PREDICTING NEURONAL PROJECTIONS FROM GENE EXPRESSION UNDER INCOMPLETE DETECTION

Anonymous authors
Paper under double-blind review

ABSTRACT

A central challenge in neuroscience is to connect what a neuron is molecularly to where it sends its axons in the brain. Predicting neuronal projections from gene expression could make this connection computationally, extending sparse anatomical tracing to the large cell populations captured by molecular brain atlases. Joint measurements of gene expression and projections make this goal increasingly accessible, but assays miss existing projections and measure different subsets of genes and projection targets. As a result, an assay zero can mean either that a projection is absent or that it was simply missed, so models trained directly on assay calls can learn detectability rather than the underlying probability of projection. We introduce Gene2Wire, a multi-target positive-unlabeled model that separates projection probability from assay detectability while sharing information across incompletely observed targets. A paired training subset estimates detection sensitivity, and a calibrated likelihood uses these estimates to learn projection probabilities from incomplete readouts. Gene2Wire further combines molecular effects shared across projection targets with target-specific effects, allowing related targets to borrow statistical strength without forcing distinct projection patterns to follow the same molecular rule. We show theoretically how errors in detection calibration propagate to projection estimates and how this shared-plus-specific parameterization exploits cross-target structure without restricting the predictor to be low rank. Controlled simulations show that Gene2Wire adapts to target relationships ranging from distinct molecular effects to shared predictive structure. Across five real projection–transcriptomic assay panels, Gene2Wire outperforms previous state-of-the-art models in both projection ranking and probability estimation across varying gene coverage, target coverage, and detection sensitivity.

1 INTRODUCTION

Where a neuron sends its axons determines which distant populations it can influence and is therefore central to its role in a brain circuit. Large-scale molecular atlases now catalog neuronal identities and spatial organization across the brain at single-cell resolution (Shi et al., 2023; Yao et al., 2023). Yet transcriptomic identity does not uniquely specify long-range wiring, and linking cell identity to axonal projections remains a major challenge (Gao et al., 2022; Sorensen et al., 2026). Recent technologies, including BARseq, BARseq2, MERGE-seq, Projection-TAGs, and SPIDER-Seq, have begun to bridge this gap by measuring gene expression together with projection targets in the same cells (Chen et al., 2019; Sun et al., 2021; Xu et al., 2024; Yang et al., 2025; Sun et al., 2026). These measurements make it possible to ask a natural predictive question: given a neuron's gene-expression profile, can we infer its probability of projecting to each candidate target region (Figure 1a)? A reliable predictor could extend projection information from experimentally traced neurons to the much larger populations represented in molecular atlases and help prioritize which projections to measure next. The challenge is that the assays used to train such a predictor provide only an imperfect view of the underlying wiring.

Challenges. Projection readouts depend on the efficiency of labeling and molecular detection, so an existing projection can yield a zero readout (Yang et al., 2025). In the retained Projection-TAGs cohort, the standard readout detects 7,072 positive cell-target pairs, while the union of standard and target-amplified readouts contains 14,020 (Appendix C.4). Nearly half of these union positives are missed by the standard readout. Under the one-sided detection model used here, an observed positive

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.

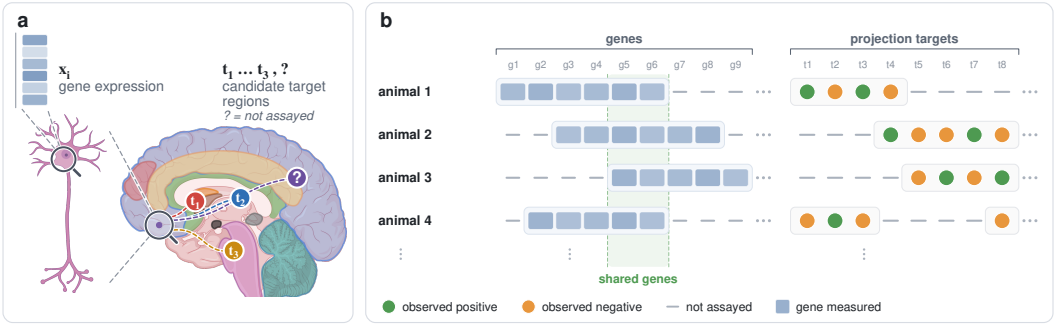

Figure 1: **Predicting neuronal projections from incomplete measurements. (a)** Gene expression informs predictions for known target regions; *?* marks an unassayed target. The evaluated features are specified in Section 3. **(b)** Gene and target panels can vary across animals and assays. Green marks a detected positive; orange marks an assayed non-detection, which can conceal an existing projection; a dash marks an unassayed pair.

implies an existing projection, whereas a non-detection can conceal one. If p is projection probability and e is detection sensitivity, the positive-readout probability is $q = ep$. Recovering projection probabilities from this readout distribution therefore requires information about sensitivity.

Missed positives also reduce the information available for learning molecular effects. This loss of information is compounded by heterogeneous coverage: animals and assays measure different gene and target panels, leaving cells with different inputs and targets with uneven supervision (Figure 1b). Unassayed pairs supply no labels, and assayed zeros can conceal positives. Sharing molecular patterns across related targets can help when individual targets have few detections. Excessive sharing can obscure distinct relationships when targets depend on different genes. A useful predictor must therefore do two things at once: recover projection probabilities from imperfect assay readouts, and adapt how strongly information is shared across targets as supervision becomes sparse and uneven.

Gene2Wire. Gene2Wire addresses these two problems through detection calibration and adaptive sharing across projection targets. A paired training subset estimates detection sensitivity. These estimates enter a calibrated likelihood over assayed neuron–target pairs, while an assay mask excludes pairs that were never measured. The coefficient matrix is decomposed into shared factors and a dense residual. The shared factors pool evidence across related targets; the residual captures target-specific predictive effects. Because the residual spans the full coefficient space, the overall predictor is not restricted to low rank. Factor width controls how many predictive directions are eligible for sharing across targets. Validation selects among direct, shared-only, and joint predictors under the same calibrated detector (Appendix B.7). Gene-observation indicators distinguish unmeasured genes from observed values.

Contributions. We develop a calibrated learning and selection procedure, Gene2Wire, for neuronal projection prediction from gene expression under incomplete measurements. We establish the projection probability targeted by a supplied sensitivity estimate and bound prediction error through readout-fitting and calibration terms (Sec. 3 and Sec. 4). We also derive the exact finite-width penalty induced by shared factors and a dense residual, explaining how strong coefficient directions receive reduced shrinkage while all directions remain representable (Proposition 2 and Appendix A). Controlled simulations test adaptation to changing target relationships, where Gene2Wire achieves the highest mean AUPRC and lowest mean log loss across all 27 sharing-by-measurement settings. Across five real assay panels, it achieves the highest mean AUPRC in 42 of 45 measurement conditions and the lowest mean log loss in 43 of 45 against the available external comparators, spanning reduced gene coverage, target coverage, and detection sensitivity (Sec. 5).

2 Related Work

Molecular prediction of connectivity. Molecular measurements support connectivity prediction at several scales. At the single-cell level, MERGE-seq uses transcriptomic profiles to predict binary projection patterns (Xu et al., 2024), while spatial transcriptomic data have been integrated with projection-labeled single-cell profiles to infer projection identities in molecular atlases (Bhattacher-

jee et al., 2023). SPIDER-Seq predicts projection targets from transcriptomic and spatial features (Sun et al., 2026), and recent work linking single-cell transcriptomes to whole-neuron projectomes shows that gene-expression patterns together with anatomical location can predict long-range projection targets of individual neurons (Sorensen et al., 2026). Bilinear models relate pre- and postsynaptic expression to connectivity between neuronal types (Qiao, 2024). At larger anatomical scales, SPERRFY reconstructs connectivity from expression gradients at projection endpoints (Koike et al., 2026), and the Spatiomolecular Transformer predicts human functional connectivity from regional expression and coordinates (Ratzan et al., 2025). These studies establish that molecular features contain predictive information about neuronal wiring. Gene2Wire focuses on a complementary measurement problem: projection readouts can be incompletely detected, while gene and target panels vary across measurements.

Positive-unlabeled learning and exposure models. Positive-unlabeled (PU) learning treats unlabeled observations as a mixture of positive and negative outcomes (Elkan & Noto, 2008; du Plessis et al., 2014; Kiryo et al., 2017). Selected-at-random PU (SAR-PU) learning allows labeling propensity to depend on observed attributes (Bekker & Davis, 2018; Bekker et al., 2020). PU matrix completion combines one-sided observations with low-rank structure and extends to feature-informed prediction (Hsieh et al., 2015). GenEML jointly models latent probabilities and exposure in multilabel learning (Jain et al., 2017). Gene2Wire uses paired reference positives to estimate sensitivity and an assay mask to specify where readouts are available. The calibrated observation model places all candidate predictors on a common scale for learning and selection.

Shared predictive structure. Multitask learning pools information across related tasks (Caruana, 1997). Shared predictive subspaces have been combined with task-specific coefficients (Ando & Zhang, 2005; Chen et al., 2009), and shared-plus-individual sparse regularization accommodates heterogeneous task relationships (Jalali et al., 2010). Reduced-rank generalized linear models (Yee & Hastie, 2003) and multidimensional item response theory (MIRT) (Reckase, 2006) motivate the shared-only predictor. Gene2Wire places feature-derived shared coordinates and a dense ridge residual within a calibrated PU likelihood, selecting their use through readout validation. Proposition 2 gives the exact finite-width penalty for this parameterization and relates its allocation of effects to spectral thresholding (Mazumder et al., 2010).

3 The Gene2Wire Model

Gene2Wire comprises two components: an observation model linking projection outcomes to assay readouts and a predictor linking encoded molecular features to projection outcomes. Detection calibration estimates the observation process; shared and target-specific coefficients parameterize the predictor.

3.1 Observation Model

For neurons $i \in \{1, \dots, N\}$ and targets $t \in \{1, \dots, T\}$, let $P_{it} \in \{0, 1\}$ indicate whether neuron i projects to target t. Let $M_{it} \in \{0, 1\}$ indicate whether the pair is assayed, and let $D_{it} \in \{0, 1\}$ be its readout when $M_{it} = 1$. The observed data have three states:

$$(M_{it}, D_{it}) = \begin{cases} (1, 1), & \text{detected positive,} \\ (1, 0), & \text{assayed non-detection,} \\ (0, \text{undefined}), & \text{unassayed pair.} \end{cases}$$

We assume one-sided detection: an observed positive implies $P_{it} = 1$. A non-detection can conceal a positive outcome, while an unassayed pair has no readout and contributes no likelihood term. Let Z collect the neuron's encoded features, target identity, and assay information. On assayed pairs, projection probability p_*, detection sensitivity e_*, and positive-readout probability q_* satisfy

$$\begin{aligned} p_*(Z) &= \Pr(P = 1 \mid Z, M = 1), \\ e_*(Z) &= \Pr(D = 1 \mid P = 1, Z, M = 1), \\ q_*(Z) &= \Pr(D = 1 \mid Z, M = 1) = e_*(Z) p_*(Z). \end{aligned} \tag{1}$$

The readout distribution therefore identifies only the product $q_* = e_* p_*$, not e_* and p_* separately.

Estimation of detection sensitivity. Under this model, paired observations of projection status P_{it} and assay readout D_{it} in a calibration subset provide information about detection sensitivity e_*. We estimate sensitivity from paired calibration observations with $P_{it} = 1$ using a regularized logistic detector, allowing sensitivity to vary by target, assay setting, and their interaction. The estimated detection sensitivity of (i, t) is denoted by $\widehat{e}_{it}$. Appendix B.4 specifies the detector covariates and fitting procedure; Appendix A.1 states the support assumptions.

3.2 Gene2Wire: Prediction with Shared and Target-Specific Effects

Gene2Wire estimates a projection probability p_{it} for each neuron–target pair. Its coefficient matrix combines patterns shared across targets with effects specific to individual targets:

$$p_{it} = \sigma(\eta_{it}), \qquad \eta = \mathbf{1}b^\top + XB, \qquad B = LA^\top + R. \tag{2}$$

Here $\sigma(z) = (1 + e^{-z})^{-1}$, $X \in \mathbb{R}^{N\times d}$ is the encoded feature matrix, $b \in \mathbb{R}^T$ contains target intercepts, and $B \in \mathbb{R}^{d\times T}$ is the full coefficient matrix. The columns of $L \in \mathbb{R}^{d\times K}$ describe K shared feature patterns, and $A \in \mathbb{R}^{T\times K}$ specifies each target's weights for these patterns. The residual $R \in \mathbb{R}^{d\times T}$ supplies target-specific coefficients throughout the feature space.

On an assayed pair, the fitted positive-readout probability is $\widehat{e}_{it}p_{it}$. Let $\Omega = \{(i, t) : i \in \text{train}, M_{it} = 1\}$ denote the assayed training pairs. The objective is

$$\mathcal{L}(b, L, A, R) = \frac{1}{|\Omega|}\sum_{(i,t)\in\Omega} \ell(D_{it}, \widehat{e}_{it}\sigma(\eta_{it})) + \frac{\lambda}{2}(\|L\|_F^2 + \|A\|_F^2) + \frac{\mu}{2}\|R\|_F^2, \tag{3}$$

where $\ell(y, q) = -y \log q - (1 - y) \log(1 - q)$ and $\lambda, \mu > 0$. The likelihood fits the available assay readouts through the calibrated observation model. The two penalties λ and μ control shared and target-specific effects.

Choosing the predictive structure. Gene2Wire selects among a direct PU predictor ($LA^\top = 0$), a shared-only PU-MIRT predictor ($R = 0$), and joint predictors with both components. All candidates use the same fitted detector. Validation log loss of $\widehat{q} = \widehat{e}\widehat{p}$ determines the structure and its regularization. Under correct sensitivity, expected readout loss is uniquely minimized by the true projection probabilities p_* (Corollary 1). Appendix B.7 specifies the candidate search and optimization.

Interpreting the fitted probabilities. For a new cell, $\widehat{p}$ estimates its probability of projecting to each target, while $\widehat{q} = \widehat{e}\widehat{p}$ estimates the probability of a positive readout under the specified assay. Given projection probability p and sensitivity e, the probability of a missed projection after a non-detection is

$$h(p, e) := \Pr(P = 1 \mid D = 0, Z, M = 1) = \frac{(1 - e)p}{1 - ep}. \tag{4}$$

At $p = 1/2$, this probability decreases from $4/9$ when $e = 1/5$ to $1/6$ when $e = 4/5$. A non-detection from a more sensitive assay therefore provides stronger evidence against a projection. We use $h(\widehat{p}, \widehat{e})$ to rank candidate missed projections (Lemma 1).

4 Theoretical Analysis of Gene2Wire

The two components address complementary consequences of incomplete detection. Calibration determines which projection probabilities the readout loss recovers. Sharing determines how molecular evidence can be pooled when detection reduces the information available for each target. Section 4.1 studies how sensitivity error affects the fitted probabilities; Section 4.2 characterizes how the penalty allocates shared and target-specific effects.

4.1 How Sensitivity Estimation Error Affects Projection Probabilities

Among assayed pairs, the readouts identify the detection probability $q_* = e_*p_*$, but do not separately identify detection sensitivity e_* and projection probability p_*. For example, $q_* = 0.2$ is consistent with $p_* = 0.25,\ e_* = 0.8$ and with $p_* = 0.5,\ e_* = 0.4$. Paired calibration observations

with $P = 1$ provide information about e_* under the conditions in Appendix A.1; the procedure described in Section 3.1 yields the estimate $\widehat{e}$.

To quantify the effect of sensitivity estimation error, consider a new assayed pair, and let ν denote the distribution of its covariates Z. Conditional on the training data, treat the fitted sensitivity $\widehat{e}$ as fixed. For a proposed projection-probability function p, define its expected readout loss as

$$\mathcal{R}_{\widehat{e}}(p) = \mathbb{E}[\ell(D, \widehat{e}(Z)p(Z)) \mid M = 1]\,.$$

Equivalently, this risk first averages over $D \mid Z, M = 1$ and then over $Z \sim \nu$; $\|\cdot\|_2$ denotes the $L^2(\nu)$ norm. The following result identifies the population minimizer and decomposes projection error into fitting error and sensitivity-estimation error.

Proposition 1 (Calibration error and projection probability)**.** *Suppose p_*, e_*, q_*, and $\widehat{e}$ are measurable functions of Z satisfying $q_* = e_* p_*$, $0 \le e_*, p_* \le 1$, and $\epsilon \le \widehat{e} \le 1$ ν-almost surely for some $\epsilon > 0$. Over all measurable functions $p : Z \mapsto [0, 1]$, the unique minimizer of $\mathcal{R}_{\widehat{e}}$, up to ν-null sets, is*

$$p^\dagger = \min\left\{\frac{q_*}{\widehat{e}}, 1\right\}. \tag{5}$$

Define $\Delta_{\widehat{e}}(p) = \mathcal{R}_{\widehat{e}}(p) - \mathcal{R}_{\widehat{e}}(p^\dagger)$. For every such p,

$$\Delta_{\widehat{e}}(p) \ge \tfrac{1}{2}\mathbb{E}_\nu[\widehat{e}(p - p^\dagger)^2], \tag{6}$$

$$\|p - p_*\|_2 \le \sqrt{2\Delta_{\widehat{e}}(p)/\epsilon} + \left\|\frac{p_*(\widehat{e} - e_*)}{\widehat{e}}\right\|_2. \tag{7}$$

Here ϵ is a uniform lower bound on the fitted sensitivity $\widehat{e}$ over the assayed-pair distribution.

Proof sketch. Bernoulli cross-entropy is minimized by the closest feasible readout probability, $\min(q_*, \widehat{e})$, which gives $p^\dagger$. Appendix A shows that excess readout loss lower-bounds the sensitivity-weighted squared distance between p and $p^\dagger$. The triangle inequality and clipping onto $[0, 1]$ then give the projection-error decomposition. □

With misspecified sensitivity, expected readout loss is minimized at $p^\dagger$, which may differ from p_*. The bound in (7) decomposes projection error into a readout-fitting term and a sensitivity-estimation term. The first term scales as $\epsilon^{-1/2}$, so a smaller sensitivity floor ϵ yields a larger upper bound on $\|p - p^\dagger\|_2$. Equivalently, obtaining the same projection-probability error guarantee at a smaller ϵ requires a smaller excess readout loss. The second is the p_*-weighted relative sensitivity error. Overestimating sensitivity lowers the optimal projection estimate; underestimating it raises the estimate, with saturation at one when $\widehat{e} < q_*$.

Corollary 1 (Correctly calibrated detection)**.** *If $\widehat{e} = e_* > 0$ ν-almost surely, then $p^\dagger = p_*$ and*

$$\mathcal{R}_{e_*}(p) - \mathcal{R}_{e_*}(p_*) \ge \tfrac{1}{2}\mathbb{E}_\nu[e_*(p - p_*)^2]. \tag{8}$$

The risk is uniquely minimized by p_* whenever $e_* > 0$ ν-almost surely; no uniform positive lower bound is required.

4.2 How the Penalty Allocates Shared and Target-Specific Effects

When sensitivity is below one, a non-detection may leave projection status unresolved even when sensitivity is known. The readout then contains less information about projection probability than observing P directly. For a pair with $p = \sigma(\eta) \in (0, 1)$ and fixed sensitivity $e \in (0, 1]$, the Fisher information in its readout D about the projection log-odds η is

$$\mathcal{I}_{\mathrm{det}}(\eta) = \frac{ep(1-p)^2}{1 - ep}, \qquad \frac{\mathcal{I}_{\mathrm{det}}}{\mathcal{I}_P} = \frac{e(1-p)}{1 - ep} \le 1, \tag{9}$$

where $\mathcal{I}_P = p(1 - p)$ is the information from observing P directly. At $p = 1/2$ and $e = 1/5$, the readout supplies $1/9$ as much information about this pair's log-odds η as observing P. This reduction provides one reason to consider sharing information across related targets. Gene2Wire does so through $B = LA^\top + R$, while R retains target-specific coefficients. The next proposition characterizes how the penalty allocates these components for a fixed B and factor width K.

Proposition 2 (Finite-width allocation of shared effects)**.** *Let $B \in \mathbb{R}^{d\times T}$, $\lambda, \mu > 0$, and K be a nonnegative integer. Set $m = \min(d, T)$, $k = \min(K, m)$, and let $\sigma_1 \geq \cdots \geq \sigma_m \geq 0$ be the singular values of B. Write $\|LA^\top\|_*$ for the sum of its singular values, and $(x)_+ = \max\{x, 0\}$. For factors of width K,*

$$\begin{aligned}\Omega_K(B) &:= \inf_{LA^\top + R = B} \left\{ \frac{\lambda}{2}(\|L\|_F^2 + \|A\|_F^2) + \frac{\mu}{2}\|R\|_F^2 \right\} \\ &= \min_{\operatorname{rank}(LA^\top)\leq K} \left\{ \lambda\|LA^\top\|_* + \frac{\mu}{2}\|B - LA^\top\|_F^2 \right\} \\ &= \frac{\mu}{2}\|B\|_F^2 - \frac{\mu}{2}\sum_{j=1}^{k}\left(\sigma_j - \frac{\lambda}{\mu}\right)_+^2 . \end{aligned} \tag{10}$$

A minimizer can be chosen to share left and right singular vectors with B, with singular values $(\sigma_j - \lambda/\mu)_+$ for $j \leq k$ and zero otherwise. Its rank is $\min\{K, \#\{j : \sigma_j > \lambda/\mu\}\}$.

Proof sketch. For a fixed shared matrix $S = LA^\top$, minimizing over its factorizations gives the nuclear-norm penalty $\lambda\|S\|_*$. For fixed singular values of S, aligning its singular directions with those of B minimizes $\|B - S\|_F$. The remaining problem separates across the largest k singular values, each of which is soft-thresholded at λ/μ. Appendix A provides the full proof. □

Because R can represent any B, the factors do not impose a rank constraint on the overall coefficient matrix. Instead, they reduce the penalty on at most K directions whose singular values exceed λ/μ. Thus K limits how many directions receive the shared penalty, while λ/μ sets the threshold; all remaining coefficient structure is carried by R.

The three candidate families differ in how they represent target effects: direct PU uses target-specific coefficients, shared-only PU-MIRT uses shared factors, and the joint predictor allows both. Section 5 evaluates their predictive behavior under missing genes, missing targets, and missed positives.

5 Experiments: Adaptive Sharing under Measurement Loss

Table 1: **Evaluation settings.** G is the number of source genes and $d = 2G + 2$ the encoded input dimension. Three outer folds preserve the listed groups. MERGE-seq includes label-free analysis rows. In BARseq, A1 and M1 denote primary auditory and primary motor cortex.

Setting	Cells	Targets	G	d	Held-out grouping
Projection-TAGs	37,827	7	32	66	Animal
SPIDER-seq spatial	45,835	15	32	66	AP slices
MERGE-seq	7,147	5	128	258	Sample
BARseq A1	593	11	23	48	Within-animal depth
BARseq M1	749	35	23	48	Within-animal depth
Simulation	400	36	16	34	Spatial slices

Settings and measurement design. Each gene is standardized using its measured values in the training split. Unmeasured entries are filled with zero, and one observation indicator per gene distinguishes missing entries from measured values at the training mean. The benchmark constructs three virtual assays with overlapping panels and adds two assay-membership indicators, using the first assay as the reference level. A source pool of G genes therefore gives $d = 2G + 2$ features. Gene indices are shared across panels, so overlapping measurements contribute to one coefficient matrix (Appendix B.3).

We use controlled measurement loss to evaluate projection prediction when molecular inputs, target labels, or positive detections are incomplete. Each real assay panel is fitted separately (Table 1).

The reference is the standard-or-amplified union for Projection-TAGs and the original assay calls for the other panels. Artificial thinning creates incomplete observations of these references, allowing the consequences of each measurement constraint to be assessed on held-out cells.

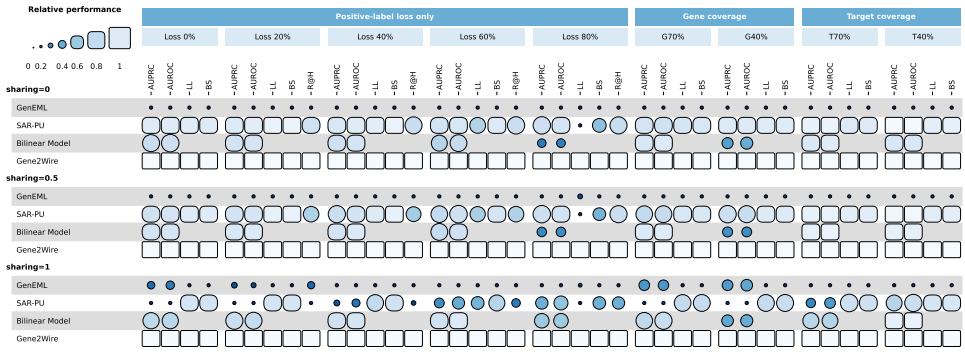

Figure 2: **Simulation benchmark across target structure and measurement conditions.** Each block compares GenEML, SAR-PU, Bilinear, and Gene2Wire at a fixed sharing strength. Larger glyphs indicate better performance within each condition and metric after min–max scaling; loss metrics are reversed. Blank cells denote unavailable or inapplicable scores. Bilinear supplies detection scores and has no reference log loss (LL) or Brier score (BS). Appendix B.10 specifies metric support, and Appendix C.3 gives absolute values.

Each dataset is divided into three virtual assays with overlapping panels that collectively cover the source gene and target sets. Gene and target coverage each take values in $\{1, .7, .4\}$, and positive retention takes values in $\{1, .8, .6, .4, .2\}$. The main comparisons use nine single-factor conditions: a common full-measurement control, four levels of positive-label loss, two reduced gene panels, and two reduced target panels. Gene removal limits the molecular information available for each prediction. Target removal limits supervision while preserving the known target set. Positive thinning introduces ambiguous non-detections. Appendix C.2 gives an example on the BARseq dataset of the full 45-condition grid and combined-factorial summaries.

Positive retention varies across assay–target pairs around the requested average. An outcome-independent 20% subset of development cells supplies paired references for sensitivity estimation. During selection, only its inner-training cells enter calibration; at full retention, sensitivity is known to be one. All Gene2Wire candidates use the same detector. Development refitting updates the selected predictor and, when estimated, the detector before outer-test evaluation. Appendix B specifies the splits and thinning mechanism.

Comparisons and evaluation. Direct PU, shared-only PU-MIRT, and Gene2Wire use the same detection sensitivities. GenEML (Jain et al., 2017) and SAR-PU (Bekker et al., 2020) infer exposure from the observed detections. Bilinear (Qiao, 2024) uses a target-identity adaptation of the published bilinear predictor. All methods receive the same physical panels, folds, and thinning draws; Appendix B.9 specifies their calibration access and tuning.

Area under the precision–recall curve (AUPRC) evaluates projection ranking; area under the ROC curve (AUROC) is reported as a secondary ranking metric. Log loss (LL) and Brier score (BS) evaluate the assigned probabilities against the benchmark outcomes. Recall at H (R@H) measures retrieval of artificially hidden positives: the fraction recovered among the top H ranked candidates on the fitted target panel, with H set equal to the number of positives hidden. Because H grows with the fraction hidden, R@H compares methods within a loss level rather than across levels, and it is defined only for the 20–80% positive-label-loss conditions. Outer-test references are used only for evaluation on a fixed native assay mask, including entries withheld by artificial target panels. Simulation repetitions generate new datasets; real-data repetitions vary panel configurations within a fixed cohort. Appendix B.10 details aggregation and metric support.

5.1 Simulation benchmarking

Generating process. Simulation lets us vary target relationships independently of measurement loss. Shared signal strength $\rho \in \{0, .5, 1\}$ ranges from target-specific to fully shared molecular

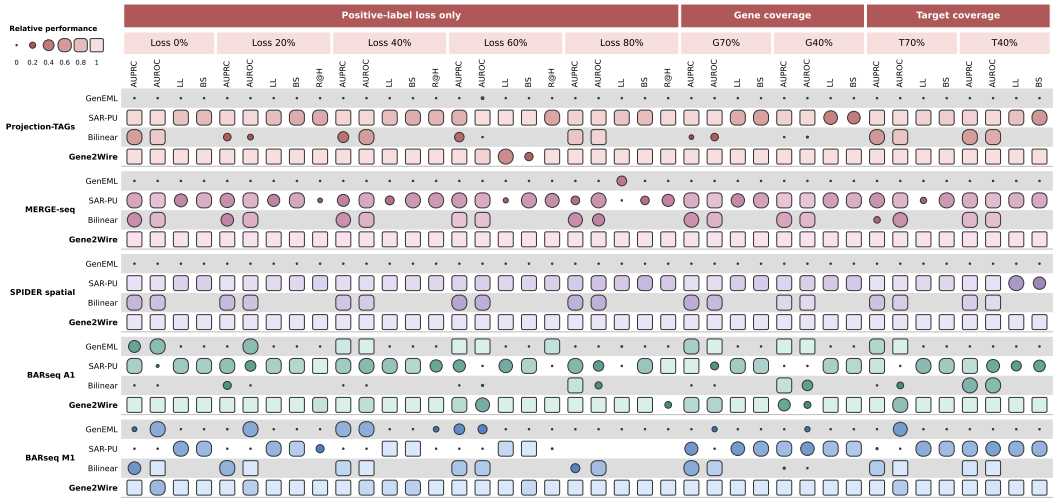

Figure 3: **Prediction across five real assay panels and three measurement axes.** Columns vary positive-label loss, gene coverage, or target coverage; Loss 0% is the common full-measurement control. The plotting style is the same as Figure 2. Dataset hue identifies the five panels. Appendix C.3 reports absolute values.

effects, with total signal scale held fixed. Each repetition draws 400 cells in 20 spatial slices, 36 targets and 16 genes. The shared signal has rank two. For each target, the shared and target-specific signals have unit variance and are orthogonal, so mixing them with weights $\sqrt{\rho}$ and $\sqrt{1-\rho}$ holds total signal variance fixed. Target intercepts set prevalence evenly across 0.1–0.3, and projections are Bernoulli draws from the resulting probabilities. Detections are censored by the same assay–target-dependent rule as the real panels. Base draws are reused across ρ to pair the comparisons, and generating probabilities are withheld from the fitted models (Appendix B.6).

As shown in Figure 2, Gene2Wire achieves the highest mean AUPRC in all 27 sharing-by-condition settings against GenEML, SAR-PU, and Bilinear, and the lowest mean log loss (LL) in all 27 settings against GenEML and SAR-PU.

5.2 Five Real Assay Panels: Breadth across Assays and Measurement Axes

The real panels test whether the same learning procedure remains effective across different molecular measurements and projection targets. Gene2Wire has the highest mean AUPRC in 42 of 45 conditions and the lowest mean LL in 43 of 45 against the available external comparators (Figure 3). At 80% positive-label loss, it leads both metrics on every panel. In MERGE-seq, for example, AUPRC is 0.3943 versus 0.3494 for Bilinear, while LL is 0.4016 versus 0.6465 for GenEML, the strongest external comparator for each metric in this condition.

Severe detection loss reveals complementary gains in ranking and probability quality. In BARseq M1, Gene2Wire's AUPRC margin over the strongest available external comparator grows from 0.011 at 40% loss to 0.026 at 60% and 0.069 at 80%. At the last setting, AUPRC is 0.335 versus 0.267, a relative increase of 25.8%. At 80% loss, BARseq A1 shows a different pattern: its AUPRC of 0.2726 is close to Bilinear's 0.2706, while its LL of 0.474 compares with 0.738 for GenEML, the strongest probability comparator. The ranking and probability comparisons reveal complementary strengths relevant to prioritizing projections and interpreting prediction confidence.

Gene and target removal identify different information constraints. In M1, Gene2Wire's AUPRC changes from 0.366 at full coverage to 0.265 at 40% gene coverage and 0.354 at 40% target coverage, with full positive retention throughout. The larger decrease under gene removal shows the importance of retaining informative molecular inputs in this panel. The fixed native test mask makes this a comparison of reduced features and reduced supervision on the same evaluation targets. Appendix C.3 provides all means and unavailable entries.

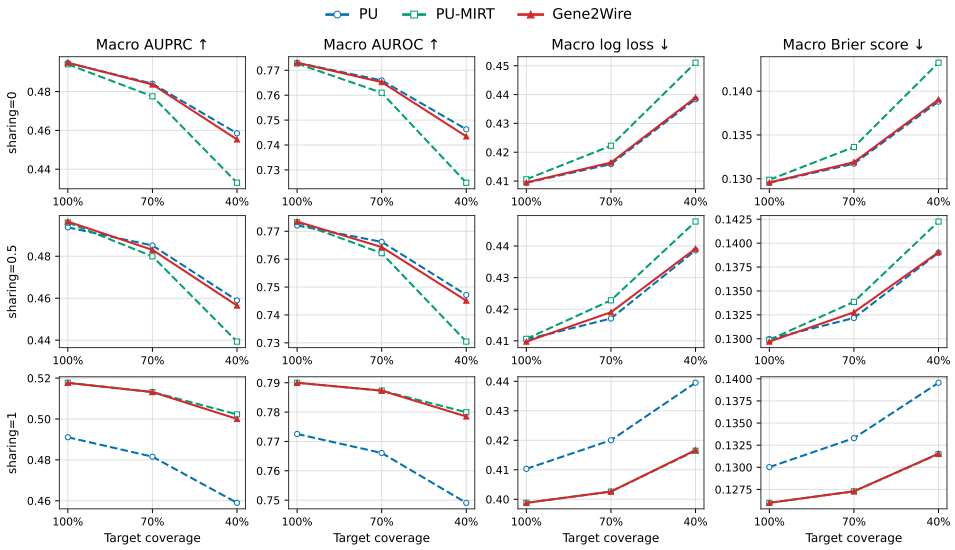

Figure 4: **Ablation study for PU, PU-MIRT and Gene2Wire.** Direct PU, shared-only PU-MIRT, and Gene2Wire use a common detector as target coverage decreases, with full gene coverage and positive retention. Rows set the generating shared-signal strength $\rho = 0, .5, 1$; columns report ranking and probability metrics. Gene2Wire selects among direct, shared-only, and joint candidates.

5.3 Ablation Study: Comparing Gene2Wire with PU and PU-MIRT

The family ablation examines whether adapting the coefficient structure is useful when the generating target relationships change. Figure 4 compares direct PU ($LA^\top = 0$), shared-only PU-MIRT ($R = 0$), and Gene2Wire as target coverage decreases, using the same detection calibration for all three. PU-MIRT's search includes full factor width, so its behavior reflects the selected parameterization, regularization, and width.

With target-specific generating effects ($\rho = 0$), Gene2Wire follows the stronger direct PU predictor as target coverage falls, while shared-only PU-MIRT loses more accuracy. With fully shared effects ($\rho = 1$), the ordering reverses and Gene2Wire follows the stronger shared-only predictor. At $\rho = .5$, the leading endpoint depends on the metric and coverage level. Validation lets the full procedure respond to changes in target relationships and supervision. Proposition 2 explains the joint candidate's regularization preference; these comparisons evaluate the selected predictor across all three families.

6 Discussion and Conclusion

Predicting a projection requires accounting for both molecular evidence and assay sensitivity. Gene2Wire combines a calibrated readout likelihood with shared and target-specific effects. The theory links projection error to sensitivity estimation and shows that sharing changes regularization without restricting the full coefficient space. Simulations and five real assay panels support the procedure under incomplete measurements.

Limitations and future work. Gene2Wire addresses estimation under incomplete detection, but identifiability still rests on calibration: the assay reveals $q = ep$, not p alone. Our references are assay-defined, and each panel is fitted separately, so the present results establish robustness within measurement settings rather than transfer across them. A natural next step is to treat multiple imperfect assays as complementary views of a common projectome and evaluate whether projection probabilities transfer across assay settings.

Reproducibility Statement

Appendix A provides complete proofs. Appendix B specifies datasets, preprocessing, evaluation splits, calibration, optimization, model selection, and comparator implementations. Appendix C reports the numerical results underlying the main comparisons.

Statement of AI Use

OpenAI Codex and Claude assisted with writing and editing, literature retrieval and related-work discovery, research ideation and execution, drafting portions of the manuscript, developing code for synthetic simulations, and checking mathematical claims and proofs. The authors reviewed and verified all AI-assisted content and take responsibility for the final submission.

References

Rie Kubota Ando and Tong Zhang. A framework for learning predictive structures from multiple tasks and unlabeled data. *Journal of Machine Learning Research*, 6(61):1817–1853, 2005. URL https://jmlr.org/papers/v6/ando05a.html.

Jessa Bekker and Jesse Davis. Learning from positive and unlabeled data under the selected at random assumption. In *Proceedings of the Second International Workshop on Learning with Imbalanced Domains: Theory and Applications*, volume 94 of *Proceedings of Machine Learning Research*, pp. 8–22. PMLR, 2018. URL https://proceedings.mlr.press/v94/bekker18a.html.

Jessa Bekker, Pieter Robberechts, and Jesse Davis. Beyond the selected completely at random assumption for learning from positive and unlabeled data. In *Machine Learning and Knowledge Discovery in Databases: ECML PKDD 2019*, volume 11907 of *Lecture Notes in Computer Science*, pp. 71–85. Springer, 2020. doi: 10.1007/978-3-030-46147-8_5.

Aritra Bhattacherjee, Chao Zhang, Brianna R. Watson, Mohamed Nadhir Djekidel, Jeffrey R. Moffitt, and Yi Zhang. Spatial transcriptomics reveals the distinct organization of mouse prefrontal cortex and neuronal subtypes regulating chronic pain. *Nature Neuroscience*, 26(11):1880–1893, 2023. doi: 10.1038/s41593-023-01455-9.

Rich Caruana. Multitask learning. *Machine Learning*, 28(1):41–75, 1997. doi: 10.1023/A:1007379606734.

Jianhui Chen, Lei Tang, Jun Liu, and Jieping Ye. A convex formulation for learning shared structures from multiple tasks. In *Proceedings of the 26th International Conference on Machine Learning*, pp. 137–144, 2009. doi: 10.1145/1553374.1553392.

Xiaoyin Chen, Yu-Chi Sun, Huiqing Zhan, Justus M Kebschull, Stephan Fischer, Katherine Matho, Z Josh Huang, Jesse Gillis, and Anthony M Zador. High-throughput mapping of long-range neuronal projection using in situ sequencing. *Cell*, 179(3):772–786.e19, 2019. doi: 10.1016/j.cell.2019.09.023.

Marthinus C du Plessis, Gang Niu, and Masashi Sugiyama. Analysis of learning from positive and unlabeled data. In *Advances in Neural Information Processing Systems*, volume 27, 2014.

Charles Elkan and Keith Noto. Learning classifiers from only positive and unlabeled data. In *Proceedings of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining*, pp. 213–220, 2008. doi: 10.1145/1401890.1401920.

Le Gao, Sang Liu, Lingfeng Gou, et al. Single-neuron projectome of mouse prefrontal cortex. *Nature Neuroscience*, 25:515–529, 2022. doi: 10.1038/s41593-022-01041-5.

Tilmann Gneiting and Adrian E. Raftery. Strictly proper scoring rules, prediction, and estimation. *Journal of the American Statistical Association*, 102(477):359–378, 2007. doi: 10.1198/016214506000001437.

Cho-Jui Hsieh, Nagarajan Natarajan, and Inderjit Dhillon. PU learning for matrix completion. In *Proceedings of the 32nd International Conference on Machine Learning*, volume 37 of *Proceedings of Machine Learning Research*, pp. 2445–2453. PMLR, 2015. URL https://proceedings.mlr.press/v37/hsiehb15.html.

Vikas Jain, Nirbhay Modhe, and Piyush Rai. Scalable generative models for multi-label learning with missing labels. In *Proceedings of the 34th International Conference on Machine Learning*, volume 70 of *Proceedings of Machine Learning Research*, pp. 1636–1644. PMLR, 2017. URL https://proceedings.mlr.press/v70/jain17a.html.

Ali Jalali, Pradeep Ravikumar, Sujay Sanghavi, and Chao Ruan. A dirty model for multi-task learning. In *Advances in Neural Information Processing Systems*, volume 23, 2010.

Ryuichi Kiryo, Gang Niu, Marthinus C du Plessis, and Masashi Sugiyama. Positive-unlabeled learning with non-negative risk estimator. In *Advances in Neural Information Processing Systems*, volume 30, 2017.

Jigen Koike, Ken Nakae, Riichiro Hira, Yuichiro Yada, and Honda Naoki. A data-driven framework linking the connectome to spatial gene expression gradients inspired by chemoaffinity theory. *Proceedings of the National Academy of Sciences*, 123(10):e2516572123, 2026. doi: 10.1073/pnas.2516572123.

Rahul Mazumder, Trevor Hastie, and Robert Tibshirani. Spectral regularization algorithms for learning large incomplete matrices. *Journal of Machine Learning Research*, 11(80):2287–2322, 2010. URL https://jmlr.org/papers/v11/mazumder10a.html.

Mu Qiao. Deciphering the genetic code of neuronal type connectivity through bilinear modeling. *eLife*, 12:RP91532, 2024. doi: 10.7554/eLife.91532.

Alexander Ratzan, Sidharth Goel, Junhao Wen, Christos Davatzikos, and Erdem Varol. Predicting functional brain connectivity with context-aware deep neural networks. In *Advances in Neural Information Processing Systems*, volume 38, pp. 87693–87728, 2025. doi: 10.52202/085713-2933.

Mark D Reckase. Multidimensional item response theory. In *Handbook of Statistics*, volume 26, pp. 607–642. Elsevier, 2006. doi: 10.1016/S0169-7161(06)26018-8.

Hailing Shi, Yichun He, Yiming Zhou, Jiahao Huang, Kamal Maher, Brandon Wang, Zefang Tang, Shuchen Luo, Peng Tan, Morgan Wu, Zuwan Lin, Jingyi Ren, Yaman Thapa, Xin Tang, Ken Y. Chan, Benjamin E. Deverman, Hao Shen, Albert Liu, Jia Liu, and Xiao Wang. Spatial atlas of the mouse central nervous system at molecular resolution. *Nature*, 622:552–561, 2023. doi: 10.1038/s41586-023-06569-5.

Staci A. Sorensen, Nathan W. Gouwens, Yun Wang, et al. Connecting single-cell transcriptomes to projectomes in the mouse visual cortex. *Nature*, 2026. doi: 10.1038/s41586-026-10424-8.

Leqiang Sun, Hu Zheng, Yayu Huang, Xuehuan Huang, Keji Yan, Zhongchao Wang, Liyao Yang, Yiping Yue, Xiaojuan Gou, Guohua Du, Yang Wang, Xiaofeng Wu, Huazhen Liu, Hang Chen, Daqing Ma, Yunyun Han, Jinxia Dai, and Gang Cao. Organization of mouse prefrontal cortex subnetwork revealed by spatial single-cell multi-omic analysis of SPIDER-Seq. *National Science Review*, 13(5):nwag004, 2026. doi: 10.1093/nsr/nwag004.

Yu-Chi Sun, Xiaoyin Chen, Stephan Fischer, Shaina Lu, Huiqing Zhan, Jesse Gillis, and Anthony M Zador. Integrating barcoded neuroanatomy with spatial transcriptional profiling enables identification of gene correlates of projections. *Nature Neuroscience*, 24(6):873–885, 2021. doi: 10.1038/s41593-021-00842-4.

Wei Wang, Dong-Dong Wu, Ming Li, Jingxiong Zhang, Gang Niu, and Masashi Sugiyama. Accessible, realistic, and fair evaluation of positive-unlabeled learning algorithms. In *International Conference on Learning Representations*, 2026. URL https://proceedings.iclr.cc/paper_files/paper/2026/hash/cbffadf96101c42be845d3e275c4b7e9-Abstract-Conference.html.

Peibo Xu, Jian Peng, Tingli Yuan, Zhaoqin Chen, Hui He, Ziyan Wu, Ting Li, Xiaodong Li, Luyue Wang, Le Gao, Jun Yan, Wu Wei, Chengyu T Li, Zhen-Ge Luo, and Yuejun Chen. High-throughput mapping of single-neuron projection and molecular features by retrograde barcoded labeling. *eLife*, 13:e85419, 2024. doi: 10.7554/eLife.85419.

Lite Yang, Fang Liu, Hannah Hahm, Takao Okuda, Xiaoyue Li, Yufen Zhang, Vani Kalyanaraman, Monique R Heitmeier, and Vijay K Samineni. Projection-TAGs enable multiplex projection tracing and multi-modal profiling of projection neurons. *Nature Communications*, 16(1):5557, 2025. doi: 10.1038/s41467-025-60360-w.

Zizhen Yao, Cindy T. J. van Velthoven, Michael Kunst, et al. A high-resolution transcriptomic and spatial atlas of cell types in the whole mouse brain. *Nature*, 624:317–332, 2023. doi: 10.1038/s41586-023-06812-z.

Thomas W Yee and Trevor J Hastie. Reduced-rank vector generalized linear models. *Statistical Modelling*, 3(1):15–41, 2003. doi: 10.1191/1471082X03st045oa.

APPENDIX GUIDE

The appendices follow the argument from statistical target to fitted estimator to experimental evidence. Appendix A proves the calibration and spectral-allocation results. Appendix B specifies the data construction, fitting procedure, and evaluation. Appendix C gives the full numerical comparisons and additional figures. The notation and workflow below connect these three parts.

Notation. There are N cells, T projection targets, and G source genes. The feature dimension is $d = 2G + 2$ in the evaluated design. The projection outcome P, assay readout D, and observation mask M have different roles: P is the prediction target, D supplies the response used for fitting, and M determines whether that response is available.

Notation	Meaning
P_{it}, D_{it}	Binary projection outcome and assay readout for cell i and target t. The readout is defined only on assayed entries, where $D_{it} \in \{0, 1\}$ and $D_{it} \leq P_{it}$. Section 5 defines the dataset-specific benchmark outcome used in evaluation.
$M_{it}^{\text{native}}, M_{it}$	Native assay availability and the fitting mask after applying the target panel for the current measurement condition. An entry with $M_{it} = 0$ supplies no likelihood term.
O_{ig}, x_i, X	Gene-observation indicator, encoded feature vector in $\mathbb{R}^d$, and design matrix in $\mathbb{R}^{N \times d}$. Missing standardized values are zero; their indicators remain in x_i.
p_{it}, e_{it}, q_{it}	Projection probability, sensitivity conditional on $P_{it} = 1$, and positive-readout probability, with $q_{it} = e_{it} p_{it}$.
η, b, B	Logits $\eta = \mathbf{1}b^\top + XB$, target intercepts $b \in \mathbb{R}^T$, and coefficients $B \in \mathbb{R}^{d \times T}$; $p = \sigma(\eta)$ entrywise.
L, A, R, K	Shared factors $L \in \mathbb{R}^{d \times K}$ and $A \in \mathbb{R}^{T \times K}$, residual coefficients $R \in \mathbb{R}^{d \times T}$, and factor width K, with $B = LA^\top + R$.
$h(p, e)$	For an assayed non-detection, $h(p, e) = (1-e)p/(1-ep)$ is the posterior probability of a missed projection, where p is the projection probability before observing the non-detection.

From training data to held-out predictions. Each outer fold uses the following sequence. The split defines an inner-training set, a validation set, and an outer-test set; training and validation together form the development set.

1. **Construct features and calibrate detection.** Fit preprocessing on visible inner-training gene values. Select the stratified, outcome-blind 20% paired subset of development cells, and use its inner-training intersection to estimate $\widehat{e}$ from paired calibration entries with $P_{it} = 1$. At full positive retention, use the known sensitivity $e = 1$.
2. **Select the predictor.** Hold this detector fixed while fitting direct PU, shared-only PU-MIRT, and joint candidates to measured inner-training readouts. Select the converged candidate with the best measured-entry validation log loss of $\widehat{q} = \widehat{e}\widehat{p}$, following the search and tie rules in Appendix B.7.
3. **Refit on development data.** Refit preprocessing on visible development values and, when sensitivity is estimated, refit the detector on paired development cells. Refit the selected predictor configuration to the development readouts using this detector.
4. **Evaluate on the outer test set.** Score $\widehat{p}$ against the dataset-specific benchmark outcomes on natively assayed test entries, including entries withheld by the artificial target panel. For missed-positive retrieval, rank assayed test non-detections within the fitting panel by $h(\widehat{p}, \widehat{e})$. Appendix B.10 specifies each metric and its eligible support.

Outer-test benchmark outcomes are reserved for scoring the fitted estimator. This separation preserves a common prediction target across the measurement conditions while allowing each condition to change the information available for fitting.

A Theory: Assumptions and Proofs

A.1 Observation model and calibration assumptions

All population statements concern an assayed pair, so probabilities condition on $M = 1$. The covariates Z include the target index, the panel-aware predictor features, and any additional detector covariates. Under one-sided detection, $D = 1$ implies $P = 1$, and therefore

$$\Pr(D = 1 \mid Z, M = 1) = \Pr(D = 1 \mid P = 1, Z, M = 1) \Pr(P = 1 \mid Z, M = 1) = e_*(Z)p_*(Z).$$

Readouts alone identify q_*, whereas separating p_* from e_* requires information about detection sensitivity.

Let C indicate inclusion in the paired calibration sample. Estimating sensitivity on the relevant $P = 1$ covariate support requires two conditions:

$$\Pr(C = 1 \mid P = 1, Z, M = 1) > 0, \qquad \Pr(D = 1 \mid P = 1, Z, M = 1, C = 1) = e_*(Z) \quad (11)$$

The first condition provides calibration coverage. The second requires the conditional detection mechanism in the paired calibration sample to be the same as the mechanism in the prediction population. Correct estimation further requires an appropriate detector specification. Proposition 1 allows the supplied estimate $\widehat{e}$ to differ from e_*. When $e_* = 0$, the readout distribution contains no information about p_* at the corresponding covariate values.

Conditional on the training data, $\widehat{e}$ and a proposed predictor p are treated as fixed functions and evaluated on a fresh assayed pair. As in Section 4.1, ν denotes the distribution of Z among such pairs, and $\|f\|_2^2 = \mathbb{E}_\nu[f(Z)^2]$. Uniqueness up to ν-null sets permits differences only at covariate values with zero probability under ν.

For $a, r \in [0, 1]$, define Bernoulli entropy and cross-entropy by

$$H(a) = -a \log a - (1 - a) \log(1 - a), \qquad \mathrm{CE}(a, r) = -a \log r - (1 - a) \log(1 - r).$$

We use natural logarithms, $0 \log 0 = 0$, and the extended-real convention that assigning probability zero to an event with positive probability gives infinite loss. Also, $0 \le H(a) \le \log 2$. Here $\mathrm{Ber}(a)$ is a Bernoulli distribution with success probability a, and KL is Kullback–Leibler divergence. Consequently,

$$\mathrm{CE}(a, r) - H(a) = \mathrm{KL}(\mathrm{Ber}(a) \| \mathrm{Ber}(r)) \,. \quad (12)$$

A.2 Proof of Proposition 1

Proof. We establish the conditional optimum, the excess-risk inequality, and the sensitivity-error bound in turn.

Conditional optimum and finite optimal risk. All quantities in the pointwise calculations below are evaluated at the same covariate value; their dependence on Z is suppressed. The conditional expected readout loss of a candidate p is $\mathrm{CE}(q_*, \widehat{e}p)$. For $0 < p < 1$, and hence $0 < \widehat{e}p < 1$,

$$\frac{\partial}{\partial p} \mathrm{CE}(q_*, \widehat{e}p) = \frac{\widehat{e}p - q_*}{p(1 - \widehat{e}p)}.$$

The sign of this derivative, together with the boundary cases, shows that the unique minimizer over $p \in [0, 1]$ satisfies

$$\widehat{e}p^\dagger = \min\{q_*, \widehat{e}\}, \qquad p^\dagger = \min\left\{\frac{q_*}{\widehat{e}}, 1\right\}.$$

The optimal conditional loss is $H(q_*) \le \log 2$ when $q_* \le \widehat{e}$. When $q_* > \widehat{e}$, it is $\mathrm{CE}(q_*, \widehat{e})$, which is at most $\max\{\log 2, -\log \epsilon\}$: as a function of $q_* \in [\widehat{e}, 1]$, this cross-entropy is affine, and its endpoint values are $H(\widehat{e})$ and $-\log \widehat{e}$. Thus $\mathcal{R}_{\widehat{e}}(p^\dagger) < \infty$, so the excess risk is well defined.

Excess loss and KL divergence. If $q_* \le \widehat{e}$, then $\widehat{e}p^\dagger = q_*$, and the conditional excess loss equals

$$\mathrm{CE}(q_*, \widehat{e}p) - \mathrm{CE}(q_*, \widehat{e}p^\dagger) = \mathrm{KL}(\mathrm{Ber}(\widehat{e}p^\dagger) \| \,\mathrm{Ber}(\widehat{e}p)). \quad (13)$$

If $q_* > \widehat{e}$, then $\widehat{e}p^\dagger = \widehat{e} < 1$. For $0 < p \le 1$, direct expansion gives

$$\mathrm{CE}(q_*, \widehat{e}p) - \mathrm{CE}(q_*, \widehat{e}) - \mathrm{KL}(\mathrm{Ber}(\widehat{e})\|\mathrm{Ber}(\widehat{e}p)) = (q_* - \widehat{e}) \log \frac{\widehat{e}(1-\widehat{e}p)}{\widehat{e}p(1-\widehat{e})} \ge 0. \tag{14}$$

The final inequality follows from $q_* > \widehat{e}$ and $p \le 1$. At $p = 0$, both the conditional excess loss and the KL divergence are infinite, so their inequality still holds.

By the definition of the expected readout risk and the law of iterated expectation,

$$\begin{aligned}
\mathcal{R}_{\widehat{e}}(p) &= \mathbb{E}[\ell(D, \widehat{e}(Z)p(Z)) \mid M = 1] \\
&= \mathbb{E}_\nu[\mathbb{E}[\ell(D, \widehat{e}(Z)p(Z)) \mid Z, M = 1]] \\
&= \mathbb{E}_\nu[q_*(Z)\ell(1, \widehat{e}(Z)p(Z)) + (1 - q_*(Z))\ell(0, \widehat{e}(Z)p(Z))] \\
&= \mathbb{E}_\nu\big[\mathrm{CE}\big(q_*(Z), \widehat{e}(Z)p(Z)\big)\big] .
\end{aligned}$$

The third equality uses $D \mid Z, M = 1 \sim \mathrm{Ber}(q_*(Z))$. The same calculation applies to $p^\dagger$. Suppressing the arguments Z from this point onward, and using the finite optimal risk established above, we obtain

$$\begin{aligned}
\Delta_{\widehat{e}}(p) &= \mathcal{R}_{\widehat{e}}(p) - \mathcal{R}_{\widehat{e}}(p^\dagger) \\
&= \mathbb{E}_\nu\big[\mathrm{CE}(q_*, \widehat{e}p) - \mathrm{CE}(q_*, \widehat{e}p^\dagger)\big] \\
&\ge \mathbb{E}_\nu \, \mathrm{KL}\big(\mathrm{Ber}(\widehat{e}p^\dagger)\big\|\mathrm{Ber}(\widehat{e}p)\big) .
\end{aligned} \tag{15}$$

The final inequality applies (13) when $q_* \le \widehat{e}$ and (14) when $q_* > \widehat{e}$, followed by averaging over ν.

Curvature in projection probability. For fixed $0 < \widehat{e} \le 1$ and $p \in (0, 1)$, define

$$F(u) = \mathrm{KL}(\mathrm{Ber}(\widehat{e}u)\|\mathrm{Ber}(\widehat{e}p)) , \qquad 0 < u < 1.$$

Then $F(p) = F'(p) = 0$, and

$$F''(u) = \frac{\widehat{e}}{u(1 - \widehat{e}u)} \ge \widehat{e}.$$

Taylor's theorem applied around $u = p$ therefore gives

$$F(u) = \mathrm{KL}(\mathrm{Ber}(\widehat{e}u)\|\mathrm{Ber}(\widehat{e}p)) \ge \frac{\widehat{e}}{2}(u - p)^2. \tag{16}$$

The inequality extends to boundary values by taking limits whenever the KL divergence is finite. If $p = 0 < u$, or if $\widehat{e} = p = 1$ and $u < 1$, the KL divergence is $+\infty$, so the bound holds under the extended-real convention. Applying the inequality with $u = p^\dagger$ to (15) yields

$$\Delta_{\widehat{e}}(p) \ge \mathbb{E}_\nu \, \mathrm{KL}\big(\mathrm{Ber}(\widehat{e}p^\dagger)\big\|\mathrm{Ber}(\widehat{e}p)\big) \ge \frac{1}{2}\mathbb{E}_\nu[\widehat{e}(p - p^\dagger)^2] \ge \frac{\epsilon}{2}\|p - p^\dagger\|_2^2. \tag{17}$$

Calibration bias. Finally, because $p_* \in [0, 1]$ and $q_* = e_* p_*$, clipping $q_*/\widehat{e}$ at one cannot increase its distance from p_*. Hence

$$|p^\dagger - p_*| \le \left|\frac{q_*}{\widehat{e}} - p_*\right| = \frac{p_*|\widehat{e} - e_*|}{\widehat{e}}.$$

Combining this inequality with $\|p - p_*\|_2 \le \|p - p^\dagger\|_2 + \|p^\dagger - p_*\|_2$ proves (7). □

A.3 Proof of Corollary 1

Proof. If $\widehat{e} = e_* > 0$ ν-almost surely, then $q_*/e_* = p_*$, so the pointwise minimizer is $p^\dagger = p_*$. Thus the conditional expected loss of a candidate p is $\mathrm{CE}(q_*, e_*p)$, whereas at p_* it is $\mathrm{CE}(q_*, q_*) = H(q_*)$. By the cross-entropy–KL identity in (12), we have

$$\mathcal{R}_{e_*}(p) - \mathcal{R}_{e_*}(p_*) = \mathbb{E}_\nu \, \mathrm{KL}(\mathrm{Ber}(e_*p_*)\|\mathrm{Ber}(e_*p)) .$$

The quadratic KL inequality established in (16) gives

$$\mathcal{R}_{e_*}(p) - \mathcal{R}_{e_*}(p_*) \ge \frac{1}{2}\mathbb{E}_\nu[e_*(p - p_*)^2].$$

Since $e_* > 0$ ν-almost surely, zero excess risk implies $p = p_*$ ν-almost surely; no uniform sensitivity floor is required.

□

A.4 Posterior gradient and Fisher information

Consider an assayed pair with $p = \sigma(\eta) \in (0,1)$ and $e \in (0,1]$, holding e fixed as η varies. One-sided detection gives

$$\Pr(P=1, D=0 \mid Z, M=1) = p(1-e), \qquad \Pr(D=0 \mid Z, M=1) = 1-ep.$$

Bayes' rule yields the missed-positive posterior used in (4):

$$h(p,e) = \Pr(P=1 \mid D=0, Z, M=1) = \frac{(1-e)p}{1-ep}.$$

Direct differentiation gives

$$\frac{\partial}{\partial \eta}\ell(D, e\sigma(\eta)) = \begin{cases} p-1, & D=1, \\ \dfrac{ep(1-p)}{1-ep} = p - h(p,e), & D=0. \end{cases}$$

Let

$$m_D = D + (1-D)h(p,e) = \mathbb{E}[P \mid D, Z, M=1].$$

Then

$$\partial_\eta \ell(D, e\sigma(\eta)) = p - m_D = p - \{D + (1-D)h(p,e)\}. \tag{18}$$

The negative-log-likelihood gradient $p - m_D$ has conditional mean zero. By the law of total variance, its conditional variance, and hence the Fisher information about η, is

$$\begin{aligned} \mathcal{I}_{\mathrm{det}} &= \mathbb{E}\big[(p-m_D)^2 \mid Z, M=1\big] \\ &= \mathrm{Var}(m_D \mid Z, M=1) \\ &= \mathrm{Var}(P \mid Z, M=1) - \mathbb{E}[\mathrm{Var}(P \mid D, Z, M=1) \mid Z, M=1] \\ &= \mathrm{Var}(D + (1-D)h(p,e) \mid Z, M=1) \\ &= p(1-p) - (1-ep)h(1-h) \\ &= \frac{ep(1-p)^2}{1-ep}. \end{aligned}$$

Observing P directly gives information $\mathcal{I}_P = p(1-p)$. Their ratio is

$$\frac{\mathcal{I}_{\mathrm{det}}}{\mathcal{I}_P} = \frac{e(1-p)}{1-ep} \le 1,$$

with equality at $e=1$. For fixed p, $h(p,e) \to p$ and $\mathcal{I}_{\mathrm{det}} \to 0$ as $e \to 0$.

A.5 Proof of Proposition 2

Proof. Write $S = LA^\top$, so $\mathrm{rank}(S) \le K$ and $R = B - S$. Let $S = U\,\mathrm{diag}(s_1, \dots, s_r)V^\top$ be a compact singular value decomposition (SVD). Every factorization $S = LA^\top$ satisfies

$$\|S\|_* = \mathrm{tr}(U^\top S V) = \langle U^\top L, V^\top A\rangle_F \le \|U^\top L\|_F \|V^\top A\|_F \le \|L\|_F \|A\|_F \le \tfrac{1}{2}(\|L\|_F^2 + \|A\|_F^2).$$

Equality is attained by the balanced factors

$$L = U\,\mathrm{diag}(\sqrt{s_1}, \dots, \sqrt{s_r}), A = V\,\mathrm{diag}(\sqrt{s_1}, \dots, \sqrt{s_r})$$

with zero columns added when $r < K$. Zero factors cover $S = 0$, including $K = 0$.

The preceding inequality shows that every factorization of S has factor penalty at least $\lambda\|S\|_*$, and the balanced factors attain equality. Since $R = B - S$, minimizing first over factorizations of S and then over rank-K matrices S gives

$$\Omega_K(B) := \inf_{LA^\top + R = B} \left\{ \frac{\lambda}{2}\left(\|L\|_F^2 + \|A\|_F^2\right) + \frac{\mu}{2}\|R\|_F^2 \right\} = \min_{\mathrm{rank}(S) \le K} \left\{ \lambda\|S\|_* + \frac{\mu}{2}\|B - S\|_F^2 \right\}. \tag{19}$$

Let $\sigma_1 \geq \cdots \geq \sigma_m$ and $s_1 \geq \cdots \geq s_m$ denote the singular values of B and S, respectively. Rank feasibility requires $s_j = 0$ for $j > k = \min(K, m)$. Von Neumann's trace inequality implies $\langle B, S\rangle_F \leq \sum_j \sigma_j s_j$, with equality for aligned singular vectors. Consequently, the objective in (19) is bounded below by

$$\begin{aligned}\lambda\|S\|_* + \frac{\mu}{2}\|B - S\|_F^2 &= \lambda \sum_j s_j + \frac{\mu}{2}\big(\|B\|_F^2 + \|S\|_F^2 - 2\langle B, S\rangle_F\big) \\ &\geq \sum_{j=1}^{m} \left\{\lambda s_j + \frac{\mu}{2}(\sigma_j - s_j)^2\right\}.\end{aligned}$$

Equality is attained by choosing S with the same left and right singular vectors as B; only its singular values then remain to be optimized.

For $j \leq k$, the scalar objective $f_j(s) = \lambda s + \mu(\sigma_j - s)^2/2$ is strictly convex on $s \geq 0$. Its derivative is $\lambda + \mu(s - \sigma_j)$, so

$$s_j^* = (\sigma_j - \lambda/\mu)_+, \qquad f_j(s_j^*) = \frac{\mu}{2}\sigma_j^2 - \frac{\mu}{2}(\sigma_j - \lambda/\mu)_+^2.$$

For $j > k$, rank feasibility sets $s_j = 0$ and contributes $\mu\sigma_j^2/2$. The selected s_j are nonincreasing, so they satisfy the singular-value ordering constraint. Alignment attains the preceding lower bound. Substituting these values into (19) proves (10):

$$\begin{aligned}\Omega_K(B) &= \frac{\mu}{2}\sum_{j=1}^{m}\sigma_j^2 - \frac{\mu}{2}\sum_{j=1}^{k}\left(\sigma_j - \frac{\lambda}{\mu}\right)_+^2 \\ &= \frac{\mu}{2}\|B\|_F^2 - \frac{\mu}{2}\sum_{j=1}^{k}\left(\sigma_j - \frac{\lambda}{\mu}\right)_+^2.\end{aligned}$$

The resulting shared matrix has rank $\min\{K, \#\{j : \sigma_j > \lambda/\mu\}\}$, as asserted. □

A.6 Missed-Positive Ranking

Lemma 1 (Posterior ranking)**.** *For an assayed non-detection, let*

$$h(p, e) := \Pr(P = 1 \mid D = 0, Z, M = 1) = \frac{(1 - e)p}{1 - ep}, \qquad 0 < p < 1, \quad 0 < e \leq 1.$$

For $0 < e < 1$*,*

$$\frac{\partial h}{\partial p} = \frac{1 - e}{(1 - ep)^2} > 0, \qquad \frac{\partial h}{\partial e} = -\frac{p(1 - p)}{(1 - ep)^2} < 0.$$

For constant $0 < e < 1$ *within a target,* h *strictly preserves the ordering of* p*; for fixed* p*, it decreases as sensitivity increases. At* $e = 1$*,* $h(p, e) = 0$*.*

Proof. The posterior formula follows from Bayes' rule:

$$h(p, e) = \frac{\Pr(D = 0 \mid P = 1, Z, M = 1)\Pr(P = 1 \mid Z, M = 1)}{\Pr(D = 0 \mid Z, M = 1)} = \frac{(1 - e)p}{1 - ep}.$$

The quotient rule gives

$$\begin{aligned}\partial_p h &= \frac{(1 - e)(1 - ep) + e(1 - e)p}{(1 - ep)^2} = \frac{1 - e}{(1 - ep)^2}, \\ \partial_e h &= \frac{-p(1 - ep) + (1 - e)p^2}{(1 - ep)^2} = -\frac{p(1 - p)}{(1 - ep)^2}.\end{aligned}$$

The first derivative is positive for $e < 1$, while the second is negative. These signs prove the stated monotonicity properties, and direct substitution gives $h(p, 1) = 0$. □

Table 2: Analysis cohorts and evaluation splits. Gene counts refer to the source pool; the constructed predictor has $2G + 2$ columns. BARseq A1 and M1 are fitted separately.

Dataset	Cells	Genes	Targets	Assayed pairs	Unassayed pairs	Split
Projection-TAGs	37,827	32	7	245,825	18,964	Animal grouped
SPIDER spatial	45,835	32	15	687,525	0	Spatial slices
MERGE-seq	7,147	128	5	14,455	21,280	Sample grouped
BARseq A1	593	23	11	6,523	0 within panel	Depth blocks
BARseq M1	749	23	35	26,215	0 within panel	Depth blocks
Simulation	400/rep.	16	36	14,400	0	Spatial holdout

B Experimental Protocol

B.1 Dataset Construction

Projection-TAGs. We first matched cell identities between the standard and amplified-union releases, then retained non-Parse cells (cells not processed on the Parse Biosciences platform) annotated as L2/3 IT, L4 IT, L5 IT, L6 IT, or L5 PT. This yielded 37,827 cells. Counts and evaluation metrics include only cell–target pairs assayed according to the animal-specific injection map. The seven targets are MOp, SSp, VP, PAG, MY, SCL, and SCS. Predictors comprise 32 prespecified endogenous genes. Projection-TAG/barcode features are excluded from both the predictor and the library-size denominator. Counts are normalized to a library size of 10^4, transformed by $\log(1+x)$, and standardized using the model-training animals. SSp was not assayed in animals 1–4, producing all 18,964 structurally unassayed pairs.

SPIDER spatial. The retained cohort contains 45,835 cells, 32 genes, 15 projection targets, and 36 anterior–posterior slices. All retained target channels are measured for every retained cell. Gene counts are $\log(1 + x)$-transformed and standardized using inner-training cells. Anterior–posterior position defines the spatial outer folds.

MERGE-seq. We use the retained excitatory analysis cohort aligned to the four recorded prefrontal cortex (PFC) samples and the five native targets AI, DMS, MD, BLA, and LH. The adapter sets $M_{it}^{\text{native}} = 1$ for all five targets of cells recorded as barcoded and zero for cells recorded as non-barcoded. This mask encodes projection-label availability; it does not establish whether a non-barcoded cell was experimentally unassayed or anatomically non-projecting. The barcoded rows define the evaluation population. The measurement-degradation adapter reserves 20% of non-barcoded, all-mask-zero cells within each sample as a label-free design subset. It selects a fixed 128-gene source pool by variability in those design cells after excluding assay-derived features, including their contribution to the library-size denominator. Design cells are removed before splitting, calibration, fitting, and evaluation. The analysis-cohort size is reported after this removal. Scaling uses visible training entries only. The analysis uses three outer folds with whole samples kept together.

BARseq. The retained cohort contains 593 A1 neurons with 11 targets and 749 M1 neurons with 35 targets; both panels have the same 23 measured genes. A1 and M1 are fitted separately, so off-panel entries in a combined display matrix are not training examples or evaluated predictions.

B.2 Splits and Reference Isolation

Projection-TAGs. The outer test-animal pairs are $\{1, 4\}$, $\{2, 6\}$, and $\{3, 5\}$. The corresponding validation animals are 2, 4, and 1, leaving model-training animals $\{3, 5, 6\}$, $\{1, 3, 5\}$, and $\{2, 4, 6\}$. Preprocessing and detection calibration use the three training animals during model selection. The validation animal then joins the training animals to form the development set, on which prepro-

cessing, detection calibration, and the selected predictor are refitted. Paired reference labels enter detection calibration; predictor fitting uses the observed detections.

SPIDER spatial. The 36 slices are ordered by median anterior–posterior coordinate and divided into three contiguous 12-slice test blocks. In each fold, four approximately spaced slices among the remaining 24 form validation, leaving 20 training slices.

MERGE-seq. The four samples are partitioned into three count-balanced outer test groups. Within each outer fold, a separate validation sample is chosen from the remaining samples using outcome-independent design information.

BARseq. Within each animal, neurons are ordered by cortical depth and divided into 12 approximately equal-count bins. Each of three folds uses the same contiguous four-bin test block in every animal. Two spaced bins among the remaining eight form validation and the remaining six form training.

Simulation. We partition the 20 slices into three contiguous outer test blocks. Each outer fold uses one block for testing and four spaced slices from the remaining slices for validation, giving inner-training/validation/test sizes of 180/80/140, 180/80/140, and 200/80/120. A fresh dataset is generated in every repetition and evaluated with three outer folds; the generated dataset is the independent Monte Carlo unit, with its three folds retained as dependent evaluations.

B.3 Panel-Aware Feature Construction

For a source pool of G genes, let O_{ig} indicate observation of gene g and g_{ig} its adapter-transformed value. Visible training rows determine its mean c_g and standard deviation s_g; constant columns use $s_g = 1$. With virtual assay $a_i \in \{A, B, C\}$, the evaluated design is

$$x_i = \left[\{O_{ig}(g_{ig} - c_g)/s_g\}_{g=1}^{G}, \{O_{ig}\}_{g=1}^{G}, \mathbf{1}\{a_i = B\}, \mathbf{1}\{a_i = C\}\right]^\top, \quad d = 2G + 2. \tag{20}$$

The first coordinates are defined as zero where $O_{ig} = 0$. Observation indicators distinguish missing values from observed zeros. All columns share one coefficient block, and d is the number of encoded columns. The evaluated configuration includes neither location nor target descriptors. Selection fits preprocessing on inner-training rows; final refitting uses development rows.

B.4 Detection Calibration

At full positive retention, sensitivity is known and set to one. For positive-label loss, we estimate sensitivity from paired references as follows.

A label-independent 20% of development cells is selected for paired calibration. Where biological group identifiers are available, sampling is stratified jointly by biological group and virtual assay; otherwise it is stratified by virtual assay. During hyperparameter selection, only the paired cells that belong to the inner-training split are used to estimate detection sensitivity. After model selection, the detector is refitted using the paired development cells.

For virtual assay a and target t, the detector contains a pooled intercept, target effects, assay effects, and assay–target interactions. Let $\mathcal{C}_+$ denote paired, assayed reference-positive entries and γ_{-0} all nonintercept coefficients. Writing its logit as $z_{it}^\top \gamma$, we minimize

$$\widehat{\gamma} = \arg\min_{\gamma} \left\{ \sum_{(i,t)\in\mathcal{C}_+} \left[\log(1 + e^{z_{it}^\top \gamma}) - D_{it} z_{it}^\top \gamma\right] + \frac{1}{2}\|\gamma_{-0}\|_2^2 + \frac{1}{2}\left[\log(1 + e^{\gamma_0}) + \log(1 + e^{-\gamma_0})\right] \right\}. \tag{21}$$

A calibration set with no assayed reference-positive entries is recorded as a calibration failure. The fitted sensitivities are then held fixed while fitting PU, PU-MIRT, and Gene2Wire.

B.5 Controlled Measurement Degradation

SPIDER spatial, MERGE-seq, and BARseq retain their original assay calls as the reference, while Projection-TAGs uses its standard-or-amplified union; artificial gene and target panels and positive censoring are applied to these references. Let $\zeta \in \{0, 0.2, 0.4, 0.6, 0.8\}$ be the nominal positive-label loss rate. A fixed standardized assay–target score ξ_{gt} is drawn independently of projection outcomes, and cell i belongs to virtual assay $a(i)$. For $\zeta > 0$, with fixed heterogeneity strength $\delta = 1$, detection sensitivity is

$$e_{it}(\zeta) = \sigma\big\{b_t(\zeta) + \delta\xi_{a(i),t}\big\}, \qquad \delta = 1. \tag{22}$$

For each fold and target, $b_t(\zeta)$ is chosen using only measured reference-positive inner-training cells so that

$$\frac{1}{|\mathcal{I}_{t,\text{train}}|} \sum_{i \in \mathcal{I}_{t,\text{train}}} \sigma\big\{b_t(\zeta) + \delta\xi_{a(i),t}\big\} = 1 - \zeta, \tag{23}$$

where $\mathcal{I}_{t,\text{train}}$ contains the measured, reference-positive training rows for target t. If none exist, a pooled inner-training intercept is used as fallback. A reference-positive label is retained according to a Bernoulli draw with probability e_{it}; reference-negative labels are unchanged, and the fitting support Ω remains fixed when varying positive retention at a fixed target panel. For nonzero positive-label loss, fitted PU models estimate detection sensitivity from paired references; generator sensitivities are retained only for evaluation. The same uniform draws are reused across rates within a repetition, coupling the comparisons across loss rates.

For MERGE-seq, (23) is constructed separately within every outer fold using that fold's inner-training positives, then applied unchanged to validation and test cells.

Gene and target panels use a common outcome-independent cyclic design. At requested coverage c over J items, each of three virtual assays receives $\max\{\lceil J/3 \rceil, \lfloor cJ + 1/2 \rfloor\}$ consecutive items from a fixed start in a seeded circular order. Panels are nested across coverages and their union contains the declared item universe; realized integer coverages are recorded separately. Target-panel indicators are intersected with native assay availability.

Gene standardization uses only visible inner-training values during selection and is refitted on visible development values afterward. Missing standardized values are set to zero and accompanied by one observation indicator per gene and two virtual-assay indicators.

B.6 Simulation Data-Generating Process

Each repetition contains 400 cells, 36 targets, 16 gene features, 20 spatial slices, and a true shared rank of two.

Gene features. For a vector a of n observations, write $s(a)^2 = n^{-1}\sum_i (a_i - \bar{a})^2$ and $\mathcal{Z}(a) = (a - \bar{a})/s(a)$; matrix standardization acts columnwise. Each slice contains 20 cells. Its center is one of 20 evenly spaced values in $[-1, 1]$, and cell position r_i is this center plus independent $\mathcal{N}(0, 0.018^2)$ noise. Starting from independent standard-normal entries Z_{ij}, form seven cell factors with

$$F_{i1} = 0.70\mathcal{Z}(r)_i + \sqrt{1 - 0.70^2}\, Z_{i1}, \tag{24}$$

$$F_{i2} = 0.55\sin(\pi r_i) + \sqrt{1 - 0.55^2}\, Z_{i2}, \qquad F_{ij} = Z_{ij} \quad (j = 3, \ldots, 7). \tag{25}$$

With independent loadings $W \in \mathbb{R}^{7\times 16}$ having standard-normal entries and independent noise $E_{ig} \sim \mathcal{N}(0, 0.55^2)$, the complete gene matrix is $\widetilde{X} = \mathcal{Z}(FW + E)$. Spatial position thus induces structured variation in the genes. The generating logits depend on these genes; location and target descriptors are omitted from the evaluated predictors. The panel-aware design x_i is constructed from $\widetilde{X}$ using the training-only transformations in Appendix B.3. Standardizing the generated population here defines the data-generating distribution and is separate from fitting predictor preprocessing.

Shared and target-specific signals. Draw $C \in \mathbb{R}^{16\times 2}$ and $V \in \mathbb{R}^{36\times 2}$ with independent standard-normal entries, then center and scale each column of V across targets. With v_t the transpose of row

t of V, define

$$\ell_t = \frac{\widetilde{X} C v_t}{s(\widetilde{X} C v_t)}, \qquad u_t = \frac{\widetilde{X} w_t}{s(\widetilde{X} w_t)}, \qquad w_t \sim \mathcal{N}(0, I_{16}),$$

where the w_t are independent of the shared loadings. The columns ℓ_t form the rank-two shared signal. For $\langle a, b\rangle_n = n^{-1} a^\top b$, the target-specific signal is

$$d_t = \frac{u_t - \ell_t \langle u_t, \ell_t\rangle_n}{s(u_t - \ell_t \langle u_t, \ell_t\rangle_n)}.$$

Both signals have zero empirical mean and unit variance, and $\langle \ell_t, d_t\rangle_n = 0$. Their mixture therefore fixes the total signal variance as the shared fraction varies:

$$f_{it} = 1.35 \left[\sqrt{\rho}\, \ell_{it} + \sqrt{1-\rho}\, d_{it}\right], \qquad \rho \in \{0, 0.5, 1\}. \tag{26}$$

The implementation rescales the mixed signal to empirical standard deviation 1.35. Target intercepts solve $n^{-1} \sum_i \sigma(\alpha_t + f_{it}) = \pi_t$, with π_t evenly spanning 0.1–0.3, and

$$P_{it} \sim \text{Bernoulli}\{\sigma(\alpha_t + f_{it})\}.$$

For a fixed repetition, the base variables and Bernoulli uniform draws are reused across ρ, pairing the sharing-strength comparisons. A new repetition generates a new base dataset. Generated probabilities and coefficients are retained for evaluation; they are not supplied as predictor inputs.

Assay–target-dependent censoring. The sensitivity follows (22), with virtual assay and target determining detection variation. The shared real-data and simulation censoring rule is

$$e_{it}(\zeta) = \sigma\{b_t(\zeta) + \xi_{a(i),t}\}. \tag{27}$$

Each $b_t(\zeta)$ is calibrated using measured reference-positive inner-training entries according to (23). Virtual-assay assignments and the assay–target heterogeneity surface do not use projection outcomes; retention normalizers use only the designated inner-training reference positives.

Observed labels and calibration. Within the fitted target panel, positives are retained as $D_{it} = P_{it}\mathbf{1}\{U_{it} < e_{it}\}$, with $U_{it} \sim \text{Uniform}(0, 1)$; measured reference negatives remain negative. The same uniform draws are reused across conditions within a repetition. Nominal rates specify expected retention; realized loss rates fluctuate. At $\zeta = 0$, all labels are retained and $e_{it} = 1$ is treated as known. For nonzero loss, generating detection probabilities are not supplied to the fitted projection models. Detection is estimated using designated paired references separately for tuning and final refitting, as described in Appendix B.4.

B.7 Initialization and Optimization

Adaptive candidate selection. The estimator's eligible families are the separately tuned direct PU and shared-only PU-MIRT endpoints and native shared-plus-residual candidates. Native candidates use positive ranks in $\{1, 2, 4, 8, 16, \min(d, T)\}$, after removing duplicates and values above $\min(d, T)$. The search screens these ranks at $\lambda = \mu = 10^{-3}$ using measured validation log loss, then refines the best two distinct converged ranks over the grid $\{10^{-4}, 10^{-3}, 10^{-2}, 10^{-1}, 1, 10\}$. A deterministic design in total-shrinkage and penalty-ratio coordinates limits the native search to 32 candidates; the two tuned endpoints remain eligible in addition to this budget. Refinement therefore adapts to the validation outcomes of the rank screen.

Each native candidate uses two initializations that reproduce the tuned endpoints' coefficient matrices and intercepts. Given an endpoint coefficient B_{end}, the initializer forms width-K factors and sets $R = B_{\text{end}} - LA^\top$, preserving its logits exactly. A shared-only endpoint with matching rank retains its factors; other cases use a deterministic factorization and place the remaining coefficient in R. These are prediction-preserving starts for the candidate objective.

Stable likelihood and analytic gradients. Let $\text{sp}(u) = \log(1 + e^u)$. For finite η, $p = \sigma(\eta)$ and supplied $0 < e \le 1$, the implementation evaluates

$$\ell(1, ep) = -\log e + \text{sp}(-\eta), \qquad \ell(0, ep) = \text{sp}(\eta) - \text{sp}\{\eta + \log(1 - e)\}, \tag{28}$$

using stable softplus and $\log(1-e)$ evaluations. The expression at $e=1$ takes $\mathrm{sp}(-\infty)=0$. These losses and (18) are algebraically exact, with no probability clipping in the training objective.

Let G be zero outside the measured training entries and satisfy

$$G_{it} = \frac{p_{it} - D_{it} - (1-D_{it})h(p_{it}, \widehat{e}_{it})}{|\Omega|} \quad \text{on } \Omega.$$

For the ordinary joint configuration in (3), the gradients are

$$\nabla_b \mathcal{L} = G^\top \mathbf{1}, \quad \nabla_L \mathcal{L} = X^\top G A + \lambda L, \quad \nabla_A \mathcal{L} = G^\top X L + \lambda A, \quad \nabla_R \mathcal{L} = X^\top G + \mu R. \tag{29}$$

Thus the likelihood is normalized by the number of measured training entries, while the displayed factor and residual penalties have exactly coefficients $\lambda/2$ and $\mu/2$. The projection intercepts receive no penalty. The evaluated predictors place all encoded gene and panel features in this coefficient block.

Optimization and refitting. Each start is optimized in float64 with analytic gradients and L-BFGS-B, for at most 500 iterations. A finite unsuccessful fit receives one deterministic continuation from its last iterate, with at most 1000 further iterations and the same objective. Both runs use function and gradient tolerances 10^{-8} and a maximum of 30 line-search steps per iteration. Among converged starts, the lower training objective is retained. If neither converges, the candidate is excluded. Only converged candidates are eligible for selection by measured-entry validation log loss of $\widehat{q} = \widehat{e}\widehat{p}$; deterministic ties use rank, penalties, and trial order. The selected configuration is refitted on development data with development-fitted preprocessing and the development detector. The resulting direct, shared-only, or joint model is the selected configuration.

Here *converged* records the optimizer's success flag. The observed-data objective can be nonconvex even in the logit: for $e<1$, the non-detection loss has

$$\frac{\partial^2 \ell(0, e\sigma(\eta))}{\partial \eta^2} = \frac{ep(1-p)(1-2p+ep^2)}{(1-ep)^2} < 0 \quad \text{when} \quad p > \frac{1}{1+\sqrt{1-e}}. \tag{30}$$

Proposition 2 characterizes the minimum penalty across decompositions of a fixed B. The fitting objective can remain nonconvex, as (30) shows. For targets with all-zero or all-one measured responses, the loss decreases as the unpenalized projection intercept tends to $-\infty$ or $+\infty$, respectively. The detector is fitted by the separate regularized calibration objective specified in Appendix B.4. Endpoint inclusion, adaptive rank selection, and development refitting are components of the procedure evaluated on the outer-test folds.

B.8 Hyperparameter Selection

Model selection must respect the labels available at validation, a concern highlighted in recent PU benchmarks (Wang et al., 2026). Table 4 gives the rank screens and penalty pool. Direct PU searches residual penalties at rank zero. PU-MIRT searches positive ranks and shared penalties on a Cartesian grid, using a balanced subset of at most 32 candidates. The top-two refinement procedure applies to native joint candidates. Selection uses validation detection log loss, followed by development-set refitting of the selected configuration.

For PU, PU-MIRT, and Gene2Wire, the base penalty pool is $\{10^{-4}, 10^{-3}, 10^{-2}, 10^{-1}, 1, 10\}$. BARseq ranks are capped at 11 for A1 and 35 for M1. The Qiao bilinear baseline uses the same penalty pool. Its candidate ranks are $\{1, 2, 4, 8, 16, K_q\}$, capped at $K_q = \min(d+1, T)$; a balanced search retains at most 32 candidates.

B.9 Comparator Implementations

All methods use the same outer splits, available gene measurements, observed detections, and fitting mask. Table 5 specifies their supervision and probability semantics. Targets are identified by their known identities throughout; the evaluation does not supply target descriptors.

GenEML. GenEML (Jain et al., 2017) uses the pinned authors-source Python-3 port with the known-mask adaptation and estimates exposure with its own model. Its factor penalties are fixed at 10^{-3}, while rank and the feature-map penalty are selected from at most 32 balanced candidates. This port uses target-specific exposure probabilities, dense features, and full-batch updates.

SAR-PU. We use the SAR-EM algorithm introduced by Bekker & Davis (2018), calling the authors' EM implementation associated with the follow-up SAR-PU study (Bekker et al., 2020). Our wrapper fits the method separately on each target's measured rows, using outcome-independent assay/QC propensity inputs and one prespecified candidate.

Bilinear. We evaluate the squared-loss bilinear model of Qiao (2024) with known-target identity inputs. It predicts a detection-ranking score; reference-probability log loss, Brier score, and hidden-positive recall are not reported.

B.10 Metrics and Aggregation

Reference AUPRC, AUROC, log loss, and Brier score are evaluated on natively assayed outer-test entries, including entries withheld only by the artificial fitting panel. Here $P_{it} = Z_{it}^{\mathrm{ref}}$ in real data and is the generated outcome in simulation. For native-reference test entries $\mathcal{T}_t = \{i \in \text{test} : M_{it}^{\mathrm{native}} = 1\}$ of target t,

$$\mathrm{LogLoss}_t = -|\mathcal{T}_t|^{-1} \sum_{i \in \mathcal{T}_t} \left[P_{it} \log \widehat{p}_{it} + (1 - P_{it}) \log(1 - \widehat{p}_{it}) \right], \tag{31}$$

$$\mathrm{Brier}_t = |\mathcal{T}_t|^{-1} \sum_{i \in \mathcal{T}_t} (\widehat{p}_{it} - P_{it})^2. \tag{32}$$

Log loss and Brier score are proper scoring rules for evaluating predicted probabilities (Gneiting & Raftery, 2007). Reference AUPRC and AUROC use $\widehat{p}_{it}$ for reference-probability models and the detection-ranking scores for the Bilinear baseline. AUPRC is average precision (step integration). Both AUPRC and AUROC omit single-class targets by the evaluation convention. Log loss clips probabilities to $[\epsilon_{\mathrm{mach}}, 1 - \epsilon_{\mathrm{mach}}]$ with float64 machine epsilon ϵ_{mach}; Brier score uses unclipped probabilities. Macro metrics average target-wise values where the metric is defined.

Hidden-positive candidates are the fitted-panel test non-detections $\mathcal{U}_t = \{i \in \mathcal{T}_t : M_{it} = 1,\, D_{it} = 0\}$. Let $H_{it} = \mathbf{1}\{P_{it} = 1, D_{it} = 0\}$ and $H_t = \sum_{i \in \mathcal{U}_t} H_{it}$. Then

$$\mathrm{Recall@H}_t = \frac{\sum_{i \in \mathrm{Top}_{H_t}(\mathcal{U}_t)} H_{it}}{H_t}. \tag{33}$$

Ties at the cutoff use their expected fractional contribution. Targets with $H_t = 0$ have undefined R@H and are excluded from its target-macro average. Reference-probability models use the plug-in estimate of (4) as the primary recovery score; R@H is unavailable for Bilinear. Recall@H uses the retrospective count H_t to evaluate hidden-positive recovery. A prospective retrieval budget is specified before reference outcomes are available.

Aggregation and variation. For each condition and model, metrics are first averaged across targets where defined, then across available outer folds within a repetition, and finally across repetitions. BARseq A1 and M1 are aggregated separately. Real-data summaries use ten prespecified panel configurations within a fixed cohort, which coincide at full gene and target coverage and do not represent independent cohorts. Simulation summaries use ten generated datasets per sharing strength, with each generated dataset serving as the independent Monte Carlo unit.

Metric support. The BARseq A1 probability metrics use common finite repetition support across the probability models. For BARseq M1 and Projection-TAGs, log loss, Brier score, and R@H average the available values for each model. AUPRC and AUROC use common finite repetition support across Gene2Wire, Bilinear, and GenEML; SAR-PU is included where a nonempty common intersection exists. A dash denotes an undefined or unavailable metric and is excluded from comparative counts.

The main figures emphasize macro AUPRC and reference log loss; the complete tables also report AUROC, Brier score, and defined Hidden Recall@H.

C Additional Experimental Results

This appendix reports absolute numerical results and additional simulation comparisons. Each favorable-condition count in the main text compares Gene2Wire with the best available external

mean for that condition and metric. Probability-metric comparisons include methods that supply reference-probability predictions. These counts summarize performance across the evaluated measurement conditions; Appendix B.10 specifies aggregation and finite metric support.

C.1 Cross-Assay Performance at Severe Detection Loss

Table 6 provides absolute values for the 80% loss columns of Figure 3. Gene and target coverage are full, so this slice isolates the effect of missed detections. The four-method comparison covers all five assay panels.

Table 3: Realized panel sizes. Each cell gives measured items / source-pool items for one virtual assay; integer rounding explains differences from requested percentage labels. The full requested factorial is listed separately.

Dataset	Panel	100%	70%	40%
BARseq A1	Genes	23/23	16/23	9/23
BARseq A1	Targets	11/11	8/11	4/11
BARseq M1	Genes	23/23	16/23	9/23
BARseq M1	Targets	35/35	25/35	14/35
MERGE-seq	Genes	128/128	90/128	51/128
MERGE-seq	Targets	5/5	4/5	2/5
Projection-TAGs	Genes	32/32	22/32	13/32
Projection-TAGs	Targets	7/7	5/7	3/7
Simulation	Genes	16/16	11/16	6/16
Simulation	Targets	36/36	25/36	14/36
SPIDER spatial	Genes	32/32	22/32	13/32
SPIDER spatial	Targets	15/15	11/15	6/15

Table 4: Gene2Wire-family rank screens and search settings. The search column describes joint candidates in the real-data rows. Joint candidates first screen positive ranks, retain two, then refine total shrinkage and the residual/shared penalty ratio within a budget of 32 native joint candidates. The two tuned endpoint candidates are additional.

Dataset	Search	Candidate ranks	Base penalty pool	Joint native budget
Projection-TAGs	Top-2 refinement	(1,2,4,7)	$10^{-4}, 10^{-3}, 10^{-2}, 10^{-1}, 1, 10$	32
SPIDER spatial	Top-2 refinement	(1,2,4,8,15)	$10^{-4}, 10^{-3}, 10^{-2}, 10^{-1}, 1, 10$	32
MERGE-seq	Top-2 refinement	(1,2,4,5)	$10^{-4}, 10^{-3}, 10^{-2}, 10^{-1}, 1, 10$	32
BARseq A1	Top-2 refinement	(1,2,4,8,11)	$10^{-4}, 10^{-3}, 10^{-2}, 10^{-1}, 1, 10$	32
BARseq M1	Top-2 refinement	(1,2,4,8,16,35)	$10^{-4}, 10^{-3}, 10^{-2}, 10^{-1}, 1, 10$	32
Simulation PU-MIRT	Bounded grid	(1,2,4,8,16,34)	$10^{-4}, 10^{-3}, 10^{-2}, 10^{-1}, 1, 10$	–
Simulation Gene2Wire	Top-2 refinement	(1,2,4,8,16,34)	$10^{-4}, 10^{-3}, 10^{-2}, 10^{-1}, 1, 10$	32

Table 5: Comparator information access. Paired reference labels are used only to estimate sensitivity in the indicated methods. Every method fits observed detections on the known measurement mask; no method receives outer-test reference labels.

Method	Sensitivity and supervision	Prediction and implementation
PU / PU-MIRT / Gene2Wire	Common $\widehat{e}$; paired calibration for nonzero loss	Reference probability p; direct, shared, or selected structure
GenEML	Own target-level exposure model; no paired references	Reference probability p; authors-source Python-3 port with known-mask support
SAR-PU	Own assay/QC propensity model; no paired references	Reference probability p; authors' SAR-EM kernel applied to each target's measured rows
Bilinear	Observed detections; no paired references	Detection-ranking score; target-identity adaptation of the bilinear model

Table 6: **Cross-assay results at 80% label loss within the four-model summary.** Gene and target coverage are full. This set contains Gene2Wire (G2W), GenEML, SAR-PU, and Bilinear; the latter supplies detection scores only. Means are descriptive; complete four-decimal values are in the appendix.

	AUPRC ↑				Log loss ↓		
Dataset	G2W	GenEML	SAR	Bilinear	G2W	GenEML	SAR
Projection-TAGs	0.179	0.072	0.171	0.163	0.191	0.303	0.208
MERGE-seq	0.394	0.310	0.332	0.349	0.402	0.646	0.754
SPIDER spatial	0.305	0.207	0.290	0.274	0.330	0.432	0.347
BARseq A1	0.273	0.252	0.263	0.271	0.474	0.738	0.741
BARseq M1	0.335	0.240	—	0.267	0.367	0.547	—

C.2 Combined Factorial Summaries

For gene coverage g, target coverage t, and positive retention r, the combined display uses

$$\kappa(g,t,r) = \frac{3}{g^{-1} + t^{-1} + r^{-1}}. \tag{34}$$

Equal κ values pool different physical settings. These two displays summarize their panels' full factorial designs, while the one-factor results identify the effect of each intervention. Integer panel sizes appear in Appendix B.5.

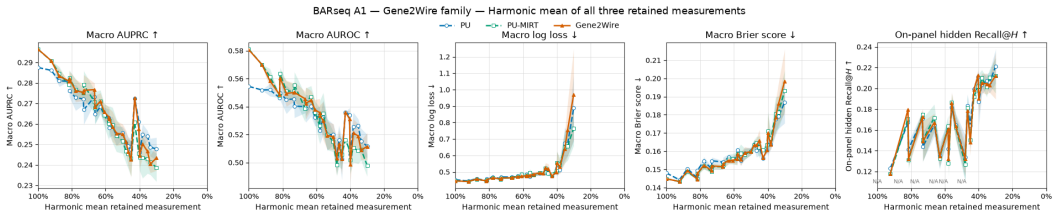

Figure 5: **BARseq A1: combined-factorial performance.** Lines show PU, PU-MIRT, and Gene2Wire scores against $\kappa(g,t,r)$. Shading denotes nominal 95% Student-t intervals across panel configurations within the same cohort; these intervals describe conditional panel-design variation. R@H is absent where pooled support includes conditions without hidden positives.

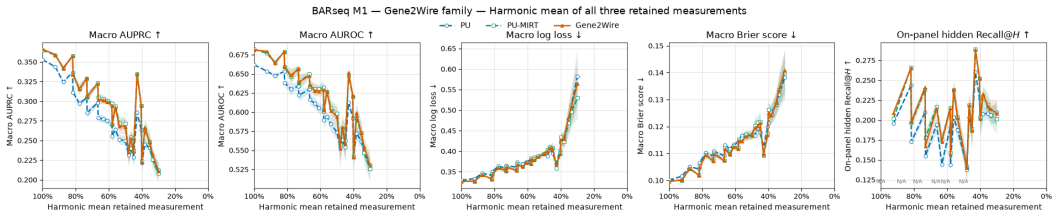

Figure 6: **BARseq M1: combined-factorial performance.** Lines show PU, PU-MIRT, and Gene2Wire scores against $\kappa(g,t,r)$. Shading denotes nominal 95% Student-t intervals across panel configurations within the same cohort; these intervals describe conditional panel-design variation. R@H is absent where pooled support includes conditions without hidden positives.

Table 7: Full factorial of gene coverage, target coverage, and positive retention. g and t are per-assay gene and target coverage, r is retained positive fraction, and $\kappa = 3/(g^{-1} + t^{-1} + r^{-1})$. Every row enters the combined summary. The all-full row is shared by all one-factor curves.

g	t	r	κ	One-factor slice
1.00	1.00	1.00	1.0000	positive loss
1.00	1.00	0.80	0.9231	positive loss
1.00	1.00	0.60	0.8182	positive loss
1.00	1.00	0.40	0.6667	positive loss
1.00	1.00	0.20	0.4286	positive loss
1.00	0.70	1.00	0.8750	target
1.00	0.70	0.80	0.8155	combined only
1.00	0.70	0.60	0.7326	combined only
1.00	0.70	0.40	0.6087	combined only
1.00	0.70	0.20	0.4038	combined only
1.00	0.40	1.00	0.6667	target
1.00	0.40	0.80	0.6316	combined only
1.00	0.40	0.60	0.5806	combined only
1.00	0.40	0.40	0.5000	combined only
1.00	0.40	0.20	0.3529	combined only
0.70	1.00	1.00	0.8750	gene
0.70	1.00	0.80	0.8155	combined only
0.70	1.00	0.60	0.7326	combined only
0.70	1.00	0.40	0.6087	combined only
0.70	1.00	0.20	0.4038	combined only
0.70	0.70	1.00	0.7778	combined only
0.70	0.70	0.80	0.7304	combined only
0.70	0.70	0.60	0.6632	combined only
0.70	0.70	0.40	0.5600	combined only
0.70	0.70	0.20	0.3818	combined only
0.70	0.40	1.00	0.6087	combined only
0.70	0.40	0.80	0.5793	combined only
0.70	0.40	0.60	0.5362	combined only
0.70	0.40	0.40	0.4667	combined only
0.70	0.40	0.20	0.3360	combined only
0.40	1.00	1.00	0.6667	gene
0.40	1.00	0.80	0.6316	combined only
0.40	1.00	0.60	0.5806	combined only
0.40	1.00	0.40	0.5000	combined only
0.40	1.00	0.20	0.3529	combined only
0.40	0.70	1.00	0.6087	combined only
0.40	0.70	0.80	0.5793	combined only
0.40	0.70	0.60	0.5362	combined only
0.40	0.70	0.40	0.4667	combined only
0.40	0.70	0.20	0.3360	combined only
0.40	0.40	1.00	0.5000	combined only
0.40	0.40	0.80	0.4800	combined only
0.40	0.40	0.60	0.4500	combined only
0.40	0.40	0.40	0.4000	combined only
0.40	0.40	0.20	0.3000	combined only

C.3 Complete One-Factor Numerical Results

The following tables give every available one-factor value at four-decimal precision. Missing entries are left blank or marked with a dash. The BARseq results are followed by the four-method real-data and simulation values underlying Figures 3 and 2. These are absolute scores: the within-cell glyph scaling used in the heatmaps is solely a visual encoding.

Table 8: BARseq A1: entries are mean (SE). Ranking summaries pair finite repetitions among eligible displayed models. LL and BS use common probability-model support. Dashes denote unavailable or inapplicable values.

Condition	Model	AUPRC ↑	AUROC ↑	LL ↓	BS ↓	R@H ↑
Loss 0%	Gene2Wire	0.2964 (0.0295)	0.5806 (0.0159)	0.4455 (0.0212)	0.1449 (0.0081)	—
	Bilinear	0.2823 (0.0388)	0.5497 (0.0304)	—	—	—
	GenEML	0.2877 (0.0305)	0.5675 (0.0092)	0.6881 (0.0245)	0.2349 (0.0143)	—
	SAR-PU	0.2928 (0.0277)	0.5513 (0.0079)	0.5231 (0.0146)	0.1691 (0.0059)	—
Loss 20%	Gene2Wire	0.2909 (0.0303)	0.5702 (0.0148)	0.4415 (0.0203)	0.1433 (0.0076)	0.1179 (0.0245)
	Bilinear	0.2748 (0.0409)	0.5292 (0.0460)	—	—	—
	GenEML	0.2696 (0.0310)	0.5527 (0.0094)	0.7244 (0.0188)	0.2437 (0.0080)	0.0702 (0.0166)
	SAR-PU	0.2827 (0.0258)	0.5433 (0.0031)	0.5154 (0.0281)	0.1637 (0.0088)	0.1240 (0.0209)
Loss 40%	Gene2Wire	0.2809 (0.0336)	0.5486 (0.0205)	0.4453 (0.0213)	0.1446 (0.0079)	0.1796 (0.0362)
	Bilinear	0.2602 (0.0377)	0.5053 (0.0469)	—	—	—
	GenEML	0.2798 (0.0332)	0.5556 (0.0112)	0.6966 (0.0451)	0.2355 (0.0210)	0.1406 (0.0413)
	SAR-PU	0.2747 (0.0302)	0.5333 (0.0184)	0.5367 (0.0285)	0.1625 (0.0105)	0.1561 (0.0281)
Loss 60%	Gene2Wire	0.2807 (0.0355)	0.5462 (0.0230)	0.4518 (0.0204)	0.1461 (0.0072)	0.2081 (0.0366)
	Bilinear	0.2601 (0.0362)	0.5310 (0.0276)	—	—	—
	GenEML	0.2850 (0.0277)	0.5649 (0.0054)	0.7104 (0.0593)	0.2366 (0.0283)	0.2056 (0.0280)
	SAR-PU	0.2700 (0.0316)	0.5300 (0.0207)	0.5953 (0.0362)	0.1668 (0.0106)	0.1867 (0.0265)
Loss 80%	Gene2Wire	0.2726 (0.0335)	0.5361 (0.0120)	0.4745 (0.0283)	0.1563 (0.0116)	0.1987 (0.0364)
	Bilinear	0.2706 (0.0351)	0.5240 (0.0201)	—	—	—
	GenEML	0.2517 (0.0264)	0.5208 (0.0131)	0.7381 (0.0789)	0.2320 (0.0325)	0.1949 (0.0224)
	SAR-PU	0.2629 (0.0325)	0.5259 (0.0187)	0.7410 (0.0453)	0.1825 (0.0131)	0.2145 (0.0310)
G70%	Gene2Wire	0.2779 (0.0314)	0.5581 (0.0096)	0.4605 (0.0181)	0.1509 (0.0068)	—
	Bilinear	0.2699 (0.0338)	0.5393 (0.0286)	—	—	—
	GenEML	0.2770 (0.0329)	0.5634 (0.0097)	0.7153 (0.0178)	0.2486 (0.0096)	—
	SAR-PU	0.2796 (0.0267)	0.5451 (0.0075)	0.5355 (0.0183)	0.1741 (0.0070)	—
G40%	Gene2Wire	0.2632 (0.0268)	0.5356 (0.0071)	0.4749 (0.0146)	0.1568 (0.0054)	—
	Bilinear	0.2652 (0.0284)	0.5370 (0.0100)	—	—	—
	GenEML	0.2656 (0.0273)	0.5442 (0.0110)	0.7230 (0.0281)	0.2515 (0.0134)	—
	SAR-PU	0.2617 (0.0242)	0.5334 (0.0062)	0.5389 (0.0180)	0.1747 (0.0063)	—
T70%	Gene2Wire	0.2889 (0.0341)	0.5587 (0.0170)	0.4526 (0.0224)	0.1479 (0.0085)	—
	Bilinear	0.2791 (0.0361)	0.5478 (0.0288)	—	—	—
	GenEML	0.2868 (0.0309)	0.5693 (0.0073)	0.6538 (0.0349)	0.2236 (0.0153)	—
	SAR-PU	0.2881 (0.0285)	0.5427 (0.0096)	0.5401 (0.0197)	0.1715 (0.0071)	—
T40%	Gene2Wire	0.2863 (0.0320)	0.5551 (0.0172)	0.4652 (0.0206)	0.1517 (0.0076)	—
	Bilinear	0.2733 (0.0340)	0.5406 (0.0269)	—	—	—
	GenEML	0.2598 (0.0277)	0.5223 (0.0125)	1.0031 (0.1174)	0.2920 (0.0138)	—
	SAR-PU	0.2825 (0.0328)	0.5362 (0.0149)	0.8320 (0.0203)	0.2395 (0.0054)	—

Table 9: BARseq M1: entries are mean (SE). Ranking summaries pair finite repetitions among eligible displayed models. LL, BS, and recovery use model-specific available support. Dashes denote unavailable or inapplicable values.

Condition	Model	AUPRC ↑	AUROC ↑	LL ↓	BS ↓	R@H ↑
Loss 0%	Gene2Wire	0.3661 (0.0384)	0.6814 (0.0304)	0.3262 (0.0168)	0.0996 (0.0052)	—
	Bilinear	0.3532 (0.0370)	0.6955 (0.0234)	—	—	—
	GenEML	0.3469 (0.0417)	0.6776 (0.0292)	0.4028 (0.0176)	0.1235 (0.0057)	—
	SAR-PU	0.3441 (0.0396)	0.6546 (0.0195)	0.3607 (0.0180)	0.1078 (0.0039)	—

Condition	Model	AUPRC ↑	AUROC ↑	LL ↓	BS ↓	R@H ↑
Loss 20%	Gene2Wire	0.3592 (0.0385)	0.6794 (0.0333)	0.3274 (0.0185)	0.1001 (0.0057)	0.2091 (0.0268)
	Bilinear	0.3468 (0.0375)	0.6883 (0.0261)	—	—	—
	GenEML	0.3272 (0.0322)	0.6672 (0.0245)	0.4093 (0.0125)	0.1283 (0.0038)	0.1488 (0.0211)
	SAR-PU	0.3270 (0.0360)	0.6412 (0.0128)	0.3584 (0.0302)	0.1058 (0.0070)	0.1636 (0.0297)
Loss 40%	Gene2Wire	0.3575 (0.0383)	0.6789 (0.0326)	0.3321 (0.0202)	0.1019 (0.0064)	0.2651 (0.0229)
	Bilinear	0.3460 (0.0312)	0.6850 (0.0237)	—	—	—
	GenEML	0.3233 (0.0358)	0.6613 (0.0256)	0.4175 (0.0257)	0.1301 (0.0074)	0.2016 (0.0283)
	SAR-PU	0.2701 (0.0073)	0.6286 (0.0185)	0.3248 (0.0183)	0.0960 (0.0045)	0.1901 (0.0157)
Loss 60%	Gene2Wire	0.3496 (0.0368)	0.6769 (0.0349)	0.3411 (0.0250)	0.1040 (0.0079)	0.2940 (0.0207)
	Bilinear	0.3241 (0.0287)	0.6643 (0.0181)	—	—	—
	GenEML	0.2828 (0.0422)	0.6347 (0.0248)	0.4591 (0.0576)	0.1399 (0.0156)	0.2031 (0.0454)
	SAR-PU	0.2486 (0.0065)	0.6181 (0.0170)	0.3534 (0.0244)	0.1000 (0.0050)	0.2041 (0.0151)
Loss 80%	Gene2Wire	0.3355 (0.0358)	0.6492 (0.0279)	0.3668 (0.0377)	0.1095 (0.0099)	0.2847 (0.0302)
	Bilinear	0.2666 (0.0151)	0.6294 (0.0119)	—	—	—
	GenEML	0.2402 (0.0429)	0.5848 (0.0276)	0.5468 (0.0985)	0.1518 (0.0233)	0.2018 (0.0428)
	SAR-PU	—	—	—	—	—
G70%	Gene2Wire	0.3208 (0.0368)	0.6510 (0.0246)	0.3546 (0.0225)	0.1080 (0.0065)	—
	Bilinear	0.3060 (0.0302)	0.6469 (0.0181)	—	—	—
	GenEML	0.2869 (0.0369)	0.6313 (0.0227)	0.4259 (0.0111)	0.1316 (0.0038)	—
	SAR-PU	0.3015 (0.0354)	0.6281 (0.0158)	0.3937 (0.0192)	0.1181 (0.0039)	—
G40%	Gene2Wire	0.2652 (0.0317)	0.5991 (0.0149)	0.3773 (0.0279)	0.1150 (0.0085)	—
	Bilinear	0.2473 (0.0227)	0.5847 (0.0116)	—	—	—
	GenEML	0.2468 (0.0336)	0.5867 (0.0125)	0.4870 (0.0402)	0.1545 (0.0152)	—
	SAR-PU	0.2594 (0.0318)	0.5911 (0.0136)	0.4106 (0.0201)	0.1239 (0.0046)	—
T70%	Gene2Wire	0.3626 (0.0385)	0.6773 (0.0317)	0.3314 (0.0193)	0.1009 (0.0059)	—
	Bilinear	0.3533 (0.0351)	0.6888 (0.0251)	—	—	—
	GenEML	0.3189 (0.0420)	0.6584 (0.0247)	0.4330 (0.0202)	0.1340 (0.0071)	—
	SAR-PU	0.3200 (0.0395)	0.6322 (0.0242)	0.3741 (0.0207)	0.1113 (0.0040)	—
T40%	Gene2Wire	0.3545 (0.0377)	0.6760 (0.0329)	0.3419 (0.0221)	0.1030 (0.0068)	—
	Bilinear	0.3498 (0.0356)	0.6798 (0.0279)	—	—	—
	GenEML	0.2260 (0.0357)	0.5584 (0.0125)	0.9821 (0.0986)	0.2712 (0.0213)	—
	SAR-PU	0.3088 (0.0387)	0.6174 (0.0189)	0.5984 (0.0345)	0.1673 (0.0011)	—

Table 10: **Projection-TAGs: raw values underlying Figure 3.** Four displayed models; 10 panel-seed repetitions. Entries are mean (SE). Values are not heatmap-normalized. A dash denotes unavailable or inapplicable values. R@H is shown only for positive-label loss of 20–80%; gene- and target-only rows do not hide positives.

Condition	Model	AUPRC ↑	AUROC ↑	LL ↓	BS ↓	R@H ↑
Positive-label loss only						
Loss 0%	GenEML	0.1185 (0.0252)	0.6185 (0.0594)	0.3084 (0.0964)	0.0921 (0.0337)	—
	SAR-PU	0.1864 (0.0206)	0.8021 (0.0204)	0.2068 (0.0389)	0.0595 (0.0128)	—
	Bilinear	0.1618 (0.0170)	0.7838 (0.0309)	—	—	—
	Gene2Wire	0.1888 (0.0207)	0.8061 (0.0205)	0.1873 (0.0403)	0.0505 (0.0115)	—
Loss 20%	GenEML	0.1132 (0.0230)	0.6397 (0.0305)	0.2948 (0.0871)	0.0848 (0.0293)	0.0394 (0.0117)
	SAR-PU	0.1843 (0.0191)	0.8010 (0.0211)	0.2084 (0.0397)	0.0606 (0.0134)	0.0748 (0.0145)
	Bilinear	0.1302 (0.0425)	0.6664 (0.0837)	—	—	—
	Gene2Wire	0.1883 (0.0203)	0.8061 (0.0205)	0.1877 (0.0405)	0.0506 (0.0118)	0.0907 (0.0177)
Loss 40%	GenEML	0.1058 (0.0201)	0.6236 (0.0256)	0.2922 (0.0796)	0.0826 (0.0279)	0.0544 (0.0175)
	SAR-PU	0.1838 (0.0190)	0.8008 (0.0220)	0.2076 (0.0392)	0.0597 (0.0130)	0.1171 (0.0177)
	Bilinear	0.1344 (0.0289)	0.7353 (0.0349)	—	—	—
	Gene2Wire	0.1879 (0.0207)	0.8071 (0.0208)	0.1874 (0.0397)	0.0505 (0.0115)	0.1376 (0.0200)
Loss 60%	GenEML	0.1032 (0.0158)	0.5978 (0.0327)	0.2186 (0.0522)	0.0526 (0.0140)	0.0729 (0.0155)
	SAR-PU	0.1635 (0.0209)	0.8212 (0.0045)	0.1605 (0.0239)	0.0438 (0.0078)	0.1299 (0.0259)
	Bilinear	0.1277 (0.0408)	0.5866 (0.1440)	—	—	—
	Gene2Wire	0.1860 (0.0214)	0.8044 (0.0214)	0.1887 (0.0391)	0.0504 (0.0112)	0.1597 (0.0202)
Loss 80%	GenEML	0.0722 (0.0135)	0.5362 (0.0138)	0.3026 (0.1080)	0.0806 (0.0348)	0.0665 (0.0136)
	SAR-PU	0.1707 (0.0175)	0.7876 (0.0236)	0.2083 (0.0396)	0.0575 (0.0120)	0.1686 (0.0231)
	Bilinear	0.1632 (0.0191)	0.7848 (0.0297)	—	—	—
	Gene2Wire	0.1789 (0.0196)	0.8032 (0.0220)	0.1915 (0.0404)	0.0516 (0.0116)	0.1721 (0.0226)
Gene coverage						
G70%	GenEML	0.1072 (0.0217)	0.5943 (0.0553)	0.3125 (0.0957)	0.0925 (0.0341)	—
	SAR-PU	0.1639 (0.0183)	0.7861 (0.0218)	0.2300 (0.0352)	0.0661 (0.0124)	—
	Bilinear	0.1141 (0.0300)	0.6384 (0.0817)	—	—	—
	Gene2Wire	0.1665 (0.0189)	0.7904 (0.0211)	0.1946 (0.0419)	0.0514 (0.0123)	—
G40%	GenEML	0.0945 (0.0190)	0.5633 (0.0478)	0.3176 (0.0950)	0.0930 (0.0337)	—
	SAR-PU	0.1376 (0.0178)	0.7553 (0.0224)	0.2667 (0.0302)	0.0763 (0.0112)	—
	Bilinear	0.0946 (0.0279)	0.5677 (0.1048)	—	—	—
	Gene2Wire	0.1412 (0.0187)	0.7612 (0.0203)	0.2049 (0.0447)	0.0527 (0.0125)	—
Target coverage						
T70%	GenEML	0.1116 (0.0140)	0.6160 (0.0335)	0.4539 (0.1772)	0.1393 (0.0621)	—
	SAR-PU	0.1856 (0.0201)	0.8013 (0.0206)	0.2089 (0.0392)	0.0601 (0.0128)	—
	Bilinear	0.1550 (0.0171)	0.7757 (0.0351)	—	—	—
	Gene2Wire	0.1878 (0.0202)	0.8054 (0.0204)	0.1883 (0.0408)	0.0506 (0.0118)	—
T40%	GenEML	0.0809 (0.0186)	0.5610 (0.0099)	1.4649 (0.2694)	0.2448 (0.0353)	—
	SAR-PU	0.1779 (0.0186)	0.7950 (0.0225)	0.4153 (0.0703)	0.1332 (0.0259)	—
	Bilinear	0.1443 (0.0285)	0.7327 (0.0394)	—	—	—
	Gene2Wire	0.1858 (0.0206)	0.8037 (0.0212)	0.1896 (0.0408)	0.0507 (0.0117)	—

Table 11: **MERGE-seq: raw values underlying Figure 3.** Four displayed models; 10 panel-seed repetitions. Entries are mean (SE). Values are not heatmap-normalized. A dash denotes unavailable or inapplicable values. R@H is shown only for positive-label loss of 20–80%; gene- and target-only rows do not hide positives.

Condition	Model	AUPRC ↑	AUROC ↑	LL ↓	BS ↓	R@H ↑
Positive-label loss only						
Loss 0%	GenEML	0.3673 (0.0425)	0.6925 (0.0359)	0.5203 (0.0621)	0.1679 (0.0219)	—
	SAR-PU	0.4097 (0.0128)	0.7666 (0.0211)	0.4649 (0.0596)	0.1390 (0.0082)	—
	Bilinear	0.3998 (0.0070)	0.7608 (0.0238)	—	—	—
	Gene2Wire	0.4359 (0.0170)	0.7862 (0.0271)	0.3899 (0.0285)	0.1248 (0.0076)	—
Loss 20%	GenEML	0.3586 (0.0401)	0.6865 (0.0316)	0.5085 (0.0584)	0.1592 (0.0182)	0.1340 (0.0353)
	SAR-PU	0.3982 (0.0127)	0.7572 (0.0188)	0.4610 (0.0607)	0.1375 (0.0084)	0.1440 (0.0252)
	Bilinear	0.3924 (0.0057)	0.7563 (0.0239)	—	—	—
	Gene2Wire	0.4411 (0.0173)	0.7876 (0.0227)	0.3879 (0.0276)	0.1243 (0.0073)	0.2192 (0.0135)
Loss 40%	GenEML	0.3421 (0.0370)	0.6752 (0.0311)	0.5335 (0.0483)	0.1626 (0.0135)	0.1889 (0.0404)
	SAR-PU	0.3774 (0.0133)	0.7446 (0.0155)	0.4945 (0.0742)	0.1430 (0.0103)	0.2501 (0.0009)
	Bilinear	0.3861 (0.0038)	0.7516 (0.0235)	—	—	—
	Gene2Wire	0.4330 (0.0178)	0.7837 (0.0226)	0.3914 (0.0280)	0.1254 (0.0078)	0.3026 (0.0124)
Loss 60%	GenEML	0.2980 (0.0407)	0.6341 (0.0423)	0.5933 (0.0583)	0.1796 (0.0161)	0.2087 (0.0488)
	SAR-PU	0.3513 (0.0148)	0.7212 (0.0179)	0.5680 (0.1148)	0.1522 (0.0138)	0.2629 (0.0198)
	Bilinear	0.3770 (0.0036)	0.7392 (0.0233)	—	—	—
	Gene2Wire	0.3960 (0.0077)	0.7669 (0.0214)	0.4035 (0.0313)	0.1286 (0.0086)	0.3255 (0.0144)
Loss 80%	GenEML	0.3096 (0.0213)	0.6502 (0.0245)	0.6465 (0.0799)	0.1819 (0.0155)	0.2699 (0.0275)
	SAR-PU	0.3320 (0.0188)	0.7109 (0.0155)	0.7540 (0.1651)	0.1678 (0.0163)	0.3058 (0.0159)
	Bilinear	0.3494 (0.0204)	0.6979 (0.0056)	—	—	—
	Gene2Wire	0.3943 (0.0073)	0.7713 (0.0219)	0.4016 (0.0275)	0.1283 (0.0081)	0.3605 (0.0119)
Gene coverage						
G70%	GenEML	0.3527 (0.0378)	0.6802 (0.0367)	0.5252 (0.0600)	0.1699 (0.0214)	—
	SAR-PU	0.3991 (0.0143)	0.7580 (0.0209)	0.4679 (0.0533)	0.1443 (0.0084)	—
	Bilinear	0.3867 (0.0038)	0.7493 (0.0197)	—	—	—
	Gene2Wire	0.4219 (0.0156)	0.7791 (0.0248)	0.3952 (0.0275)	0.1270 (0.0076)	—
G40%	GenEML	0.3332 (0.0335)	0.6656 (0.0359)	0.5479 (0.0498)	0.1787 (0.0185)	—
	SAR-PU	0.3832 (0.0128)	0.7477 (0.0205)	0.4625 (0.0442)	0.1453 (0.0078)	—
	Bilinear	0.3711 (0.0069)	0.7393 (0.0171)	—	—	—
	Gene2Wire	0.3945 (0.0139)	0.7658 (0.0223)	0.4001 (0.0247)	0.1288 (0.0075)	—
Target coverage						
T70%	GenEML	0.3624 (0.0324)	0.6883 (0.0276)	0.5077 (0.0544)	0.1641 (0.0196)	—
	SAR-PU	0.3997 (0.0126)	0.7585 (0.0204)	0.4862 (0.0679)	0.1433 (0.0094)	—
	Bilinear	0.3760 (0.0138)	0.7356 (0.0129)	—	—	—
	Gene2Wire	0.4351 (0.0181)	0.7858 (0.0242)	0.3875 (0.0255)	0.1252 (0.0075)	—
T40%	GenEML	0.2652 (0.0153)	0.5790 (0.0284)	1.4502 (0.3406)	0.3056 (0.0123)	—
	SAR-PU	0.3652 (0.0120)	0.7259 (0.0168)	0.7329 (0.0923)	0.1909 (0.0076)	—
	Bilinear	0.3740 (0.0145)	0.7310 (0.0118)	—	—	—
	Gene2Wire	0.4145 (0.0126)	0.7745 (0.0220)	0.3950 (0.0229)	0.1278 (0.0071)	—

Table 12: **SPIDER spatial: raw values underlying Figure 3.** Four displayed models; 10 panel-seed repetitions. Entries are mean (SE). Values are not heatmap-normalized. A dash denotes unavailable or inapplicable values. R@H is shown only for positive-label loss of 20–80%; gene- and target-only rows do not hide positives.

Condition	Model	AUPRC ↑	AUROC ↑	LL ↓	BS ↓	R@H ↑
Positive-label loss only						
Loss 0%	GenEML	0.2417 (0.0201)	0.6703 (0.0025)	0.4435 (0.0173)	0.1386 (0.0078)	—
	SAR-PU	0.3111 (0.0257)	0.7241 (0.0108)	0.3397 (0.0122)	0.1027 (0.0052)	—
	Bilinear	0.2917 (0.0274)	0.7131 (0.0109)	—	—	—
	Gene2Wire	0.3118 (0.0255)	0.7254 (0.0113)	0.3270 (0.0100)	0.0978 (0.0040)	—
Loss 20%	GenEML	0.2213 (0.0151)	0.6451 (0.0058)	0.4912 (0.0597)	0.1537 (0.0226)	0.0893 (0.0097)
	SAR-PU	0.3091 (0.0253)	0.7228 (0.0105)	0.3374 (0.0126)	0.1019 (0.0054)	0.1610 (0.0160)
	Bilinear	0.2871 (0.0265)	0.7085 (0.0113)	—	—	—
	Gene2Wire	0.3108 (0.0256)	0.7257 (0.0107)	0.3276 (0.0103)	0.0980 (0.0041)	0.1688 (0.0180)
Loss 40%	GenEML	0.2073 (0.0135)	0.6227 (0.0074)	0.4952 (0.0574)	0.1543 (0.0201)	0.1279 (0.0074)
	SAR-PU	0.3051 (0.0255)	0.7183 (0.0111)	0.3355 (0.0128)	0.1012 (0.0055)	0.2277 (0.0167)
	Bilinear	0.2897 (0.0247)	0.7071 (0.0110)	—	—	—
	Gene2Wire	0.3092 (0.0244)	0.7231 (0.0108)	0.3276 (0.0111)	0.0980 (0.0043)	0.2375 (0.0186)
Loss 60%	GenEML	0.2099 (0.0168)	0.6231 (0.0082)	0.4628 (0.0310)	0.1393 (0.0114)	0.1664 (0.0120)
	SAR-PU	0.2984 (0.0260)	0.7095 (0.0114)	0.3399 (0.0146)	0.1022 (0.0061)	0.2622 (0.0199)
	Bilinear	0.2691 (0.0235)	0.6895 (0.0110)	—	—	—
	Gene2Wire	0.3084 (0.0249)	0.7225 (0.0100)	0.3283 (0.0113)	0.0982 (0.0044)	0.2765 (0.0188)
Loss 80%	GenEML	0.2072 (0.0183)	0.6229 (0.0102)	0.4323 (0.0121)	0.1201 (0.0071)	0.2053 (0.0130)
	SAR-PU	0.2897 (0.0267)	0.6993 (0.0130)	0.3471 (0.0160)	0.1046 (0.0066)	0.2886 (0.0201)
	Bilinear	0.2736 (0.0273)	0.6880 (0.0141)	—	—	—
	Gene2Wire	0.3047 (0.0252)	0.7187 (0.0115)	0.3301 (0.0109)	0.0989 (0.0042)	0.3062 (0.0202)
Gene coverage						
G70%	GenEML	0.2160 (0.0212)	0.6335 (0.0088)	0.4688 (0.0156)	0.1486 (0.0063)	—
	SAR-PU	0.2774 (0.0243)	0.6943 (0.0100)	0.3560 (0.0126)	0.1081 (0.0054)	—
	Bilinear	0.2576 (0.0235)	0.6801 (0.0102)	—	—	—
	Gene2Wire	0.2795 (0.0242)	0.6971 (0.0105)	0.3384 (0.0105)	0.1014 (0.0042)	—
G40%	GenEML	0.1909 (0.0181)	0.5958 (0.0111)	0.5298 (0.0452)	0.1709 (0.0168)	—
	SAR-PU	0.2393 (0.0224)	0.6534 (0.0102)	0.3801 (0.0115)	0.1159 (0.0053)	—
	Bilinear	0.2398 (0.0215)	0.6579 (0.0086)	—	—	—
	Gene2Wire	0.2442 (0.0222)	0.6621 (0.0102)	0.3495 (0.0107)	0.1049 (0.0044)	—
Target coverage						
T70%	GenEML	0.2175 (0.0136)	0.6356 (0.0060)	0.5028 (0.0530)	0.1583 (0.0196)	—
	SAR-PU	0.3107 (0.0256)	0.7233 (0.0109)	0.3398 (0.0122)	0.1027 (0.0052)	—
	Bilinear	0.2911 (0.0269)	0.7125 (0.0109)	—	—	—
	Gene2Wire	0.3117 (0.0254)	0.7255 (0.0111)	0.3270 (0.0100)	0.0978 (0.0040)	—
T40%	GenEML	0.1633 (0.0069)	0.5452 (0.0101)	0.8155 (0.1192)	0.2522 (0.0361)	—
	SAR-PU	0.3047 (0.0248)	0.7211 (0.0107)	0.5640 (0.0180)	0.1900 (0.0067)	—
	Bilinear	0.2903 (0.0268)	0.7117 (0.0110)	—	—	—
	Gene2Wire	0.3095 (0.0255)	0.7236 (0.0110)	0.3277 (0.0099)	0.0980 (0.0039)	—

Table 13: **BARseq A1: raw values underlying Figure 3.** Four displayed models; 10 panel-seed repetitions. Entries are mean (SE). Values are not heatmap-normalized. A dash denotes unavailable or inapplicable values. R@H is shown only for positive-label loss of 20–80%; gene- and target-only rows do not hide positives.

Condition	Model	AUPRC ↑	AUROC ↑	LL ↓	BS ↓	R@H ↑
Positive-label loss only						
Loss 0%	GenEML	0.2877 (0.0305)	0.5675 (0.0092)	0.6881 (0.0245)	0.2349 (0.0143)	—
	SAR-PU	0.2928 (0.0277)	0.5513 (0.0079)	0.5231 (0.0146)	0.1691 (0.0059)	—
	Bilinear	0.2823 (0.0388)	0.5497 (0.0304)	—	—	—
	Gene2Wire	0.2964 (0.0295)	0.5806 (0.0159)	0.4455 (0.0212)	0.1449 (0.0081)	—
Loss 20%	GenEML	0.2696 (0.0310)	0.5527 (0.0094)	0.7244 (0.0188)	0.2437 (0.0080)	0.0702 (0.0166)
	SAR-PU	0.2827 (0.0258)	0.5433 (0.0031)	0.5154 (0.0281)	0.1637 (0.0088)	0.1240 (0.0209)
	Bilinear	0.2748 (0.0409)	0.5292 (0.0460)	—	—	—
	Gene2Wire	0.2909 (0.0303)	0.5702 (0.0148)	0.4415 (0.0203)	0.1433 (0.0076)	0.1179 (0.0245)
Loss 40%	GenEML	0.2798 (0.0332)	0.5556 (0.0112)	0.6966 (0.0451)	0.2355 (0.0210)	0.1406 (0.0413)
	SAR-PU	0.2747 (0.0302)	0.5333 (0.0184)	0.5367 (0.0285)	0.1625 (0.0105)	0.1561 (0.0281)
	Bilinear	0.2602 (0.0377)	0.5053 (0.0469)	—	—	—
	Gene2Wire	0.2809 (0.0336)	0.5486 (0.0205)	0.4453 (0.0213)	0.1446 (0.0079)	0.1796 (0.0362)
Loss 60%	GenEML	0.2850 (0.0277)	0.5649 (0.0054)	0.7104 (0.0593)	0.2366 (0.0283)	0.2056 (0.0280)
	SAR-PU	0.2700 (0.0316)	0.5300 (0.0207)	0.5953 (0.0362)	0.1668 (0.0106)	0.1867 (0.0265)
	Bilinear	0.2601 (0.0362)	0.5310 (0.0276)	—	—	—
	Gene2Wire	0.2807 (0.0355)	0.5462 (0.0230)	0.4518 (0.0204)	0.1461 (0.0072)	0.2081 (0.0366)
Loss 80%	GenEML	0.2517 (0.0264)	0.5208 (0.0131)	0.7381 (0.0789)	0.2320 (0.0325)	0.1949 (0.0224)
	SAR-PU	0.2629 (0.0325)	0.5259 (0.0187)	0.7410 (0.0453)	0.1825 (0.0131)	0.2145 (0.0310)
	Bilinear	0.2706 (0.0351)	0.5240 (0.0201)	—	—	—
	Gene2Wire	0.2726 (0.0335)	0.5361 (0.0120)	0.4745 (0.0283)	0.1563 (0.0116)	0.1987 (0.0364)
Gene coverage						
G70%	GenEML	0.2770 (0.0329)	0.5634 (0.0097)	0.7153 (0.0178)	0.2486 (0.0096)	—
	SAR-PU	0.2796 (0.0267)	0.5451 (0.0075)	0.5355 (0.0183)	0.1741 (0.0070)	—
	Bilinear	0.2699 (0.0338)	0.5393 (0.0286)	—	—	—
	Gene2Wire	0.2779 (0.0314)	0.5581 (0.0096)	0.4605 (0.0181)	0.1509 (0.0068)	—
G40%	GenEML	0.2656 (0.0273)	0.5442 (0.0110)	0.7230 (0.0281)	0.2515 (0.0134)	—
	SAR-PU	0.2617 (0.0242)	0.5334 (0.0062)	0.5389 (0.0180)	0.1747 (0.0063)	—
	Bilinear	0.2652 (0.0284)	0.5370 (0.0100)	—	—	—
	Gene2Wire	0.2632 (0.0268)	0.5356 (0.0071)	0.4749 (0.0146)	0.1568 (0.0054)	—
Target coverage						
T70%	GenEML	0.2868 (0.0309)	0.5693 (0.0073)	0.6538 (0.0349)	0.2236 (0.0153)	—
	SAR-PU	0.2881 (0.0285)	0.5427 (0.0096)	0.5401 (0.0197)	0.1715 (0.0071)	—
	Bilinear	0.2791 (0.0361)	0.5478 (0.0288)	—	—	—
	Gene2Wire	0.2889 (0.0341)	0.5587 (0.0170)	0.4526 (0.0224)	0.1479 (0.0085)	—
T40%	GenEML	0.2598 (0.0277)	0.5223 (0.0125)	1.0031 (0.1174)	0.2920 (0.0138)	—
	SAR-PU	0.2825 (0.0328)	0.5362 (0.0149)	0.8320 (0.0203)	0.2395 (0.0054)	—
	Bilinear	0.2733 (0.0340)	0.5406 (0.0269)	—	—	—
	Gene2Wire	0.2863 (0.0320)	0.5551 (0.0172)	0.4652 (0.0206)	0.1517 (0.0076)	—

Table 14: **BARseq M1: raw values underlying Figure 3.** Four displayed models; 10 panel-seed repetitions. Entries are mean (SE). Values are not heatmap-normalized. A dash denotes unavailable or inapplicable values. R@H is shown only for positive-label loss of 20–80%; gene- and target-only rows do not hide positives.

Condition	Model	AUPRC ↑	AUROC ↑	LL ↓	BS ↓	R@H ↑
Positive-label loss only						
Loss 0%	GenEML	0.3469 (0.0417)	0.6776 (0.0292)	0.4028 (0.0176)	0.1235 (0.0057)	—
	SAR-PU	0.3441 (0.0396)	0.6546 (0.0195)	0.3607 (0.0180)	0.1078 (0.0039)	—
	Bilinear	0.3532 (0.0370)	0.6955 (0.0234)	—	—	—
	Gene2Wire	0.3661 (0.0384)	0.6814 (0.0304)	0.3262 (0.0168)	0.0996 (0.0052)	—
Loss 20%	GenEML	0.3272 (0.0322)	0.6672 (0.0245)	0.4093 (0.0125)	0.1283 (0.0038)	0.1488 (0.0211)
	SAR-PU	0.3270 (0.0360)	0.6412 (0.0128)	0.3584 (0.0302)	0.1058 (0.0070)	0.1636 (0.0297)
	Bilinear	0.3468 (0.0375)	0.6883 (0.0261)	—	—	—
	Gene2Wire	0.3592 (0.0385)	0.6794 (0.0333)	0.3274 (0.0185)	0.1001 (0.0057)	0.2091 (0.0268)
Loss 40%	GenEML	0.3233 (0.0358)	0.6613 (0.0256)	0.4175 (0.0257)	0.1301 (0.0074)	0.2016 (0.0283)
	SAR-PU	0.2701 (0.0073)	0.6286 (0.0185)	0.3248 (0.0183)	0.0960 (0.0045)	0.1901 (0.0157)
	Bilinear	0.3460 (0.0312)	0.6850 (0.0237)	—	—	—
	Gene2Wire	0.3575 (0.0383)	0.6789 (0.0326)	0.3321 (0.0202)	0.1019 (0.0064)	0.2651 (0.0229)
Loss 60%	GenEML	0.2828 (0.0422)	0.6347 (0.0248)	0.4591 (0.0576)	0.1399 (0.0156)	0.2031 (0.0454)
	SAR-PU	0.2486 (0.0065)	0.6181 (0.0170)	0.3534 (0.0244)	0.1000 (0.0050)	0.2041 (0.0151)
	Bilinear	0.3241 (0.0287)	0.6643 (0.0181)	—	—	—
	Gene2Wire	0.3496 (0.0368)	0.6769 (0.0349)	0.3411 (0.0250)	0.1040 (0.0079)	0.2940 (0.0207)
Loss 80%	GenEML	0.2402 (0.0429)	0.5848 (0.0276)	0.5468 (0.0985)	0.1518 (0.0233)	0.2018 (0.0428)
	SAR-PU	—	—	—	—	—
	Bilinear	0.2666 (0.0151)	0.6294 (0.0119)	—	—	—
	Gene2Wire	0.3355 (0.0358)	0.6492 (0.0279)	0.3668 (0.0377)	0.1095 (0.0099)	0.2847 (0.0302)
Gene coverage						
G70%	GenEML	0.2869 (0.0369)	0.6313 (0.0227)	0.4259 (0.0111)	0.1316 (0.0038)	—
	SAR-PU	0.3015 (0.0354)	0.6281 (0.0158)	0.3937 (0.0192)	0.1181 (0.0039)	—
	Bilinear	0.3060 (0.0302)	0.6469 (0.0181)	—	—	—
	Gene2Wire	0.3208 (0.0368)	0.6510 (0.0246)	0.3546 (0.0225)	0.1080 (0.0065)	—
G40%	GenEML	0.2468 (0.0336)	0.5867 (0.0125)	0.4870 (0.0402)	0.1545 (0.0152)	—
	SAR-PU	0.2594 (0.0318)	0.5911 (0.0136)	0.4106 (0.0201)	0.1239 (0.0046)	—
	Bilinear	0.2473 (0.0227)	0.5847 (0.0116)	—	—	—
	Gene2Wire	0.2652 (0.0317)	0.5991 (0.0149)	0.3773 (0.0279)	0.1150 (0.0085)	—
Target coverage						
T70%	GenEML	0.3189 (0.0420)	0.6584 (0.0247)	0.4330 (0.0202)	0.1340 (0.0071)	—
	SAR-PU	0.3200 (0.0395)	0.6322 (0.0242)	0.3741 (0.0207)	0.1113 (0.0040)	—
	Bilinear	0.3533 (0.0351)	0.6888 (0.0251)	—	—	—
	Gene2Wire	0.3626 (0.0385)	0.6773 (0.0317)	0.3314 (0.0193)	0.1009 (0.0059)	—
T40%	GenEML	0.2260 (0.0357)	0.5584 (0.0125)	0.9821 (0.0986)	0.2712 (0.0213)	—
	SAR-PU	0.3088 (0.0387)	0.6174 (0.0189)	0.5984 (0.0345)	0.1673 (0.0011)	—
	Bilinear	0.3498 (0.0356)	0.6798 (0.0279)	—	—	—
	Gene2Wire	0.3545 (0.0377)	0.6760 (0.0329)	0.3419 (0.0221)	0.1030 (0.0068)	—

Table 15: **Simulation, $\rho = 0$: raw values underlying Figure 2.** Ten generated-data repetitions per sharing strength. Entries are mean (SE). Values are not heatmap-normalized. A dash denotes unavailable or inapplicable scores. R@H is unavailable for Bilinear.

Condition	Model	AUPRC ↑	AUROC ↑	LL ↓	BS ↓	R@H ↑
Positive-label loss only						
Loss 0%	GenEML	0.4315 (0.0068)	0.7245 (0.0055)	0.5692 (0.0128)	0.1885 (0.0049)	—
	SAR-PU	0.4871 (0.0046)	0.7676 (0.0018)	0.4273 (0.0021)	0.1364 (0.0008)	—
	Bilinear	0.4802 (0.0041)	0.7630 (0.0015)	—	—	—
	Gene2Wire	0.4949 (0.0042)	0.7730 (0.0017)	0.4095 (0.0019)	0.1296 (0.0008)	—
Loss 20%	GenEML	0.3710 (0.0091)	0.6680 (0.0088)	0.6137 (0.0120)	0.2061 (0.0051)	0.1459 (0.0080)
	SAR-PU	0.4717 (0.0041)	0.7565 (0.0018)	0.4279 (0.0020)	0.1356 (0.0008)	0.2109 (0.0045)
	Bilinear	0.4662 (0.0059)	0.7511 (0.0033)	—	—	—
	Gene2Wire	0.4865 (0.0043)	0.7664 (0.0017)	0.4159 (0.0018)	0.1318 (0.0007)	0.2263 (0.0062)
Loss 40%	GenEML	0.3253 (0.0071)	0.6164 (0.0081)	0.6143 (0.0088)	0.2067 (0.0038)	0.1733 (0.0051)
	SAR-PU	0.4501 (0.0036)	0.7416 (0.0021)	0.4432 (0.0025)	0.1387 (0.0008)	0.2729 (0.0041)
	Bilinear	0.4434 (0.0054)	0.7337 (0.0027)	—	—	—
	Gene2Wire	0.4728 (0.0038)	0.7578 (0.0018)	0.4250 (0.0019)	0.1349 (0.0008)	0.3029 (0.0052)
Loss 60%	GenEML	0.2867 (0.0049)	0.5708 (0.0054)	0.6071 (0.0047)	0.2011 (0.0017)	0.1900 (0.0052)
	SAR-PU	0.4241 (0.0038)	0.7177 (0.0030)	0.4890 (0.0043)	0.1477 (0.0012)	0.3103 (0.0033)
	Bilinear	0.4058 (0.0051)	0.7014 (0.0043)	—	—	—
	Gene2Wire	0.4543 (0.0037)	0.7435 (0.0021)	0.4392 (0.0023)	0.1397 (0.0009)	0.3432 (0.0033)
Loss 80%	GenEML	0.2755 (0.0040)	0.5597 (0.0054)	0.6141 (0.0045)	0.1869 (0.0027)	0.2128 (0.0053)
	SAR-PU	0.3819 (0.0038)	0.6815 (0.0033)	0.6148 (0.0046)	0.1658 (0.0009)	0.3189 (0.0037)
	Bilinear	0.3154 (0.0070)	0.6048 (0.0078)	—	—	—
	Gene2Wire	0.4105 (0.0038)	0.7061 (0.0039)	0.4727 (0.0033)	0.1510 (0.0013)	0.3503 (0.0033)
Gene coverage						
G70%	GenEML	0.3907 (0.0061)	0.6912 (0.0054)	0.6164 (0.0119)	0.2067 (0.0047)	—
	SAR-PU	0.4356 (0.0033)	0.7294 (0.0021)	0.4599 (0.0019)	0.1475 (0.0008)	—
	Bilinear	0.4370 (0.0044)	0.7321 (0.0028)	—	—	—
	Gene2Wire	0.4448 (0.0042)	0.7381 (0.0027)	0.4309 (0.0023)	0.1368 (0.0009)	—
G40%	GenEML	0.3368 (0.0073)	0.6363 (0.0082)	0.6723 (0.0087)	0.2313 (0.0041)	—
	SAR-PU	0.3988 (0.0041)	0.6969 (0.0031)	0.4723 (0.0031)	0.1518 (0.0012)	—
	Bilinear	0.3678 (0.0053)	0.6705 (0.0051)	—	—	—
	Gene2Wire	0.4078 (0.0055)	0.7053 (0.0037)	0.4474 (0.0029)	0.1423 (0.0011)	—
Target coverage						
T70%	GenEML	0.3712 (0.0122)	0.6665 (0.0110)	0.5938 (0.0235)	0.1996 (0.0098)	—
	SAR-PU	0.4759 (0.0042)	0.7600 (0.0019)	0.4388 (0.0023)	0.1402 (0.0009)	—
	Bilinear	0.4680 (0.0031)	0.7541 (0.0017)	—	—	—
	Gene2Wire	0.4837 (0.0041)	0.7652 (0.0020)	0.4164 (0.0022)	0.1319 (0.0008)	—
T40%	GenEML	0.2719 (0.0041)	0.5543 (0.0049)	0.7593 (0.0118)	0.2681 (0.0032)	—
	SAR-PU	0.4547 (0.0044)	0.7434 (0.0022)	0.4685 (0.0032)	0.1490 (0.0011)	—
	Bilinear	0.4287 (0.0117)	0.7164 (0.0123)	—	—	—
	Gene2Wire	0.4553 (0.0058)	0.7435 (0.0031)	0.4390 (0.0029)	0.1390 (0.0011)	—

Table 16: **Simulation, $\rho = 0.5$: raw values underlying Figure 2.** Ten generated-data repetitions per sharing strength. Entries are mean (SE). Values are not heatmap-normalized. A dash denotes unavailable or inapplicable scores. R@H is unavailable for Bilinear.

Condition	Model	AUPRC ↑	AUROC ↑	LL ↓	BS ↓	R@H ↑
Positive-label loss only						
Loss 0%	GenEML	0.4272 (0.0059)	0.7222 (0.0041)	0.5924 (0.0119)	0.1970 (0.0044)	—
	SAR-PU	0.4835 (0.0038)	0.7648 (0.0015)	0.4299 (0.0019)	0.1374 (0.0007)	—
	Bilinear	0.4842 (0.0041)	0.7655 (0.0017)	—	—	—
	Gene2Wire	0.4965 (0.0042)	0.7735 (0.0016)	0.4098 (0.0015)	0.1297 (0.0005)	—
Loss 20%	GenEML	0.3838 (0.0057)	0.6825 (0.0058)	0.6170 (0.0087)	0.2071 (0.0033)	0.1544 (0.0068)
	SAR-PU	0.4689 (0.0038)	0.7537 (0.0021)	0.4306 (0.0019)	0.1365 (0.0007)	0.2069 (0.0053)
	Bilinear	0.4730 (0.0042)	0.7563 (0.0017)	—	—	—
	Gene2Wire	0.4878 (0.0039)	0.7669 (0.0018)	0.4164 (0.0015)	0.1320 (0.0005)	0.2307 (0.0065)
Loss 40%	GenEML	0.3477 (0.0059)	0.6418 (0.0067)	0.6007 (0.0065)	0.1991 (0.0032)	0.1925 (0.0056)
	SAR-PU	0.4474 (0.0043)	0.7371 (0.0027)	0.4462 (0.0024)	0.1396 (0.0007)	0.2702 (0.0037)
	Bilinear	0.4542 (0.0046)	0.7425 (0.0024)	—	—	—
	Gene2Wire	0.4758 (0.0045)	0.7582 (0.0023)	0.4257 (0.0022)	0.1351 (0.0008)	0.3123 (0.0042)
Loss 60%	GenEML	0.3045 (0.0071)	0.5935 (0.0076)	0.5988 (0.0042)	0.1949 (0.0020)	0.2026 (0.0085)
	SAR-PU	0.4214 (0.0042)	0.7153 (0.0035)	0.4915 (0.0041)	0.1483 (0.0008)	0.3055 (0.0043)
	Bilinear	0.4228 (0.0041)	0.7160 (0.0030)	—	—	—
	Gene2Wire	0.4604 (0.0044)	0.7460 (0.0025)	0.4385 (0.0018)	0.1395 (0.0005)	0.3519 (0.0037)
Loss 80%	GenEML	0.2794 (0.0050)	0.5641 (0.0054)	0.6084 (0.0055)	0.1843 (0.0016)	0.2199 (0.0057)
	SAR-PU	0.3850 (0.0030)	0.6814 (0.0031)	0.6211 (0.0074)	0.1667 (0.0017)	0.3212 (0.0029)
	Bilinear	0.3260 (0.0076)	0.6147 (0.0106)	—	—	—
	Gene2Wire	0.4129 (0.0041)	0.7066 (0.0033)	0.4748 (0.0044)	0.1509 (0.0015)	0.3506 (0.0037)
Gene coverage						
G70%	GenEML	0.3884 (0.0043)	0.6911 (0.0035)	0.6390 (0.0055)	0.2166 (0.0021)	—
	SAR-PU	0.4303 (0.0030)	0.7260 (0.0019)	0.4629 (0.0018)	0.1487 (0.0008)	—
	Bilinear	0.4370 (0.0035)	0.7320 (0.0014)	—	—	—
	Gene2Wire	0.4444 (0.0032)	0.7373 (0.0015)	0.4326 (0.0014)	0.1374 (0.0005)	—
G40%	GenEML	0.3523 (0.0049)	0.6556 (0.0062)	0.6793 (0.0111)	0.2333 (0.0049)	—
	SAR-PU	0.3972 (0.0026)	0.6957 (0.0019)	0.4759 (0.0025)	0.1534 (0.0010)	—
	Bilinear	0.3739 (0.0032)	0.6748 (0.0023)	—	—	—
	Gene2Wire	0.4093 (0.0040)	0.7077 (0.0026)	0.4472 (0.0021)	0.1423 (0.0007)	—
Target coverage						
T70%	GenEML	0.3671 (0.0090)	0.6649 (0.0116)	0.6376 (0.0223)	0.2177 (0.0092)	—
	SAR-PU	0.4747 (0.0034)	0.7586 (0.0016)	0.4409 (0.0020)	0.1409 (0.0007)	—
	Bilinear	0.4734 (0.0045)	0.7581 (0.0019)	—	—	—
	Gene2Wire	0.4830 (0.0040)	0.7643 (0.0014)	0.4190 (0.0014)	0.1328 (0.0005)	—
T40%	GenEML	0.2736 (0.0047)	0.5508 (0.0048)	0.7509 (0.0094)	0.2645 (0.0030)	—
	SAR-PU	0.4515 (0.0033)	0.7422 (0.0021)	0.4707 (0.0027)	0.1500 (0.0009)	—
	Bilinear	0.4423 (0.0063)	0.7304 (0.0071)	—	—	—
	Gene2Wire	0.4565 (0.0037)	0.7451 (0.0019)	0.4392 (0.0020)	0.1391 (0.0007)	—

Table 17: **Simulation, $\rho = 1$: raw values underlying Figure 2.** Ten generated-data repetitions per sharing strength. Entries are mean (SE). Values are not heatmap-normalized. A dash denotes unavailable or inapplicable scores. R@H is unavailable for Bilinear.

Condition	Model	AUPRC ↑	AUROC ↑	LL ↓	BS ↓	R@H ↑
Positive-label loss only						
Loss 0%	GenEML	0.4876 (0.0062)	0.7712 (0.0035)	0.5883 (0.0038)	0.1944 (0.0015)	—
	SAR-PU	0.4782 (0.0049)	0.7636 (0.0027)	0.4315 (0.0022)	0.1381 (0.0008)	—
	Bilinear	0.5069 (0.0056)	0.7827 (0.0023)	—	—	—
	Gene2Wire	0.5177 (0.0061)	0.7900 (0.0024)	0.3988 (0.0015)	0.1260 (0.0006)	—
Loss 20%	GenEML	0.4706 (0.0067)	0.7562 (0.0056)	0.5642 (0.0050)	0.1862 (0.0018)	0.2055 (0.0070)
	SAR-PU	0.4636 (0.0052)	0.7528 (0.0030)	0.4315 (0.0020)	0.1370 (0.0007)	0.1914 (0.0039)
	Bilinear	0.5023 (0.0056)	0.7793 (0.0025)	—	—	—
	Gene2Wire	0.5135 (0.0062)	0.7875 (0.0024)	0.4036 (0.0016)	0.1276 (0.0006)	0.2549 (0.0078)
Loss 40%	GenEML	0.4324 (0.0106)	0.7192 (0.0108)	0.5553 (0.0082)	0.1827 (0.0032)	0.2565 (0.0084)
	SAR-PU	0.4435 (0.0051)	0.7365 (0.0033)	0.4462 (0.0021)	0.1399 (0.0007)	0.2629 (0.0036)
	Bilinear	0.4935 (0.0061)	0.7740 (0.0032)	—	—	—
	Gene2Wire	0.5082 (0.0061)	0.7841 (0.0026)	0.4100 (0.0015)	0.1298 (0.0006)	0.3420 (0.0063)
Loss 60%	GenEML	0.3677 (0.0077)	0.6512 (0.0092)	0.5687 (0.0067)	0.1838 (0.0026)	0.2599 (0.0068)
	SAR-PU	0.4161 (0.0042)	0.7138 (0.0031)	0.4918 (0.0034)	0.1485 (0.0009)	0.2986 (0.0032)
	Bilinear	0.4743 (0.0059)	0.7599 (0.0040)	—	—	—
	Gene2Wire	0.4991 (0.0063)	0.7765 (0.0037)	0.4240 (0.0025)	0.1344 (0.0007)	0.3894 (0.0062)
Loss 80%	GenEML	0.2893 (0.0081)	0.5714 (0.0092)	0.6156 (0.0086)	0.1872 (0.0021)	0.2280 (0.0086)
	SAR-PU	0.3744 (0.0041)	0.6758 (0.0033)	0.6201 (0.0062)	0.1661 (0.0013)	0.3101 (0.0047)
	Bilinear	0.3948 (0.0121)	0.6809 (0.0154)	—	—	—
	Gene2Wire	0.4568 (0.0039)	0.7447 (0.0030)	0.4653 (0.0057)	0.1467 (0.0015)	0.3898 (0.0036)
Gene coverage						
G70%	GenEML	0.4481 (0.0077)	0.7411 (0.0069)	0.6080 (0.0046)	0.2031 (0.0020)	—
	SAR-PU	0.4276 (0.0047)	0.7255 (0.0036)	0.4639 (0.0022)	0.1492 (0.0009)	—
	Bilinear	0.4643 (0.0058)	0.7553 (0.0038)	—	—	—
	Gene2Wire	0.4770 (0.0056)	0.7638 (0.0034)	0.4164 (0.0021)	0.1321 (0.0008)	—
G40%	GenEML	0.4147 (0.0055)	0.7155 (0.0062)	0.6237 (0.0068)	0.2094 (0.0027)	—
	SAR-PU	0.3940 (0.0045)	0.6948 (0.0040)	0.4782 (0.0041)	0.1542 (0.0017)	—
	Bilinear	0.4175 (0.0051)	0.7167 (0.0046)	—	—	—
	Gene2Wire	0.4438 (0.0050)	0.7389 (0.0036)	0.4319 (0.0033)	0.1374 (0.0012)	—
Target coverage						
T70%	GenEML	0.4449 (0.0078)	0.7342 (0.0072)	0.6005 (0.0086)	0.2008 (0.0035)	—
	SAR-PU	0.4683 (0.0049)	0.7564 (0.0030)	0.4427 (0.0020)	0.1419 (0.0008)	—
	Bilinear	0.4937 (0.0116)	0.7714 (0.0094)	—	—	—
	Gene2Wire	0.5132 (0.0061)	0.7873 (0.0026)	0.4026 (0.0013)	0.1273 (0.0005)	—
T40%	GenEML	0.3168 (0.0115)	0.5969 (0.0124)	0.7622 (0.0195)	0.2611 (0.0065)	—
	SAR-PU	0.4465 (0.0048)	0.7391 (0.0034)	0.4732 (0.0030)	0.1511 (0.0011)	—
	Bilinear	0.4848 (0.0056)	0.7669 (0.0038)	—	—	—
	Gene2Wire	0.5001 (0.0070)	0.7785 (0.0034)	0.4167 (0.0019)	0.1315 (0.0007)	—

C.4 Natural Paired Projection-TAGs Readouts

A standard positive indicates that the standard readout detected a barcode. The union reference is positive when either the standard or target-amplified readout is positive. Thus standard positives are contained in the union, while the additional union positives quantify what amplification reveals beyond the standard readout. The target-wise counts below describe these natural paired readouts, separately from the artificial thinning experiments.

Table 18: Projection-TAGs paired-label audit. "Revealed" denotes standard-nondetected but union-positive pairs. The final column is the standard-positive count divided by the union-positive count.

Target	Assayed	Standard +	Union +	Revealed +	Detection ratio
MOp	37,827	1,709	3,999	2,290	0.427
SSp	18,863	823	1,721	898	0.478
VP	37,827	2,136	3,971	1,835	0.538
PAG	37,827	1,348	2,288	940	0.589
MY	37,827	560	1,090	530	0.514
SCL	37,827	229	489	260	0.468
SCS	37,827	267	462	195	0.578
Total	245,825	7,072	14,020	6,948	0.504

C.5 Additional Simulation Comparisons

The family comparisons below vary gene coverage and positive-label loss, complementing the target-coverage comparison in Figure 4. Rows span target-specific, partially shared, and fully shared generating signals. Together, the three intervention axes show how direct PU, shared-only PU-MIRT, and the selected Gene2Wire procedure respond to target structure and measurement availability.

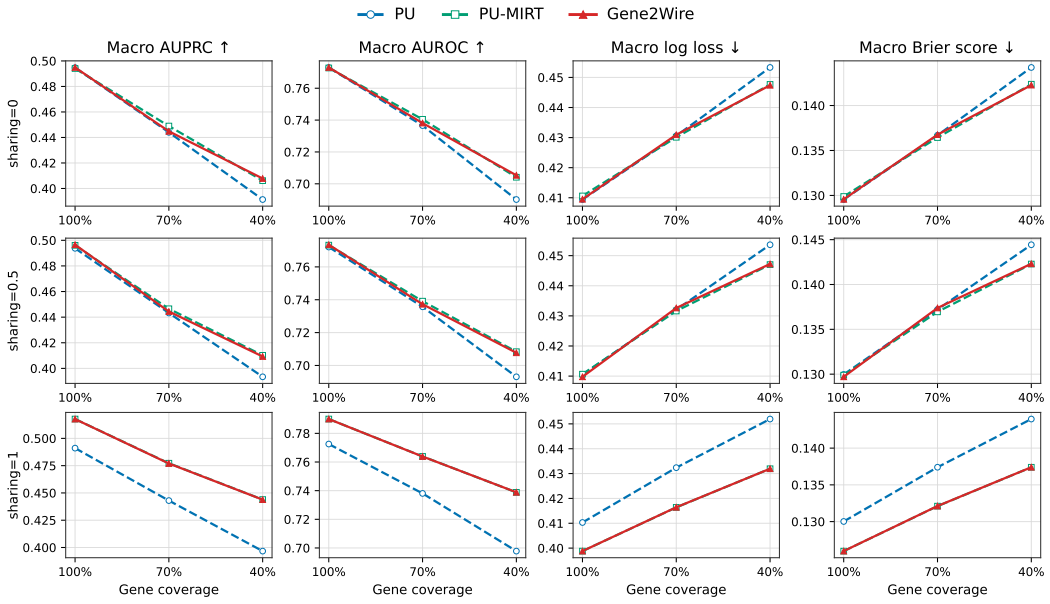

Figure 7: **Structural adaptation under gene removal.** Target coverage and positive retention are full. Rows vary generating sharing strength; columns report ranking and probability metrics for PU, PU-MIRT, and Gene2Wire.

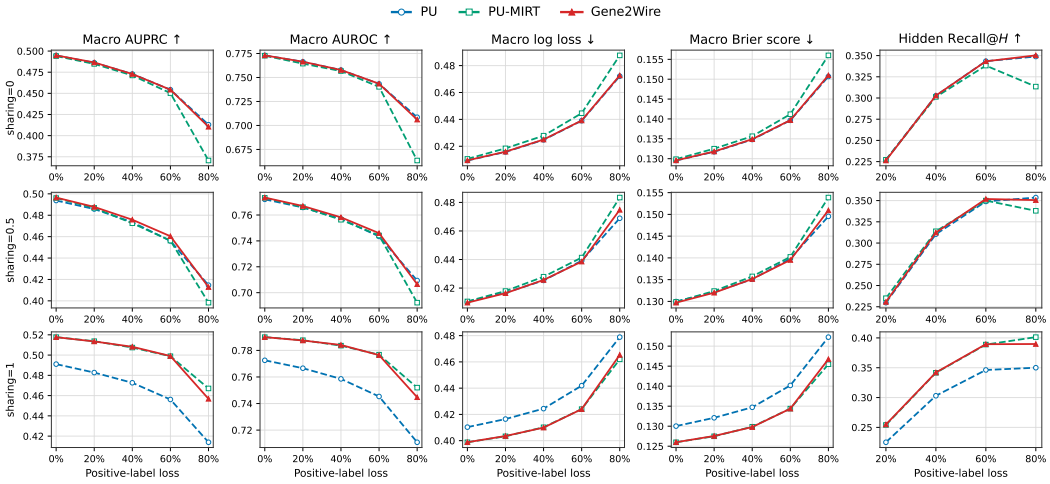

Figure 8: **Structural adaptation under positive-label loss.** Gene and target coverage are full. Rows vary generating sharing strength; columns report ranking and probability metrics for PU, PU-MIRT, and Gene2Wire.