = 54$ members (6 pairs $\times$ 9 bases). To minimize the loss L we use Polyak subgradient with momentum, driven by the signed subgradient g of the active slice,

$$w \leftarrow \beta\, w - \frac{L(u)}{\|g\|^2}\, g, \qquad u \leftarrow \frac{u+w}{\|u+w\|}, \qquad \beta = 0.5.$$

Here $g = 2\sum_k \operatorname{sign}\big(\delta_k^{\tau_\star}(u)\big)\, \Pi_k^{\tau_\star} u$ for the maximizing slice $\tau_\star$, an element of the Clarke subdifferential $\partial_C L(u)$. Initial points are uniform on the sphere, and the baseline factored least squares (FGD) is run from the same data and the same initial point. Data are generated as in Assumption 2.1. Detection efficiency and drift are not simulated—they motivate the weight restriction (§1), and the restricted measurement is treated as ideal multinomial counting.

Results. We first compare, on exact data ($R = \infty$), four losses that differ only in how the residuals are aggregated (Table 1). The ℓ^2 sum of Pauli residuals is trapped at a spurious local minimum ($F \approx 0$) in 25 of 100 trials, and replacing the residuals by cell distributions, or by total variation, does not remove the traps. They disappear when the sum is replaced by a max. What matters is not the choice of distance but the change of aggregation from sum to max.

The finite-shot comparison is given in Table 2. The max-sliced loss has no traps at $R = 10^4, 10^5$ (minimum fidelity 0.994 and 0.999), but because of the statistical floor its final accuracy is slightly below the successful runs of ℓ^2. A *two-stage* scheme—descend with the max-sliced loss, then polish with ℓ^2—combines the absence of traps with the accuracy and is best at every shot budget. At $R = 10^3$ the max-sliced loss alone does not reach the threshold $F > 0.99$ (median 0.972), and a few of the 40 runs stall in a near-tie band.

5 Discussion

Theorem 3.5 classifies the stationary points on the stratum where the maximizing slice is unique, and Corollary 3.6 confines spurious local minima to the tie set, where several slices share the maximum. In practice an exact tie is never attained because of numerical error, and the silent zone shows up as oscillation rather than as stationarity—an oscillation that momentum largely removes: in the setting of §4 the success rate improves from $72/100$ without momentum to $100/100$ with it. A rigorous analysis of the stationary points on the tie set is left for future work.

Reproducibility statement

All theoretical claims are stated with their assumptions in the main text and proved in full in Appendix A: Lemma 3.2, Theorem 3.5 and Corollary 3.6 are restated there under the same numbers, and the four steps of the proof sketch of §3 correspond to Lemmas A.1–A.4. For the experiments, the target states are defined explicitly by (15) and (16), the measurement model and the data generation are fixed by Assumption 2.1, and the optimizer, its step rule and the initialization are given in §4; the reported numbers are aggregates over the stated numbers of independent trials on identical data and identical initial points. Code reproducing every table, together with the random seeds, is provided as anonymous supplementary material.

AI use statement

In this work, we used generative AI tools for implementing the numerical experiments, for translating and copy-editing the manuscript, and for producing the figure sources. We have not used generative AI tools for research ideation, for formulating the theoretical claims, or for constructing their proofs; the definitions, the statements and the proofs in §3 and Appendix A were written and verified by the author. We have reviewed all AI-assisted work: the experiment code was inspected and re-run by the author, every number reported in §4 was reproduced from that code, and all references were checked against their primary sources. We take responsibility for the final content of this work, including text, claims or artifacts produced with the aid of generative AI.

References

Charles H. Baldwin, Ivan H. Deutsch, and Amir Kalev. Strictly-complete measurements for bounded-rank quantum-state tomography. *Physical Review A*, 93(5):052105, 2016.

Srinadh Bhojanapalli, Behnam Neyshabur, and Nathan Srebro. Global optimality of local search for low rank matrix recovery. In *Advances in Neural Information Processing Systems*, volume 29, 2016.

Robin Blume-Kohout. Hedged maximum likelihood estimation. *Physical Review Letters*, 105(20):200504, 2010.

Samuel Burer and Renato D. C. Monteiro. A nonlinear programming algorithm for solving semidefinite programs via low-rank factorization. *Mathematical Programming*, 95(2):329–357, 2003.

Jianxin Chen, Hillary Dawkins, Zhengfeng Ji, Nathaniel Johnston, David Kribs, Frederic Shultz, and Bei Zeng. Uniqueness of quantum states compatible with given measurement results. *Physical Review A*, 88:012109, 2013.

Yuejie Chi, Yue M. Lu, and Yuxin Chen. Nonconvex optimization meets low-rank matrix factorization: an overview. *IEEE Transactions on Signal Processing*, 67(20):5239–5269, 2019.

Frank H. Clarke. *Optimization and Nonsmooth Analysis*, volume 5 of *Classics in Applied Mathematics*. SIAM, Philadelphia, 1990.

Jens Eisert, Dominik Hangleiter, Nathan Walk, Ingo Roth, Damian Markham, Rhea Parekh, Ulysse Chabaud, and Elham Kashefi. Quantum certification and benchmarking. *Nature Reviews Physics*, 2(7):382–390, 2020.

Steven T. Flammia, David Gross, Yi-Kai Liu, and Jens Eisert. Quantum tomography via compressed sensing: error bounds, sample complexity and efficient estimators. *New Journal of Physics*, 14(9):095022, 2012.

Rong Ge, Jason D. Lee, and Tengyu Ma. Matrix completion has no spurious local minimum. In *Advances in Neural Information Processing Systems*, volume 29, pages 2981–2989, 2016.

Rong Ge, Chi Jin, and Yi Zheng. No spurious local minima in nonconvex low rank problems: a unified geometric analysis. In *Proceedings of the 34th International Conference on Machine Learning*, volume 70 of *Proceedings of Machine Learning Research*, pages 1233–1242, 2017.

David Gross, Yi-Kai Liu, Steven T. Flammia, Stephen Becker, and Jens Eisert. Quantum state tomography via compressed sensing. *Physical Review Letters*, 105(15):150401, 2010.

Ming-Chien Hsu, En-Jui Kuo, Wei-Hsuan Yu, Jian-Feng Cai, and Min-Hsiu Hsieh. Quantum state tomography via nonconvex Riemannian gradient descent. *Physical Review Letters*, 132(24):240804, 2024.

Nick S. Jones and Noah Linden. Parts of quantum states. *Physical Review A*, 71:012324, 2005.

Amir Kalev, Robert L. Kosut, and Ivan H. Deutsch. Quantum tomography protocols with positivity are compressed sensing protocols. *npj Quantum Information*, 1:15018, 2015.

Junhyung Lyle Kim, George Kollias, Amir Kalev, Ken X. Wei, and Anastasios Kyrillidis. Fast quantum state reconstruction via accelerated non-convex programming. *Photonics*, 10(2):116, 2023.

Richard Kueng, Holger Rauhut, and Ulrich Terstiege. Low rank matrix recovery from rank one measurements. *Applied and Computational Harmonic Analysis*, 42(1):88–116, 2017.

Anastasios Kyrillidis, Amir Kalev, Dohyung Park, Srinadh Bhojanapalli, Constantine Caramanis, and Sujay Sanghavi. Provable compressed sensing quantum state tomography via non-convex methods. *npj Quantum Information*, 4(1):36, 2018.

Noah Linden, Sandu Popescu, and William K. Wootters. Almost every pure state of three qubits is completely determined by its two-particle reduced density matrices. *Physical Review Letters*, 89:207901, 2002.

Noah Linden and William K. Wootters. The parts determine the whole in a generic pure quantum state. *Physical Review Letters*, 89:277906, 2002.

Yi-Kai Liu. Universal low-rank matrix recovery from Pauli measurements. In *Advances in Neural Information Processing Systems*, volume 24, pages 1638–1646, 2011.

Travis L. Scholten and Robin Blume-Kohout. Behavior of the maximum likelihood in quantum state tomography. *New Journal of Physics*, 20(2):023050, 2018.

Scott N. Walck and David W. Lyons. Only n-qubit Greenberger–Horne–Zeilinger states are undetermined by their reduced density matrices. *Physical Review Letters*, 100:050501, 2008.

Richard Y. Zhang, Somayeh Sojoudi, and Javad Lavaei. Sharp restricted isometry bounds for the inexistence of spurious local minima in nonconvex matrix recovery. *Journal of Machine Learning Research*, 20(114):1–34, 2019.

Wanchen Zhang, Fei Shi, and Xiande Zhang. Almost all even-particle pure states are determined by their half-body marginals. *Journal of Physics A: Mathematical and Theoretical*, 57(49):495302, 2024.

Zhihui Zhu, Qiuwei Li, Gongguo Tang, and Michael B. Wakin. Global optimality in low-rank matrix optimization. *IEEE Transactions on Signal Processing*, 66(13):3614–3628, 2018.

A Proof of the landscape theorem

In this appendix we restate Lemma 3.2, Theorem 3.5 and Corollary 3.6 of the main text under their own numbers and prove them. The organization follows the order of the proof sketch in §3: after the preliminaries and the proof of Lemma 3.2, Step 1 (localization) is Lemma A.1, Step 2 (mass flows realizable at first order) is Lemma A.2, Step 3 (the support condition fails) is Lemma A.3 and Step 4 (the support condition holds) is Lemma A.4; combining them proves Theorem 3.5 and Corollary 3.6.

A.1 PRELIMINARIES

Throughout, $\mathbb{S} = \{u \in \mathbb{C}^d : \|u\| = 1\}$, and the *real tangent space* at $u \in \mathbb{S}$ is

$$T_u = \big\{v \in \mathbb{C}^d : \ \mathrm{Re}\langle u, v\rangle = 0\big\},$$

the set of directions orthogonal to u in the real part of the complex inner product, a real linear space of dimension $2d-1$. A "unit tangent direction orthogonal to u" in Definition 3.1 is a unit vector of T_u, and for such a v the great circle $\gamma_v(t) = \cos t\, u + \sin t\, v$ lies on $\mathbb{S}$ because $\|\gamma_v(t)\|^2 = \cos^2 t + \sin^2 t + 2\sin t \cos t\, \mathrm{Re}\langle u, v\rangle = 1$. The one-sided directional derivative $f'(u; v) = \lim_{t\downarrow 0}\big[f(\gamma_v(t)) - f(u)\big]/t$ of $f : \mathbb{S} \to \mathbb{R}$ and the notion of a stationary point are as in Definition 3.1, and we use the same words for each slice loss ℓ_τ as well as for L.

Fix a slice τ, write its projectors as $\{\Pi_k^\tau\}_k$ and its cell probabilities as $p_k^\tau(u) = u^\dagger \Pi_k^\tau u$. Since Π_k^τ is Hermitian and idempotent, $p_k^\tau(u) = \|\Pi_k^\tau u\|^2$, and the resolution of the identity (9) gives the Parseval-type identity $\sum_k \|\Pi_k^\tau x\|^2 = \|x\|^2$ for every $x \in \mathbb{C}^d$. In particular $\sum_k p_k^\tau(u) = \|u\|^2 = 1$, while q^τ consists of relative frequencies so that $\sum_k q_k^\tau = 1$; subtracting gives the mass conservation (13). The *support* of u is

$$\mathrm{supp}(u) = \{k : \ p_k^\tau(u) > 0\},$$

and the first-order change of the cell probabilities induced by $v \in T_u$ is written

$$dp_k^\tau(v) = 2\,\mathrm{Re}\langle \Pi_k^\tau u,\ v\rangle.$$

Each dp_k^τ is a real linear functional on T_u; that it really is the first-order change is verified in Lemma A.2. The partition of the cells by the sign of the residual $\delta_k^\tau(u) = p_k^\tau(u) - q_k^\tau$ into $\mathcal{P} = \{k : \delta_k^\tau(u) > 0\}$, $\mathcal{Z} = \{k : \delta_k^\tau(u) = 0\}$ and $\mathcal{D} = \{k : \delta_k^\tau(u) < 0\}$, and the support condition $\Pi_k^\tau u = 0$ for all $k \in \mathcal{D}$, are as in Definition 3.4, always referring to the residuals of the slice in question. Since $p_k^\tau(u) = \|\Pi_k^\tau u\|^2$, the support condition is equivalent to $p_k^\tau(u) = 0$ for every cell of $\mathcal{D}$.

A.2 PROOF OF LEMMA 3.2

The loss L is not differentiable, so stationarity cannot be written with a gradient; the tool we use instead is Clarke's nonsmooth analysis (Clarke, 1990). For a locally Lipschitz F and a direction v, the *Clarke generalized directional derivative* and the (ordinary) *one-sided directional derivative* are

$$F^\circ(u; v) = \limsup_{w \to u,\ t\downarrow 0} \frac{F(w + tv) - F(w)}{t}, \qquad F'(u; v) = \lim_{t\downarrow 0} \frac{F(u + tv) - F(u)}{t}.$$

The former is an upper limit in which the base point w also moves; Lipschitz continuity makes it finite and convex in v, and it supports the *Clarke subdifferential* $\partial_C F(u) = \{\xi : \langle \xi, v\rangle \le F^\circ(u; v)\ \forall v\}$, which is nonempty, compact and convex with $F^\circ(u; v) = \max_{\xi \in \partial_C F(u)} \langle \xi, v\rangle$. The latter is the naive derivative at a fixed base point and need not exist. If $F'(u; v)$ exists in every direction at u and $F'(u; v) = F^\circ(u; v)$ for all v, then F is said to be *Clarke regular* at u. Regularity is what makes the two notions of stationarity agree: if F is regular then

$$0 \in \partial_C F(u) \iff F'(u; v) \ge 0 \ \ (\forall v),$$

so the stationarity of Definition 3.1, defined through one-sided directional derivatives, can be read directly as the standard stationarity $0 \in \partial_C F(u)$ of nonsmooth optimization. Without regularity the equivalence fails: for $F(x) = -|x|$ at $x = 0$ we have $F'(0; \pm 1) = -1 < 0$, so the point is not stationary in the directional sense, yet $F^\circ(0; v) = |v|$ gives $\partial_C F(0) = [-1, 1] \ni 0$, so it is Clarke stationary.

Lemma 3.2. *L is 2-Lipschitz on $\mathbb{S}_d$. Moreover L is Clarke regular.*

Proof. We first prove the Lipschitz property. Since L is the maximum of $\{\ell_\tau\}_\tau$ we have $|L(u) - L(w)| \le \max_\tau |\ell_\tau(u) - \ell_\tau(w)|$, so it suffices to show that each ℓ_τ is 2-Lipschitz; this fact is used again in Lemma A.1. Fix τ. The triangle inequality for norms gives $|\ell_\tau(u) - \ell_\tau(w)| \le \|p^\tau(u) - p^\tau(w)\|_1$, so we bound the right-hand side. As $p_k^\tau(u) = \|\Pi_k^\tau u\|^2$, factoring the difference of squares gives

$$\big|p_k^\tau(u) - p_k^\tau(w)\big| = \big|\ \|\Pi_k^\tau u\| - \|\Pi_k^\tau w\|\ \big| \cdot \big(\|\Pi_k^\tau u\| + \|\Pi_k^\tau w\|\big).$$

Summing over k and applying the Cauchy–Schwarz inequality to the two factors,

$$\|p^\tau(u) - p^\tau(w)\|_1 \le \Big(\sum_k \big|\, \|\Pi_k^\tau u\| - \|\Pi_k^\tau w\| \,\big|^2\Big)^{1/2} \Big(\sum_k \big(\|\Pi_k^\tau u\| + \|\Pi_k^\tau w\|\big)^2\Big)^{1/2}.$$

The first factor is at most $\|u - w\|$ by the reverse triangle inequality $\big|\|\Pi_k^\tau u\| - \|\Pi_k^\tau w\|\big| \le \|\Pi_k^\tau(u - w)\|$ together with the Parseval-type identity of §A.1, and the second is at most 2 because $(a+b)^2 \le 2(a^2+b^2)$ and the same identity give $\sum_k(\|\Pi_k^\tau u\| + \|\Pi_k^\tau w\|)^2 \le 2(\|u\|^2 + \|w\|^2) = 4$. Together, $\|p^\tau(u) - p^\tau(w)\|_1 \le 2\|u - w\|$. The number of cells is absorbed by the Parseval-type identity, so the constant does not depend on it, and a pointwise maximum inherits the largest of the individual Lipschitz constants, so the constant for L does not depend on the number of slices either. This estimate never uses what q^τ is, so it holds even when the empirical distribution is not generated by any state; and since the geodesic distance on $\mathbb{S}_d$ dominates the chordal distance $\|u - w\|$, L is 2-Lipschitz for the geodesic distance as well.

We next prove Clarke regularity. Rewrite the absolute values as a maximum over sign patterns: let $m = 2^N$ be the number of cells of a slice and, for a sign vector $s \in \{\pm 1\}^m$, put

$$f_{\tau,s}(u) \;=\; \sum_k s_k\big(p_k^\tau(u) - q_k^\tau\big).$$

Then $\sum_k |\delta_k^\tau(u)| = \max_s f_{\tau,s}(u)$, so

$$L(u) \;=\; \max_\tau \max_{s\in\{\pm1\}^m} f_{\tau,s}(u) \tag{17}$$

exhibits L as a maximum of finitely many ($|\Sigma|\, 2^m$) functions. Each $f_{\tau,s}$ is, upon identifying u with the real $2d$-dimensional vector $(\operatorname{Re} u, \operatorname{Im} u)$, a real quadratic form with constant coefficients, hence a C^∞ function with gradient $\nabla f_{\tau,s}(u) = 2\sum_k s_k \Pi_k^\tau u$. A C^∞ function is Clarke regular, so the proposition on pointwise maxima (Clarke, 1990, Prop. 2.3.12)—if all active components are regular then the maximum is regular and its subdifferential is the convex hull of the subdifferentials of the active components—applies to (17): L is Clarke regular at every point and, with the active set $A(u) = \{(\tau, s) : f_{\tau,s}(u) = L(u)\}$,

$$\partial_C L(u) \;=\; \operatorname{co}\big\{\nabla f_{\tau,s}(u) : (\tau, s) \in A(u)\big\}.$$

In particular the Clarke generalized directional derivative coincides with the one-sided directional derivative and $L'(u; v) = \max_{(\tau,s)\in A(u)} \langle \nabla f_{\tau,s}(u), v\rangle$. Neither ℓ_τ nor L is C^1—both bend at cells where a residual changes sign—but regularity coexists with this nonsmoothness: all of it comes from the maximum in (17), and the functions underneath the maximum are smooth. □

A.3 Proof of the main theorem

Step 1: localization

Lemma A.1 (Localization). *Let $u \in \mathbb{S}$ satisfy $L(u) > 0$ and let the maximizing slice $\tau_\star$ be unique, and put $\Delta = \ell_{\tau_\star}(u) - \max_{\tau\neq\tau_\star} \ell_\tau(u) > 0$. Then $\tau_\star$ is the unique maximizing slice at every $w \in \mathbb{S}$ with $\|w - u\| < \Delta/4$, and $L(w) = \ell_{\tau_\star}(w)$ there. Consequently, at u the one-sided directional derivatives of L and of $\ell_{\tau_\star}$ agree for every unit vector $v \in T_u$ (one exists if and only if the other does), and u is a stationary point of L if and only if it is one of $\ell_{\tau_\star}$, and a local minimum of L if and only if it is one of $\ell_{\tau_\star}$.*

Proof. As shown in the proof of Lemma 3.2, every ℓ_τ is 2-Lipschitz, so $\|w - u\| < \Delta/4$ implies $|\ell_\tau(w) - \ell_\tau(u)| < \Delta/2$ for every τ. Hence

$$\ell_{\tau_\star}(w) \;>\; \ell_{\tau_\star}(u) - \tfrac{\Delta}{2} \;=\; \max_{\tau\neq\tau_\star} \ell_\tau(u) + \tfrac{\Delta}{2} \;>\; \max_{\tau\neq\tau_\star} \ell_\tau(w),$$

so $\tau_\star$ is still the unique maximizing slice at w and $L(w) = \ell_{\tau_\star}(w)$. Since $\gamma_v(t) \to u$ as $t \downarrow 0$, for small $t > 0$ the point $\gamma_v(t)$ lies in this neighbourhood and the difference quotients defining the one-sided directional derivatives of L and $\ell_{\tau_\star}$ coincide. Stationarity is defined by the sign of those derivatives and therefore agrees, and local minimality compares function values in a neighbourhood of u, on which $L = \ell_{\tau_\star}$, so it agrees as well. □

STEP 2: MASS FLOWS REALIZABLE AT FIRST ORDER

Lemma A.2 (Directional derivative and realizable mass flows). *Fix $u \in \mathbb{S}$ and a slice τ. First, for a unit vector $v \in T_u$ and each cell,*

$$p_k^\tau(\gamma_v(t)) = p_k^\tau(u) + t\, dp_k^\tau(v) + O(t^2),$$

with a remainder uniform in k. Second, the one-sided directional derivative $\ell'_\tau(u; v)$ exists and

$$\ell'_\tau(u; v) = \sum_{k\in\mathcal{P}} dp_k^\tau(v) - \sum_{k\in\mathcal{D}} dp_k^\tau(v) + \sum_{k\in\mathcal{Z}} \big|dp_k^\tau(v)\big|, \tag{18}$$

and $L'(u; v)$ exists as well, with $L'(u; v) = \max_{\tau:\, \ell_\tau(u)=L(u)} \ell'_\tau(u; v)$ (this is the existence announced in Definition 3.1). Third, the set of mass flows realizable at first order is exactly the set of zero-sum vectors on the support:

$$\big\{\, (dp_k^\tau(v))_k \;:\; v \in T_u \,\big\} \;=\; \Big\{\, (\alpha_k)_k \;:\; \sum_k \alpha_k = 0,\;\; \alpha_k = 0 \;\; (\forall k \notin \mathrm{supp}(u)) \,\Big\}. \tag{19}$$

In particular the first-order mass conservation $\sum_k dp_k^\tau(v) = 0$ holds for every $v \in T_u$, cells with $p_k^\tau(u) = 0$ have $dp_k^\tau \equiv 0$, and conversely every zero-sum vector (α_k) on the support is realized as $dp_k^\tau(v) = \alpha_k$ for all k by

$$v = \sum_{k\in\mathrm{supp}(u)} \frac{\alpha_k}{2\, p_k^\tau(u)}\, \Pi_k^\tau u \;\in\; T_u.$$

Proof. First. Since Π_k^τ is Hermitian and idempotent and $\mathrm{Re}\langle u, v\rangle = 0$, along $\gamma_v(t) = \cos t\, u + \sin t\, v$,

$p_k^\tau(\gamma_v(t)) = \cos^2 t\, p_k^\tau(u) + 2 \sin t \cos t\, \mathrm{Re}\langle \Pi_k^\tau u, v\rangle + \sin^2 t\, \langle v, \Pi_k^\tau v\rangle = p_k^\tau(u) + t\, dp_k^\tau(v) + O(t^2)$.
The bounds $0 \le p_k^\tau \le 1$, $|dp_k^\tau(v)| \le 2$ and $0 \le \langle v, \Pi_k^\tau v\rangle \le 1$ do not depend on k, so the remainder is $O(t^2)$ uniformly in k.

Second. On cells with $\delta_k^\tau \neq 0$, the residual $\delta_k^\tau(\gamma_v(t)) = \delta_k^\tau + t\, dp_k^\tau(v) + O(t^2)$ keeps the sign of δ_k^τ for small $t > 0$, so the one-sided derivative of $|\delta_k^\tau|$ is $\mathrm{sgn}(\delta_k^\tau)\, dp_k^\tau(v)$; on cells with $\delta_k^\tau = 0$ we have $|\delta_k^\tau(\gamma_v(t))|/t = |dp_k^\tau(v) + O(t)| \to |dp_k^\tau(v)|$. As there are finitely many cells, sum and limit commute and (18) follows. For L: a slice that is not active ($\ell_\tau(u) < L(u)$) does not contribute to the maximum for small t by continuity, so $L(\gamma_v(t)) = \max_{\tau \text{ active}} \ell_\tau(\gamma_v(t))$, and since $\ell_\tau(\gamma_v(t)) = L(u) + t\, \ell'_\tau(u; v) + o(t)$ for each active τ, taking the maximum of finitely many such expressions gives $L(\gamma_v(t)) - L(u) = t \max_{\tau \text{ active}} \ell'_\tau(u; v) + o(t)$, so $L'(u; v)$ exists and equals the stated expression.

Third. For the inclusion $\subseteq$: the resolution of the identity (9) and the tangency condition give

$$\sum_k dp_k^\tau(v) = 2\,\mathrm{Re}\big\langle \big(\textstyle\sum_k \Pi_k^\tau\big)u,\; v\big\rangle = 2\,\mathrm{Re}\langle u, v\rangle = 0,$$

so the vector is zero-sum; and $p_k^\tau(u) = \|\Pi_k^\tau u\|^2$ means that $p_k^\tau(u) = 0$ implies $\Pi_k^\tau u = 0$, whence $dp_k^\tau(v) = 2\,\mathrm{Re}\langle \Pi_k^\tau u, v\rangle = 0$ for every v, i.e. the components outside the support vanish. For $\supseteq$: take any element (α_k) of the right-hand side and let v be as in the statement. By orthogonality of the projectors, $\{\Pi_k^\tau u\}_{k\in\mathrm{supp}(u)}$ are mutually orthogonal with $\|\Pi_k^\tau u\|^2 = p_k^\tau(u)$, so

$$\mathrm{Re}\langle u, v\rangle = \sum_{k\in\mathrm{supp}(u)} \frac{\alpha_k}{2p_k^\tau(u)}\, \mathrm{Re}\langle u, \Pi_k^\tau u\rangle = \sum_k \frac{\alpha_k}{2} = 0,$$

that is, $v \in T_u$. Furthermore $dp_j^\tau(v) = 2\,\mathrm{Re}\langle \Pi_j^\tau u, v\rangle = \alpha_j$ for $j \in \mathrm{supp}(u)$, while $\Pi_j^\tau u = 0$ gives $dp_j^\tau(v) = 0 = \alpha_j$ otherwise. Hence $(dp_k^\tau(v))_k = (\alpha_k)_k$, and both inclusions give the equality (19). □

The equality (19) supports the argument below in two directions. The inclusion $\subseteq$ provides the *restriction* that no mass can be poured into an empty cell at first order, which is the reason for stationarity at points satisfying the support condition (Step 4). The inclusion $\supseteq$ provides the *freedom* that any zero-sum flow on the support is realizable, which is used to construct the descent direction moving mass from a surplus cell to a deficient one when the support condition fails (Step 3). Note also that dp_k^τ is linear in v and the right-hand side of (18) is positively homogeneous in v, so even if the realizing v is not a unit vector, the sign of the one-sided directional derivative along $v/\|v\|$ is unchanged.

STEP 3: WHEN THE SUPPORT CONDITION FAILS

Lemma A.3 (A point violating the support condition is not stationary)**.** *Fix a slice τ and let $u \in \mathbb{S}$ satisfy $\ell_\tau(u) > 0$. If the support condition fails, i.e. if some deficient cell $k_- \in \mathcal{D}$ has $p^\tau_{k_-}(u) > 0$, then there is a unit vector $v \in T_u$ with $\ell'_\tau(u; v) < 0$. In particular u is neither a stationary point nor a local minimum of ℓ_τ.*

Proof. Mass conservation (13) together with $\ell_\tau(u) = \sum_k |\delta^\tau_k| > 0$ implies that $\mathcal{P}$ is nonempty, and on each cell of $\mathcal{P}$ we have $p^\tau_k > q^\tau_k \geq 0$, hence $p^\tau_k > 0$. Pick a surplus cell $k_+ \in \mathcal{P}$; both $k_\pm$ lie in $\mathrm{supp}(u)$, so by the construction in the third part of Lemma A.2 there is $w \in T_u$ realizing the zero-sum flow $\alpha_{k_+} = -\varepsilon$, $\alpha_{k_-} = +\varepsilon$ (with $\varepsilon > 0$) and 0 elsewhere. For this w we have $dp^\tau_{k_+}(w) = -\varepsilon$, $dp^\tau_{k_-}(w) = +\varepsilon$ and $dp^\tau_j(w) = 0$ on every other cell—on $\mathcal{P} \setminus \{k_+\}$, on $\mathcal{D} \setminus \{k_-\}$ and on all of $\mathcal{Z}$. Taking the unit vector $v = w/\|w\|$ and using linearity of dp^τ_k, substitution into (18) leaves only two of the three sums:

$$\ell'_\tau(u; v) = dp^\tau_{k_+}(v) - dp^\tau_{k_-}(v) = \frac{(-\varepsilon) - (+\varepsilon)}{\|w\|} = -\frac{2\varepsilon}{\|w\|} < 0.$$

The amount removed from the surplus cell decreases $\delta^\tau_{k_+} > 0$, and the amount poured into the deficient cell decreases $|\delta^\tau_{k_-}|$—the descent comes from both sides at once. Hence u is not stationary, and since ℓ_τ decreases strictly along $\gamma_v(t)$ for small $t > 0$, it is not a local minimum either. □

STEP 4: WHEN THE SUPPORT CONDITION HOLDS

Lemma A.4 (Consequences of the support condition)**.** *Fix a slice τ and let $u \in \mathbb{S}$ satisfy $\ell_\tau(u) > 0$ and the support condition. Then the following hold. First, u is a stationary point of ℓ_τ. Second, mass conservation (13) gives*

$$\sum_{k \in \mathcal{P}} \delta^\tau_k = \sum_{k \in \mathcal{D}} q^\tau_k = \tfrac{1}{2}\, \ell_\tau(u) > 0,$$

every cell k_0 of $\mathcal{D}$ has $q^\tau_{k_0} > 0$ (and $p^\tau_{k_0} = 0$), and the surplus mass satisfies

$$P_\mathcal{P} := \sum_{k \in \mathcal{P}} p^\tau_k(u) \;\geq\; \tfrac{1}{2}\, \ell_\tau(u) \;>\; 0.$$

Third, for any deficient cell $k_0 \in \mathcal{D}$ and any unit vector $v \in \operatorname{ran} \Pi^\tau_{k_0}$, along the great circle $u_t = \gamma_v(t) = \cos t\, u + \sin t\, v \in \mathbb{S}$ and in the range

$$\sin^2 t \;\leq\; h_0 := \min\Big\{ \min_{k \in \mathcal{P}} \frac{\delta^\tau_k}{p^\tau_k},\; q^\tau_{k_0} \Big\} \;>\; 0,$$

the exact identity

$$\ell_\tau(u_t) = \ell_\tau(u) - 2 P_\mathcal{P}\, \sin^2 t \tag{20}$$

holds. In particular u is not a local minimum of ℓ_τ.

Proof. We prove the second claim first. The support condition gives $p^\tau_k = 0$ on $\mathcal{D}$, i.e. $\delta^\tau_k = p^\tau_k - q^\tau_k = -q^\tau_k$, and $\delta^\tau_k < 0$ then forces $q^\tau_k > 0$ for every $k \in \mathcal{D}$. Mass conservation (13) reads $\sum_\mathcal{P} \delta^\tau_k = -\sum_\mathcal{D} \delta^\tau_k = \sum_\mathcal{D} q^\tau_k$, so $\ell_\tau(u) = \sum_\mathcal{P} \delta^\tau_k + \sum_\mathcal{D} q^\tau_k = 2 \sum_\mathcal{D} q^\tau_k > 0$. On each cell of $\mathcal{P}$ we have $p^\tau_k > q^\tau_k \geq 0$, hence $p^\tau_k \geq \delta^\tau_k > 0$, and summing gives $P_\mathcal{P} \geq \sum_\mathcal{P} \delta^\tau_k = \ell_\tau(u)/2 > 0$; in particular $h_0 > 0$.

Next the first claim. By the support condition $p^\tau_k = 0$ on all of $\mathcal{D}$, and together with the cells of $\mathcal{Z} \cap \{p^\tau_k = 0\}$, the third part of Lemma A.2 gives $dp^\tau_k \equiv 0$ on those cells. The sum over $\mathcal{D}$ in (18) therefore vanishes and the sum over $\mathcal{Z}$ reduces to the sum over $\mathcal{Z} \cap \{p^\tau_k > 0\}$ (the cells dropped contribute 0; they are not ignored):

$$\ell'_\tau(u; v) = \sum_{k \in \mathcal{P}} dp^\tau_k(v) + \sum_{k \in \mathcal{Z},\, p^\tau_k > 0} |dp^\tau_k(v)|.$$

For the same reason the first-order mass conservation $\sum_k dp_k^\tau(v) = 0$ reads $\sum_{k\in\mathcal{P}} dp_k^\tau(v) = -\sum_{k\in\mathcal{Z},\, p_k^\tau>0} dp_k^\tau(v)$ on the surviving cells. Substituting,

$$\ell'_\tau(u;v) = \sum_{k\in\mathcal{Z},\, p_k^\tau>0} \big(|dp_k^\tau(v)| - dp_k^\tau(v)\big) \;\geq\; 0$$

for every unit vector $v \in T_u$, i.e. u is stationary. In words: by Lemma A.2 the mass flows realizable at first order are the zero-sum flows on the support of u, but under the support condition that support consists of $\mathcal{P}$ and $\mathcal{Z}\cap\{p_k^\tau > 0\}$ only, i.e. of cells with nonnegative residual. Lowering the loss at first order would require pouring mass into a deficient cell, yet the deficient cells are empty and inflow into an empty cell begins only at second order in t because $p_k^\tau = \|\Pi_k^\tau u\|^2$ is quadratic. Moving mass inside the support only trades a decrease at the source for an increase at the destination, with no first-order gain; that is the content of the inequality above.

Finally the third claim. Three basic facts. Since $k_0 \in \mathcal{D}$, the support condition gives $p_{k_0}^\tau(u) = 0$, and with $p_{k_0}^\tau(u) = \|\Pi_{k_0}^\tau u\|^2$ this means $\Pi_{k_0}^\tau u = 0$. As $v \in \operatorname{ran}\Pi_{k_0}^\tau$ means $v = \Pi_{k_0}^\tau v$, self-adjointness gives $\langle u, v\rangle = \langle u, \Pi_{k_0}^\tau v\rangle = \langle \Pi_{k_0}^\tau u, v\rangle = 0$ (complex orthogonality; in particular $v \in T_u$), and orthogonality of the projectors (9) gives $\Pi_k^\tau v = \Pi_k^\tau \Pi_{k_0}^\tau v = 0$ for $k \neq k_0$. Hence, first, the cross term vanishes and

$$\|u_t\|^2 = \cos^2 t\, \|u\|^2 + 2\cos t\, \sin t\, \mathrm{Re}\langle u, v\rangle + \sin^2 t\, \|v\|^2 = \cos^2 t + \sin^2 t = 1,$$

so u_t lies on $\mathbb{S}$. Second, applying Π_k^τ linearly to u_t,

$$\Pi_k^\tau u_t = \cos t\, \Pi_k^\tau u + \sin t\, \Pi_k^\tau v = \begin{cases} \cos t\, \Pi_k^\tau u & (k \neq k_0;\ \Pi_k^\tau v = 0), \\ \sin t\, v & (k = k_0;\ \Pi_{k_0}^\tau u = 0,\ \Pi_{k_0}^\tau v = v), \end{cases}$$

so that, applying $p_k^\tau(w) = \|\Pi_k^\tau w\|^2$ to each line and abbreviating $h = \sin^2 t$, exactly

$$p_k^\tau(u_t) = \cos^2 t\, \|\Pi_k^\tau u\|^2 = (1-h)\, p_k^\tau(u) \;\; (k \neq k_0), \qquad p_{k_0}^\tau(u_t) = \sin^2 t\, \|v\|^2 = h.$$

Count the change of the absolute residuals in the range $h \leq h_0$. For $k \in \mathcal{P}$ the new residual $\delta_k^\tau - h\, p_k^\tau$ stays nonnegative (because $h \leq \delta_k^\tau / p_k^\tau$) and changes by $-h\, p_k^\tau$. For $k \in \mathcal{Z}$ with $p_k^\tau > 0$ (where $q_k^\tau = p_k^\tau$) the change is an increase $|-h\, p_k^\tau| = h\, p_k^\tau$. Cells of $\mathcal{Z}$ with $p_k^\tau = 0$ stay at 0, and cells of $\mathcal{D} \setminus \{k_0\}$ are unchanged since p remains 0 there. At k_0 the residual is $|h - q_{k_0}^\tau| = q_{k_0}^\tau - h$ (because $h \leq q_{k_0}^\tau$), a change of $-h$. Summing: the whole mass of p sits on $\mathcal{P}$ and $\mathcal{Z} \cap \{p_k^\tau > 0\}$ (the cells of $\mathcal{D}$ and of $\mathcal{Z} \cap \{p_k^\tau = 0\}$ carry no mass), so writing $P_{\mathcal{Z}_+}$ for the mass of the latter we have $P_\mathcal{P} + P_{\mathcal{Z}_+} = 1$ and

$$\ell_\tau(u_t) - \ell_\tau(u) = -h\, P_\mathcal{P} + h\, P_{\mathcal{Z}_+} - h = -2h\, P_\mathcal{P},$$

which is (20). Since $P_\mathcal{P} > 0$ by the second claim, $\ell_\tau(u_t) < \ell_\tau(u)$ for $0 < h \leq h_0$, and $u_t \to u$ as $t \downarrow 0$, so u is not a local minimum. □

PROOF OF THEOREM 3.5

Theorem 3.5 (Landscape theorem). *Let $L(u) > 0$ and let the maximizing slice be unique. Then u is a stationary point if and only if the support condition holds. Moreover, when the support condition holds, u is not a local minimum: for any deficient cell $k_0 \in \mathcal{D}$ and any unit vector $v \in \operatorname{ran}\Pi_{k_0}^{\tau_\star}$, along the great circle γ_v and for all sufficiently small $t > 0$,*

$$L(\gamma_v(t)) \;=\; L(u) \;-\; 2P_\mathcal{P}(u)\, \sin^2 t, \qquad P_\mathcal{P}(u) = \sum_{k\in\mathcal{P}} p_k^{\tau_\star}(u) \;>\; 0.$$

Proof. Write $\tau_\star$ for the maximizing slice and apply the four steps in order.

Step 1 (localization). By Lemma A.1, the one-sided directional derivatives, stationarity and local minimality of L at u agree with those of $\ell_{\tau_\star}$; below we argue about $\ell_{\tau_\star}$ only. Note $\ell_{\tau_\star}(u) = L(u) > 0$.

Step 2 (mass flows realizable at first order). By Lemma A.2, $\ell'_{\tau_\star}(u;v)$ exists for every unit vector $v \in T_u$ and is given by (18), and the flows realizable at first order form exactly the set (19) of zero-sum flows on the support.

Step 3 (the support condition fails). If it fails, Lemma A.3 provides a unit vector v with $\ell'_{\tau_\star}(u; v) < 0$. Hence u is not a stationary point of $\ell_{\tau_\star}$, and by Step 1 not one of L either.

Step 4 (the support condition holds). If it holds, the first claim of Lemma A.4 makes u a stationary point of $\ell_{\tau_\star}$, hence of L by Step 1. Moreover, by the third claim of Lemma A.4, for any $k_0 \in \mathcal{D}$ and any unit vector $v \in \operatorname{ran} \Pi^{\tau_\star}_{k_0}$ we have $\ell_{\tau_\star}(\gamma_v(t)) = \ell_{\tau_\star}(u) - 2P_{\mathcal{P}} \sin^2 t$ in the range $\sin^2 t \le h_0$. Since $\|\gamma_v(t) - u\| \to 0$ as $t \downarrow 0$, for sufficiently small $t > 0$ the point $\gamma_v(t)$ lies in the neighbourhood of Lemma A.1 (the ball of radius $\Delta/4$), so that $L(\gamma_v(t)) = \ell_{\tau_\star}(\gamma_v(t))$, and $\sin^2 t \le h_0$ holds as well; the claimed identity therefore holds for L. By the second claim of Lemma A.4, $P_{\mathcal{P}} = P_{\mathcal{P}}(u) \ge \ell(u)/2 > 0$, so the identity gives a strict descent of L and u is not a local minimum of L.

Combining: the support condition implies stationarity (Step 4) and its failure implies non-stationarity (Step 3), so stationarity is equivalent to the support condition; and when it holds, u is not a local minimum, as shown in Step 4. □

If, in addition, $p^{\tau_\star}_k(u) > 0$ on every cell of $\tau_\star$ (interior condition), then the nonempty $\mathcal{D}$ has mass on all of its cells, so the support condition cannot hold and u is not even a stationary point.

Proof of Corollary 3.6

Corollary 3.6 (Tie set and identification)**.** *Every spurious local minimum of L (a local minimum with positive value) lies in the* tie set*, the set of points at which several slices share the maximum. Moreover, under exact data, if the truth is uniquely determined among pure states by its weight-$\le N$ marginals, then the global minimum of L is the truth up to a global phase (Zhang et al., 2024).*

Proof. The first statement is immediate from Theorem 3.5: a point with $L(u) > 0$ at which the maximizing slice is unique is not a local minimum, whether the support condition holds or fails, so every local minimum with $L > 0$ is a point at which the maximizing slice is not unique, i.e. lies in the tie set.

For the second statement, $L(u) = 0$ is by the definition (12) equivalent to $p^\tau(u) = q^\tau$ for all τ, which under exact data $q^\tau = p^\tau(\psi)$ reads $u^\dagger \Pi^\tau_k u = \psi^\dagger \Pi^\tau_k \psi$ for all τ and k. By the substring expansion (7), each cell projector Π^τ_k of a setting $\tau = (S, b)$ is a linear combination of the Pauli substrings $P_{b,T}$ ($T \subseteq S$) belonging to that slice, and conversely each $P_{b,T}$ is a signed sum of the $\{\Pi^\tau_k\}_k$; hence agreement of the cell probabilities is equivalent to $\operatorname{tr}(P\, uu^\dagger) = \operatorname{tr}(P\, \psi\psi^\dagger)$ for every Pauli string P belonging to a slice. Under Assumption 2.1 the settings $\tau = (S, b)$ run over all blocks with $|S| = N$ and all labels $b \in \{x, y, z\}^S$, so every Pauli string of weight at most N belongs to some slice; therefore $L(u) = 0$ is equivalent to the agreement of all Pauli expectations of weight at most N, i.e. of all N-body reduced density matrices of the two states. If the truth ψ is determined uniquely, up to a global phase, among pure states by these expectations, then $L(u) = 0$ implies $u = e^{i\theta}\psi$; since $L \ge 0$ and $L(\psi) = 0$, the global minimizers of L are exactly $e^{i\theta}\psi$. That almost every pure state has this property when $n = 2N$ is Proposition 2.2. □