# Supplementary Information for "Gradient-estimator design overcomes trainability barriers in neural-network-based variational optimization"

Yi-Ran Xue,[1,2] Rui Wang,[1,3,4,5,*] Baigeng Wang,[1,3,4,†] and Chenan Wei[1,2,6,7,‡]

[1] *National Laboratory of Solid State Microstructures and Department of Physics, Nanjing University, Nanjing 210093, China*

[2] *Department of Physics, University of Massachusetts, Amherst, Massachusetts 01003, USA*

[3] *Collaborative Innovation Center of Advanced Microstructures, Nanjing University, Nanjing 210093, China*

[4] *Jiangsu Physical Science Research Center, Nanjing 210093, China*

[5] *Hefei National Laboratory, Hefei 230088, People's Republic of China*

[6] *A. Alikhanyan National Science Laboratory, Br. Alikhanian 2, Yerevan 0036, Armenia*

[7] *Institute of Theoretical Physics, Faculty of Physics, University of Warsaw, Warsaw, Poland*

(Dated: October 21, 2026)

This Supplementary Information provides the technical derivations and supporting analysis for the main text, including the amplitude–phase parameterization, the direct-estimator construction, the unbiasedness and variance-minimization proofs, the descent-bound analysis, the coupled-network extension and additional numerical results.

## S1. NUMERICAL PROTOCOLS AND HYPERPARAMETERS

The main text retains the estimator definitions, benchmark Hamiltonians and comparison principles. The optimization settings needed to interpret or reproduce the numerical comparisons are collected here. The lattice rows use minSR [1], and the molecular calculations use Adam [2].

* Corresponding author: rwang89@nju.edu.cn

† Corresponding author: bgwang@nju.edu.cn

‡ Corresponding author: chenanwei@umass.edu

| Benchmark | Ansatz and optimizer | Batch or walkers | Updates |
|---|---|---|---|
| Flux ladder, baseline | Separate width-128, depth-2 tanh MLPs; minSR | 1,024 walkers; 48 exchange sweeps | 50k |
| Flux ladder, MLP controls | Widths 192 or 256 at depth 2; width 128 at depth 3; minSR | 1,024 walkers; 48 exchange sweeps | 50k |
| Flux ladder, ResNet | 24-channel stem, six two-layer residual blocks and spatial sum pooling; minSR | Direct: 256; standard: 512; 48 sweeps | 50k |
| Flux ladder, post-training VMC evaluation | MCMC sampling without network-parameter updates | 2,048 walkers; 64 sweeps | 5k |
| $8 \times 8$ flux square | Shared width-256, depth-2 tanh trunk with two heads; minSR | 1,024 walkers; 48 sweeps | 50k |
| Chiral chain | Shared width-512, depth-2 tanh network; minSR | 1,024 walkers; 64 sweeps | 50k |
| Haldane sphere | Two attention layers, four width-64 heads, one determinant; KFAC[3] | 1,120 walkers; 10 Metropolis moves | 100k ($N = 6$); 200k ($N = 8$) |
| $N_2$ bond stretch | Jointly trained width-256, depth-3 tanh trunk with amplitude-residual and phase heads; Adam | Batch 256 | 10k |
| $N_2$ batch sweep | Same jointly trained amplitude–phase network; Adam | Batches 64, 128, 256, 512 | 10k |
| $I_2$ with spin–orbit coupling | Fixed-amplitude width-256, depth-3 phase-channel MLP; Adam | Batches 1024, 2048, 4096, 8192 and 16384 | 20k |
| $I_2$, joint amplitude–phase | Width-256, depth-3 shared amplitude–phase MLP; Adam | Batch 16384 | 20k |
| $C_2$ width–batch grid | Depth-3 tanh phase MLP, widths 64, 128, 256; Adam | Batches 128, 256, 512 | 10k |
| $H_2O$ and LiH | Width-256 phase MLP; Adam | Batch 128 | 10k |
| Honeycomb zero-set control | $100 \times 100$ continuous-twist phase ansatz with two width-64, four-head Transformer blocks; Adam | Batch 6,400 | 1k |

TABLE S1. **Optimization and post-training evaluation protocols used for the reported numerical comparisons.** Within each row, all compared estimators share the listed model, sampling and optimizer settings unless an endpoint-specific batch is stated explicitly.

| Benchmark | Molecular specification | Determinant space and optimization |
|---|---|---|
| $N_2$ bond stretch | STO-3G[4]; ten spatial orbitals; fourteen electrons; no frozen core; $R = 1.12, 1.50, 1.70, 1.90, 2.10, 2.50$ Å | Fixed SCI-ranked 512-determinant support; amplitude and phase heads optimized jointly with the shared trunk |
| $N_2$ MLP batch sweep | Same electronic space; $R = 2.10$ Å | Same fixed support and jointly trained amplitude–phase model |
| $I_2$ | Explicit spin–orbit Hamiltonian; $R = 2.666$ Å | Fixed amplitudes; complex-valued neural phase optimization |
| $I_2$, joint amplitude–phase | Explicit spin–orbit Hamiltonian; $R = 2.666$ Å | Full 184756-determinant variational space; amplitude and phase optimized jointly |
| $C_2$ width–batch grid | $R = 1.24$ Å | 1,024-determinant subspace; width–batch phase-model grid |
| $H_2O$ | O–H distances 0.958, 1.50 and 2.00 Å at 104.5° | Complete 441-determinant fixed-spin sector |
| LiH | $R = 3.00$ Å | Complete 225-determinant fixed-spin sector |

TABLE S2. **Molecular systems and determinant spaces.** Phase parameterizations and optimization protocols are listed in Table S1.

For the flux ladder, the five standard controls multiply only the phase component after the common minSR solve. GPU hours are measured from wall-clock training time on NVIDIA A100 GPUs.

## S2. AMPLITUDE–PHASE PARAMETERIZATION

For a discrete basis configuration $x$, we write a complex neural quantum state as

$$\psi_\theta(x) = \exp[u_{\theta_1}(x) + \mathrm{i}\varphi_{\theta_2}(x)], \tag{S1}$$

where $u_{\theta_1}$ is the log-amplitude network and $\varphi_{\theta_2}$ is the phase network. The local energy is

$$E_{\mathrm{loc}}(x) = \sum_y H_{xy} \frac{\psi_\theta(y)}{\psi_\theta(x)}. \tag{S2}$$

The standard VMC gradient for a real parameter $\theta$ is

$$\partial_\theta E = 2\,\mathrm{Re}\left[\langle (E_{\mathrm{loc}} - E) O_\theta^*(x)\rangle_{|\psi|^2}\right], \tag{S3}$$

where $O_\theta(x) = \partial_\theta \log \psi_\theta(x)$. The symbol $O_\varphi$ in Fig. 1 denotes the real phase score $O_\varphi \equiv \partial_\beta \varphi$, rather than the complex logarithmic derivative $O_\theta$. For the separated networks this gives the population force vectors

$$\mathbf{g}_{\theta_1}^{\mathrm{amp}} = 2\,\langle [\mathrm{Re}\, E_{\mathrm{loc}}(x) - E]\, \nabla_{\theta_1} u_{\theta_1}(x)\rangle\,, \tag{S4}$$

$$\mathbf{g}_{\theta_2}^{\mathrm{std}} = 2\,\langle \mathrm{Im}\, E_{\mathrm{loc}}(x)\, \nabla_{\theta_2} \varphi_{\theta_2}(x)\rangle\,. \tag{S5}$$

Two clarifications frame what follows. First, the standard phase estimator is unbiased: its expectation equals the exact phase force, so the difficulty studied in this work is statistical, not a formal bias. Concretely, resolving a gradient component from $N_{\mathrm{samp}}$ samples requires $N_{\mathrm{samp}} \gtrsim \mathrm{Var}(\hat{g})/|g|^2$, and for the phase sector this ratio can grow without bound, because the force is a covariance that can become small through cancellation while the per-sample fluctuations of $\mathrm{Im}\, E_{\mathrm{loc}}$ remain large. Equations in the separated-network derivation use the uncentered per-configuration score variable $2\,\mathrm{Im}\, E_{\mathrm{loc}}\, \partial_{\theta_2}\varphi$. The coupled population derivation in Sec. S8 uses the population means $\mu_R = \mathbb{E}_{p_\theta}[E_R]$ and $\mu_I = \mathbb{E}_{p_\theta}[E_I] = 0$, whereas the finite-batch implementation replaces them by empirical batch means $\widehat{\mu}_{R,t}$ and $\widehat{\mu}_{I,t}$. These conventions target the same population force but have different finite-batch fluctuations. Second, this noise problem is pre-asymptotic. At an exact eigenstate the local energy is constant over configurations, so $\mathrm{Im}\, E_{\mathrm{loc}} \equiv 0$ pointwise and the standard phase estimator is then itself zero-variance, consistent with the zero-variance principle [5]; practical optimization operates far from that limit, and the crossover between the two regimes is made quantitative in the variance-mixing subsection below.

## S3. THE DIRECT LOCAL-ENERGY DERIVATIVE

The direct estimator differentiates the local energy at fixed sampled configuration:

$$\partial_\theta E_{\mathrm{loc}}(x) = \sum_{y \neq x} H_{xy} \frac{\psi_\theta(y)}{\psi_\theta(x)} \left[O_\theta(y) - O_\theta(x)\right], \tag{S6}$$

because the diagonal term $H_{xx}$ is independent of $\theta$. For phase parameters, $O_{\theta_2}(x) = \mathrm{i}\,\partial_{\theta_2}\varphi(x)$, so Eq. (S6) directly probes how changing the phase modifies the off-diagonal interference ratios. The real per-configuration direct phase-force vector used in the separated phase block is $\widehat{\mathbf{g}}_{\theta_2}^{\mathrm{dir}}(x) = \mathrm{Re}[\nabla_{\theta_2} E_{\mathrm{loc}}(x)]$.

The relation to the phase force follows from a Hermiticity identity that holds for any parameter and any basis[6]. Differentiating $E_{\mathrm{loc}} = \psi^{-1} H \psi$ at fixed configuration and taking the $|\psi|^2$ expectation,

$$\langle \partial_\theta E_{\mathrm{loc}}\rangle = \langle (E_{\mathrm{loc}}^* - E_{\mathrm{loc}})\, O_\theta\rangle = -2\,\langle \mathrm{Im}\, E_{\mathrm{loc}}\, \mathrm{i} O_\theta\rangle\,. \tag{S7}$$

For separated phase parameters, $O_{\theta_2} = \mathrm{i}\,\partial_{\theta_2}\varphi$, so the expectation is purely real,

$$\langle \partial_{\theta_2} E_{\mathrm{loc}}\rangle = 2\,\langle \mathrm{Im}\, E_{\mathrm{loc}}\, \partial_{\theta_2}\varphi\rangle\,, \tag{S8}$$

and the direct local-energy derivative targets exactly the same phase force as the standard phase gradient, with a different Monte Carlo estimator.

Equation (S7) also delimits the scope of the method. For a parameter that affects the amplitude as well, $\partial_\theta u \neq 0$, the exact gradient decomposes as

$$\partial_\theta E = 2\,\langle [\mathrm{Re}\, E_{\mathrm{loc}} - E]\, \partial_\theta u\rangle + \mathrm{Re}\,\langle \partial_\theta E_{\mathrm{loc}}\rangle\,, \tag{S9}$$

and the direct estimator alone omits the first (score) contribution: its bias relative to the full gradient is exactly $-g_\theta^{\mathrm{amp}}$. Thus a direct local-energy derivative by itself is not a valid full-gradient estimator for any parameter that also changes the amplitude. In the separated phase block this issue is absent because $\partial_{\theta_2} u = 0$. In a coupled real

two-head network the obstruction can instead be removed by restoring the amplitude/Pulay force and restricting the pathwise derivative to the phase output, with the amplitude values held fixed; this construction is derived in Sec. S8. A genuinely tied or holomorphic single-complex parameterization is not covered unless such a stopped-amplitude, active-phase derivative is a well-defined operation.

Eq. (S6) gives the direct estimator two structural properties.

**Diagonal decoupling.** The diagonal term drops out exactly: the $y = x$ ratio is identically one, so $\partial_\theta H_{xx}$-type contributions vanish sample by sample; in continuum notation the same statement reads $\partial_\theta V(\mathbf{R}) = 0$, so the direct estimator is sample-wise independent of diagonal (potential) fluctuations.

The diagonal-energy statement should be interpreted narrowly: the diagonal matrix element $H_{xx}$ drops out of the direct local-energy derivative in Eq. (S6). For a purely separated phase network, a real diagonal term also does not directly enter the standard phase-gradient expectation, which depends on $\mathrm{Im}\, E_{\mathrm{loc}}$. The variance advantage is therefore more accurately stated as an estimator-level reduction observed in phase-structured systems. By contrast, in shared or single-complex-network parameterizations the score-function estimator couples every parameter to $\mathrm{Re}\, E_{\mathrm{loc}}$ fluctuations through Eq. (S9), and diagonal (potential) noise does leak into phase directions; that observation was the original motivation for this work.

**Global gauge invariance.** The direct estimator depends on the phase network only through differences $\partial_{\theta_2}\varphi(y) - \partial_{\theta_2}\varphi(x)$ on connected pairs. Hence, global-phase directions vanish identically sample by sample. In an uncentered score-function implementation, the corresponding standard phase variable contains the pure-noise term $2\,\mathrm{Im}\, E_{\mathrm{loc}} \times$ const; a centered covariance implementation can remove this particular global mode. The broader variance comparison below therefore concerns the full standard/direct random variables, not only this gauge direction.

## S4. UNBIASED PHASE ESTIMATORS AND VARIANCE MIXING

For direct comparison with the main text, this section denotes the basis configuration by $R$ and the wavefunction by $\Psi$; these are the same objects denoted by $x$ and $\psi_\theta$ in Secs. S2 and S3. The estimator comparison can be made explicit for each phase parameter $\beta$ and each sampled configuration $R$. Write

$$\partial_\beta \log \Psi(R) = \mathrm{i}\, \partial_\beta \varphi(R), \tag{S10}$$

and define two real per-configuration quantities

$$A(R) = \mathrm{Im}\, E_{\mathrm{loc}}(R)\, \partial_\beta \varphi(R), \tag{S11}$$

$$B(R) = \sum_{R'} \mathrm{Im}\left[ H_{RR'} \frac{\Psi(R')}{\Psi(R)} \right] \partial_\beta \varphi(R'). \tag{S12}$$

The standard score estimator and the direct local-energy-derivative estimator can then be written as

$$\hat{g}_\beta^{\mathrm{std}}(R) = 2A(R), \qquad \hat{g}_\beta^{\mathrm{dir}}(R) = A(R) - B(R). \tag{S13}$$

Hermiticity gives $\langle B \rangle = -\langle A \rangle$, so both estimators have the same expectation,

$$\langle \hat{g}_\beta^{\mathrm{std}} \rangle = \langle \hat{g}_\beta^{\mathrm{dir}} \rangle = 2\langle A \rangle. \tag{S14}$$

Thus the direct estimator does not change the exact variational gradient; it changes the finite-sample random variable used to estimate the same phase force.

Let

$$a^2 = \langle A^2 \rangle, \qquad b^2 = \langle B^2 \rangle, \qquad c = \langle AB \rangle. \tag{S15}$$

Because the two estimators have the same mean, comparing their variances gives the simple criterion

$$\mathrm{Var}\big(\hat{g}_\beta^{\mathrm{dir}}\big) < \mathrm{Var}\big(\hat{g}_\beta^{\mathrm{std}}\big) \quad \Longleftrightarrow \quad b^2 - 2c < 3a^2. \tag{S16}$$

This condition is not a universal inequality; it is an instance-dependent statement about the phase-gradient noise, and the two regimes of training sit on opposite sides of it.

*Pre-asymptotic regime.* The phase force is a covariance, $2\langle A \rangle = 2\,\mathrm{Cov}(\mathrm{Im}\, E_{\mathrm{loc}}, \partial_\beta \varphi)$, since $\langle \mathrm{Im}\, E_{\mathrm{loc}} \rangle = 0$ for Hermitian $H$. The force can therefore be small through cancellation while the noise floor $a^2 = \langle A^2 \rangle$ remains at the

scale of $\langle(\mathrm{Im}\, E_{\mathrm{loc}})^2\rangle$: the signal-to-noise ratio (SNR) of the standard estimator can collapse long before the energy converges. This is the regime in which the direct estimator helps.

*Eigenstate limit.* At an exact eigenstate, $\mathrm{Im}\, E_{\mathrm{loc}} \equiv 0$ pointwise, so $\hat{g}_\beta^{\mathrm{std}}$ is a zero-variance estimator. The direct estimator $\hat{g}_\beta^{\mathrm{dir}} = -B$ retains a generally nonzero variance, because only the unweighted sum over connected configurations vanishes at the eigenstate, not its $\partial_\beta\varphi(R')$-weighted counterpart. In this componentwise limit $v_s \to 0$ and $c_{sd} \to 0$, so $\lambda_\beta^\star \to 0$: the population mixture returns to the standard endpoint exactly when the zero-variance principle begins to favor it. The matched-state chiral-chain trajectory in Supplementary Fig. S2 illustrates that the finite-batch standard–direct variance ordering can change during optimization; it is not presented as an exact-eigenstate limit.

The same decomposition suggests a variance-aware interpolation. Convex combinations of unbiased estimators chosen to minimize variance are well established in Monte Carlo and machine-learning gradient estimation [7–9]. Here we apply this construction to NQS-VMC: we derive a direct local-energy-derivative phase estimator and combine it adaptively with the standard phase-gradient estimator. This is useful because the relative variance ordering of the two estimators can change during training, so a fixed choice of estimator is not obviously optimal. Since both endpoints are unbiased, any convex combination

$$\hat{g}_{\beta,\lambda} = (1-\lambda)\hat{g}_\beta^{\mathrm{std}} + \lambda \hat{g}_\beta^{\mathrm{dir}}, \qquad 0 \le \lambda \le 1, \tag{S17}$$

targets the same population force for every deterministic $\lambda$. For one component $\beta$, denote

$$v_s = \mathrm{Var}(\hat{g}_\beta^{\mathrm{std}}), \quad v_d = \mathrm{Var}(\hat{g}_\beta^{\mathrm{dir}}), \quad c_{sd} = \mathrm{Cov}(\hat{g}_\beta^{\mathrm{std}}, \hat{g}_\beta^{\mathrm{dir}}). \tag{S18}$$

The unconstrained and clipped componentwise optima are

$$\lambda_{\mathrm{unc}}^{(\beta)} = \frac{v_s - c_{sd}}{v_s + v_d - 2c_{sd}}, \qquad \lambda_\beta^\star = \mathrm{clip}_{[0,1]}\left(\lambda_{\mathrm{unc}}^{(\beta)}\right). \tag{S19}$$

The denominator is itself a variance, $v_s + v_d - 2c_{sd} = \mathrm{Var}(\hat{g}_\beta^{\mathrm{std}} - \hat{g}_\beta^{\mathrm{dir}}) \ge 0$, so the mixture variance $V_\beta(\lambda) \equiv \mathrm{Var}(\hat{g}_{\beta,\lambda})$ is a convex quadratic on $[0,1]$; convexity here is a theorem, not an assumption. (If the denominator vanishes, the two estimators differ by a constant and every $\lambda$ is equivalent.) Because the endpoints $\lambda = 0, 1$ lie in the feasible interval, the population minimum over convex mixtures obeys

$$\mathrm{Var}\left(\hat{g}_{\beta,\lambda_\beta^\star}\right) \le \min\left[\mathrm{Var}(\hat{g}_\beta^{\mathrm{std}}), \mathrm{Var}(\hat{g}_\beta^{\mathrm{dir}})\right]. \tag{S20}$$

For an estimated coefficient, the same quadratic gives a sensitivity bound. Completing the square gives the exact identity

$$V_\beta(\lambda) = V_\beta\left(\lambda_{\mathrm{unc}}^{(\beta)}\right) + (v_s + v_d - 2c_{sd})\left(\lambda - \lambda_{\mathrm{unc}}^{(\beta)}\right)^2, \tag{S21}$$

so a deterministic or independently estimated coefficient whose value is $\widetilde{\lambda}$ has componentwise variance exceeding the unconstrained population minimum by $(v_s + v_d - 2c_{sd})\,(\widetilde{\lambda} - \lambda_{\mathrm{unc}}^{(\beta)})^2$ for that realization, and by $(v_s + v_d - 2c_{sd})\,\mathbb{E}[(\widetilde{\lambda} - \lambda_{\mathrm{unc}}^{(\beta)})^2]$ after averaging over the coefficient estimate. Thus Eq. (S20) is an exact raw-population statement, while a finite coefficient estimate can deviate from the population optimum by an amount controlled by Eq. (S21). For the full phase-force vector, define the trace-summed population statistics

$$V_s = \mathrm{tr}\,\mathrm{Cov}(\hat{\mathbf{g}}_{\mathrm{std}}^{\mathrm{ph}}), \quad V_d = \mathrm{tr}\,\mathrm{Cov}(\hat{\mathbf{g}}_{\mathrm{dir}}^{\mathrm{ph}}), \quad C = \mathrm{tr}\,\mathrm{Cov}(\hat{\mathbf{g}}_{\mathrm{std}}^{\mathrm{ph}}, \hat{\mathbf{g}}_{\mathrm{dir}}^{\mathrm{ph}}).$$

The production coefficient is the batch plug-in estimate

$$\widehat{\lambda}_t = \mathrm{clip}_{[0,1]}\left(\frac{\widehat{V}_s - \widehat{C}}{\widehat{V}_s + \widehat{V}_d - 2\widehat{C}}\right), \tag{S22}$$

where the hats on $V_s, V_d, C$ denote estimates from the current covariance batch, and the coefficient is clipped to $[0,1]$. In the separated-network experiments, both endpoints estimate the same phase force and the plug-in $\widehat{\lambda}_t$ chooses the interpolation used for that step; residual dependence between $\widehat{\lambda}_t$ and estimator noise is a finite-sample implementation effect, not a change in the population force being targeted. We therefore monitor the actual finite-batch behavior through matched-state covariance diagnostics. In the chiral-chain analysis in Supplementary Fig. S2, the three phase-estimator trace covariances are evaluated on common walker configurations at the same AMVP checkpoints. The

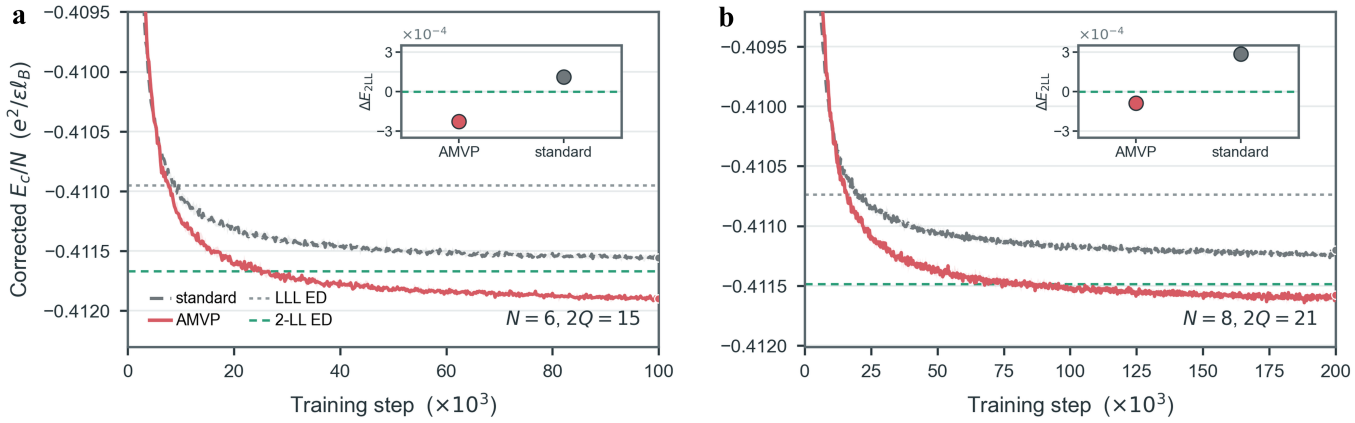



FIG. S1. **Training trajectories for continuum fractional quantum Hall states.** Corrected Coulomb energy per particle for the $\nu = 1/3$ Haldane-sphere systems with (a) $N = 6$ ($2Q = 15$, left) and (b) $N = 8$ ($2Q = 21$, right). Curves and bands show the median and interquartile range for the standard estimator (grey dashed) and AMVP (red). Insets show the converged energies relative to the corresponding two-Landau-level ED result (mean ± s.e.m., in units of $10^{-4}$). Horizontal dotted and green dashed lines denote the projected lowest-Landau-level ED and two-Landau-level ED reference energies, respectively.

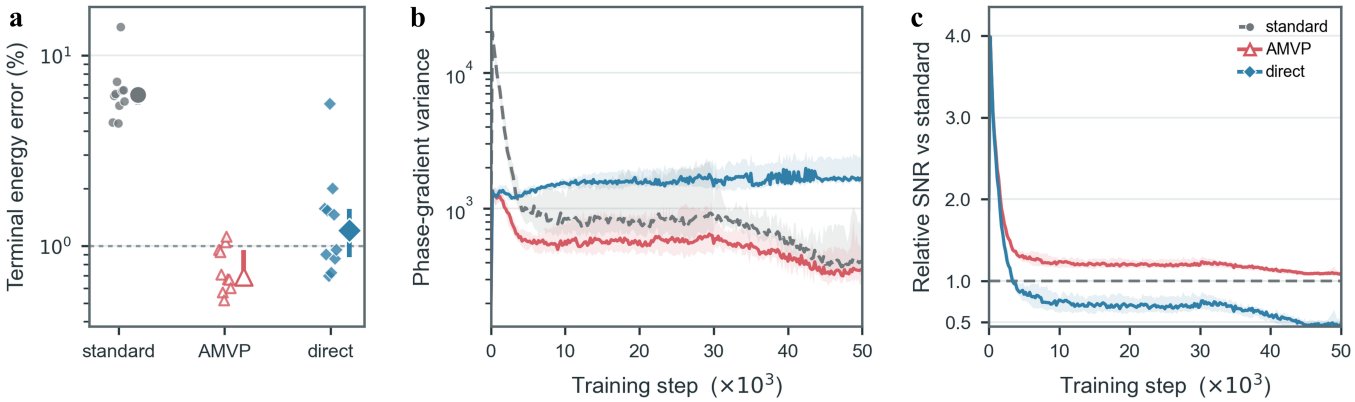



FIG. S2. **Energy accuracy and phase-estimator statistics for the chiral spin chain.** Results for the $L = 50$ spin-1/2 chiral XXX chain with $J = \alpha = 1$. (a) terminal relative energy error with respect to $E_{\mathrm{DMRG}} = -23.748041$; small markers show individual runs and large markers with bars show the median and interquartile range. (b) phase-row variance evaluated on matched configurations along the AMVP trajectories. (c) relative SNR. Curves and bands show the median and interquartile range.

accompanying relative SNR is defined only as $\sqrt{V_{\mathrm{std}}/V_m}$ under a shared-signal convention and is consequently an algebraic restatement of those variance estimates, not an independent or absolute SNR measurement. Supplementary Fig. S1 reports FQH energy trajectories only; the other lattice matched-state variance diagnostics are shown in main Fig. 4. With a single scalar coefficient, the population statement applies to the summed (trace) phase-gradient variance rather than to each component individually. The coefficient in Eq. (S22) is the one used for the separated amplitude–phase experiments. For a coupled two-head network the object whose variance is minimized is the full shared-parameter force, not the phase block alone; the corresponding coefficient contains amplitude–phase cross-covariances and is given in Sec. S8.

## S5. VARIANCE SETS THE ATTAINABLE DESCENT SPEED

The practical consequence of estimator variance can be understood through an idealized unpreconditioned phase-sector SGD step; we restate the standard argument [10] in the present setting. This subsection should be read as a regime statement under smoothness assumptions, not as a global convergence theorem for the nonconvex, preconditioned VMC dynamics used in production.

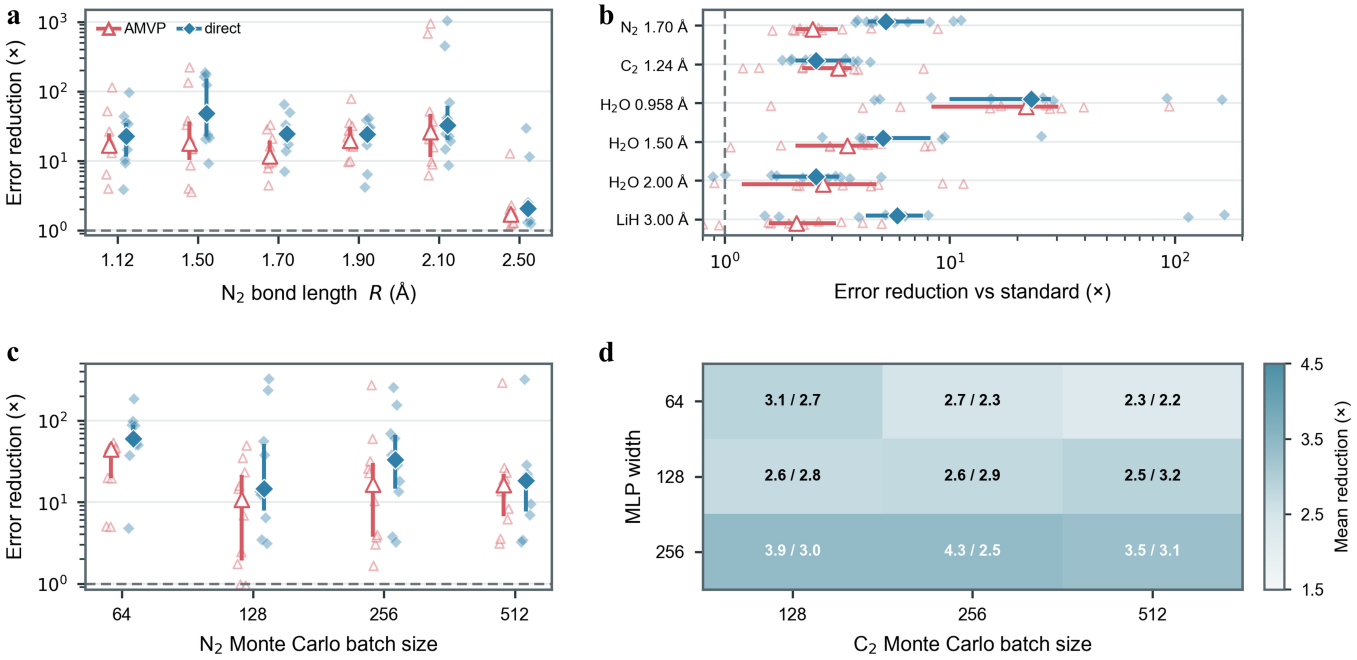



FIG. S3. **Robustness of molecular neural-state learning.** For each paired run $s$, the error-reduction factor is $r_{m,s} = \epsilon_{\mathrm{std},s}/\epsilon_{m,s}$ for $m \in \{\mathrm{AMVP}, \mathrm{dir}\}$, where $\epsilon_{m,s}$ is the FCI error; $r_{m,s} > 1$ indicates improvement over the standard estimator. In (a)–(c), small symbols denote individual runs, and large symbols with bars show the median and interquartile range. (a) The jointly optimized shared amplitude–phase $N_2$ MLP. (b) Fixed-amplitude phase-learning controls for $N_2$, $C_2$, three $H_2O$ geometries and LiH. (c) The same jointly optimized $N_2$ model at $R = 2.10$ Å across batch sizes. (d) The $C_2$ width–batch grid; cell text gives the direct/AMVP median factors and cell color their mean.

Assume the energy restricted to the phase sector (amplitude parameters held fixed at their current values) is $L_{\mathrm{sm}}$-smooth. For an unbiased vector estimator $\widehat{\mathbf{g}}$ of the sector gradient $\nabla E$, define $V = \mathbb{E}\|\widehat{\mathbf{g}} - \nabla E\|^2$, so that $\mathbb{E}\|\widehat{\mathbf{g}}\|^2 = \|\nabla E\|^2 + V$. One stochastic-gradient step of size $\eta$ obeys

$$\mathbb{E}[E_{t+1}] - E_t \le -\eta\|\nabla E\|^2 + \tfrac{1}{2}L_{\mathrm{sm}}\eta^2\left(\|\nabla E\|^2 + V\right). \tag{S23}$$

Optimizing this bound over $\eta$ gives $\eta^\star = \|\nabla E\|^2/[L_{\mathrm{sm}}(\|\nabla E\|^2 + V)]$ and the maximal expected one-step decrease

$$\Delta E_{\max} = -\frac{1}{2L_{\mathrm{sm}}}\,\frac{\|\nabla E\|^4}{\|\nabla E\|^2 + V}. \tag{S24}$$

The bound separates into two regimes. When $V \ll \|\nabla E\|^2$, the decrease is the deterministic $-\|\nabla E\|^2/(2L_{\mathrm{sm}})$. When $V \gg \|\nabla E\|^2$, it collapses to $-\|\nabla E\|^4/(2L_{\mathrm{sm}}V)$: a quartic stall in which progress is inversely proportional to the gradient variance. Combined with the pre-asymptotic signal-to-noise collapse identified above, this places noisy wavefunction-phase optimization in the stalled regime, while the amplitude sector — whose residual $(\mathrm{Re}\,E_{\mathrm{loc}} - E)^2$ is suppressed by the variational principle itself — can descend at the deterministic rate.

There are two corollaries connecting Eq. (S24) to the experiments. First, in this idealized unpreconditioned SGD model, if the optimizer is in the stalled regime $V \gg \|\nabla E\|^2$, a $\kappa$-fold reduction of the phase-gradient variance raises the optimized one-step descent bound by the same factor $\kappa$. The logged variance ratios in Fig. 4 should therefore be read as variance-controlled descent-speed headroom, rather than as an exact convergence theorem for the full minSR update. Second, rescaling the phase-network learning rate cannot remove this raw-gradient variance ceiling: Eq. (S24) is already optimized over a scalar step size, so amplifying $\eta_\varphi$ rescales signal and noise together, and steps beyond $\eta^\star$ violate the descent condition. This is consistent with the phase-network learning-rate baselines, which improve through $\eta_\varphi/\eta = 4$ and deteriorate at multiplier 5, while remaining above the direct estimator.

## S6. WHY THE ESTIMATOR CAN IMPROVE WAVEFUNCTION-PHASE OPTIMIZATION

The mechanism is quantitative at the estimator level: the $A/B$ decomposition gives the instance criterion of Eq. (S16), the diagonal and gauge structure of $\widehat{\mathbf{g}}^{\mathrm{ph}}_{\mathrm{dir}}$ removes identifiable noise channels exactly, and Eq. (S24) explains how measured variance ratios translate into descent-speed headroom in an idealized raw-gradient model. We therefore formulate the mechanism as a conditional proposition rather than an unconditional convergence theorem: when the phase network is expressive enough, sampling is adequate, and the optimization does not become trapped in unrelated variational basins, replacing the standard phase-gradient estimator by the direct/adaptive-mixture local-energy derivative can reduce phase-gradient variance and accelerate descent.

A complementary, less formal viewpoint is that finite-batch optimization does not follow the population vector field directly. Let $\mathcal{B}$ denote a sampled batch and write

$$\overline{\mathbf{g}}_{s,\mathcal{B}} = \nabla_{\theta_2} E + \boldsymbol{\xi}_{s,\mathcal{B}}, \qquad \overline{\mathbf{g}}_{d,\mathcal{B}} = \nabla_{\theta_2} E + \boldsymbol{\xi}_{d,\mathcal{B}}, \tag{S25}$$

for the standard and direct phase-force estimates on $\mathcal{B}$. In the infinite-sample limit, $\mathbb{E}_{\mathcal{B}}[\boldsymbol{\xi}_{s,\mathcal{B}}] = \mathbb{E}_{\mathcal{B}}[\boldsymbol{\xi}_{d,\mathcal{B}}] = 0$, so the two estimators target the same phase force. At finite batch, however, $\boldsymbol{\xi}_{s,\mathcal{B}}$ and $\boldsymbol{\xi}_{d,\mathcal{B}}$ are not the same random vectors. The standard estimator samples fluctuations of a score-function form, while the direct estimator samples fluctuations of link-local phase derivatives. The adaptive mixture therefore produces

$$\overline{\mathbf{g}}_{\lambda,\mathcal{B}} = \nabla_{\theta_2} E + (1 - \widehat{\lambda}_{\mathcal{B}})\boldsymbol{\xi}_{s,\mathcal{B}} + \widehat{\lambda}_{\mathcal{B}}\boldsymbol{\xi}_{d,\mathcal{B}}, \tag{S26}$$

which is a different finite-batch stochastic vector field, not merely the same field with a rescaled noise amplitude. If $\lambda$ were fixed and one looked at a single batch, a cancellation between the two endpoint noises could make the mixed force small even when neither endpoint is small. Such a cancellation is fragile: a stable small-force region for a mixture whose coefficient varies over an interval generally requires both endpoint finite-batch drifts to be small or aligned in a compatible way. This gives an intuitive reason why the adaptive mixture can avoid basins created by the standard estimator alone, while also avoiding commitment to a direct-only basin. The statement is intentionally qualitative; the true dynamics include Markov-chain correlations, same-batch coefficient estimates, stochastic reconfiguration and neural-network nonconvexity.

## S7. SEPARATE AMPLITUDE AND PHASE UPDATES

The separated-network experiments use independent amplitude and phase parameter blocks. In that setting, we keep the conventional amplitude update and replace only the phase estimator:

$$\Delta\theta_1 = -\eta_u\, \mathbf{g}^{\mathrm{amp}}_{\theta_1}, \qquad \Delta\theta_2 = -\eta_\varphi\, \mathbf{g}^{\mathrm{dir}}_{\theta_2}. \tag{S27}$$

This avoids using the direct estimator as a biased amplitude update. At the level of exact population forces for a strictly separated amplitude–phase ansatz, the update combines the conventional amplitude force with a phase force whose expectation is identical for the standard, direct, and adaptive-mixture estimators. Because the two forces act on disjoint parameter blocks, the exact composite force vanishes if and only if both sector forces vanish. Thus, before finite-sample preconditioning or same-batch estimates of $\widehat{\lambda}_t$ are introduced, the stationary points of the sector-separated raw-gradient update coincide with those of the exact variational energy gradient; the estimator change does not introduce new population-level stationary points.

The estimator identities in this section concern raw, population gradient estimators. The production runs use the same minSR-style preconditioner for every compared method: amplitude and phase parameters are concatenated into one tangent-space system, with amplitude scores entering through the real log-amplitude derivatives and phase scores entering through the imaginary phase derivatives. The sampled configurations, MCMC schedule, and linear-solve pipeline are identical for the standard, direct, and adaptive-mixture estimators; only the phase-force vector is changed. Because the preconditioner and the adaptive mixing coefficient are sample-dependent, the raw-gradient expectation equivalence is a population statement rather than a literal identity for every finite-sample preconditioned update. The theorem identifies the force targeted by the separated estimator, and the numerical comparisons test the complete preconditioned pipeline.

The stationarity statement above comes with two caveats so the table-ansatz control below is built around them. First, the coincidence is with the stationary points of the parameter-space energy $E(\theta)$, which for a neural ansatz is weaker than reaching an eigenstate: the sector forces are the Hilbert-space gradient projected onto the network's

tangent directions, so they can vanish at spurious critical points where the Hilbert-space gradient is nonzero but annihilated by the network Jacobian. No uniqueness or global-convergence claim is therefore made for arbitrary neural parameterizations. Second, the table ansatz over the full finite Hilbert space removes exactly this confounder: there every Hilbert-space direction is a parameter direction, the representation can express the exact eigenvector, and critical points on the full-support stratum are eigenstates, so table-ansatz experiments isolate estimator behavior from network expressivity. For neural networks the corresponding statement is empirical: under the tested protocols, the direct estimator often improves finite-budget energy and the observed run-level convergence. The ordering remains state- and protocol-dependent, so no universal improvement is claimed for every non-trivial complex phase structure.

## S8. GENERALIZATION TO COUPLED TWO-HEAD NETWORKS

The preceding estimator identities are used in this paper primarily for separated amplitude and phase networks. A closely related, but distinct, construction applies to a coupled real two-head ansatz

$$\psi_\theta(x) = \exp[u_\theta(x) + \mathrm{i}\varphi_\theta(x)], \qquad p_\theta(x) = \frac{e^{2u_\theta(x)}}{\sum_z e^{2u_\theta(z)}}, \tag{S28}$$

where the same real parameter vector $\theta$ can affect both $u_\theta$ and $\varphi_\theta$.

Fix a current parameter value and an infinitesimal direction $v$. Write

$$\dot{u}_v(x) = \nabla_\theta u_\theta(x)\cdot v, \qquad \dot{\varphi}_v(x) = \nabla_\theta \varphi_\theta(x)\cdot v, \tag{S29}$$

and denote $E_R(x) = \mathrm{Re}\, E_{\mathrm{loc}}(x)$ and $E_I(x) = \mathrm{Im}\, E_{\mathrm{loc}}(x)$ at the current $\theta$. Differentiating

$$E(\theta) = \sum_x p_\theta(x) E_R(x) \tag{S30}$$

along $v$ gives

$$DE(\theta)[v] = \sum_x \dot{p}(x) E_R(x) + \sum_x p_\theta(x)\dot{E}_R(x). \tag{S31}$$

The first term is the derivative of the sampling law. Since $p_\theta \propto e^{2u_\theta}$,

$$\dot{p}(x) = 2p_\theta(x)\left[\dot{u}_v(x) - \langle \dot{u}_v\rangle\right], \tag{S32}$$

and therefore

$$\sum_x \dot{p}(x) E_R(x) = 2\,\mathrm{Cov}_{p_\theta}(E_R, \dot{u}_v). \tag{S33}$$

This is the amplitude/Pulay contribution. Its Monte Carlo summand is the state-dependent coefficient

$$2\left[E_R(x) - \mu_R\right]\nabla_\theta u_\theta(x), \qquad \mu_R = \langle E_R\rangle, \tag{S34}$$

so the sampling derivative is included precisely when the amplitude force is included.

The explicit local-energy derivative in Eq. (S31) splits as

$$\dot{E}_R(x) = D_u E_R(x)[\dot{u}_v] + D_\varphi E_R(x)[\dot{\varphi}_v]. \tag{S35}$$

For Hermitian, $\theta$-independent $H$, the explicit $u$-path term has zero real expectation. To see this, define

$$\mathcal{K}_{xy} = p_\theta(x) H_{xy}\frac{\psi_\theta(y)}{\psi_\theta(x)} = \frac{\psi_\theta(x)^* H_{xy}\psi_\theta(y)}{\sum_z |\psi_\theta(z)|^2}. \tag{S36}$$

Hermiticity gives $\mathcal{K}_{yx} = \overline{\mathcal{K}_{xy}}$. Hence

$$\langle D_u E_R[\dot{u}_v]\rangle = \mathrm{Re}\sum_{x,y}\mathcal{K}_{xy}\left[\dot{u}_v(y) - \dot{u}_v(x)\right]. \tag{S37}$$

The contribution of the pair $(x, y)$ and $(y, x)$ is

$$\mathrm{Re}\left[\left(\mathcal{K}_{xy} - \overline{\mathcal{K}_{xy}}\right)[\dot{u}_v(y) - \dot{u}_v(x)]\right] = 0, \tag{S38}$$

because $\mathcal{K}_{xy} - \overline{\mathcal{K}_{xy}}$ is purely imaginary while $\dot{u}_v(y) - \dot{u}_v(x)$ is real. Thus

$$\langle D_u E_R[\dot{u}_v]\rangle = 0. \tag{S39}$$

The explicit phase term gives the usual phase force. Indeed,

$$D_\varphi E_{\mathrm{loc}}(x)[\dot{\varphi}_v] = \mathrm{i}\sum_y H_{xy}\frac{\psi_\theta(y)}{\psi_\theta(x)}[\dot{\varphi}_v(y) - \dot{\varphi}_v(x)]. \tag{S40}$$

Using the same Hermitian pairing as above,

$$\langle D_\varphi E_R[\dot{\varphi}_v]\rangle = 2\,\langle E_I\,\dot{\varphi}_v\rangle = 2\,\mathrm{Cov}_{p_\theta}(E_I, \dot{\varphi}_v), \tag{S41}$$

where the last equality uses $\langle E_I\rangle = 0$ for Hermitian $H$.

Combining Eqs. (S33)–(S41), for every direction $v$,

$$DE(\theta)[v] = 2\,\mathrm{Cov}_{p_\theta}(E_R, \nabla_\theta u_\theta\cdot v) + \langle D_\varphi E_R[\nabla_\theta\varphi_\theta\cdot v]\rangle\,. \tag{S42}$$

Equivalently, in vector form,

$$\nabla_\theta E = \mathbf{g}_{\mathrm{amp}} + \mathbf{g}_{\mathrm{ph}}, \qquad \mathbf{g}_{\mathrm{amp}} = 2\,\mathrm{Cov}_{p_\theta}(E_R, \nabla_\theta u_\theta), \qquad \mathbf{g}_{\mathrm{ph}} = 2\,\mathrm{Cov}_{p_\theta}(E_I, \nabla_\theta\varphi_\theta). \tag{S43}$$

The phase force $\mathbf{g}_{\mathrm{ph}}$ can equivalently be estimated by a stopped-amplitude pathwise derivative,

$$\mathbf{g}_{\mathrm{ph}} = \left\langle \nabla_\theta E_R(x)\big|_{\varphi\text{-path},\,u\text{ detached}}\right\rangle. \tag{S44}$$

Thus the coupled-network analogue of the direct estimator is not the bare pathwise derivative through all of $\theta$. The reconstructed per-configuration mixed force is

$$\widehat{\mathbf{g}}^{\mathrm{coupled}}_\lambda(x) = \widehat{\mathbf{g}}_{\mathrm{amp}}(x) + (1-\lambda)\widehat{\mathbf{g}}^{\mathrm{ph}}_{\mathrm{std}}(x) + \lambda\widehat{\mathbf{g}}^{\mathrm{ph}}_{\mathrm{dir}}(x), \tag{S45}$$

where $\mu_R = \mathbb{E}_{p_\theta}[E_R]$ and $\mu_I = \mathbb{E}_{p_\theta}[E_I] = 0$ are population means, and

$$\widehat{\mathbf{g}}_{\mathrm{amp}}(x) = 2[E_R(x) - \mu_R]\,\nabla_\theta u_\theta(x), \tag{S46}$$

$$\widehat{\mathbf{g}}^{\mathrm{ph}}_{\mathrm{std}}(x) = 2[E_I(x) - \mu_I]\,\nabla_\theta\varphi_\theta(x), \tag{S47}$$

$$\widehat{\mathbf{g}}^{\mathrm{ph}}_{\mathrm{dir}}(x) = \nabla_\theta\mathrm{Re}\left[\sum_y H_{xy}\exp(\mathrm{sg}\,[u_\theta(y) - u_\theta(x)] + \mathrm{i}[\varphi_\theta(y) - \varphi_\theta(x)])\right]. \tag{S48}$$

Here sg denotes a stop-gradient operation in the direct summand. For a finite batch $\mathcal{B} = \{x_n\}_{n=1}^B$, the implementation replaces $\mu_R$ and $\mu_I$ by the empirical means $\widehat{\mu}_{R,\mathcal{B}} = B^{-1}\sum_n E_R(x_n)$ and $\widehat{\mu}_{I,\mathcal{B}} = B^{-1}\sum_n E_I(x_n)$. Define $\overline{\mathbf{g}}_{\mathrm{amp}} = B^{-1}\sum_n \widehat{\mathbf{g}}_{\mathrm{amp}}(x_n)$, $\overline{\mathbf{g}}^{\mathrm{ph}}_{\mathrm{std}} = B^{-1}\sum_n \widehat{\mathbf{g}}^{\mathrm{ph}}_{\mathrm{std}}(x_n)$, and $\overline{\mathbf{g}}^{\mathrm{ph}}_{\mathrm{dir}} = B^{-1}\sum_n \widehat{\mathbf{g}}^{\mathrm{ph}}_{\mathrm{dir}}(x_n)$. The corresponding batch-mean force is

$$\overline{\mathbf{g}}^{\mathrm{coupled}}_\lambda = \overline{\mathbf{g}}_{\mathrm{amp}} + (1-\lambda)\overline{\mathbf{g}}^{\mathrm{ph}}_{\mathrm{std}} + \lambda\overline{\mathbf{g}}^{\mathrm{ph}}_{\mathrm{dir}}. \tag{S49}$$

With exact current-law sampling, fixed population baselines, and deterministic $\lambda$, Eqs. (S42)–(S49) give

$$\mathbb{E}\left[\overline{\mathbf{g}}^{\mathrm{coupled}}_\lambda\right] = \nabla_\theta E \qquad \text{for every fixed } 0 \le \lambda \le 1. \tag{S50}$$

This is the raw-population identity. In the finite adaptive implementation, the plug-in coefficient $\widehat{\lambda}_t$ is estimated from samples, the population means $\mu_R$ and $\mu_I$ are replaced by the empirical batch means $\widehat{\mu}_{R,t}$ and $\widehat{\mu}_{I,t}$, and the force is additionally affected by stale Markov chains and stochastic-reconfiguration preconditioning. These are implementation approximations around the same estimator-design principle; in coupled networks, where the amplitude and phase paths share parameters, they should be interpreted more cautiously than the separated-block identity.

Parameter sharing changes the variance optimization. For the separated phase block, the adaptive coefficient in Eq. (S22) minimizes the trace variance of the phase estimator alone. For a coupled network the full per-configuration random vector is

$$\widehat{\mathbf{g}}_{\mathrm{amp}}(x) + (1-\lambda)\widehat{\mathbf{g}}^{\mathrm{ph}}_{\mathrm{std}}(x) + \lambda\widehat{\mathbf{g}}^{\mathrm{ph}}_{\mathrm{dir}}(x). \tag{S51}$$

Let

$$\begin{aligned} V_s &= \operatorname{tr}\operatorname{Var}[\widehat{\mathbf{g}}^{\mathrm{ph}}_{\mathrm{std}}(x)], & V_d &= \operatorname{tr}\operatorname{Var}[\widehat{\mathbf{g}}^{\mathrm{ph}}_{\mathrm{dir}}(x)], & C &= \operatorname{tr}\operatorname{Cov}[\widehat{\mathbf{g}}^{\mathrm{ph}}_{\mathrm{std}}(x), \widehat{\mathbf{g}}^{\mathrm{ph}}_{\mathrm{dir}}(x)], \\ A_s &= \operatorname{tr}\operatorname{Cov}[\widehat{\mathbf{g}}_{\mathrm{amp}}(x), \widehat{\mathbf{g}}^{\mathrm{ph}}_{\mathrm{std}}(x)], & A_d &= \operatorname{tr}\operatorname{Cov}[\widehat{\mathbf{g}}_{\mathrm{amp}}(x), \widehat{\mathbf{g}}^{\mathrm{ph}}_{\mathrm{dir}}(x)]. \end{aligned} \tag{S52}$$

Then the trace-variance-minimizing unconstrained coefficient and its clipped population value are

$$\lambda^{\mathrm{coupled}}_{\mathrm{unc}} = \frac{(V_s - C) + (A_s - A_d)}{V_s + V_d - 2C}, \qquad \lambda^{\star}_{\mathrm{coupled}} = \operatorname{clip}_{[0,1]}\left(\lambda^{\mathrm{coupled}}_{\mathrm{unc}}\right). \tag{S53}$$

The coefficient in Eq. (S53) is written for the convention of Fig. 1, where $\lambda$ multiplies the direct estimator. If one instead writes $\widehat{\mathbf{g}}_{\mathrm{amp}}(x) + \lambda\widehat{\mathbf{g}}^{\mathrm{ph}}_{\mathrm{std}}(x) + (1-\lambda)\widehat{\mathbf{g}}^{\mathrm{ph}}_{\mathrm{dir}}(x)$, the corresponding standard-estimator weight is $(V_d - C + A_d - A_s)/(V_s + V_d - 2C)$. The separated formula is recovered when $A_s = A_d$, in particular when the amplitude and phase forces have disjoint parameter supports. The new quantities $A_s$ and $A_d$ are the diagnostic signature of coupling: a direct phase estimator can reduce the phase-channel variance while still failing to reduce the full-force variance if its covariance with the amplitude force is unfavorable.

This coupled construction applies to real two-head networks for which the stopped-$u$, active-$\varphi$ derivative in Eq. (S48) is well defined. It does not apply to a tied parameterization in which no such phase-path derivative can be formed. In particular, a genuinely holomorphic complex-weight network, where $u$ and $\varphi$ are Cauchy–Riemann tied as functions of the parameters, would require a separate analysis; mixing a biased full-pathwise derivative with the standard estimator would generally be biased for $\lambda > 0$.

## S9. IDEALIZED DESCENT PATH IN THE TABLE-ANSATZ LIMIT

It is useful to separate two statements that are often conflated. In the infinite-sample limit, $\mathbf{g}^{\mathrm{dir}}_{\theta_2}$ and the standard phase-force vector have the same expectation for separated phase parameters; therefore the direct estimator does not define a different variational energy surface. Its advantage is the estimator-level change in noise and conditioning analyzed above. A stronger convergence statement is available only in an idealized full-Hilbert-space limit, under assumptions that are standard in the nonconvex-optimization literature but that we do not prove for the sampled dynamics.

Consider a table ansatz on a finite Hilbert space, with independent log-amplitude and phase parameters $(u_x, \varphi_x)$ for every basis configuration, and let $p(x) = |\psi(x)|^2 / \sum_z |\psi(z)|^2$. The exact sector forces coincide with the coordinate partial derivatives,

$$\frac{\partial E}{\partial u_x} = 2\,p(x)\left[\operatorname{Re} E_{\mathrm{loc}}(x) - E\right], \qquad \frac{\partial E}{\partial \varphi_x} = 2\,p(x)\operatorname{Im} E_{\mathrm{loc}}(x), \tag{S54}$$

so the exact separated update

$$\dot\theta_1 = -\mathbf{g}^{\mathrm{amp}}_{\theta_1}, \qquad \dot\theta_2 = -\mathbf{g}^{\mathrm{dir}}_{\theta_2} \tag{S55}$$

is a genuine gradient flow of the Rayleigh quotient $E[\psi] = \langle\psi|H|\psi\rangle / \langle\psi|\psi\rangle$ in amplitude–phase coordinates, with

$$\frac{dE}{dt} = -\left\|\nabla_\theta E\right\|^2 \le 0. \tag{S56}$$

Relative to the Hilbert-space metric, the coordinate flow is preconditioned by a positive diagonal weight proportional to $p(x)$, which degenerates as $p(x) \to 0$.

On the full-support stratum, the stationary points of the Rayleigh quotient are eigenstates up to gauge, and every excited eigenstate is a strict saddle: if $|n\rangle$ has energy $E_n$ and $|m\rangle$ is any lower-energy eigenstate, then

$$|\psi(\epsilon)\rangle = \sqrt{1-\epsilon^2}\,|n\rangle + \epsilon|m\rangle \quad \Rightarrow \quad E[\psi(\epsilon)] = E_n + \epsilon^2(E_m - E_n), \tag{S57}$$

exactly, which decreases for $E_m < E_n$. Convergence of the idealized flow to the ground-state manifold then rests on three assumptions: (A1) exact expectations replace Monte Carlo estimates; (A2) the support stays bounded away from zero along the trajectory, so the $p(x)$ weight never degenerates; and (A3) generic initializations avoid the measure-zero strict-saddle manifolds, as established for gradient methods on smooth nonconvex landscapes [11], with convergence of the iterates for analytic costs following from Łojasiewicz-type arguments [12]; Rayleigh-quotient flows are a classical instance of this theory [13]. These are standard assumptions rather than statements we prove for the sampled dynamics.

We emphasize that (A2) is a genuine restriction, not a technicality: by Eq. (S54) the force on $u_x$ is proportional to $p(x)$, so configurations whose weight collapses stop receiving updates, and the flow possesses additional quasi-stationary states at the support boundary that are not eigenstates of the full Hamiltonian. This boundary failure mode — distinct from interior saddles — is the idealized counterpart of the bad-basin trajectories observed in finite-sample runs, and it is not removed by either phase estimator.

For neural networks this picture becomes an assumption rather than a theorem: the network image must contain a lift of a descending trajectory, and the parameter landscape must not introduce additional spurious local minima. The role of the numerical experiments is precisely to test how well this idealized estimator-level picture survives in finite-width neural networks and finite-sample VMC.

## S10. SUPPLEMENTARY FIGURES

The three supplementary figures provide continuum fractional quantum Hall training energies (Fig. S1), chiral-chain terminal energies and matched-state phase variances (Fig. S2), and paired 10,000-update molecular endpoint controls (Fig. S3). Each displayed condition uses the same ten-run scope as the corresponding main-text analysis.


[1] A. Chen and M. Heyl, Empowering deep neural quantum states through efficient optimization, Nature Physics **20**, 1476 (2024).

[2] D. P. Kingma and J. Ba, Adam: A method for stochastic optimization, in *International Conference on Learning Representations (ICLR)* (2015) arXiv:1412.6980.

[3] J. Martens and R. Grosse, Optimizing neural networks with kronecker-factored approximate curvature, in *Proceedings of the 32nd International Conference on Machine Learning*, Proceedings of Machine Learning Research, Vol. 37 (PMLR, 2015) pp. 2408–2417.

[4] W. J. Hehre, R. F. Stewart, and J. A. Pople, Self-consistent molecular-orbital methods. i. use of gaussian expansions of slater-type atomic orbitals, The Journal of Chemical Physics **51**, 2657 (1969).

[5] R. Assaraf and M. Caffarel, Zero-variance principle for Monte Carlo algorithms, Physical Review Letters **83**, 4682 (1999).

[6] A. Sinibaldi, C. Giuliani, G. Carleo, and F. Vicentini, Unbiasing time-dependent Variational Monte Carlo by projected quantum evolution, Quantum **7**, 1131 (2023).

[7] E. Greensmith, P. L. Bartlett, and J. Baxter, Variance reduction techniques for gradient estimates in reinforcement learning, Journal of Machine Learning Research **5**, 1471 (2004).

[8] S. Mohamed, M. Rosca, M. Figurnov, and A. Mnih, Monte Carlo gradient estimation in machine learning, Journal of Machine Learning Research **21**, 1 (2020).

[9] L. Metz, N. Maheswaranathan, J. Nixon, D. Freeman, and J. Sohl-Dickstein, Understanding and correcting pathologies in the training of learned optimizers, in *Proceedings of the 36th International Conference on Machine Learning*, Proceedings of Machine Learning Research, Vol. 97 (2019) pp. 4556–4565.

[10] L. Bottou, F. E. Curtis, and J. Nocedal, Optimization methods for large-scale machine learning, SIAM Review **60**, 223 (2018).

[11] J. D. Lee, M. Simchowitz, M. I. Jordan, and B. Recht, Gradient descent only converges to minimizers, in *29th Annual Conference on Learning Theory (COLT)*, Proceedings of Machine Learning Research, Vol. 49 (2016) pp. 1246–1257.

[12] P.-A. Absil, R. Mahony, and B. Andrews, Convergence of the iterates of descent methods for analytic cost functions, SIAM Journal on Optimization **16**, 531 (2005).

[13] U. Helmke and J. B. Moore, *Optimization and Dynamical Systems*, Communications and Control Engineering (Springer, London, 1994).