Ruler Diffusion with Incremental Projections and Shina: An Epure-Based Approach to Quantum Many-Body Reconstruction

Dmitrii Shultsev*[1]

[1]Independent researcher

November 22, 2025

Abstract

We introduce a hybrid classical–quantum reconstruction scheme, *Ruler Diffusion with Incremental Projections* (RD–IP), together with an adaptive step controller *Shina*, for learning consistent reduced density matrices of large quantum systems from noisy local data. Instead of reconstructing the full 2^N-dimensional wave function, RD–IP works directly in the space of overlapping k-body marginals and treats the global state as an "epure"—a smoothly stitched diagram of local reduced density matrices, inspired by structural mechanics. A small diffusion-style neural denoiser (*DiffuserHead*) cleans two-qubit (or two-orbital) marginals, while a recurrent *RulerNet* carriage enforces consistency between overlapping blocks along a chain or graph. The Shina module acts as a learned trust-region optimizer, predicting per-bond and per-edge step sizes that suppress catastrophic energy overshoot and help maintain approximate N-representability. On Ising and Heisenberg spin chains up to $N = 50$ sites, RD–IP + Shina reduces strong non-variational energy overshoot to within ~ 0.1–0.4 Hartree while preserving or improving local correlator RMSE. On a strongly correlated Fe_2S_2 molecular cluster in a CAS(10,10) active space, a graph-based variant (GraphRulerNet) achieves ~ 0.7 Hartree mean absolute energy error with ~ 0.96 local fidelity, in seconds on a consumer GPU. We argue that this represents a structural validation of epure-based reconstruction and outline the remaining gaps—notably explicit N-representability constraints—required to reach chemical accuracy on Level-3 systems such as FeMoco.

1 Introduction

Full quantum state tomography scales exponentially with the number of qubits and quickly becomes intractable beyond $\sim$ 14–20 qubits for general mixed states [1,3]. This has motivated a variety of scalable schemes that reconstruct only restricted classes of states or limited sets of observables: tensor-network methods such as DMRG and matrix product states (MPS) [11,13], classical shadow tomography and overlapping tomography [1,2], as well as neural-network-based quantum state tomography and neural quantum states (NQS) [4–7,14,16].

In noisy intermediate-scale quantum (NISQ) regimes, one is often less interested in the full state and more in local observables, energy densities and correlation functions. For Hamiltonians with k-local interactions, these are fully determined by k-body reduced density matrices (RDMs). However, directly fitting a set of local marginals faces the well-known quantum marginal or N-representability problem: not every collection of local RDMs is compatible with a global positive semidefinite density matrix.

*Email: d.shultsev@gmail.com

In this work we develop a practical framework, *Ruler Diffusion with Incremental Projections* (RD–IP), for reconstructing consistent local RDMs from noisy or incomplete measurement data. The central idea is to treat the set of marginals not as independent objects but as a one-dimensional or graph-structured *epure*: a diagram of overlapping blocks that must agree on their shared degrees of freedom. Consistency is enforced by a recurrent *carriage* (RulerNet) that sweeps along the structure, together with a learned step regulator *Shina* that prevents non-variational energy overshoot.

We validate RD–IP and Shina on three classes of problems: (i) transverse-field Ising chains, (ii) Heisenberg (XXZ) spin chains in critical and non-integrable regimes, and (iii) a strongly correlated Fe_2S_2 cluster in a CAS(10,10) active space. Across these tests, we observe that the combination of a diffusion-style denoiser, a recurrent carriage and an adaptive step controller can significantly reduce energy pathologies while preserving local correlator fidelity.

2 Background and Motivation

2.1 Local tomography and overlapping marginals

Quantum overlapping tomography [1–3] and its local variants show that in many physically relevant cases, global properties can be recovered from a carefully chosen set of overlapping local measurements. Rather than measuring all 4^N Pauli strings, one measures local observables on windows of size k and exploits overlaps to reconstruct a global description.

In neural-network-based quantum state tomography [4–7], one instead trains a neural model (CNN, MLP, or autoregressive network) to map measurement outcomes to either a wave function parameterization or directly to density matrices. Neural quantum states (NQS) [14, 16] and recurrent neural network (RNN) wave functions [15] have been particularly successful as variational ansätze for many-body ground states.

Our approach is closer in spirit to overlapping tomography, but with two key differences: (i) we work directly in the space of local RDMs (or Pauli vectors) instead of wave functions, and (ii) we *learn* the stitching map that enforces consistency between overlapping marginals, rather than solving a constrained optimization problem via semidefinite programming.

2.2 Energy overshoot and non-variational reconstruction

Variational methods such as VQE, DMRG and NQS obey a variational upper bound $E_{\mathrm{var}} \geq E_0$ by construction [11, 14]. Non-variational reconstruction schemes that directly predict density matrices or marginals do not enjoy this guarantee. If the predicted marginals are not globally N-representable, then the reconstructed energy can in principle fall below the true ground state energy ("energy overshoot") or lie arbitrarily far above it.

We observed precisely such behavior in early versions of RD–IP: when trained only on random and simple Ising data, the model could "hallucinate" unphysical patterns of correlations on a Heisenberg chain, leading to energies several Hartree below the exact ground state. In this work, we treat the magnitude and sign of the energy error as a primary diagnostic of physicality and introduce Shina as an adaptive controller that learns to avoid overshoot.

3 Epure Representation and RD–IP Architecture

3.1 Incremental projections (DIPR) and epures

Consider a one-dimensional chain of N qubits with local interactions. Instead of parameterizing the full state ρ or $|\psi\rangle$, we consider the ordered set of nearest-neighbor 2-RDMs

$$\mathcal{D} = \{\rho_{12}, \rho_{23}, \ldots, \rho_{N-1,N}\}, \tag{1}$$

together with a small number of "closing" blocks if needed for periodic boundaries or longer-range terms.

Following the *Dima's Incremental Projection Ruler* (DIPR) viewpoint, we treat this set as an *epure*: an engineering-style diagram of overlapping cross-sections. Each pair $\rho_{i,i+1}$ and $\rho_{i+1,i+2}$ shares a common site $i+1$ and must satisfy the consistency condition

$$\mathrm{Tr}_i(\rho_{i,i+1}) \approx \mathrm{Tr}_{i+2}(\rho_{i+1,i+2}). \tag{2}$$

The total energy of a 2-local Hamiltonian $H = \sum_i h_{i,i+1}$ is a linear functional of these 2-RDMs:

$$E = \mathrm{Tr}(H\rho) = \sum_i \mathrm{Tr}(h_{i,i+1}\rho_{i,i+1})\,. \tag{3}$$

The RD–IP objective is therefore twofold: (i) denoise each local block $\rho_{i,i+1}$, and (ii) adjust them collectively to minimize both local reconstruction error and global inconsistency (epure overlap) while keeping the implied energy within a physically plausible band around E_0.

3.2 DiffuserHead: local denoiser in Pauli space

We work in the Pauli basis for two-qubit RDMs. Any $\rho_{i,i+1}$ can be written as

$$\rho_{i,i+1} = \frac{1}{4} \sum_{\mu,\nu\in\{I,X,Y,Z\}} c_{\mu\nu}\, \sigma_i^\mu \otimes \sigma_{i+1}^\nu, \tag{4}$$

with a real coefficient vector $v \in \mathbb{R}^{16}$ (often stored in a reduced 15-dimensional form without the identity-identity component). Synthetic and physics-derived training data are generated by sampling clean RDMs and corrupting them with depolarizing, amplitude damping and Gaussian Pauli noise.

The *DiffuserHead* is a small MLP that maps noisy vectors v_{noisy} to cleaned vectors v_{clean}. It is trained to minimize a combination of MSE and ℓ_1 loss on the Pauli coefficients, and in multi-bank variants it is exposed to a mixture of random, Ising, Heisenberg and Level-1 spin-chain data, as well as to Fe_2S_2-derived 2-RDMs in later experiments. This is loosely inspired by classical and quantum denoising diffusion models [8, 9] but implemented as a single-step conditional denoiser rather than an explicit diffusion process.

3.3 RulerNet carriage and epure overlap

To enforce consistency along the chain, we introduce a recurrent neural network, *RulerNet*, which acts as a moving carriage along the epure. At each step i, the RNN receives the denoised local vector v_i and a hidden state h_{i-1} summarizing the left context. It outputs a corrected vector $\tilde{v}_i$ and an updated hidden state h_i.

We define an *epure overlap* penalty by comparing the single-site reductions of neighboring 2-RDMs:

$$\mathcal{L}_{\text{overlap}} = \sum_{i=2}^{N-1} \left\| \mathrm{Tr}_{i-1}(\tilde{\rho}_{i-1,i}) - \mathrm{Tr}_{i+1}(\tilde{\rho}_{i,i+1}) \right\|_1, \tag{5}$$

where $\tilde{\rho}$ is reconstructed from $\tilde{v}$. This penalty acts as the discrete analogue of enforcing continuity of bending moments in a structural epure.

4 Shina: Adaptive Step Control

4.1 Motivation and design

Early RD–IP experiments applied the RNN corrections with a fixed global step parameter α, updating

$$v_i^{\text{new}} = v_i^{\text{noisy}} + \alpha\, \Delta v_i, \tag{6}$$

where Δv_i is the correction proposed by RulerNet. This often led to energy overshoot on difficult tasks such as Heisenberg chains and critical TFIM instances.

To address this, we introduce *Shina*, a learned step-size controller that predicts a per-bond step $\alpha_i \in [0, 1]$ based on local features:

- noisy and denoised Pauli vectors $(v_{\text{noisy}}, v_{\text{raw}})$ and their difference,
- simple diagnostics of local purity and norm,
- meta-information about the Hamiltonian (couplings, anisotropy, field),
- epure overlap and (during training) local energy deltas.

The final update becomes

$$v_i^{\text{new}} = v_i^{\text{noisy}} + \alpha_i \left(v_i^{\text{raw}} - v_i^{\text{noisy}} \right), \tag{7}$$

optionally iterated for K micro-steps (predictor–corrector style) per site.

4.2 Line-search teacher and energy-aware loss

Shina is trained with a teacher-forcing scheme based on discrete line search. For each training sample and site, we evaluate a small grid of candidate step sizes $\{\alpha^{(k)}\}$ and compute the resulting local energy error and overlap inconsistency. The best candidate $\alpha^\star$ minimizes a combined objective

$$\mathcal{L}_{\text{teacher}} = \lambda_{\text{rec}} \mathcal{L}_{\text{rec}} + \lambda_E (E_\alpha - E_{\text{ref}})^2 + \lambda_O \mathcal{L}_{\text{overlap}}. \tag{8}$$

Shina is then trained to predict $\alpha^\star$ from the local features, effectively learning a data-driven trust-region step rule [17, 18].

In practice, Shina converges to regime-dependent typical step sizes: on simple Ising chains, mean $\alpha_i \sim 0.4$, while on critical or non-integrable instances it tends to smaller values ~ 0.2–0.3, significantly reducing catastrophic overshoot.

5 Numerical Experiments on Spin Chains

5.1 Transverse-field Ising chain (Level-0)

As a sanity check, we benchmark RD–IP + Shina on small transverse-field Ising chains,

$$H_{\text{TFIM}} = -J \sum_i Z_i Z_{i+1} - h \sum_i X_i, \tag{9}$$

for sizes $N \approx 8$–12 near the critical point $h/J \approx 1$. Ground states are computed by exact diagonalization. Training data consists of exact 2-RDMs corrupted by synthetic noise; test data is held out at the Hamiltonian level.

Across these tests, Shina consistently improves both energy and local correlators relative to noisy inputs: for $N = 12$, $h = 0.5$, we observe a reduction in energy error from $\Delta E \approx +4.7$ Ha (noisy) to $\approx +3.1$ Ha (RD–IP + Shina, $\alpha_{\text{global}} = 0.5$), while RMSE in $\langle Z_i Z_{i+1} \rangle$ and $\langle X_i \rangle$ is reduced by roughly a factor of two. Epure overlap drops from ~ 0.09 to ~ 0.04, indicating improved marginal consistency.

5.2 Heisenberg XXX/XXZ chain (Level-1)

We next consider the antiferromagnetic Heisenberg chain,

$$H_{\text{Heis}} = J \sum_i \left(X_i X_{i+1} + Y_i Y_{i+1} + \Delta Z_i Z_{i+1} \right), \tag{10}$$

with open boundaries and lengths up to $N = 32$. Reference ground states and 2-RDMs are obtained with high-accuracy DMRG calculations [11, 13].

Naive RD–IP (without Shina and without Heisenberg data in the training bank) exhibited severe energy overshoot, with $\Delta E \approx -3.7$ Ha and large RMSE in correlators for $N = 32$. After adding Heisenberg-derived 2-RDMs to the DiffuserHead training bank and performing joint training of RulerNet and Shina with energy-aware loss, we obtain:

- typical energy errors $\Delta E \approx -0.1$ to -0.4 Ha at moderate global α,
- improved RMSE in $\langle X_i X_{i+1} \rangle$, $\langle Y_i Y_{i+1} \rangle$, with a modest increase in $\langle Z_i Z_{i+1} \rangle$ RMSE relative to the minimal-overshoot configuration,
- a reduction in epure overlap compared to both noisy inputs and naive RD–IP, indicating better N-representability.

These results show that Shina can trade a small amount of local correlator error for a large reduction in global energy pathology.

5.3 Critical TFIM, $N = 50$ (Level-1 "boss")

As an out-of-distribution (OOD) test, we apply the same architecture to a 50-site TFIM instance with parameters chosen near criticality, representing a hard chaotic regime for variational ansätze. Here, DiffuserHead and RulerNet are not retrained specifically on this size, but Shina's line-search teacher has seen similar "Level-1" data.

While the model still exhibits negative energy error (overshoot) of order ~ -18 Ha in its best configuration, this represents a $\sim 40\%$ improvement relative to prior baselines with overshoot ~ -30 Ha. Nearest-neighbor $\langle Z_i Z_{i+1} \rangle$ correlators show improved RMSE compared to noisy data, while $\langle X_i \rangle$ RMSE is slightly worse, reflecting the model's spectral bias towards dominant interaction terms.

6 From Chains to Graphs: Fe_2S_2 as a Level-2.5 Stand

A central goal of this research program is to move beyond 1D spin chains to realistic strongly correlated molecules such as FeMoco. As an intermediate "Level-2.5" benchmark, we construct a graph-based variant of RD–IP and test it on an Fe_2S_2 cluster in a CAS(10,10) active space.

6.1 GraphRulerNet and dm_2 blocks

Instead of Pauli vectors on qubit pairs, the Fe_2S_2 stand works directly with two-electron RDM blocks $\mathrm{dm}_2^{(uv)}$ associated with pairs of active spatial orbitals (u, v). The molecular Hamiltonian is encoded via one- and two-electron integrals h_{uv} and $V_{uv,kl}$, and the total energy is recovered as a linear functional of dm_1 and dm_2.

Each edge in the orbital interaction graph carries: (i) a 4×4 dm_2 block, (ii) a 20-dimensional feature vector including atomic numbers, electronegativities, local occupations, $|h_{uv}|$, aggregate $|V_{uv}|$, and norms of the current block. A message-passing GraphRulerNet replaces the 1D RNN, maintaining node states and predicting corrections to edge blocks that improve both reconstruction loss and energy consistency.

6.2 Fe_2S_2 results

On a dataset of 25 Fe_2S_2 graphs generated by PySCF FCI in CAS(10,10), the best GraphRulerNet checkpoint (fe2s2_ruler_patch32) achieves:

- validation energy mean absolute error ~ 0.61 Ha,

- dm_2 fidelity ~ 0.96,
- training and evaluation times on the order of tens of seconds on a single RTX 4060 Ti GPU.

Subsequent ShinaEvolve experiments, which adapt the internal Shina policy via evolutionary search, confirm a persistent energy floor around ~ 0.7 Ha despite small improvements in local metrics. This suggests that the current architecture is structurally capable of learning the qualitative structure of the FCI solution but is limited energetically by the absence of explicit N-representability constraints and by the small dataset.

7 Discussion and Outlook

The RD–IP + Shina framework demonstrates that: (i) epure-based reconstruction of overlapping RDMs is a viable alternative to full state tomography for many-body systems, (ii) a learned trust-region controller (Shina) can dramatically reduce non-variational energy overshoot, and (iii) graph-based variants (GraphRulerNet) scale naturally to small strongly correlated molecular clusters.

At the same time, the Fe_2S_2 stand exposes a clear *variational gap*: high local fidelity does not guarantee accurate total energies in the absence of explicit N-representability enforcement. Future work must therefore integrate differentiable projections onto the set of valid 2-RDMs (e.g., via SDP-inspired layers) or include explicit P, Q, G-constraints in the loss, as well as expand the training banks for molecular systems.

Ultimately, we envision RD–IP + Shina as a fast "front-end" for more expensive quantum chemistry solvers: providing high-quality initial guesses, denoising experimental RDMs, and flagging unphysical data via epure overlap and energy diagnostics.

Acknowledgments

All ideas and experiments reported here were conceived and carried out by Dmitrii Shultsev. Large language models were used as tools for code prototyping and drafting: OpenAI GPT–5.1 Pro (ChatGPT) and Codex 5.1 assisted with Python and PyTorch boilerplate; Google Gemini 3 was used to help restructure early versions of the manuscript; and an in-house *ShinkaEnvolve* framework was employed to explore meta-parameter schedules for the Shina controller. Any errors or misinterpretations remain solely the responsibility of the author.

References

[1] J. Cotler and F. Wilczek, "Quantum Overlapping Tomography," *Phys. Rev. Lett.* **124**, 100401 (2020).

[2] Y. Yang *et al.*, "Experimental Demonstration of Quantum Overlapping Tomography," *Phys. Rev. Lett.* **130**, 050804 (2023).

[3] B. G. M. Araújo, M. M. Taddei, D. Cavalcanti and A. Acín, "Local Quantum Overlapping Tomography," arXiv:2112.03924 (2023).

[4] D. Koutný, L. Motka, Z. Hradil, J. Řeháček and L. L. Sánchez-Soto, "Neural-network quantum state tomography," *Phys. Rev. A* **106**, 012409 (2022).

[5] Y. Zuo *et al.*, "Optical neural network quantum state tomography," *Adv. Photonics* **4**(2), 026004 (2022).

[6] H. Ma, D. Dong and I. R. Petersen, "On how neural networks enhance quantum state tomography with limited resources," *Quantum Inf. Process.* **23**, 45 (2024).

[7] T. Schmale, L. Jerbi, E. Lutz and J. Eisert, "Efficient quantum state tomography with convolutional neural networks," *npj Quantum Inf.* **8**, 101 (2022).

[8] M. Kölle *et al.*, "Quantum Denoising Diffusion Models," arXiv:2401.07049 (2024).

[9] A. Cacioppo *et al.*, "Quantum Diffusion Models," arXiv:2311.15444 (2023).

[10] J. Preskill, "Quantum computing in the NISQ era and beyond," *Quantum* **2**, 79 (2018).

[11] S. R. White, "Density matrix formulation for quantum renormalization groups," *Phys. Rev. Lett.* **69**, 2863–2866 (1992).

[12] S. R. White, "Density-matrix algorithms for quantum renormalization groups," *Phys. Rev. B* **48**, 10345–10356 (1993).

[13] R. Orús, "A practical introduction to tensor networks: Matrix product states and projected entangled pair states," *Ann. Phys.* **349**, 117–158 (2014).

[14] G. Carleo and M. Troyer, "Solving the quantum many-body problem with artificial neural networks," *Science* **355**, 602–606 (2017).

[15] M. Hibat-Allah, M. Ganahl, L. E. Hayward, R. G. Melko and J. Carrasquilla, "Recurrent neural network wave functions," *Phys. Rev. Research* **2**, 023358 (2020).

[16] D. R. Vivas, M. F. Rey and G. Carleo, "Neural-Network Quantum States: A Systematic Review," *Quantum Sci. Technol.* **7**, 043001 (2022).

[17] J. Nocedal and S. J. Wright, *Numerical Optimization*, Springer (2006).

[18] J. Stokes, J. Izaac, N. Killoran and G. Carleo, "Quantum Natural Gradient," *Quantum* **4**, 269 (2020).