VQIL: VISIBLE QUANTUM INTERFERENCE LEARNING FOR PERTURBATION RESPONSE RANKING ON GRAPHS

Anonymous authors
Paper under double-blind review

ABSTRACT

Learning how local perturbations affect a network is important for identifying which components drive or absorb a system's response. We formulate this as seed-conditioned response ranking: given a perturbation at one seed node, ranking all other nodes by response strength. This is challenging when the model is given only the interaction topology, while the mechanism governing how perturbations propagate must be learned from data. We present VQIL (Visible Quantum Interference Learning), a trainable Szegedy quantum walk for seed-conditioned ranking. Edge weights, link phases, and per-node phase rotations are optimized end-to-end by backpropagation while the walk operator remains unitary throughout training. Node scores are obtained from a geometrically discounted coherent sum of quantum-walk amplitudes, enabling quantum interference before the final squared-norm readout. Because each learned parameter is attached to a named edge or node, the trained model provides directly inspectable hypotheses about the roles of individual graph components in the propagation mechanism. On AC power-flow perturbations over six transmission grids, with electrical parameters withheld from the model, VQIL improves Spearman rank correlation from 0.147 for the untrained walk to 0.831 and outperforms an identically trained random-walk model. On gene perturbation over five protein–protein interaction networks with Perturb-seq supervision, VQIL improves rank correlation from 0.171 to 0.620, outperforming the strongest learned classical and neural baselines. These results show that trainable quantum-walk interference provides a scalable and interpretable inductive bias for learning seed-conditioned response rankings on graphs.

1 INTRODUCTION

Graphs provide a natural abstraction for representing complex systems, where many foundational questions reduce to seed-conditioned ranking: if one node of the system is perturbed, how strongly do the other nodes respond, and through what mechanism? Given a graph whose edges encode known interactions and a perturbed seed node, the task is to rank all remaining nodes by their downstream response magnitude. Instances range from in silico gene perturbation—ranking transcriptomic shifts when a target gene is knocked down—to infrastructure resilience analysis, such as identifying which buses in a power grid re-equilibrate under localized load injection. The task is substantive because the observed topology typically records only structural connectivity, not transmission capacity or functional dynamics. Two things are consequently asked of a model. It must learn the underlying transmission mechanism from perturbation-response supervision, and keep this mechanism exposed for scientific interpretation.

Graph neural networks (GNNs) (Kipf & Welling, 2017; Veličković et al., 2018) are the dominant learning framework for graph-structured data and have ample capacity to learn complex perturbation mechanisms. However, they struggle to expose them: the seed typically enters through node attributes, and training optimizes shared transformation operators across the topology. Consequently, the learned mechanism remains distributed across continuous latent representations rather than expressed as explicit, interpretable interaction rules. For seed-conditioned ranking, however, a simpler family of methods remains highly relevant: *graph propagation*. Propagation methods spread unit mass from the seed along edges and score nodes by the mass they accumulate. Random walk with restart is the canonical example and underpins network-propagation methods across several domains

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.

(Page et al., 1999; Tong et al., 2006). Supervised extensions learn the walk itself from training queries, via per-edge conductances, feature-based edge strengths, or per-node restart probabilities (Agarwal et al., 2006; Backstrom & Leskovec, 2011; Jin et al., 2019), preserving a particularly useful correspondence between model and graph: the seed defines the initial condition, propagation occurs directly on the observed topology, and learned quantities modulate named edges or nodes. What is learned thus lives on the graph and can be read as a hypothesis about which components most strongly transmit the perturbation. This follows the philosophy of visible neural networks (Ma et al., 2018; Yu et al., 2018), where interpretability is achieved by construction rather than by post hoc explanation (Rudin, 2019). We push this correspondence to its limit by enforcing strict architectural transparency:

Definition 1 (Visibility) *A graph-learning model is* visible *if every learned parameter is a scalar attached to a named edge or node of the input graph, so that each parameter can be interpreted as a hypothesis about the role of that component in the underlying mechanism.*

The propagation dynamics themselves, however, remain a limitation. Accurately ranking perturbation responses often requires combining the contributions of many parallel regulatory paths, which a classical walk can only accumulate as non-negative probability mass. Quantum walks provide a natural extension of this propagation view. Unlike classical diffusion, where path contributions are combined monotonically, quantum walks propagate complex amplitudes that combine coherently before probabilities are formed, allowing constructive and destructive interference (Aharonov et al., 2001; Szegedy, 2004). This provides a richer hypothesis class for modeling how perturbations propagate through a graph. By making the graph-phased Szegedy quantum walk trainable (Ortega & Martin-Delgado, 2025), we turn quantum interference into a learnable propagation mechanism:

Definition 2 (Quantum interference learning) *Quantum interference learning is supervised learning over the hypothesis space of unitary walks on a given graph: gradient descent on the walk's own edge and node parameters selects the interference pattern whose readout best explains the observed responses.*

We introduce *visible quantum interference learning* (VQIL), a trainable Szegedy quantum walk for seed-conditioned response ranking. VQIL learns edge weights, edge phases, and node phases end-to-end from supervision while preserving unitarity throughout training. The resulting model combines interference-based propagation with parameter-level visibility: the walk is optimized to rank perturbation responses accurately, while its learned quantities remain attached to named graph components and can therefore be examined mechanistically. Our contributions are threefold:

- **A trainable Szegedy quantum walk.** We make the phased walk operator—per-edge weights, per-edge link phases, per-node phase rotations—fully differentiable and learnable. The standard walk provides the initialization, unitarity is preserved at every point of training, and the resulting implementation scales at $O(m)$ cost per step for m directed edges.
- **A coherent-resolvent readout.** Node scores are derived from a geometrically discounted sum of walk amplitudes, preserving rich quantum interference prior to the final squared-norm measurement. The discount replaces a hard temporal cutoff with a smooth, convergent horizon, yielding a resolvent of the walk operator and a differentiable objective across all walk parameters.
- **Visible mechanistic interpretation.** Every learned quantity is a scalar on a named edge or node of the input graph. The trained model therefore constitutes an explicit, graph-level hypothesis about the propagation mechanism. We show that the trained fields carry information about withheld properties of the underlying systems, turning the model into a hypothesis-generating tool for domain experts.

2 Problem statement

Let $G = (V, E)$ be an undirected, unweighted, connected simple graph with $n = |V|$ nodes, and let m denote the number of directed edges (where each undirected edge is represented in both directions). A *query* is a pair (s, y_s), where $s \in V$ is the seed node at which the system is perturbed,

and the label $y_s \in \mathbb{R}_{\geq 0}^{|V|-1}$ contains the absolute response magnitude measured at every other node. A model is a parameterized mapping $f_\Theta : (G, s) \mapsto x \in \mathbb{R}_{\geq 0}^{|V|}$, where Θ denotes the full parameter set. The model assigns a response score to every node conditioned on the perturbed seed. Predictive performance is evaluated by how well the rank ordering of the predicted scores x (excluding the seed) aligns with the ordering of the ground-truth measured responses y_s. Training minimizes a differentiable ranking loss over a set of training queries, while evaluation uses only queries whose seeds were held out during training.

The defining property of the task is that *the supervision carries functional dynamics that the observed graph does not natively encode*. The true response y_s is generated by an external physical or biological mechanism, whereas the input graph records only structural connectivity. Consequently, the graph defines the model's *hypothesis space*, not its answer sheet. Any predictive accuracy above a naive topological baseline must come from the learned mechanism. Because the model is strictly visible, whatever mechanism is learned is directly readable from Θ as a quantifiable statement about the underlying system.

3 Related work

Trainable propagation and quantum walks. Random-walk propagation is the standard tool for seed-conditioned prioritization: random walk with restart (Tong et al., 2006; Page et al., 1999), which underlies network propagation analyses in biology and other domains (Cowen et al., 2017; Han et al., 2021), diffuses scalar probability mass. Supervised extensions learn this diffusion from training queries (Agarwal et al., 2006; Backstrom & Leskovec, 2011; Jin et al., 2019). Quantum walks substitute this classical diffusion with unitary amplitude evolution and have been applied as fixed topological propagators for node centrality (Paparo & Martin-Delgado, 2012; Paparo et al., 2013) and disease-network analysis (Saarinen et al., 2024; Park et al., 2026). While the theoretical capacity for arbitrary per-node and link phases in Szegedy walks is described (Ortega & Martin-Delgado, 2023; 2025), alongside scalable simulators (Ortega & Martin-Delgado, 2024; Ortega & Park, 2026), existing trainable quantum walk neural networks learn coin operators that act on continuous node features (Dernbach et al., 2018; 2019). VQIL instead optimizes the walk's native per-edge and per-node parameters directly and extracts scores through a differentiable coherent resolvent.

Message-passing and geometric graph models. Standard message-passing GNNs (Kipf & Welling, 2017; Veličković et al., 2018) optimize shared filters acting on continuous node features. Consequently, perturbation mechanisms are embedded in feature-space transformations rather than in explicit graph operators. Our approach aligns more closely with geometric graph learning, specifically with sheaf neural networks (Bodnar et al., 2022), which learn per-edge restriction maps to modulate message passing; their $O(2)$ variant, which learns planar rotations, is the real-valued, feature-space analog of VQIL's complex-valued $U(1)$ link phases. To isolate the capacity of the propagation mechanism itselft, we evaluate against both families without external node attributes, so that all models rely only on the topology.

Structurally visible architectures. To keep mechanistic readouts faithful to the model, VQIL avoids post hoc explanation methods (Rudin, 2019) in favor of intrinsically interpretable architecture. Visible neural networks achieve this by structurally constraining neural connectivity to mirror known biological ontologies or hierarchies (Ma et al., 2018; Yu et al., 2018). VQIL instantiates this visibility constraint strictly at the graph level: by enforcing a bijective mapping between the learnable parameters and the physical network's nodes and edges, the optimized walk parameters act as direct hypotheses (e.g., transmission strengths or node-level response propensities) that can be audited independently.

4 Methods

4.1 Learned parameters and unitary propagation

VQIL is built on the graph-phased formulation of the Szegedy quantum walk (Ortega & Martin-Delgado, 2025). The walker's state lives on the space spanned by the m directed edges of the graph:

register 1 holds the node the walker occupies, register 2 the node it points at. One step of the walk is $U = SR$, where S swaps the two registers, and

$$R \;=\; \sum_i \left(1 - e^{i\theta_i}\right) |\psi_i\rangle\langle\psi_i| \;-\; I, \quad |\psi_i\rangle \;=\; |i\rangle_1 \otimes \sum_k e^{i\varphi_{ik}} \sqrt{p_{ki}}\, |k\rangle_2. \tag{1}$$

From this operator we define the following trainable parameters of the walk.

Edge weights. The weight p_{ji} is the probability of jumping from node i to node j. Each undirected edge $\{i, j\}$ carries a single, symmetric log-weight $\omega_{ij} = \omega_{ji} \in [-2, 2]$. The column-stochastic transition matrix is the softmax normalization over each node's incident edges,

$$p_{ji} \;=\; \frac{\exp(\omega_{ij})}{\sum_{k:\,\{i,k\}\in E} \exp(\omega_{ik})}, \tag{2}$$

so columns sum to one for every parameter value, and $\omega \equiv 0$ recovers the degree-normalized naive walk. Because each column is normalized over a different node's edges, the resulting transition probabilities remain directional ($p_{ji} \neq p_{ij}$) despite the underlying log-weights being symmetric. We enforce this symmetry because it serves as a structural regularizer that empirically yields better reproducibility than independently parameterized directed weights.

Link phases. Each edge carries a phase $\varphi_{ij} = \varphi_{ji} \in [-\pi, \pi]$ entering the walk states $|\psi_i\rangle$, with $\varphi \equiv 0$ recovering the standard walk states.

Node phases. Each node carries a phase rotation $\theta_i = \pi + \delta_i$, and we train the offsets $\delta_i \in [-\pi, \pi]$. At $\delta \equiv 0$ the reflection reduces to the standard reflection of the walk.

$U(\Theta)$ is unitary for *every* value of every parameter in $\Theta = (\omega, \varphi, \delta)$: training moves through the space of legitimate quantum walks and never leaves it. The whole construction runs in $O(m)$ per step in an edge-list representation (Ortega & Park, 2026) — a per-step cost that is optimal, since m numbers are the minimal description of the graph itself.

4.2 Coherent-resolvent readout

Starting from the seed's walk state $|\psi_s\rangle$ in (1), we accumulate the walk's evolution under a geometric discount and read out node scores by measuring the overlap with the position states $|\psi_v\rangle, v \in V$:

$$|\chi_s\rangle \;=\; \sum_{k=0}^{K} \gamma^k\, U^k |\psi_s\rangle, \qquad x_v \;=\; \left|\langle\psi_v|\chi_s\rangle\right|^2, \tag{3}$$

where K is fixed by γ so that the discarded tail is numerically negligible, $\gamma^K \leq 10^{-5}$, giving $K = 40,\ 109$ for the two discounts we study, $\gamma \in \{0.7,\ 0.9\}$. We use $\gamma = 0.9$ throughout the paper, and analog results for $\gamma = 0.7$ can be found in Appendix A.2.

Coherent. Amplitudes over all walk lengths are summed *first*, and the squared-norm is applied *once*, at the end. The alternative order, $x_v = \sum_k \gamma^k |a_k(v)|^2$ with $a_k(v) = \langle\psi_v|U^k\psi_s\rangle$, depends on the moduli alone. Summing first also keeps the cross-length products $a_j(v)\overline{a_k(v)}$, so walks of different lengths arriving at v interfere and the learned phases enter the score.

Resolvent. The geometric discount replaces a hard walk-length window, which weights all steps equally up to an abrupt cutoff, with a weighting that decay smoothly with walk length; empirically, this yields a much smoother training objective. It sets the effective horizon $1/(1-\gamma)$, the amplitude analog of a restart probability. We choose γ so that this horizon stays *below* the graph's mixing time so the walk never equilibrates within the readout and x_v retains the seed's identity rather than relaxing to a degree-dominated pattern.

Disclosure. Accumulating $|\chi_s\rangle$ across walk lengths is not norm-preserving, and the readout on the states $|\psi_v\rangle$ is not a physical measurement, so the scores are not probabilities—they are real, non-negative numbers that order the nodes, which is all a ranking requires. VQIL is a **quantum-inspired classical algorithm**: unitary quantum-walk propagation at every step, with a readout designed for classical execution, and no quantumimplementation claimed or required.

4.3 VQIL TRAINING

Variants. A VQIL configuration is a choice of which parameter sets are trainable. We study six variants. Both link phases and log-weights are always parameterized symmetrically across the two directions of an undirected edge.

1. **Naive-VQIL.** Nothing trained. The walk runs on the unweighted, phase-free graph; this is the untrained floor.
2. **Standard-VQIL.** Edge weights ω only.
3. **Link-phased-VQIL.** Edge weights ω and link phases φ.
4. **Vertex-phased-VQIL.** Edge weights ω and node phases δ.
5. **Graph-phased-VQIL.** Edge weights ω, link phases φ and node phases δ together.
6. **Phased-only-VQIL.** Phases φ and δ only, with weights ω frozen at the unweighted graph.

Algorithm 1 VQIL training step (graph-phased)

Input: edge list (m directed edges); query batch $\{(s^b, y_s^b)\}_{b=1}^{B}$; $\Theta = (\omega, \varphi, \delta)$
$p \leftarrow \mathrm{softmax}(\omega)$ over incident edges ▷ (2)
$\psi_e \leftarrow e^{i\varphi_e}\sqrt{p_e}$ for every directed edge e ▷ (1)
$|\chi^b\rangle \leftarrow |\xi^b\rangle \leftarrow |\psi_s^b\rangle$ ▷ $|\xi^b\rangle$: running walk state
for $k = 1 \dots K$ **do**
  $|\xi^b\rangle \leftarrow S\,R(\Theta)\,|\xi^b\rangle$ ▷ (1), $O(m)$
  $|\chi^b\rangle \leftarrow |\chi^b\rangle + \gamma^k |\xi^b\rangle$ ▷ coherent sum
$x_v^b \leftarrow |\langle\psi_v|\chi^b\rangle|^2$ for all v ▷ (3)
$\mathcal{L} \leftarrow \frac{1}{B}\sum_b \big(1 - \rho_\tau(\tilde{x}_s^b, y_s^b)\big)$ ▷ (4)
$\Theta \leftarrow \mathrm{AdamW}(\Theta, \nabla_\Theta \mathcal{L})$; clamp $\omega \in [-2, 2]$; $\varphi, \delta \in [-\pi, \pi]$

Loss. For a query (s, y_s), let $\tilde{x}_s$ denote the vector of log-transformed, standardized scores over the $|V| - 1$ non-seed nodes. We minimize a differentiable rank correlation, the soft-Spearman objective

$$\mathcal{L}(s) \;=\; 1 - \rho_\tau\big(\tilde{x}_s,\; y_s\big), \tag{4}$$

where ρ_τ is the Spearman correlation computed with soft ranks at temperature $\tau = 1$ (Blondel et al., 2020), so training never thresholds the graded labels.

Backpropagation. Gradients flow exactly, by automatic differentiation, through the entire construction: the softmax normalization (2), the complex phases in (1), the K applications of U, and the squared-norm readout (3). Algorithm 1 summarizes one training step.

Protocol. AdamW (learning rate 0.05), up to 300 epochs with early stopping on validation Spearman correlation (patience 10), model selection at the best validation checkpoint, float64/complex128 throughout. Weights initialize at 0 (the naive walk) and phases uniformly in $[-0.3, 0.3]$. Splits are seed-disjoint: a query's seed appears in exactly one of train/validation/test (Appendix Table 4).

4.4 BASELINES

Classical control. The matched classical control is **random walk with restart (RWR)** with learned edge weights: identical weight parameterization (2), identical loss, optimizer, splits and early stopping, with only the propagation differing (probability-level restart iteration in place of the unitary walk). Its restart probability is fixed at $r = 1 - \gamma$ so that both models weight walks of length k by γ^k. We report $\gamma = 0.9$ ($r = 0.1$) as the primary block and $\gamma = 0.7$ ($r = 0.3$) alongside. A second arm, **RWR-naive** (uniform weights, $r = 1 - \gamma$), marks the untrained propagation floor, exactly as Naive-VQIL does for the quantum walk.

Neural baselines. We compare against three message-passing architectures: GCN (Kipf & Welling, 2017), GAT (Veličković et al., 2018) and $O(2)$ neural sheaf diffusion (Bodnar et al., 2022). All three use width 64 (sheaf: 16 channels on 2-dimensional stalks), residual/diffusion updates, LayerNorm, LeakyReLU(0.1), and dropout 0.2. GAT uses single-head attention with self-loops, and the sheaf network predicts each edge's rotation angle with a per-layer MLP on the incident features. Our graphs lack node attributes, therefore we explicitly provide topological features capturing each node's relation to the query seed (Zhang & Chen, 2018; Li et al., 2020; Zhang et al., 2021): a seed indicator, normalized and log degrees, one-hot shortest-path distance to the seed (up to ≥ 13), a

16-step random-walk structural encoding (Dwivedi et al., 2022), and the first 16 non-trivial Laplacian eigenvectors. Node representations h_v are scored by a two-layer MLP on $[\, h_v,\ h_s,\ h_v \odot h_s \,]$, s the seed, followed by a softplus to keep the score positive for the ranking loss. Learning rate $\{0.005, 0.01, 0.05\}$ and depth $\{3, 6\}$ were tuned per architecture and domain (Appendix Table 5). The remaining training protocol exactly matches VQIL.

5 EXPERIMENTS

5.1 SET-UP

We evaluate seed-conditioned perturbation response ranking in two domains chosen so that the response is generated by a physical or biological mechanism withheld from the model, and to demonstrate scalability across orders of magnitude in graph size. For explicit details about datasets and trained models see Appendix A.1.

Domain 1: AC power-flow perturbation on transmission grids. A transmission grid joins *buses* (substations) by lines and transformers. At equilibrium each bus carries a voltage magnitude V and a phase angle θ. One query: solve AC power flow on the grid as shipped, inject a fixed load (+20 MW, +6 MVAr) at a seed bus, re-solve, and rank every other bus by its response. Each grid gives two instances with different physics: the angle response $|\Delta\theta|$ is real-power driven and path-mediated, whereas the voltage response $|\Delta V|$ is reactive-power-mediated and shaped by voltage regulation, since buses held at a set point by generators absorb the disturbance. The model sees only the unweighted bus graph and the seed index; reactances, transformer positions, generator identities, the slack bus, and all loads are withheld. We use six standard test networks with **30–1,354 nodes and 41–1,710 edges** (Christie, 1993; Alsac & Stott, 1974; Vittal et al., 1992; Josz et al., 2016; Fliscounakis et al., 2013; Zimmerman et al., 2011; Thurner et al., 2018), each partitioned seed-disjointly into five folds (60/20/20 train/validation/test, each query is tested exactly once).

Domain 2: In silico gene perturbation on protein–protein interaction (PPI) networks. A pooled CRISPR screen with single-cell readout perturbs one gene per cell and reports the whole transcriptome. One query: knock down the seed gene and rank every other gene by its response—$|\mathrm{logFC}|$ of mean log-normalized expression against the pooled control cells, used as a graded relevance. Five published screens (Replogle et al., 2022; Frangieh et al., 2021; Tian et al., 2021; Adamson et al., 2016), obtained in processed form from scPerturBench (Wei et al., 2026), supply the instances and span cell type, CRISPR modalities, and scales (75–1,615 perturbations). Each screen has its own graph: its gene panel (the 5,000 most variable genes plus every perturbed gene) mapped onto STRING v12.0 (Szklarczyk et al., 2023) at medium confidence ($\geq$ 400), largest component only (22–144 panel genes per dataset and five seeds in total fall outside it and are dropped rather than reconnected), giving **2,965–5,787 nodes and 29,059–218,844 edges**. The model sees only this unweighted graph and the seed index, seed-disjoint five-fold partitions as in Domain 1.

Metrics and protocol. Primary metrics: Spearman ρ between the predicted and measured response over all candidates, and AUPRC@10%, taking the strongest-responding 10% of candidates as positives. Secondary: AUROC@10% on the same positive set, and NDCG@10% with the measured response as graded relevance. Every arm selects its epoch on validation ρ and is trained under the identical protocol of §4.3. Values are mean $\pm$ sd over evaluation cells (one task instance $\times$ one seed-disjoint fold) and arms are compared paired within cells by a Wilcoxon signed-rank test.

5.2 RESULTS

Robust generalization across graph topologies and scales. On the six transmission grids (Table 1), weight training alone improves Spearman ρ from $0.312/0.147$ (angle/voltage) to $0.880/0.799$, with the best phase-bearing variants reaching $0.917/0.831$. On the PPI networks (Table 2), up to two orders of magnitude larger in scale, ρ lifts from 0.171 to 0.620. Among the trainable axes, edge weights drive the majority of the gain. Freezing the weights and training only the phases (Phased-only-VQIL) trails weights-only training by 0.128 (angle) and 0.211 (voltage) on the power grids and by 0.113 on the PPI networks. Per-node phases add a smaller but consistent gain on top of weights: $+0.031$ (angle) and $+0.032$ (voltage) in ρ on the power grids (27/30 and 28/30 cells, $p \leq 2.4\times10^{-6}$; AUPRC $+0.080$ and $+0.052$) and $+0.033$ on the PPI networks. Link phases contribute no measurable benefit: Link-phased-VQIL is indistinguishable from Standard-VQIL ($|\Delta\rho| \leq 0.0004$,

$p \geq 0.57$), and Graph-phased-VQIL mirrors Vertex-phased-VQIL ($|\Delta\rho| \leq 0.006$). We therefore treat edge weights and node phases as the two effective trainable axes and adopt Vertex-phased-VQIL as the main variant.

Compared with the matched classical control, the unitary walk demonstrates a clear post-training advantage. Prior to training, the classical walk ranks better on unweighted graphs across all three blocks, indicating the unitary walk holds no intrinsic zero-shot advantage. Post-training, this ordering decisively reverses: Vertex-phased-VQIL outperforms the classical control by 0.134 and 0.149 in ρ on the power grids, and by 0.050 on the PPI networks.

Finally, we benchmark against seed-conditioned neural baselines, which benefit from additional structural node features and a pair decoder. On the power grids, Vertex-phased-VQIL matches GCN and Sheaf-$\mathcal{O}(2)$ on the angle response, falls 0.020 below GAT, and trails all three on the voltage response. On the PPI networks the order reverses: Vertex-phased-VQIL outperforms GCN, GAT and Sheaf-$\mathcal{O}(2)$ by 0.038, 0.091 and 0.255 in ρ (25/25 cells, $p \leq 6\times10^{-8}$), leading GCN, the strongest neural baseline, on every dataset and on all four metrics. GCN in turn matches or exceeds the classical control. VQIL is also cheaper to train: $2.4\times$ to $4.6\times$ less total wall-clock time on the power grids and $3.8\times$ to $7.3\times$ less on each PPI network (Appendix Table 6). Its edge-indexed parameterization uses $4.9\times$ to over $400\times$ fewer parameters than the neural baselines on the grids but grows with graph size, reaching $0.8\times$ to $15\times$ their size on the PPI networks (Appendix Table 7).

Table 1: Power-grid perturbation-response ranking. Mean $\pm$ sd over 30 cells (six power grids $\times$ five seed-disjoint folds). VQIL variants use symmetric edge weights and link phases with discount $\gamma = 0.9$; the classical control uses the matched restart $r = 1 - \gamma = 0.1$. Results at $\gamma = 0.7$ are in Appendix Table 8. **Bold**: main VQIL variant.

	Phase-angle response $\lvert\Delta\theta\rvert$		Voltage response $\lvert\Delta V\rvert$	
Model	Spearman ρ	AUPRC@10%	Spearman ρ	AUPRC@10%
Naive-VQIL (untrained)	0.312 ± 0.269	0.541 ± 0.210	0.147 ± 0.154	0.310 ± 0.074
Standard-VQIL	0.880 ± 0.049	0.750 ± 0.107	0.799 ± 0.066	0.573 ± 0.132
Link-phased-VQIL	0.881 ± 0.048	0.749 ± 0.108	0.799 ± 0.066	0.572 ± 0.133
Vertex-phased-VQIL	$\mathbf{0.911 \pm 0.043}$	$\mathbf{0.830 \pm 0.062}$	$\mathbf{0.831 \pm 0.056}$	$\mathbf{0.624 \pm 0.144}$
Graph-phased-VQIL	0.917 ± 0.042	0.835 ± 0.065	0.829 ± 0.056	0.623 ± 0.144
Phased-only-VQIL	0.752 ± 0.144	0.757 ± 0.101	0.588 ± 0.096	0.545 ± 0.156
Naive-RWR (untrained)	0.411 ± 0.170	0.662 ± 0.109	0.228 ± 0.110	0.407 ± 0.114
Trained-RWR	0.777 ± 0.091	0.744 ± 0.083	0.681 ± 0.090	0.544 ± 0.140
GCN	0.915 ± 0.071	0.813 ± 0.079	0.920 ± 0.041	0.789 ± 0.102
GAT	0.932 ± 0.058	0.830 ± 0.076	0.920 ± 0.041	0.795 ± 0.107
Sheaf-$\mathcal{O}(2)$	0.913 ± 0.081	0.796 ± 0.095	0.908 ± 0.039	0.735 ± 0.116

The coherent resolvent readout enables smooth training. To isolate what the resolvent contributes, we factor the readout into two independent choices: how walk lengths are weighted (geometric γ^k versus a uniform window) and when the squared-norm is applied (amplitudes summed then squared, versus each step squared then summed). We train all four combinations at two horizons ($\gamma = 0.7, 0.9$) from identical per-step amplitudes (Table 3;Appendix Table 10). At $\gamma = 0.9$ the coherent geometric readout used by VQIL is the best of the four in every task: it leads its square-then-add counterpart (order) by $+0.103$ in ρ and the uniform window (weighting) by $+0.105$ on the PPI networks, and by $+0.072$/$+0.098$ (order) and $+0.208$/$+0.057$ (weighting) on the power grids (angle/voltage). Of the two geometric readouts, it is also the only readout that benefits from the longer horizon: moving from $\gamma = 0.7$ to 0.9 lifts it by $+0.063$ and $+0.127\,\rho$ on the grids and $+0.033\,\rho$ on the PPI networks, whereas square-then-add changes by -0.021 and $+0.018$ on the grids and -0.019 on the PPI networks. The benefit of coherent summation grows from -0.012 to $+0.072$ in ρ on the grid angle response and from $+0.051$ to $+0.103$ on the PPI networks.

The two components of the readout contribute to trainability in distinct ways, which we measure directly on the loss surface. Evaluating all four readouts at common parameter points (the untrained walk and the trained Vertex-phased-VQIL solution) along shared random directions, the geometric discount is what makes the objective smooth: a uniform window of the same horizon raises the curvature of the loss (mean absolute second difference relative to its range) by $22\times$ on the grids and $18\times$

Table 2: Gene-perturbation response ranking on protein–protein interaction graphs. Mean ± sd over 25 cells (five Perturb-seq datasets × five seed-disjoint folds). Model configurations match Table 1 ($\gamma = 0.9$). Results at $\gamma = 0.7$ are in Appendix Table 9. **Bold**: main VQIL variant.

	Gene expression response $\lvert\mathrm{logFC}\rvert$			
Model	Spearman ρ	AUPRC@10%	AUROC@10%	NDCG@10%
Naive-VQIL (untrained)	0.171 ± 0.050	0.155 ± 0.026	0.616 ± 0.040	0.330 ± 0.053
Standard-VQIL	0.587 ± 0.157	0.336 ± 0.080	0.824 ± 0.039	0.532 ± 0.045
Link-phased-VQIL	0.587 ± 0.157	0.336 ± 0.080	0.824 ± 0.039	0.532 ± 0.044
Vertex-phased-VQIL	$\mathbf{0.620 \pm 0.167}$	$\mathbf{0.361 \pm 0.086}$	$\mathbf{0.849 \pm 0.042}$	$\mathbf{0.552 \pm 0.045}$
Graph-phased-VQIL	0.620 ± 0.167	0.361 ± 0.085	0.849 ± 0.042	0.554 ± 0.047
Phased-only-VQIL	0.474 ± 0.108	0.270 ± 0.054	0.784 ± 0.037	0.468 ± 0.054
Naive-RWR (untrained)	0.174 ± 0.055	0.157 ± 0.029	0.617 ± 0.042	0.335 ± 0.042
Trained-RWR	0.570 ± 0.150	0.296 ± 0.073	0.808 ± 0.041	0.491 ± 0.047
GCN	0.582 ± 0.166	0.314 ± 0.077	0.815 ± 0.047	0.518 ± 0.056
GAT	0.528 ± 0.183	0.285 ± 0.088	0.782 ± 0.063	0.487 ± 0.070
Sheaf-$\mathcal{O}(2)$	0.365 ± 0.141	0.208 ± 0.039	0.696 ± 0.043	0.411 ± 0.032

on the PPI networks under coherent summation, and by $2.4\times$ to $6.4\times$ under square-then-add. Coherent summation then smooths the surface further, around trained solutions: $1.3\times$ lower curvature on the grids and $1.6\times$ on the PPI networks, and $1.7\times$ and $2.4\times$ lower detrended total variation. On the grids it also makes the gradient consistent from the outset: gradients computed on disjoint mini-batches at the untrained walk agree with median cosine 0.94 versus 0.54 for square-then-add. The coherent readout is therefore not an accuracy trade-off against trainability: its geometric weighting yields a smooth objective, and its coherence makes the longer horizon informative while producing a better-conditioned surface around the solutions it finds.

Table 3: Readout ablation at the $\gamma = 0.9$ horizon: the 2×2 factorial over step weighting ($K, T = 109$) and the order in which the squared norm is taken. All arms are Vertex-phased-VQIL trained under the protocol of Tables 1 and 2. Spearman ρ, mean ± sd over five seed-disjoint folds of six grids (30 cells) and five Perturb-seq datasets (25 cells). **Bold**: coherent resolvent readout of VQIL.

	Power grids		PPI networks
Readout	Phase-angle $\lvert\Delta\theta\rvert$	Voltage $\lvert\Delta V\rvert$	Gene expression
Uniform, add-then-square	0.703 ± 0.165	0.773 ± 0.099	0.515 ± 0.136
Uniform, square-then-add	0.782 ± 0.191	0.692 ± 0.149	0.453 ± 0.155
Geometric, add-then-square	$\mathbf{0.911 \pm 0.043}$	$\mathbf{0.831 \pm 0.056}$	$\mathbf{0.620 \pm 0.167}$
Geometric, square-then-add	0.839 ± 0.126	0.733 ± 0.093	0.517 ± 0.156

The trained fields read back as statements about the withheld mechanism. Every parameter sits on a named edge or node (Definition 1): the fitted Vertex-phased-VQIL is one weight $\omega \in [-2, 2]$ per edge and one phase $\delta \in [-\pi, \pi]$ per node, readable without post hoc attribution. Neural baselines have no such correspondence, and trained-RWR carries weights only. We therefore compare the classical and quantum ω fields on identical edges and evaluate VQIL's unique δ field against withheld annotations: branch susceptance and bus type on the power grids, and co-expression and the response propensity on the PPI network (Figure 1, Appendix Figure 5, Appendix Figure 6). Both fields are reproducible: across training folds, ω correlates at $\rho = 0.68$–0.96 and $|\delta|$ at $\rho = 0.37$–0.83. The phase field is broadly distributed (median $|\delta| \in [0.24, 0.90]$ rad) with no boundary saturation (Appendix Figure 4). The sign of δ, however, does not reproduce across folds (mean cross-fold $|\rho| < 0.03$). A simultaneous flip of all node phases ($\delta \to -\delta$) leaves the readout exactly invariant, so the overall sign is unidentifiable. Relative signs are not constrained by this symmetry but are empirically unstable, so we interpret $|\delta|$ only.

Power grids. A bus perturbation propagates in proportion to branch susceptance and is arrested at voltage-regulated nodes. The VQIL and RWR weight fields largely agree (Appendix Figure 2), but

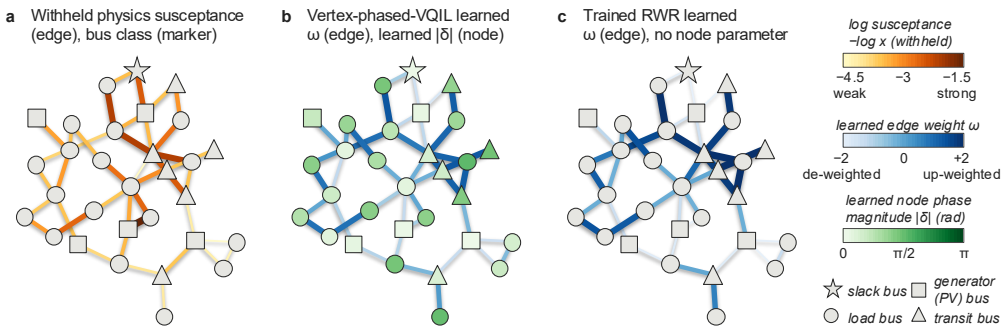


Figure 1: Trained fields on the 30-bus grid (case30; voltage response, $\gamma = 0.9$, five-fold means). (a) Ground-truth branch susceptance (edges) and bus class (markers). (b) Vertex-phased-VQIL: edge weights (ω), which follow susceptance, and node-phase magnitudes $|\delta|$, smallest at regulated buses. (c) Trained-RWR learns only edge weights (ω). Fields show 5-fold means.

RWR drives about twice as many weights to the ± 2 boundary (46% vs. 21% of edges; Appendix Figure 3). VQIL's weights also carry susceptance information beyond RWR's: after partialling out the RWR weights, the VQIL residual still correlates with withheld log-susceptance ($\rho = 0.06$, $p = 3 \times 10^{-3}$), whereas the RWR residual does not ($\rho = -0.01$, n.s.). The phase magnitudes capture voltage regulation $|\delta|$: it is lowest at voltage-regulated buses, and separates generator buses from all other buses with AUC 0.74 (degree-matched null 0.56).

PPI networks. Perturbations propagate through physical interactions modulated by cellular co-regulation. The two weight fields again agree closely (Appendix Figure 2), with RWR saturating more edges (43% vs. 25%), and both track control-cell co-expression to a similar, modest degree (VQIL $\rho = 0.11$–0.28; RWR 0.11–0.27). The phase field, which has no RWR counterpart, is larger at genes that respond strongly across many perturbations, capturing stress-responsive genes (AUC 0.71–0.87), independently of their baseline expression level. Across both domains, the learned weights encode transmission strength and the phase magnitudes track local response propensity: low at regulated buses, high at strongly responsive genes.

6 Discussion and Conclusion

In this work, we introduced VQIL, a scalable framework taking advantage of quantum interference as an alternative to conventional graph representation learning. Standard random walks and graph neural networks frequently struggle with oversmoothing, where information homogenizes around high-degree topological hubs. By introducing learned local node phases (δ), VQIL leverages interference to dynamically bypass these bottlenecks, preserving critical local signals such as responsive gene modules or regulated electrical buses. A coherent resolvent readout highlights that amplitude superposition is not merely an incidental feature, but the foundational mechanism enabling targeted signal routing.

Beyond competitive empirical performance, VQIL's primary contribution lies in mechanistic visibility. Deep neural baselines obscure predictive logic within dense, distributed latent spaces, whereas classical walks restrict interpretability exclusively to edge transitions. In contrast, VQIL maps every scalar parameter directly to an identifiable physical component. Edge weights (ω) denote primary transmission corridors, while node-phase magnitudes ($|\delta|$) mark explicit sinks where perturbations are halted. This turns the trained model into an interpretable hypothesis, which is vital for discovery-driven fields like computational biology.

Despite these strengths, several architectural limitations define clear future directions. Currently, VQIL remains a transductive ranking tool bound to fixed topologies, incapable of predicting signed, continuous expression values or generalizing to unseen graphs. Future extensions could introduce lightweight encoders to parameterize ω and δ inductively from features. Furthermore, moving beyond symmetric constraints to incorporate anti-symmetric link phases in directed networks will unlock causal modeling where detailed balance is fundamentally broken.

ETHICS STATEMENT

This work uses only previously published, de-identified data: standard IEEE/PEGASE power-grid test cases (Zimmerman et al., 2011; Josz et al., 2016; Thurner et al., 2018) and publicly released Perturb-seq screens (Replogle et al., 2022; Frangieh et al., 2021; Tian et al., 2021; Adamson et al., 2016), none of which involve new data collection from human subjects. The method ranks perturbation responses on existing networks and does not generate synthetic biological or infrastructure data. We do not foresee direct harmful applications. As with any model of infrastructure or gene-regulatory response, we caution against deploying predictions operationally without independent experimental or engineering validation.

REPRODUCIBILITY STATEMENT

The model, training objective, and baselines are fully specified in Section 4 (Equations 1–4, Algorithm 1) and Section 5 (training and evaluation protocol, §5.1). Dataset sources, graph construction, and train/validation/test splits are detailed in §5.1 and Table 4; all datasets are publicly available at the cited sources. Hyperparameters and baseline architectures are listed in Supplementary Table 5, and per-model parameter counts and training cost in Supplementary Tables 6–7. We report mean $\pm$ sd over five seed-disjoint folds for every result, with paired Wilcoxon signed-rank tests where significance is claimed.

AI DISCLOSURE

In this work, we used generative AI tools for designing methodology or experiments, implementing methods, aiding and polishing writing, drafting specific sections of the manuscript, and for retrieval and discovery of literature. We have not used generative AI tools for generating synthetic datasets, developing theoretical models, formulating or proving mathematical claims, proposing hypotheses, or interpreting results, and the rest of the required disclosure tasks are not applicable to this work.

We have thoroughly reviewed all AI-assisted work. Specifically, AI-assisted methodology designs and implemented code were rigorously tested, executed, and verified for mathematical and functional correctness by the authors prior to running the final evaluations. All AI-drafted text was critically reviewed, extensively edited, and fact-checked to ensure scientific accuracy and alignment with our experimental findings. Additionally, all related work and literature discovered by AI retrieval tools were manually verified through an independent literature survey to ensure factual correctness, contextual relevance, and proper citation. The authors take full responsibility for the final content of this work, including all text, claims, code, and artifacts produced with the aid of generative AI.

REFERENCES

Britt Adamson, Thomas M. Norman, Marco Jost, Min Y. Cho, James K. Nuñez, Yuwen Chen, Jacqueline E. Villalta, Luke A. Gilbert, Max A. Horlbeck, Marco Y. Hein, Ryan A. Pak, Andrew N. Gray, Carol A. Gross, Atray Dixit, Oren Parnas, Aviv Regev, and Jonathan S. Weissman. A multiplexed single-cell CRISPR screening platform enables systematic dissection of the unfolded protein response. *Cell*, 167(7):1867–1882.e21, 2016. doi: 10.1016/j.cell.2016.11.048.

Alekh Agarwal, Soumen Chakrabarti, and Sunny Aggarwal. Learning to rank networked entities. In *Proceedings of the 12th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD)*, pp. 14–23, 2006. doi: 10.1145/1150402.1150409.

Dorit Aharonov, Andris Ambainis, Julia Kempe, and Umesh Vazirani. Quantum walks on graphs. In *Proceedings of the 33rd Annual ACM Symposium on Theory of Computing (STOC)*, pp. 50–59, 2001. doi: 10.1145/380752.380758.

O. Alsac and B. Stott. Optimal load flow with steady-state security. *IEEE Transactions on Power Apparatus and Systems*, PAS-93(3):745–751, 1974. doi: 10.1109/TPAS.1974.293972.

Lars Backstrom and Jure Leskovec. Supervised random walks: Predicting and recommending links in social networks. In *Proceedings of the 4th ACM International Conference on Web Search and Data Mining (WSDM)*, pp. 635–644, 2011. doi: 10.1145/1935826.1935914.

Mathieu Blondel, Olivier Teboul, Quentin Berthet, and Josip Djolonga. Fast differentiable sorting and ranking. In *Proceedings of the 37th International Conference on Machine Learning (ICML)*, volume 119 of *Proceedings of Machine Learning Research*, pp. 950–959, 2020.

Cristian Bodnar, Francesco Di Giovanni, Benjamin Paul Chamberlain, Pietro Liò, and Michael M. Bronstein. Neural sheaf diffusion: A topological perspective on heterophily and oversmoothing in GNNs. In *Advances in Neural Information Processing Systems (NeurIPS)*, volume 35, pp. 18527–18541, 2022.

Richard D. Christie. Power systems test case archive. University of Washington, `https://labs.ece.uw.edu/pstca/`, 1993.

Lenore Cowen, Trey Ideker, Benjamin J. Raphael, and Roded Sharan. Network propagation: A universal amplifier of genetic associations. *Nature Reviews Genetics*, 18(9):551–562, 2017. doi: 10.1038/nrg.2017.38.

Stefan Dernbach, Arman Mohseni-Kabir, Siddharth Pal, and Don Towsley. Quantum walk neural networks for graph-structured data. In *Complex Networks and Their Applications VII*, Studies in Computational Intelligence, pp. 182–193. Springer, 2018. doi: 10.1007/978-3-030-05414-4_15.

Stefan Dernbach, Arman Mohseni-Kabir, Siddharth Pal, Miles Gepner, and Don Towsley. Quantum walk neural networks with feature dependent coins. *Applied Network Science*, 4(1):76, 2019. doi: 10.1007/s41109-019-0188-2.

Vijay Prakash Dwivedi, Anh Tuan Luu, Thomas Laurent, Yoshua Bengio, and Xavier Bresson. Graph neural networks with learnable structural and positional representations. In *International Conference on Learning Representations (ICLR)*, 2022.

Stéphane Fliscounakis, Patrick Panciatici, Florin Capitanescu, and Louis Wehenkel. Contingency ranking with respect to overloads in very large power systems taking into account uncertainty, preventive, and corrective actions. *IEEE Transactions on Power Systems*, 28(4):4909–4917, 2013. doi: 10.1109/TPWRS.2013.2251015.

Chris J. Frangieh, Johannes C. Melms, Pratiksha I. Thakore, Kathryn R. Geiger-Schuller, Patricia Ho, Adrienne M. Luoma, Brian Cleary, Livnat Jerby-Arnon, Shruti Malu, Michael S. Cuoco, Maryann Zhao, Casey R. Ager, Meri Rogava, Lila Hovey, Asaf Rotem, Chantale Bernatchez, Kai W. Wucherpfennig, Bruce E. Johnson, Orit Rozenblatt-Rosen, Dirk Schadendorf, Aviv Regev, and Benjamin Izar. Multimodal pooled Perturb-CITE-seq screens in patient models define mechanisms of cancer immune evasion. *Nature Genetics*, 53(3):332–341, 2021. doi: 10.1038/s41588-021-00779-1.

Namshik Han, Woochang Hwang, Konstantinos Tzelepis, Patrick Schmerer, Eliza Yankova, Méabh MacMahon, Winnie Lei, Nicholas M. Katritsis, Anika Liu, Ulrike Felgenhauer, Alison Schuldt, Rebecca Harris, Kathryn Chapman, Frank McCaughan, Friedemann Weber, and Tony Kouzarides. Identification of SARS-CoV-2–induced pathways reveals drug repurposing strategies. *Science Advances*, 7(27):eabh3032, 2021. doi: 10.1126/sciadv.abh3032.

Woojeong Jin, Jinhong Jung, and U Kang. Supervised and extended restart in random walks for ranking and link prediction in networks. *PLOS ONE*, 14(3):e0213857, 2019. doi: 10.1371/journal.pone.0213857.

Cédric Josz, Stéphane Fliscounakis, Jean Maeght, and Patrick Panciatici. AC power flow data in MATPOWER and QCQP format: iTesla, RTE snapshots, and PEGASE. arXiv:1603.01533, 2016.

Thomas N. Kipf and Max Welling. Semi-supervised classification with graph convolutional networks. In *International Conference on Learning Representations (ICLR)*, 2017.

Pan Li, Yanbang Wang, Hongwei Wang, and Jure Leskovec. Distance encoding: Design provably more powerful neural networks for graph representation learning. In *Advances in Neural Information Processing Systems (NeurIPS)*, volume 33, pp. 4465–4478, 2020.

Jianzhu Ma, Michael Ku Yu, Samson Fong, Keiichiro Ono, Eric Sage, Barry Demchak, Roded Sharan, and Trey Ideker. Using deep learning to model the hierarchical structure and function of a cell. *Nature Methods*, 15(4):290–298, 2018. doi: 10.1038/nmeth.4627.

Sergio A. Ortega and Miguel A. Martin-Delgado. Generalized quantum PageRank algorithm with arbitrary phase rotations. *Physical Review Research*, 5(1):013061, 2023. doi: 10.1103/PhysRevResearch.5.013061.

Sergio A. Ortega and Miguel A. Martin-Delgado. SQUWALS: A Szegedy QUantum WALks simulator. *Advanced Quantum Technologies*, 7(7):2400022, 2024. doi: 10.1002/qute.202400022.

Sergio A. Ortega and Miguel A. Martin-Delgado. Complex-phase extensions of the Szegedy quantum walk on graphs. *Physical Review A*, 111(3):032216, 2025. doi: 10.1103/PhysRevA.111.032216.

Sergio A. Ortega and Daniel K. Park. Efficient simulation of Szegedy quantum walk formulations and algorithms. arXiv:2606.14226, 2026.

Lawrence Page, Sergey Brin, Rajeev Motwani, and Terry Winograd. The PageRank citation ranking: Bringing order to the web. Technical Report 1999-66, Stanford InfoLab, 1999.

Giuseppe Davide Paparo and Miguel Angel Martin-Delgado. Google in a quantum network. *Scientific Reports*, 2:444, 2012. doi: 10.1038/srep00444.

Giuseppe Davide Paparo, Markus Müller, Francesc Comellas, and Miguel Angel Martin-Delgado. Quantum Google in a complex network. *Scientific Reports*, 3:2773, 2013. doi: 10.1038/srep02773.

Jaesub Park, Woochang Hwang, Seokjun Lee, Hyun Chang Lee, Méabh MacMahon, Matthias Zilbauer, and Namshik Han. Advancing understanding of long COVID pathophysiology through quantum walk-based network analysis. *Bioinformatics Advances*, 6(1):vbag050, 2026. doi: 10.1093/bioadv/vbag050.

Joseph M. Replogle, Reuben A. Saunders, Angela N. Pogson, Jeffrey A. Hussmann, Alexander Lenail, Alina Guna, Lauren Mascibroda, Eric J. Wagner, Karen Adelman, Gila Lithwick-Yanai, Nika Iremadze, Florian Oberstrass, Doron Lipson, Jessica L. Bonnar, Marco Jost, Thomas M. Norman, and Jonathan S. Weissman. Mapping information-rich genotype–phenotype landscapes with genome-scale Perturb-seq. *Cell*, 185(14):2559–2575.e28, 2022. doi: 10.1016/j.cell.2022.05.013.

Cynthia Rudin. Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *Nature Machine Intelligence*, 1(5):206–215, 2019. doi: 10.1038/s42256-019-0048-x.

Harto Saarinen, Mark Goldsmith, Rui-Sheng Wang, Joseph Loscalzo, and Sabrina Maniscalco. Disease gene prioritization with quantum walks. *Bioinformatics*, 40(8):btae513, 2024. doi: 10.1093/bioinformatics/btae513.

Mario Szegedy. Quantum speed-up of Markov chain based algorithms. In *Proceedings of the 45th Annual IEEE Symposium on Foundations of Computer Science (FOCS)*, pp. 32–41, 2004. doi: 10.1109/FOCS.2004.53.

Damian Szklarczyk, Rebecca Kirsch, Mikaela Koutrouli, Katerina Nastou, Farrokh Mehryary, Radja Hachilif, Annika L. Gable, Tao Fang, Nadezhda T. Doncheva, Sampo Pyysalo, Peer Bork, Lars J. Jensen, and Christian von Mering. The STRING database in 2023: Protein–protein association networks and functional enrichment analyses for any sequenced genome of interest. *Nucleic Acids Research*, 51(D1):D638–D646, 2023. doi: 10.1093/nar/gkac1000.

Leon Thurner, Alexander Scheidler, Florian Schäfer, Jan-Hendrik Menke, Julian Dollichon, Friederike Meier, Steffen Meinecke, and Martin Braun. pandapower—an open-source Python tool for convenient modeling, analysis, and optimization of electric power systems. *IEEE Transactions on Power Systems*, 33(6):6510–6521, 2018. doi: 10.1109/TPWRS.2018.2829021.

Ruilin Tian, Anthony Abarientos, Jason Hong, Sayed Hadi Hashemi, Rui Yan, Nina Dräger, Kun Leng, Mike A. Nalls, Andrew B. Singleton, Ke Xu, Faraz Faghri, and Martin Kampmann. Genome-wide CRISPRi/a screens in human neurons link lysosomal failure to ferroptosis. *Nature Neuroscience*, 24(7):1020–1034, 2021. doi: 10.1038/s41593-021-00862-0.

Hanghang Tong, Christos Faloutsos, and Jia-Yu Pan. Fast random walk with restart and its applications. In *Proceedings of the 6th IEEE International Conference on Data Mining (ICDM)*, pp. 613–622, 2006. doi: 10.1109/ICDM.2006.70.

Petar Veličković, Guillem Cucurull, Arantxa Casanova, Adriana Romero, Pietro Liò, and Yoshua Bengio. Graph attention networks. In *International Conference on Learning Representations (ICLR)*, 2018.

Vijay Vittal, D. Martin, R. Chu, J. Fish, J. C. Giri, C. K. Tang, F. Eugenio Villaseca, and R. G. Farmer. Transient stability test systems for direct stability methods. *IEEE Transactions on Power Systems*, 7(1):37–43, 1992. doi: 10.1109/59.141684.

Zhiting Wei, Yiheng Wang, Yicheng Gao, Shuguang Wang, Ping Li, Duanmiao Si, Yuli Gao, Siqi Wu, Danlu Li, Kejing Dong, Xingbo Yang, Chen Tang, Shaliu Fu, Xiaohan Chen, Wannian Li, Yuzhou You, Chen Zhang, Aibin Liang, Guohui Chuai, and Qi Liu. Benchmarking algorithms for generalizable single-cell perturbation response prediction. *Nature Methods*, 23(2):451–464, 2026. doi: 10.1038/s41592-025-02980-0.

Michael K. Yu, Jianzhu Ma, Jasmin Fisher, Jason F. Kreisberg, Benjamin J. Raphael, and Trey Ideker. Visible machine learning for biomedicine. *Cell*, 173(7):1562–1565, 2018. doi: 10.1016/j.cell.2018.05.056.

Muhan Zhang and Yixin Chen. Link prediction based on graph neural networks. In *Advances in Neural Information Processing Systems (NeurIPS)*, volume 31, 2018.

Muhan Zhang, Pan Li, Yinglong Xia, Kai Wang, and Long Jin. Labeling trick: A theory of using graph neural networks for multi-node representation learning. In *Advances in Neural Information Processing Systems (NeurIPS)*, volume 34, pp. 9061–9073, 2021.

Ray D. Zimmerman, Carlos E. Murillo-Sánchez, and Robert J. Thomas. MATPOWER: Steady-state operations, planning, and analysis tools for power systems research and education. *IEEE Transactions on Power Systems*, 26(1):12–19, 2011. doi: 10.1109/TPWRS.2010.2051168.

A SUPPLEMENTARY DATA

A.1 DATASETS AND MODELS

This appendix shows the explicit details about the datasets and models used. Table 4 summarizes the properties of each graph dataset. Table 5 shows the configuration of the graph neural networks baseline. Table 6 compare the training cost of each experiment. Finally, we also summarizes the parameters count for each domain and dataset in Table 7.

Table 4: Datasets and graphs. The model sees only the unweighted, undirected graph and the seed. Everything else is withheld. Domain 1: bus graph of in-service lines and transformers (largest connected component). A query injects $+20$ MW / $+6$ MVAr at the seed bus and records the AC power-flow response at every other bus (non-converged probes are dropped). gen./transf.: generator buses and transformer branches, both withheld from the model. Domain 2: largest connected component of STRING v12.0 human interactions with combined score ≥ 400 over the measured genes (no edges added). A query is one perturbed gene (Frangieh: control arm only). In both domains all non-seed nodes are candidates, the top 10% by response are positives, and splits are seed-disjoint five-fold partitions (per-fold query counts shown). ρ_{hop}: Spearman correlation between $-$(hop distance to the seed) and the response, averaged over queries (angle / voltage for grids); a proximity-only baseline.

Dataset	Source	Details	$\lvert V\rvert$	$\lvert E\rvert$	Avg. deg.	Density	Diam.	Queries	Train / Val / Test	ρ_{hop}
Domain 1: power grids (pandapower); responses $\lvert\Delta\theta\rvert$ and $\lvert\Delta V\rvert$										
case30	IEEE 30-bus	6 gen., 0 transf.	30	41	2.73	0.0943	6	30	18 / 6 / 6	0.344 / 0.383
case57	IEEE 57-bus	7 gen., 15 transf.	57	78	2.74	0.0489	12	52	30–32 / 10–11 / 10–11	0.213 / 0.082
case118	IEEE 118-bus	54 gen., 13 transf.	118	179	3.03	0.0259	14	117	69–71 / 23–24 / 23–24	0.571 / 0.214
case145	IEEE 145-bus	59 gen., 64 transf.	145	422	5.82	0.0404	11	144	86–87 / 28–29 / 28–29	0.339 / 0.153
case300	IEEE 300-bus	77 gen., 128 transf.	300	409	2.73	0.0091	24	281	168–169 / 56–57 / 56–57	0.485 / 0.381
case1354	PEGASE 1354-bus	312 gen., 151 transf.	1,354	1,710	2.53	0.0019	25	1,354	812–813 / 270–271 / 270–271	0.490 / 0.273
Domain 2: Perturb-seq on STRING PPI graphs; response $\lvert\text{logFC}\rvert$										
K562	Replogle et al. (2022)	CRISPRi, K562; 263k cells	5,787	218,844	75.6	0.0131	6	1,615	969 / 323 / 323	0.125
RPE1	Replogle et al. (2022)	CRISPRi, RPE1; 174k cells	5,655	194,587	68.8	0.0122	6	1,329	797–798 / 265–266 / 265–266	0.098
Frangieh	Frangieh et al. (2021)	CRISPR-KO, melanoma; 27k cells	5,023	99,522	39.6	0.0079	9	188	112–114 / 37–38 / 37–38	0.092
Tian-i	Tian et al. (2021)	CRISPRi, iPSC-derived neurons; 32k cells	2,965	29,059	19.6	0.0066	10	178	106–108 / 35–36 / 35–36	0.127
Adamson	Adamson et al. (2016)	CRISPRi, K562; 57k cells	4,950	105,599	42.7	0.0086	7	75	45 / 15 / 15	0.215

Table 5: Neural baseline configurations. Entries of the form a / b give Domain 1 / Domain 2. Architecture and training protocol are shared. Depth and learning rate were selected per domain on validation Spearman ρ from $\{0.005, 0.01, 0.05\} \times \{3, 6\}$ layers (Domain 1: case118, fold 0, mean over both responses; Domain 2: Adamson, fold 0) and then fixed for all datasets of that domain. The training protocol (lower block) is identical for every VQIL, RWR and neural arm, except that VQIL and RWR use learning rate 0.05.

	GCN	GAT	Sheaf-$\mathcal{O}(2)$
Architecture (both domains)			
Input features (50-dim)	seed indicator, degree, log-degree, hop distance to seed (one-hot, 0–12, $\geq$13), RWSE-16, LapPE-16; one RWR-landing channel held at zero		
Message passing	sym.-normalised, self-loops	single-head additive attention, self-loops	$\mathcal{O}(2)$ sheaf diffusion, 2-dim stalks
Hidden width	64	64	16 channels $\times$ 2
Layer details	residual, LayerNorm, LeakyReLU(0.1)	residual, LayerNorm, LeakyReLU(0.1)	per-edge rotation angle from 2-layer MLP (32), learned step ε
Dropout	0.2	0.2	0.2
Decoder	pair head MLP($[h_v, h_s, h_v \odot h_s]$), 64 hidden, PReLU; softplus output		
Selected per domain (Domain 1 / Domain 2)			
Layers	3 / 6	3 / 6	3 / 3
Learning rate	0.01 / 0.01	0.005 / 0.01	0.005 / 0.005
Trainable parameters	28,546 / 41,410	28,930 / 42,178	15,019 / 15,019
Validation ρ at selection	0.946 / 0.729	0.944 / 0.690	0.943 / 0.397
Training protocol (all arms, both domains)			
Loss	soft-Spearman ($\tau = 1$) on log scores		
Optimiser	AdamW, weight decay 0.01		
Batch	8 queries (evaluation 16); at most 200 training queries sampled per epoch		
Epochs	at most 300; early stopping with patience 10 on validation ρ; best-validation checkpoint		

Table 6: Training cost ($\gamma = 0.9$ for VQIL variants, $r = 0.1$ for RWR). (a) Wall-clock seconds per epoch, computed as total cell time divided by epochs run (so fixed evaluation cost is included); (b) mean total seconds per cell; (c) median epochs to early stopping (at most 300, patience 10). Means and medians over the five folds (and, for Domain 1, both responses). †Untrained models: test-set inference time only.

	Domain 1: power grids						Domain 2: PPI graphs				
Model	case30	case57	case118	case145	case300	case1354	K562	RPE1	Frangieh	Tian-i	Adamson
(a) Seconds per epoch											
Standard-VQIL	0.20	0.28	0.62	0.76	1.49	3.08	40.1	34.1	8.49	3.10	3.57
Link-phased-VQIL	0.21	0.28	0.63	0.77	1.51	3.08	40.3	34.3	8.54	3.11	3.59
Vertex-phased-VQIL	0.21	0.28	0.63	0.79	1.54	3.09	41.1	34.8	8.68	3.15	3.65
Graph-phased-VQIL	0.20	0.29	0.64	0.78	1.56	3.13	41.2	35.1	8.72	3.16	3.66
Phased-only-VQIL	0.21	0.29	0.64	0.86	1.62	3.10	40.8	34.7	8.65	3.14	3.64
Trained-RWR	0.15	0.23	0.50	0.63	1.24	2.68	12.8	11.0	2.95	1.55	1.23
GCN	0.11	0.17	0.51	0.54	1.90	7.41	68.3	59.4	16.4	7.73	6.75
GAT	0.13	0.19	0.42	0.63	1.91	9.13	87.0	75.0	20.8	8.98	8.58
Sheaf-$\mathcal{O}(2)$	0.13	0.19	0.63	0.96	2.83	14.7	104	87.8	21.6	7.27	9.11
(b) Seconds per cell (training to early stopping, including evaluation)											
Naive-VQIL (untrained)	0.02†	0.03†	0.07†	0.08†	0.17†	1.10†	12.6†	9.49†	0.82†	0.34†	0.34†
Standard-VQIL	8.84	13.8	21.6	36.0	54.9	134	3,550	3,405	693	208	382
Link-phased-VQIL	9.06	14.0	22.1	39.7	55.2	142	3,932	3,050	672	200	425
Vertex-phased-VQIL	8.58	15.0	32.1	49.2	84.7	177	1,604	1,770	420	176	228
Graph-phased-VQIL	9.86	17.0	42.2	46.5	70.3	176	1,937	1,411	436	151	240
Phased-only-VQIL	9.86	15.5	23.1	50.5	42.6	101	3,627	3,203	1,194	677	468
Naive-RWR (untrained)	0.02†	0.03†	0.07†	0.08†	0.16†	1.08†	4.87†	3.80†	0.41†	0.23†	0.17†
Trained-RWR	6.19	8.96	20.5	24.0	52.1	142	569	390	71.9	36.5	52.8
GCN	3.89	8.65	25.6	35.2	128	719	11,635	10,337	2,897	1,107	1,308
GAT	5.24	10.2	27.1	38.8	179	892	7,020	9,883	2,635	856	1,030
Sheaf-$\mathcal{O}(2)$	5.87	9.98	44.8	57.9	284	1,284	9,451	6,849	2,013	922	861
(c) Median epochs to early stopping											
Standard-VQIL	42	44	34	46	35	42	94	108	74	72	108
Link-phased-VQIL	42	44	34	46	32	45	110	88	84	66	119
Vertex-phased-VQIL	41	44	49	62	54	52	37	50	44	57	67
Graph-phased-VQIL	42	56	48	62	49	54	41	38	51	43	67
Phased-only-VQIL	47	52	35	52	18	31	88	93	140	234	133
Trained-RWR	42	35	38	34	37	53	38	37	23	23	42
GCN	38	46	48	71	76	108	172	172	197	147	203
GAT	42	36	62	58	92	104	49	131	111	93	122
Sheaf-$\mathcal{O}(2)$	44	49	66	59	101	88	98	80	95	119	95

Table 7: Trainable parameters per model and dataset ($\mathrm{K} = 10^3$). VQIL and RWR learn one weight per undirected edge, plus one phase per node (vertex-phased) and/or per edge (link-phased); Phased-only-VQIL freezes the weights and trains node and edge phases. Neural baselines have a size independent of the graph but set by the depth selected per domain (Table 5: 3 layers for Domain 1; 6 for GCN and GAT in Domain 2). Counts do not depend on γ or on the response.

		Domain 1: power grids						Domain 2: PPI graphs				
Model	Scaling	case30	case57	case118	case145	case300	case1354	K562	RPE1	Frangieh	Tian-i	Adamson
Nodes $\lvert V\rvert$		30	57	118	145	300	1,354	5,787	5,655	5,023	2,965	4,950
Edges $\lvert E\rvert$		41	78	179	422	409	1,710	218,844	194,587	99,522	29,059	105,599
Naive-VQIL (untrained)	–	0	0	0	0	0	0	0	0	0	0	0
Standard-VQIL	$\lvert E\rvert$	41	78	179	422	409	1,710	218.8K	194.6K	99.5K	29.1K	105.6K
Link-phased-VQIL	$2\lvert E\rvert$	82	156	358	844	818	3,420	437.7K	389.2K	199.0K	58.1K	211.2K
Vertex-phased-VQIL	$\lvert E\rvert + \lvert V\rvert$	71	135	297	567	709	3,064	224.6K	200.2K	104.5K	32.0K	110.5K
Graph-phased-VQIL	$2\lvert E\rvert + \lvert V\rvert$	112	213	476	989	1,118	4,774	443.5K	394.8K	204.1K	61.1K	216.1K
Phased-only-VQIL	$\lvert E\rvert + \lvert V\rvert$	71	135	297	567	709	3,064	224.6K	200.2K	104.5K	32.0K	110.5K
Naive-RWR (untrained)	–	0	0	0	0	0	0	0	0	0	0	0
Trained-RWR	$\lvert E\rvert$	41	78	179	422	409	1,710	218.8K	194.6K	99.5K	29.1K	105.6K
GCN	fixed	28.5K	28.5K	28.5K	28.5K	28.5K	28.5K	41.4K	41.4K	41.4K	41.4K	41.4K
GAT	fixed	28.9K	28.9K	28.9K	28.9K	28.9K	28.9K	42.2K	42.2K	42.2K	42.2K	42.2K
Sheaf-$\mathcal{O}(2)$	fixed	15.0K	15.0K	15.0K	15.0K	15.0K	15.0K	15.0K	15.0K	15.0K	15.0K	15.0K

A.2 RESULTS FOR $\gamma = 0.7$

In this appendix we show the training results using the horizon $\gamma = 0.7$. They are summarized in Table 8 for the transmision grids, and in Table 9 for the PPI networks. Moreover, Table 10 shows the readout ablation results at this horizon.

Table 8: Power-grid perturbation-response ranking at the shorter readout horizon, $\gamma = 0.7$ (matched restart $r = 1 - \gamma = 0.3$ for RWR); all other settings as in Table 1. Mean $\pm$ sd over 30 cells (six power grids $\times$ five seed-disjoint folds). The neural baselines have no horizon parameter, so their rows repeat Table 1. **Bold**: main VQIL variant.

	Phase-angle response $\lvert\Delta\theta\rvert$		Voltage response $\lvert\Delta V\rvert$	
Model	Spearman ρ	AUPRC@10%	Spearman ρ	AUPRC@10%
Naive-VQIL (untrained)	0.441 ± 0.160	0.715 ± 0.104	0.238 ± 0.099	0.427 ± 0.123
Standard-VQIL	0.755 ± 0.119	0.735 ± 0.094	0.620 ± 0.099	0.511 ± 0.128
Link-phased-VQIL	0.755 ± 0.119	0.735 ± 0.094	0.620 ± 0.099	0.511 ± 0.128
Vertex-phased-VQIL	$\mathbf{0.849 \pm 0.085}$	$\mathbf{0.777 \pm 0.087}$	$\mathbf{0.704 \pm 0.080}$	$\mathbf{0.548 \pm 0.137}$
Graph-phased-VQIL	0.855 ± 0.083	0.780 ± 0.082	0.704 ± 0.076	0.548 ± 0.133
Phased-only-VQIL	0.769 ± 0.113	0.752 ± 0.100	0.513 ± 0.092	0.494 ± 0.136
Naive-RWR (untrained)	0.436 ± 0.146	0.677 ± 0.097	0.245 ± 0.102	0.420 ± 0.118
Trained-RWR	0.713 ± 0.122	0.722 ± 0.096	0.581 ± 0.117	0.509 ± 0.139
GCN	0.915 ± 0.071	0.813 ± 0.079	0.920 ± 0.041	0.789 ± 0.102
GAT	0.932 ± 0.058	0.830 ± 0.076	0.920 ± 0.041	0.795 ± 0.107
Sheaf-$\mathcal{O}(2)$	0.913 ± 0.081	0.796 ± 0.095	0.908 ± 0.039	0.735 ± 0.116

Table 9: Gene-perturbation response ranking on protein–protein interaction graphs at the shorter readout horizon, $\gamma = 0.7$ (matched restart $r = 0.3$ for RWR); all other settings as in Table 2. Mean $\pm$ sd over 25 cells (five Perturb-seq datasets $\times$ five seed-disjoint folds). The neural baselines have no horizon parameter, so their rows repeat Table 2. **Bold**: main VQIL variant.

	Gene expression response $\lvert\mathrm{logFC}\rvert$			
Model	Spearman ρ	AUPRC@10%	AUROC@10%	NDCG@10%
Naive-VQIL (untrained)	0.171 ± 0.057	0.153 ± 0.029	0.615 ± 0.042	0.323 ± 0.039
Standard-VQIL	0.554 ± 0.147	0.269 ± 0.062	0.798 ± 0.041	0.454 ± 0.043
Link-phased-VQIL	0.554 ± 0.147	0.269 ± 0.062	0.798 ± 0.041	0.454 ± 0.043
Vertex-phased-VQIL	$\mathbf{0.587 \pm 0.161}$	$\mathbf{0.278 \pm 0.061}$	$\mathbf{0.819 \pm 0.042}$	$\mathbf{0.456 \pm 0.041}$
Graph-phased-VQIL	0.587 ± 0.160	0.278 ± 0.060	0.819 ± 0.041	0.456 ± 0.042
Phased-only-VQIL	0.377 ± 0.101	0.189 ± 0.034	0.711 ± 0.041	0.358 ± 0.031
Naive-RWR (untrained)	0.169 ± 0.058	0.153 ± 0.030	0.613 ± 0.041	0.327 ± 0.037
Trained-RWR	0.525 ± 0.140	0.251 ± 0.055	0.779 ± 0.042	0.441 ± 0.040
GCN	0.582 ± 0.166	0.314 ± 0.077	0.815 ± 0.047	0.518 ± 0.056
GAT	0.528 ± 0.183	0.285 ± 0.088	0.782 ± 0.063	0.487 ± 0.070
Sheaf-$\mathcal{O}(2)$	0.365 ± 0.141	0.208 ± 0.039	0.696 ± 0.043	0.411 ± 0.032

Table 10: Readout ablation at the $\gamma = 0.7$ horizon: the 2×2 factorial over step weighting ($K, T = 40$) and the order in which the squared norm is taken. Same protocol, cells and layout as Table 3.Spearman ρ, mean $\pm$ sd over five seed-disjoint folds of six grids (30 cells) and five Perturb-seq datasets (25 cells). **Bold**: coherent resolvent readout of VQIL.

	Power grids		PPI networks
Readout	Phase-angle $\lvert\Delta\theta\rvert$	Voltage $\lvert\Delta V\rvert$	Gene expression
Uniform, add-then-square	0.456 ± 0.120	0.299 ± 0.168	0.508 ± 0.144
Uniform, square-then-add	0.495 ± 0.145	0.310 ± 0.178	0.516 ± 0.141
Geometric, add-then-square	$\mathbf{0.849} \pm 0.085$	$\mathbf{0.704} \pm 0.080$	$\mathbf{0.587 \pm 0.161}$
Geometric, square-then-add	0.861 ± 0.083	0.715 ± 0.086	0.536 ± 0.153

A.3 Learned Parameter Fields

This appendix provides additional details on the learned parameter fields discussed in Section 5.2. Figure 2 compares the edge weights learned by Vertex-phased-VQIL and Trained-RWR, while Figure 3 shows their corresponding weight distributions. Figure 4 shows the distributions of the learned node-phase offsets of Vertex-phased-VQIL. Finally, Figures 5 and 6 visualize the trained fields together with the withheld physics and biology for the remaining power grids and the five PPI networks, respectively.

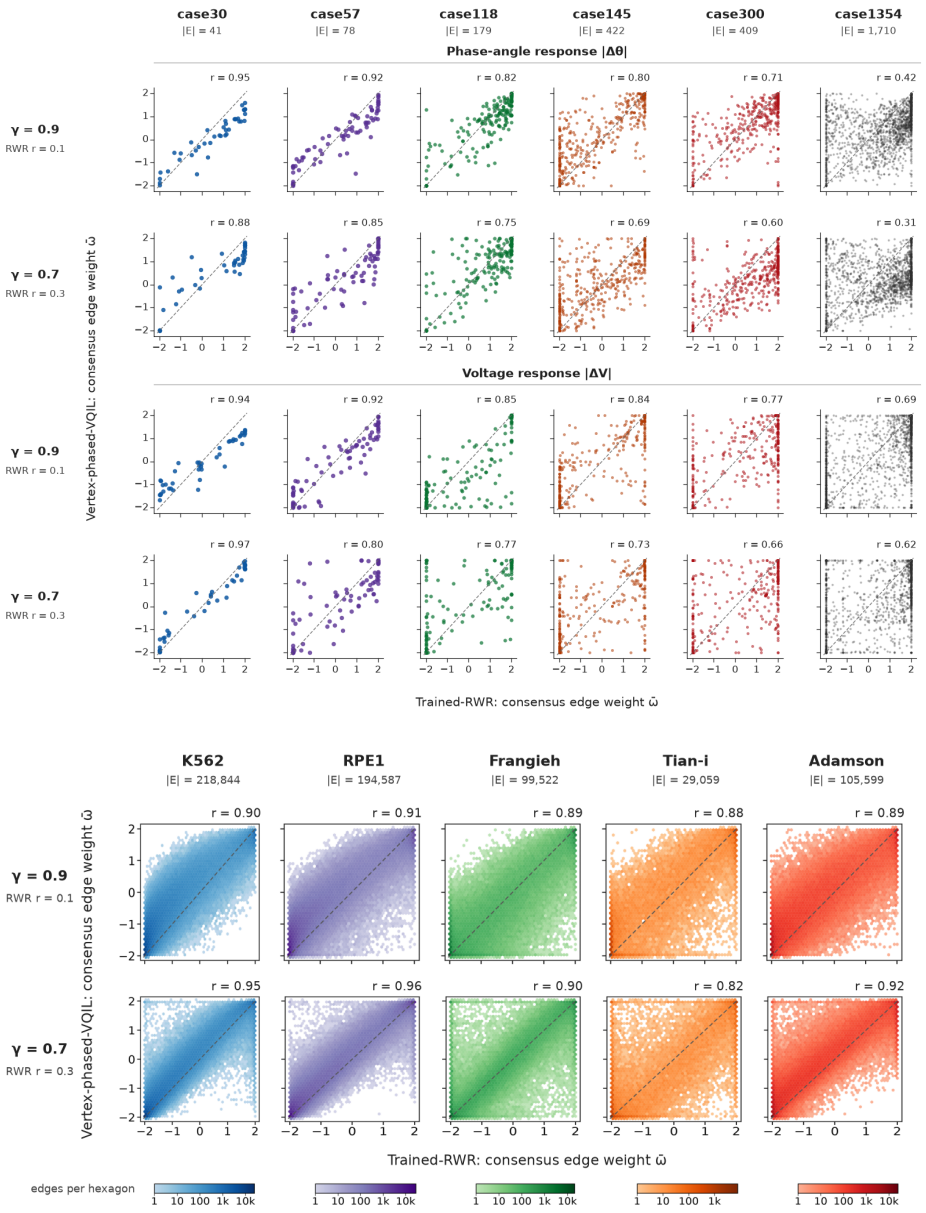


Figure 2: Agreement between the edge-weight fields learned by Vertex-phased-VQIL and Trained-RWR on identical edges. Axes show consensus weights $\bar{\omega}$ (means over the five folds; horizontal: Trained-RWR, vertical: Vertex-phased-VQIL), the dashed line is the identity, and the annotated r is the Pearson correlation between the two consensus fields. Top: power grids (one point per edge) for the phase-angle and voltage responses at $\gamma = 0.9$ (RWR restart 0.1) and $\gamma = 0.7$ (restart 0.3). Bottom: PPI networks at both horizons, as hexagonal bins colored by the number of edges per bin (log scale).

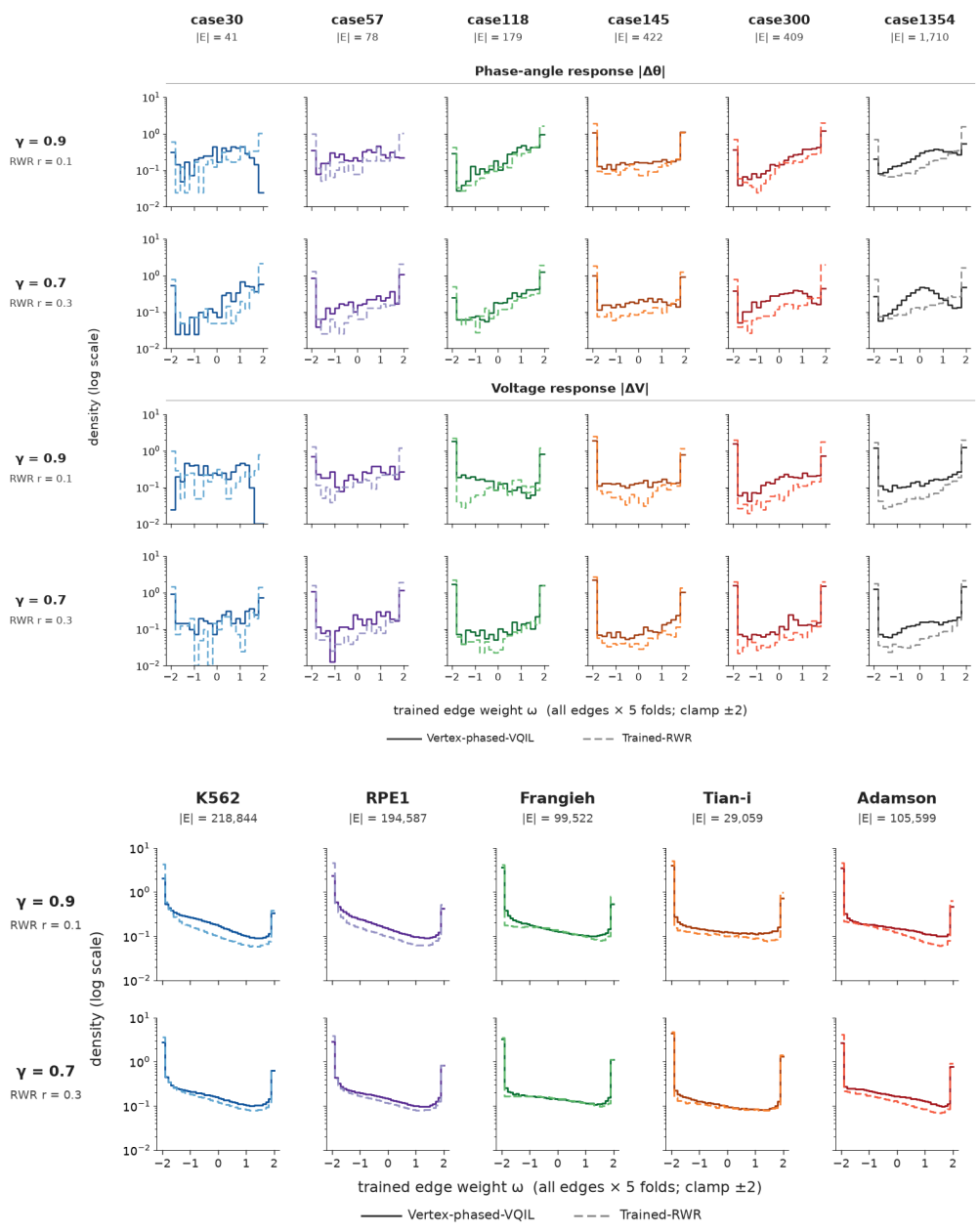


Figure 3: Distributions of the trained edge weights ω of Vertex-phased-VQIL (solid) and Trained-RWR (dashed), pooled over all edges and the five folds (log-scale density; weights are clamped to $[-2, 2]$). Top: power grids (columns) for the phase-angle and voltage responses at $\gamma = 0.9$ and $\gamma = 0.7$. Bottom: PPI networks at both horizons. At $\gamma = 0.9$, Trained-RWR places 46% (grids) and 43% (PPI networks) of its weights at the ± 2 bounds, compared with 21% and 25% for VQIL.

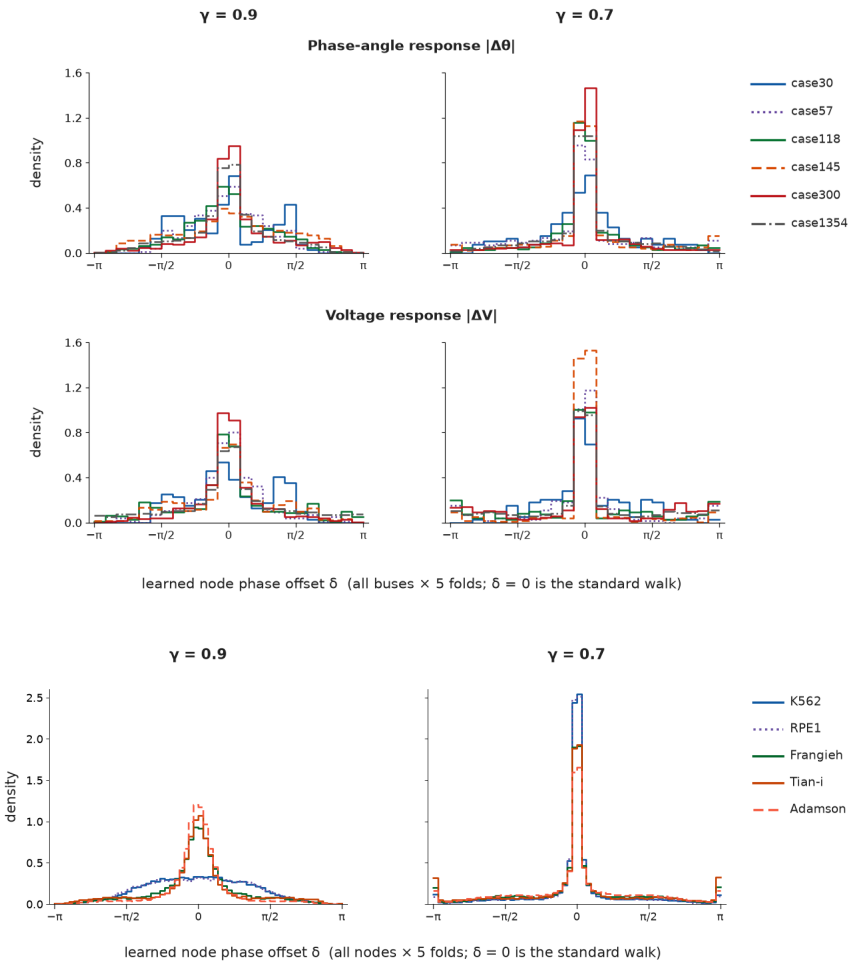


Figure 4: Distributions of the learned node-phase offsets δ of Vertex-phased-VQIL (Trained-RWR has no node parameter), pooled over all nodes and the five folds, at $\gamma = 0.9$ (left) and $\gamma = 0.7$ (right); $\delta = 0$ is the standard walk. Top: power grids (one line per grid) for the phase-angle and voltage responses. Bottom: PPI networks (one line per dataset). The distributions are nearly symmetric about zero. At $\gamma = 0.9$, 0.3% (grids) and 0.06% (PPI networks) of the nodes lie at the $\pm\pi$ bound; at $\gamma = 0.7$, a small fraction of nodes accumulates at the bound.

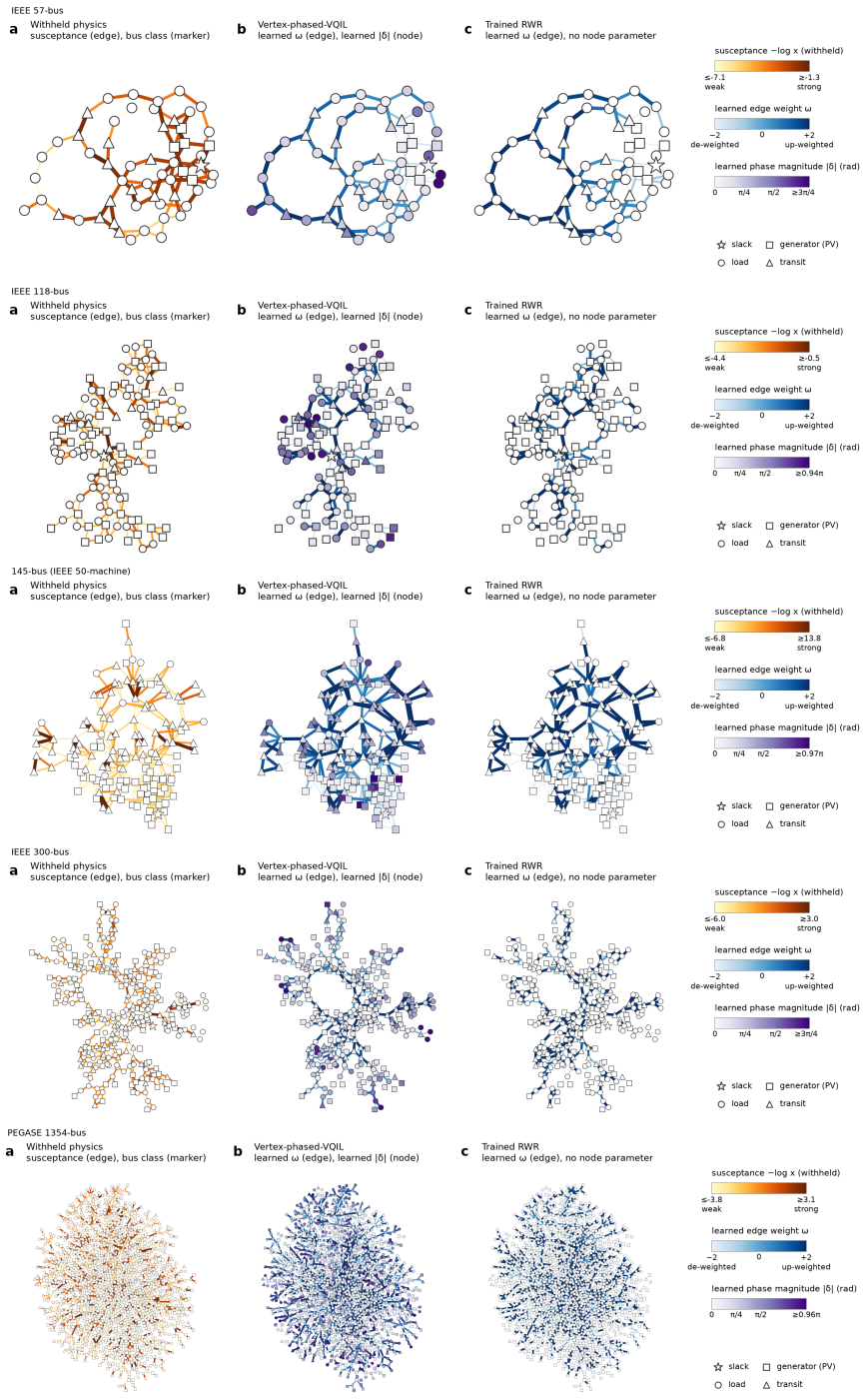


Figure 5: Trained fields on the remaining five power grids (case57, case118, case145, case300, and case1354), shown as in Figure 1 (voltage response, $\gamma = 0.9$, means over the five folds). (a) Withheld branch susceptance $-\log x$ (edge color) and bus class (marker: slack, generator, load, or transit bus). (b) Vertex-phased-VQIL: learned edge weights ω (edge color) and node-phase magnitudes $|\delta|$ (node color; the upper end of the color scale is set per grid). (c) Trained-RWR ($r = 0.1$): learned edge weights; RWR has no node parameter.

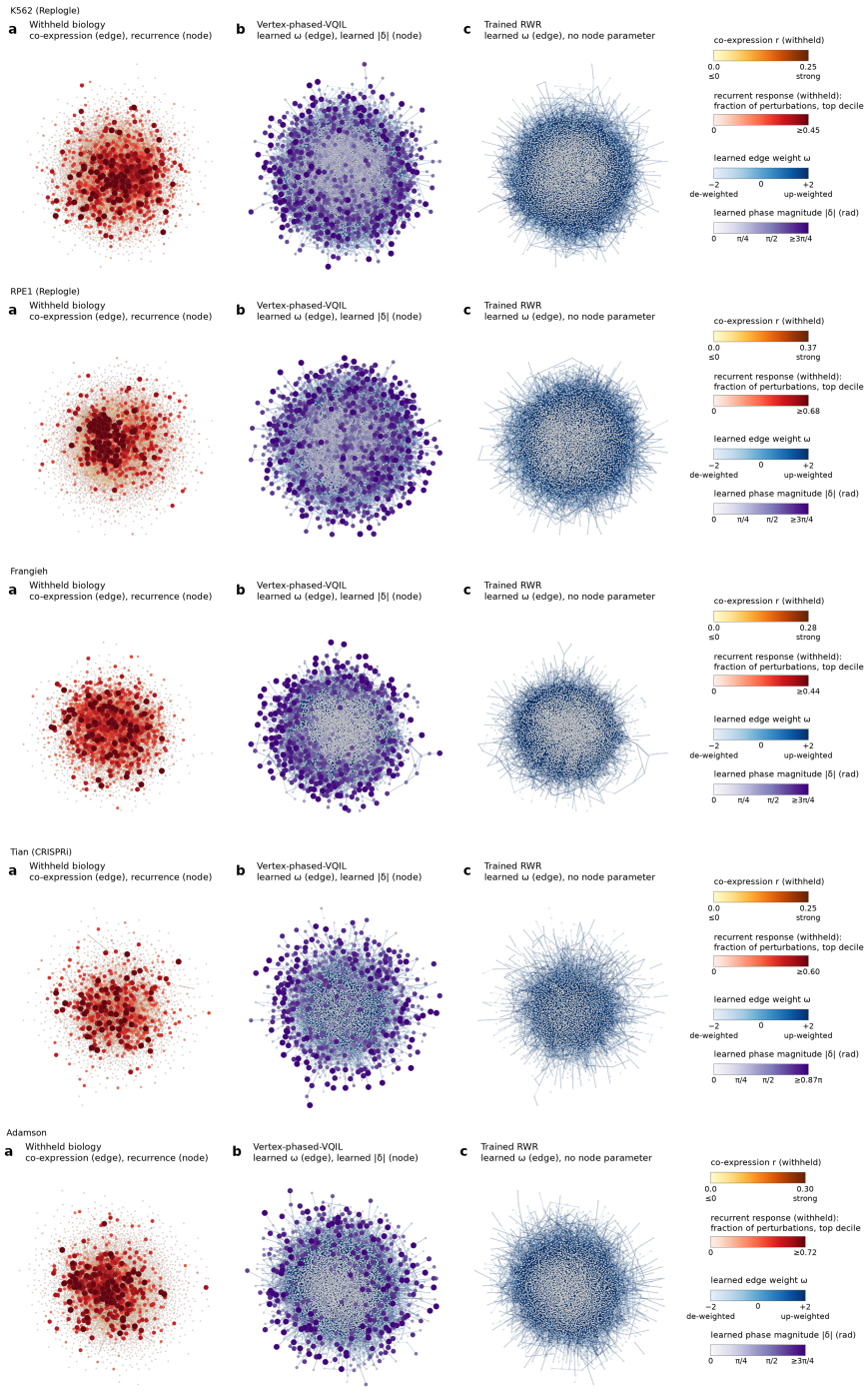


Figure 6: Trained fields on the five PPI networks (K562, RPE1, Frangieh, Tian-i, and Adamson; $\gamma = 0.9$, means over the five folds). (a) Withheld biology: control-cell co-expression of the two endpoint genes (edge color) and recurrent response, the fraction of training perturbations in which a gene is among the top 10% of responders (node color and size). (b) Vertex-phased-VQIL: learned edge weights ω (edge color) and node-phase magnitudes $|\delta|$ (node color and size). (c) Trained-RWR ($r = 0.1$): learned edge weights; RWR has no node parameter.