BAMBOOSR: SUPER-RESOLVING DARK MATTER DETECTION WITH PHYSICS-AWARE IMPLICIT NEURAL REPRESENTATION

Anonymous authors
Paper under double-blind review

ABSTRACT

Dark matter direct-detection experiments reconstruct rare particle interactions from faint optical signals. In the PandaX-4T detector, the transverse position of an interaction is encoded in a secondary-scintillation (S2) light pattern that is sparsely and irregularly sampled by 169 top-array photomultiplier tubes, with inactive channels and incomplete coverage further degrading reconstruction near the detector wall. We introduce **BambooSR**, which formulates this readout as a continuous-field super-resolution problem. Rather than mapping sensor counts directly to a position, BambooSR reconstructs an event-conditioned light field over the readout plane, guided by physics priors on direct propagation and boundary reflection, and integrates it over finite sensor footprints to produce readouts on a denser virtual layout, including responses at non-operational sensors. To enable supervised training and evaluation, we construct the first signal super-resolution benchmark for a large liquid-xenon time-projection chamber, comprising 112,890 paired low- and high-resolution readouts from photon-level simulation, together with 5,583 real PandaX-4T Run 2 wall events for downstream evaluation. BambooSR outperforms implicit neural representation and geometric baselines, with the largest gains near the detector wall and around the inactive sensor. On real PandaX-4T wall events, BambooSR improves calibrated position-reconstruction resolution by 39.6%. We will release the code and dataset.

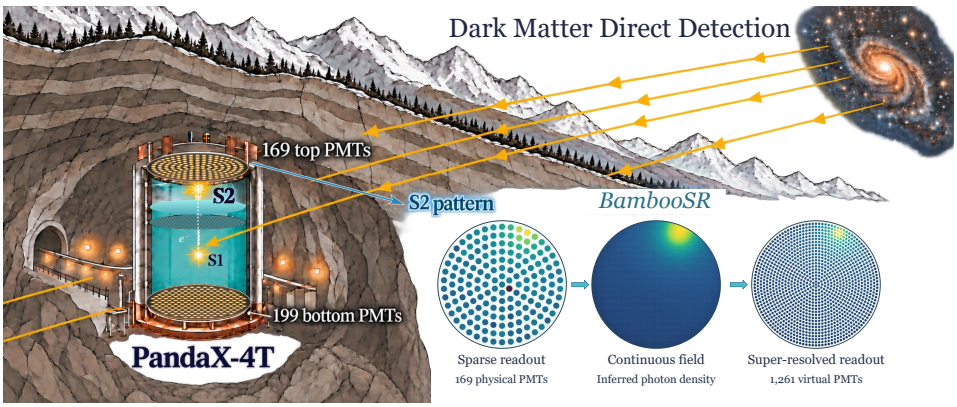


Figure 1: Overview of the scientific setting and the BambooSR pipeline. Sparse top-array S2 readout is reconstructed as a continuous light field and integrated over a virtual dense PMT layout. Detector geometry and signal amplitudes are schematic.

1 INTRODUCTION

Astronomical and cosmological observations indicate that dark matter constitutes roughly 85% of the matter in the Universe (Planck Collaboration, 2020), yet a confirmed direct detection of dark-matter particles remains elusive (Aalbers et al., 2022; PandaX-4T Collaboration, 2021). Searching

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.

for these rare interactions requires reconstructing faint signals recorded by large scientific instruments, a challenge shared with other rare-event measurements (Aalbers et al., 2022). These experiments motivate a broader computational problem: recovering a continuous signal field from weak measurements collected by sparse, irregular sensor arrays. Unlike images, such readouts lie on no regular grid, and the target sensor layout may differ from the physical one in sensor size, number, and placement, which rules out grid-based super-resolution (SR). Implicit neural representation (INR) instead model signals as continuous functions that can be queried at arbitrary locations (Chen et al., 2021b; Zhu et al., 2025; Jiao et al., 2024), making them a natural fit for this setting. We study this AI for Science task in PandaX-4T, using its detector readouts as a concrete setting for signal SR. Fig. 1 summarizes the scientific setting and reconstruction pipeline.

In PandaX-4T, weak secondary scintillation (S2) signals are recorded by the 169 top-array photomultiplier tubes (PMTs) (Li et al., 2024; Qian et al., 2025). These measurements are spatially sparse and aggregated over finite photocathodes, and the physical low-resolution (LR) and virtual high-resolution (HR) layouts do not form nested pixel grids, as illustrated in Fig. 2. Counting fluctuations, inactive channels, and boundary reflections further challenge reconstruction (PandaX-4T Collaboration, 2021; Qian et al., 2025). With hardware replacement or expansion impractical during ongoing operation, these limitations motivate a physics-aware INR that decouples the continuous signal from heterogeneous sensor layouts and uses physical priors to guide recovery in weakly observed regions.

Our goal is to estimate the expected S2 signal density as a continuous field and integrate it over a denser virtual layout to produce a 1,261-channel HR readout. More faithful readouts can benefit position reconstruction, which supports background rejection and spatial-response corrections used in energy estimation (XENON Collaboration, 2019b; Qian et al., 2025; PandaX Collaboration, 2024). We evaluate signal fidelity in simulation and downstream position reconstruction on real data.

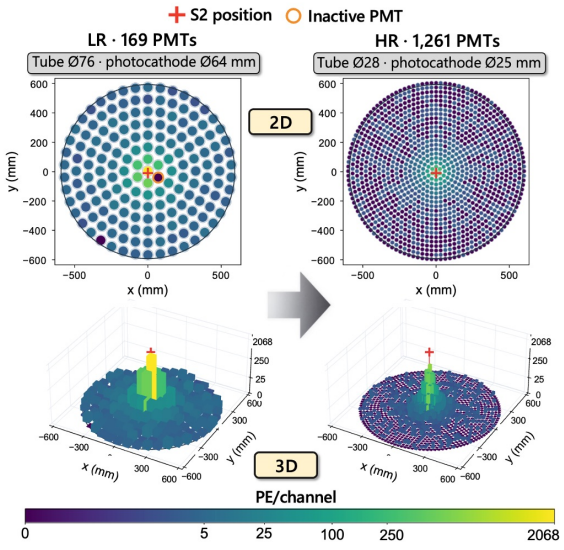


Figure 2: Heterogeneous LR–HR sensing geometry. The physical 169-PMT LR array and the virtual 1,261-PMT HR layout differ in sensor size, number, and placement rather than forming nested grids. ***Top:*** 2D layouts. ***Bottom:*** 3D views. Colors indicate PE/channel; the red cross and orange circle mark the S2 position and the inactive PMT.

Existing approaches address related but distinct reconstruction targets. Response-model and neural estimators infer interaction positions directly from measured patterns, rather than returning HR signal representations for subsequent analysis (Qian et al., 2025; LUX Collaboration, 2018; XENON Collaboration, 2025a; Liang et al., 2022). INRs support continuous SR and scientific field reconstruction, including approaches that incorporate physical structure (Chen et al., 2021b; Zhu et al., 2025; Shen et al., 2025). Particle-detector studies have also explored INR-based forward-response surrogates and SR of discrete sensor or calorimeter outputs (Lei et al., 2022; Yu et al., 2025; Erdmann et al., 2023). However, adopting a continuous representation alone does not specify how weak counts, incomplete coverage, and heterogeneous receiving areas should guide reconstruction. The challenge is to combine the measurement geometry with informative physics priors while retaining the flexibility needed to model event-dependent signal structure.

To address these challenges, we introduce BambooSR, a physics-aware framework for continuous-field signal super-resolution. It conditions an INR on the observed event and incorporates physics priors describing direct propagation and boundary reflection into the learned representation. These priors provide structure without restricting the output to a fixed analytical template, while integration over virtual photocathodes converts the nonnegative field into HR count predictions. Since real HR

readouts cannot be measured, we train on paired LR–HR readouts derived from the same photon realizations in photon-level simulation, and reserve real PandaX-4T wall events for evaluation. Our contributions are summarized as follows:

- We propose **BambooSR**, the first framework incorporating physics priors into implicit neural representations for particle-detector signal super-resolution.
- We construct a PandaX-4T signal SR benchmark comprising 112,890 paired simulated LR–HR readouts, documented sensor geometries, and real wall-event observations for downstream evaluation. We will publicly release the benchmark and evaluation protocols to support reproducible research on sparse, heterogeneous detector measurements.
- Experiments demonstrate that our method achieves superior performance on HR count-reconstruction error. Controlled ablations validate the contribution of physics priors, and evaluation on real PandaX-4T data yields a narrower calibrated wall-event distribution as a proxy for downstream position reconstruction performance.

2 Related Work

2.1 Implicit Neural Representations for Continuous Super-Resolution

Implicit neural representations (INRs) model signals as continuous coordinate-based functions, and their ability to recover fine spatial structure has been improved through several complementary directions. Early work addressed spectral bias through Fourier-feature coordinate encodings and periodic activations (Tancik et al., 2020; Sitzmann et al., 2020), followed by multiplicative filtering, band-limited architectures, wavelet and variable-periodic activations, spatial adaptation, and optimization strategies for high-frequency recovery (Fathony et al., 2021; Lindell et al., 2022; Saragadam et al., 2023; Liu et al., 2024; Feng et al., 2026; Shi et al., 2025). Continuous implicit functions have also become a practical representation for arbitrary-scale super-resolution, with LIIF providing an influential formulation for querying signals beyond a fixed output grid (Chen et al., 2021b). This paradigm has since expanded across scientific domains, including spatial transcriptomics, volumetric and biomedical imaging, and continuous reconstruction of physical flow fields and time-varying simulation data (Zhu et al., 2025; Wu et al., 2023; Xu et al., 2023; Ye et al., 2023; Fang et al., 2024; Jiang et al., 2020; Jiao et al., 2024; Tang & Wang, 2024; Shen et al., 2025). Together, these studies establish INR-based super-resolution and continuous reconstruction as a broadly useful paradigm for scientific measurements; our work extends this direction to weak, sparse particle-detector signals acquired under heterogeneous sensing layouts and incomplete observations.

2.2 Particle-Detector Signal Reconstruction and Super-Resolution

Large xenon time-projection chambers have become established rare-event observatories, enabling leading dark-matter searches and measurements of other rare processes (Liu et al., 2017; XENON Collaboration, 2019a; PandaX Collaboration, 2023a). Position reconstruction in these detectors has traditionally relied on calibrated photon- or light-response models (LUX Collaboration, 2018; PandaX Collaboration, 2021; Qian et al., 2025), while learned approaches directly infer interaction positions from S2 patterns using fully connected, convolutional, and graph-constrained neural architectures (XENON Collaboration, 2019b; 2025a; Liang et al., 2022). These methods primarily target downstream position reconstruction rather than reconstructing a denser signal representation. Related work in particle detection has also explored INRs as differentiable forward surrogates for photon propagation and super-resolution models that predict virtual sensor responses or finer calorimeter images (Lei et al., 2022; Yu et al., 2025; Erdmann et al., 2023). These approaches respectively model forward detector response or discrete sensor/image outputs. In contrast, BambooSR reconstructs an event-conditioned continuous signal field from sparse PMT measurements, incorporates physics priors directly into the INR, and integrates the recovered field over a heterogeneous virtual sensor layout for super-resolution.

Table 1: Comparison with representative data resources from large liquid-xenon TPC experiments. Our paired LR–HR supervision is constructed from photon-level simulation[3], whereas real PandaX-4T wall events are provided separately for downstream evaluation.

Resource	Main use	Real data	Paired LR–HR	Signal SR
XENONnT (XENON Collaboration, 2025b)	Physics analysis	Yes	No	No
PandaX-4T (PandaX-4T Collaboration, 2021)	Physics analysis	Yes	No	No
LZ (LUX-ZEPLIN Collaboration, 2023)	Physics analysis	Yes	No	No
Ours	**Signal SR**	**Yes**	**Yes**	**Yes**

3 PandaX-4T Signal Super-Resolution Benchmark

Existing public data resources from large liquid-xenon time-projection chambers (TPCs) primarily support physics analyses and reinterpretation rather than sensor-level super-resolution (XENON Collaboration, 2025b; PandaX-4T Collaboration, 2021; LUX-ZEPLIN Collaboration, 2023). To our knowledge, we introduce the first benchmark specifically designed for signal super-resolution in a large liquid-xenon TPC, providing paired event-level LR–HR sensor readouts together with real detector data for downstream evaluation. PandaX-4T records S2 signals with a fixed top array of 169 three-inch Hamamatsu R11410-23 PMTs (Yan et al., 2026; Luo et al., 2024); therefore, a corresponding real HR measurement cannot be acquired for the same physical event. We instead construct paired supervision using BambooMC (Chen et al., 2021a), the Geant4-based simulation framework developed for the PandaX experiments.[1] For each simulated event, photon landing positions at the readout plane are aggregated independently over the physical LR and virtual HR photocathode geometries. The HR target is therefore obtained from the same photon realization rather than through interpolation or synthetic downsampling of an LR detector image. Moreover, the virtual HR layout follows the footprint of a physically available smaller PMT, the Hamamatsu R7057 (Hamamatsu Photonics), yielding 1,261 channels over the same readout region instead of imposing a conventional $2\times$ or $4\times$ image-scale factor.

The resulting benchmark contains 112,890 paired events over 11,289 simulated source positions, with ten independent photon realizations per position. We split the data by source position into 90,330 training, 11,280 validation, and 11,280 test events, preventing repeated realizations from the same spatial location from crossing data partitions. To complement simulated supervision, we additionally include 5,583 selected PandaX-4T Run 2 wall events for downstream position reconstruction. These events contain real 169-channel detector measurements but have neither HR readout ground truth nor truth interaction coordinates, and are therefore reserved for real-data evaluation rather than supervised training. We will release the paired benchmark, sensor geometries, fixed data splits, selected real wall-event sample, and evaluation protocols.

4 BambooSR

BambooSR combines an event-conditioned INR with a physics-feature branch and a finite-area readout operator. Fig. 3 summarizes the model and its training with HR count supervision.

4.1 Problem Formulation

For event e, let $\widetilde{\mathbf{q}}_e = \mathbf{m} \odot \mathbf{q}_e$ denote the LR readout after applying the fixed operational mask $\mathbf{m}$, and let $Q_e = \sum_i \widetilde{q}_{ei}$ denote its total observed signal. We seek a nonnegative, event-conditioned signal density $\lambda_\theta(\mathbf{u} \mid \widetilde{\mathbf{q}}_e)$ over the readout plane, measured in PE/mm^2. For a virtual sensor with photocathode region $\mathcal{B}_j$, the expected count is

$$\mu_{e,j} = \int_{\mathcal{B}_j} \lambda_\theta(\mathbf{u} \mid \widetilde{\mathbf{q}}_e)\, \mathrm{d}A. \tag{1}$$

The target $y_{e,j}$ is the simulated photon count within the same region. Supervision therefore constrains finite-area measurements rather than pointwise ground-truth field values, separating the con-

[1] BambooSR is named after BambooMC (Chen et al., 2021a), which provides the photon-level simulation used to construct its paired supervision.

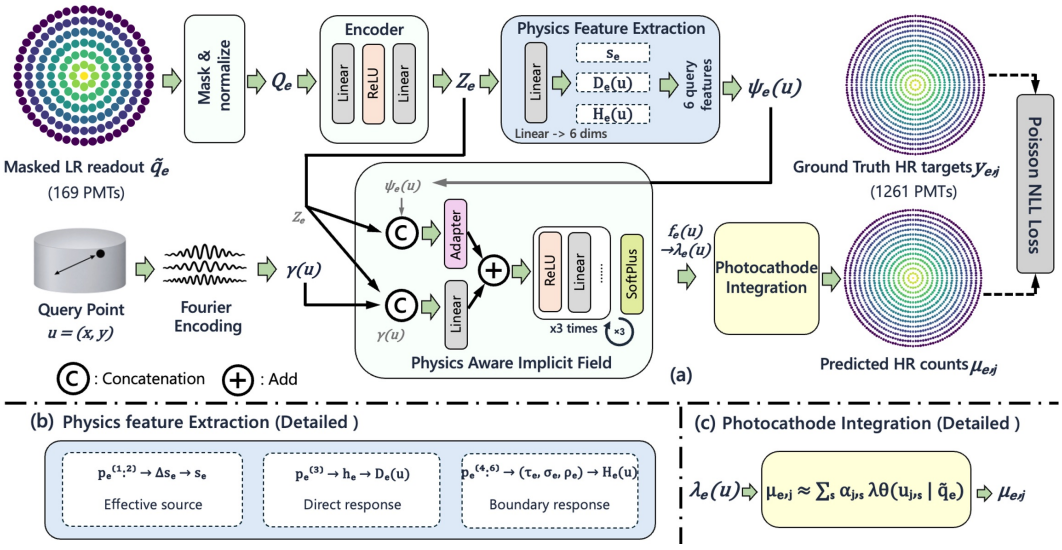


Figure 3: Overview of BambooSR. The physics branch constructs effective-source, direct-response, and boundary-response features that adapt the implicit field, while finite-area photocathode integration produces HR count predictions. Detailed model design is provided in Appendix B.

tinuous representation from the discrete layout on which it is read out. Throughout this section, θ denotes all trainable model parameters.

4.2 Physics-Aware Conditional Implicit Field

Event conditioning. A shared encoder combines normalized channel measurements with the log-transformed total signal,

$$\mathbf{z}_e = E_\phi\left(\left[\frac{\widetilde{\mathbf{q}}_e}{\max(Q_e, \epsilon)}, \frac{\log(1+Q_e)}{10}\right]\right). \tag{2}$$

Spatial queries use fixed Gaussian Fourier features $\gamma(\mathbf{u})$ (Tancik et al., 2020). The event representation $\mathbf{z}_e$ conditions both the coordinate decoder and the physics-feature branch. The same network is shared across events, rather than fitted separately to each event.

Physics features. Sparse observations provide limited guidance in weakly measured regions. We augment the coordinate representation with event-dependent features describing source geometry, direct response, and boundary response. A linear head followed by fixed output transformations predicts a componentwise bounded source correction, positive response scales h_e, τ_e, σ_e, and a mixing coefficient $\rho_e \in (0, 1)$. The effective source is anchored to the observed charge centroid:

$$\mathbf{s}_e = \frac{\sum_i \widetilde{q}_{ei}\mathbf{c}_i}{\max(Q_e, \epsilon)} + \Delta\mathbf{s}(\mathbf{z}_e), \tag{3}$$

where $\mathbf{c}_i$ is the center of LR sensor i. These are learned effective parameters; no ground-truth source coordinates are supplied as inputs or supervision.

The direct-response and boundary-response features are

$$\begin{aligned} D_e(\mathbf{u}) &= \frac{h_e}{2\pi\left(\|\mathbf{u}-\mathbf{s}_e\|^2 + h_e^2\right)^{3/2}}, \\ \pi_{e,k} &= \mathrm{softmax}_k\left(-\frac{\|\mathbf{b}_k-\mathbf{s}_e\|^2}{2\tau_e^2}\right), \\ H_e(\mathbf{u}) &= (1-\rho_e)\sum_{k=1}^{K}\pi_{e,k}\frac{\exp\left[-\|\mathbf{u}-\mathbf{b}_k\|^2/(2\sigma_e^2)\right]}{2\pi\sigma_e^2} + \frac{\rho_e}{\pi R^2}. \end{aligned} \tag{4}$$

Here, $\mathbf{b}_k$ are fixed anchors on a reference circle of radius R. The scales τ_e and σ_e control source-dependent anchor weighting and spatial spreading, respectively; the final term provides a spatially constant component. These are phenomenological features, not calibrated transport probabilities or imposed boundary conditions.

For each query, we assemble the six-dimensional feature vector

$$\boldsymbol{\psi}_e(\mathbf{u}) = \left[\frac{\mathbf{u} - \mathbf{s}_e}{S}, \log\big(1 + S^2 D_e(\mathbf{u})\big), \log\big(1 + S^2 H_e(\mathbf{u})\big), d_R(\mathbf{u}), d_R(\mathbf{s}_e) \right], \tag{5}$$

where S is a coordinate scale and $d_R(\mathbf{x}) = (R - \|\mathbf{x}\|)/S$ is the signed radial gap relative to the reference circle. The features combine source-relative geometry, compressed response terms, and radial context.

Feature-level adaptation. A compact adapter maps the event representation and physics features to an additive correction at the first hidden preactivation:

$$\begin{aligned} \mathbf{a}_e(\mathbf{u}) &= W_0[\gamma(\mathbf{u}), \mathbf{z}_e] + \mathbf{b}_0 + A_\eta([\mathbf{z}_e, \boldsymbol{\psi}_e(\mathbf{u})]), \\ f_\theta(\mathbf{u} \mid \widetilde{\mathbf{q}}_e) &= \mathrm{softplus}(G_\omega(\mathrm{ReLU}(\mathbf{a}_e(\mathbf{u})))), \end{aligned} \tag{6}$$

where G_ω comprises the remaining decoder layers. Only the adapter's final linear layer is zero-initialized, preserving the base conditional INR at initialization. Joint optimization subsequently learns the contribution of the physics-feature branch. Because the priors modify hidden features rather than prescribe the output profile, the decoder retains the flexibility to compensate for their approximations.

4.3 Finite-Area Readout and Training Loss

The internal field f_θ is nonnegative but is not constrained to integrate to one. Its physical density is $\lambda_\theta = (Q_e/A_{\mathrm{ref}}) f_\theta$, where A_{ref} is the fixed total nominal LR photocathode area, including the masked channel. Using precomputed quadrature points $\mathbf{u}_{js}$ and area weights α_{js}, we evaluate Equation 1 as

$$\begin{aligned} \mu_{e,j} &\simeq \sum_s \alpha_{js}\, \lambda_\theta(\mathbf{u}_{js} \mid \widetilde{\mathbf{q}}_e) \\ &= Q_e \sum_s w_{js}\, f_\theta(\mathbf{u}_{js} \mid \widetilde{\mathbf{q}}_e), \qquad w_{js} = \frac{\alpha_{js}}{A_{\mathrm{ref}}}. \end{aligned} \tag{7}$$

This aggregation accounts for the finite sensor footprints. Scaling by Q_e does not enforce equality between the predicted HR total and the observed LR total.

We learn conditional expected counts with an HR Poisson negative log-likelihood (NLL) loss. For an event minibatch $\mathcal{E}_b$ and a shared sampled subset of HR channels $\mathcal{J}_b$,

$$\mathcal{L}_{\mathrm{HR}} = \frac{1}{|\mathcal{E}_b|\,|\mathcal{J}_b|} \sum_{e \in \mathcal{E}_b} \sum_{j \in \mathcal{J}_b} \left[\mu_{e,j} - y_{e,j} \log \max(\mu_{e,j}, \epsilon) \right], \tag{8}$$

where $\epsilon > 0$ ensures numerical stability and the target-only log-factorial term is omitted. All trainable components are optimized jointly using HR count supervision, without an LR-consistency term or auxiliary physics loss. Thus, physical information enters through the representation rather than an additional supervision loss. Validation evaluates the complete HR layout, and checkpoint selection uses validation NLL.

5 Experiments

We evaluate BambooSR at two levels: HR signal reconstruction on paired simulation data and downstream position reconstruction on real PandaX-4T wall events. Controlled ablations examine the contributions of the physics features within the proposed architecture.

Implementation details. We use the source-position-disjoint benchmark split described in Section 3, evaluating all methods on the same 11,280 test events and 1,261 HR channels. Geometric baselines comprise area integration, nearest-neighbor mapping, and inverse-distance weighting (IDW) (Shepard, 1968). The INR baselines are FFN, SIREN, WIRE, and SASNet, adapted to event-conditioned count prediction with the same target sensing geometry (Tancik et al., 2020; Sitzmann et al., 2020; Saragadam et al., 2023; Feng et al., 2026). For BambooSR and its ablations, we train from scratch with Adam at a learning rate of 10^{-4}, batch size 8, and 32 sampled HR channels per training step. Training runs for at most 160 epochs, with checkpoint selection by validation NLL and an early-stopping patience of 15 epochs. The reported runs use seed 42; the comparisons below are point estimates rather than averages over repeated training seeds.

Evaluation metrics. We report eight complementary metrics. RMSE, MAE, and mean Poisson deviance quantify channel-level count accuracy. Total relative MAE is the mean absolute event-total error divided by the corresponding target total, expressed as a percentage. Absolute total bias instead measures the magnitude of the aggregate signed count error, normalized by the aggregate target count. Shape TV is the mean total variation distance between predicted and target channel distributions after event-wise normalization; events with zero target total are excluded from the relative-total and shape averages. PMT3-region RMSE is evaluated over HR channels whose centers lie within 32 mm of the inactive LR PMT center. Edge RMSE is evaluated on events with simulated source radius $r \geq 520$ mm. RMSE and MAE, including their regional variants, are reported in PE/channel; lower values are better for all eight metrics. Best results are shaded as first and second.

5.1 Signal Super-Resolution

Table 2: Signal super-resolution on the simulated test set. PMT3-region and edge errors measure performance near the inactive channel and for near-wall source positions, respectively.

Method	Pixel-level Fidelity			Event-level Fidelity			Challenging Regions	
	RMSE↓ [PE/ch]	MAE↓ [PE/ch]	Poisson Dev.↓	Total Rel. MAE↓ [%]	|Total Bias|↓ [%]	Shape TV↓	PMT3 RMSE↓ [PE/ch]	Edge RMSE↓ [PE/ch]
Area Integration	15.252	1.709	16.384	53.845	53.654	0.483	18.555	13.295
Nearest Neighbor	12.423	1.683	2.423	7.131	1.615	0.391	16.459	11.353
IDW	12.227	1.581	1.922	4.882	2.418	0.369	15.033	11.027
SIREN Sitzmann et al. (2020)	3.180	0.690	0.797	3.616	0.806	0.183	4.669	4.296
FFN Tancik et al. (2020)	2.546	0.664	0.770	3.017	1.139	0.175	3.327	2.713
WIRE Saragadam et al. (2023)	2.411	0.645	0.762	3.690	1.746	**0.169**	3.096	2.888
SASNet Feng et al. (2026)	3.528	0.703	0.801	4.386	2.645	0.185	4.143	4.617
BambooSR (Ours)	**2.185**	**0.638**	**0.756**	**2.580**	**0.052**	**0.169**	**2.745**	**2.311**

Pixel-level fidelity. Table 2 compares BambooSR with the INR and geometric baselines. BambooSR achieves an RMSE of 2.185 PE/channel, compared with 2.411 for WIRE, the strongest competing INR on this metric, corresponding to an approximately 9.4% reduction. It also achieves lower MAE and Poisson deviance: 0.638 and 0.756, respectively, compared with 0.645 and 0.762 for WIRE. The geometric baselines yield substantially larger count errors, showing that spatial mapping alone is insufficient for the reconstruction task represented by this benchmark.

Event-level and challenging-region fidelity. Event-level metrics evaluate whether the reconstruction preserves both the total signal and its normalized spatial distribution, complementing channel-level errors. BambooSR achieves the best total relative MAE and absolute total bias, while tying WIRE for the best Shape TV at 0.169. Challenging-region metrics isolate reconstruction near the inactive PMT and detector boundary, where sparse or incomplete sensing is most pronounced; BambooSR achieves the lowest RMSE in both regions.

5.2 Real-Data Position Reconstruction

We select BambooSR, WIRE, and FFN for downstream evaluation based on both competitive upstream signal-reconstruction accuracy and coherent event-level total-PE calibration, including in the

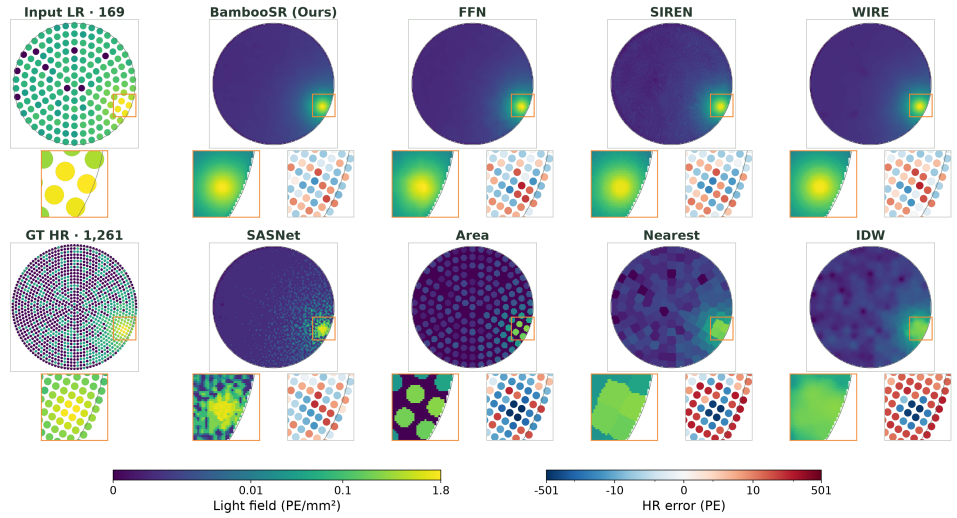


Figure 4: Representative boundary event comparing reconstructed signal fields and local HR errors.

low-PE regime (Appendix E.3), and evaluate them on 5,583 selected PandaX-4T Run 2 wall events. Their reconstructed HR readouts are passed to the same charge-centroid (COG) position estimator and compared with the original 169-channel LR readout.

Because interaction-position ground truth is unavailable, we assess wall-position resolution from the calibrated radial wall-residual distribution rather than a supervised two-dimensional position error. With two-fold, 24-sector radial cross-calibration, we define $\Delta r_e = r_{e,\mathrm{corr}} - R_{\mathrm{wall}}(\varphi_e)$. We use the calibrated inward width σ_{L} obtained from a split-normal fit as the primary resolution metric. To complement this fit-based measure, we additionally report the fit-free 68% containment half-width $w_{68} = (Q_{84} - Q_{16})/2$ and the mean absolute error (MAE), both computed from the same calibrated residuals.

Fig. 5 compares the calibrated wall-residual distributions. All three SR readouts produce narrower inward distributions than the original LR readout. Table 3 reports the corresponding resolution metrics and σ_{L} reductions relative to LR COG. BambooSR achieves the smallest σ_{L}, reducing it from 11.01 to 6.65 mm (39.6%), while also yielding the smallest fit-free 68% containment half-width of 4.19 mm and the lowest MAE of 3.57 mm. These calibrated wall-residual metrics should be interpreted as proxies for wall-position resolution rather than absolute two-dimensional position errors.

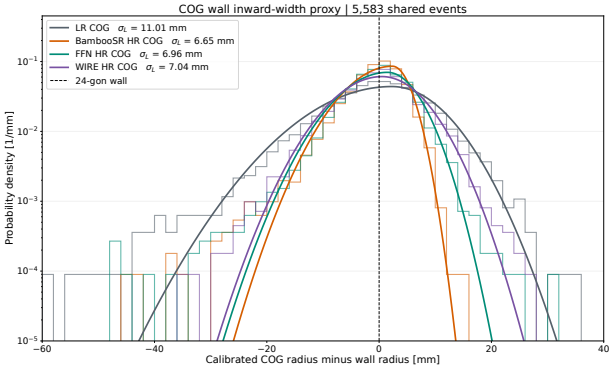


Figure 5: Calibrated wall-residual distributions of 4 readouts for 5,583 shared PandaX-4T Run 2 wall events.

Table 3: Real-data wall-position resolution after two-fold, 24-sector radial cross-calibration. σ_{L} red. is relative to LR COG.

Readout	σ_{L} [mm]↓	σ_{L} red. [%]↑	w_{68} [mm]↓	MAE [mm]↓
LR COG	11.01	–	8.16	6.88
WIRE	7.04	36.1	5.85	4.95
FFN	6.96	36.8	4.83	4.18
Ours	**6.65**	**39.6**	**4.19**	**3.57**

6 Ablation Study

We compare the full model with two controlled variants to assess the explicit boundary-response feature and the use of physical rather than generic learned features. All three variants are trained from scratch using the same data split and training protocol.

Table 4: **Ablation of the physics-feature branch.** Lower is better for all metrics. *Direct-only* removes the explicit boundary-response feature while retaining the other geometric features. *Neural control* replaces the physical features with generic event- and coordinate-dependent features while preserving the head and adapter dimensions.

Variant	Pixel-level Fidelity			Event-level Fidelity			Challenging Regions	
	RMSE↓ [PE/ch]	MAE↓ [PE/ch]	Poisson Dev.↓	Total Rel. MAE↓ [%]	|Total Bias|↓ [%]	Shape TV↓	PMT3 RMSE↓ [PE/ch]	Edge RMSE↓ [PE/ch]
Direct-only	2.576	0.669	0.769	2.699	0.628	0.178	2.965	2.707
Neural control	2.421	0.650	0.766	2.785	0.609	0.172	3.093	2.501
BambooSR (Full)	**2.185**	**0.638**	**0.756**	**2.580**	**0.052**	**0.169**	**2.745**	**2.311**

Boundary-response feature. Direct-only retains the effective-source construction, direct propagation, source-relative coordinates, and radial-gap features, but removes the explicit boundary-response term. As shown in Table 4, restoring this term reduces overall RMSE from 2.576 to 2.185 PE/channel, edge RMSE from 2.707 to 2.311 PE/channel, and PMT3-region RMSE from 2.965 to 2.745 PE/channel. These results support the value of the boundary-response feature within the proposed representation, without implying that its phenomenological kernels constitute an exact transport model.

Physics features vs. generic adaptation. Neural control retains the same head and adapter dimensions but replaces the six physics-derived features with generic event- and coordinate-dependent features. The full model reduces RMSE from 2.421 to 2.185 PE/channel and improves all eight reported metrics relative to this control. Because the adaptation capacity is preserved, these gains indicate that the physics-informed feature content itself contributes beyond the addition of an auxiliary adaptation branch. Together with the Direct-only comparison, this supports the importance of both the explicit boundary-response term and the broader physics-aware representation.

7 Conclusion

We presented BambooSR, a physics-aware implicit neural representation for signal super-resolution from sparse and heterogeneous particle-detector readouts. BambooSR reconstructs an event-conditioned continuous signal field, incorporates physics priors through feature-level adaptation, and integrates the recovered field over finite virtual photocathodes to predict HR sensor counts. We further introduced the PandaX-4T Signal Super-Resolution Benchmark, combining photon-level paired LR–HR simulation with real Run 2 wall events for downstream evaluation.

On the simulated test set, BambooSR reduces RMSE by approximately 9.4% relative to the strongest evaluated INR baseline and reduces edge RMSE by approximately 14.8% relative to the strongest competing baseline in the near-wall regime. On 5,583 real PandaX-4T wall events, BambooSR improves the wall-event position resolution from $\sigma_{\mathrm{L}} = 11.01$ mm for the original LR readout to 6.65 mm, corresponding to a 39.6% reduction. Together, these results show that improved signal reconstruction in simulation transfers to enhanced position reconstruction on real detector data.

Limitations. The paired HR supervision is constructed from photon-level simulation, and the current study evaluates a fixed LR–HR sensing configuration rather than cross-layout generalization. The physics features are phenomenological rather than a calibrated transport model. For the real wall events, event-level ground-truth interaction coordinates are unavailable; therefore, the reported position resolution is determined from the calibrated radial spread of wall events rather than direct two-dimensional position errors. Future work will extend the continuous formulation to additional sensing geometries and broader real-detector validation.

AI USE STATEMENT

Generative AI tools were used to assist with research ideation and experimental planning, literature discovery, code development and debugging, scientific figure preparation, manuscript drafting, and language polishing. All AI-assisted outputs were reviewed and verified by the authors against the underlying source code, experimental results, and cited literature where applicable. The authors take full responsibility for the final content of this work.

REFERENCES

J Aalbers et al. A next-generation liquid xenon observatory for dark matter and neutrino physics. *Journal of Physics G: Nuclear and Particle Physics*, 50(1):013001, dec 2022. doi: 10.1088/1361-6471/ac841a. URL https://doi.org/10.1088/1361-6471/ac841a.

X. Chen, C. Cheng, M. Fu, F. Giuliani, J. Liu, X. Lu, X. Ji, Z. Qian, H. Qiao, Q. Wang, J. Xia, P. Xie, Y. Yao, and H. Zhang. BambooMC—a Geant4-based simulation program for the PandaX experiments. *Journal of Instrumentation*, 16:T09004, Sep 2021a. doi: 10.1088/1748-0221/16/09/T09004. URL https://doi.org/10.1088/1748-0221/16/09/T09004.

Yinbo Chen, Sifei Liu, and Xiaolong Wang. Learning continuous image representation with local implicit image function. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 8628–8638, June 2021b.

Johannes Erdmann, Aaron van der Graaf, Florian Mausolf, and Olaf Nackenhorst. SR-GAN for SR-gamma: Super resolution of photon calorimeter images at collider experiments. *The European Physical Journal C*, 83(11):1001, 2023. doi: 10.1140/epjc/s10052-023-12178-3. URL https://doi.org/10.1140/epjc/s10052-023-12178-3.

Wei Fang, Yuxing Tang, Heng Guo, Mingze Yuan, Tony C. W. Mok, Ke Yan, Jiawen Yao, Xin Chen, Zaiyi Liu, Le Lu, Ling Zhang, and Minfeng Xu. Cycleinr: Cycle implicit neural representation for arbitrary-scale volumetric super-resolution of medical data. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 11631–11641, June 2024.

Rizal Fathony, Anit Kumar Sahu, Devin Willmott, and J Zico Kolter. Multiplicative filter networks. In *International Conference on Learning Representations*, 2021. URL https://openreview.net/forum?id=OmtmcPkkhT.

Haoan Feng, Diana Aldana, Tiago Novello, and Leila De Floriani. Sasnet: Spatially-adaptive sinusoidal networks for inrs. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 41964–41973, June 2026.

Hamamatsu Photonics. Photomultiplier tube R7057. URL https://www.hamamatsu.com/us/en/product/optical-sensors/pmt/pmt_tube-alone/head-on-type/R7057.html. Accessed: 2026-09-22.

Chiyu "Max" Jiang, Soheil Esmaeilzadeh, Kamyar Azizzadenesheli, Karthik Kashinath, Mustafa Mustafa, Hamdi A. Tchelepi, Philip Marcus, Mr Prabhat, and Anima Anandkumar. Meshfreeflownet: A physics-constrained deep continuous space-time super-resolution framework. In *SC20: International Conference for High Performance Computing, Networking, Storage and Analysis*, pp. 1–15, 2020.

Chenyue Jiao, Chongke Bi, and Lu Yang. FFEINR: Flow feature-enhanced implicit neural representation for spatiotemporal super-resolution. *Journal of Visualization*, 27(2):273–289, 2024. doi: 10.1007/s12650-024-00959-1. URL https://doi.org/10.1007/s12650-024-00959-1.

Minjie Lei, Ka Vang Tsang, Sean Gasiorowski, Chuan Li, Youssef Nashed, Gianluca Petrillo, Olivia Piazza, Daniel Ratner, and Kazuhiro Terao. Implicit neural representation as a differentiable surrogate for photon propagation in a monolithic neutrino detector, 2022. URL https://arxiv.org/abs/2211.01505.

Jiafu Li et al. Waveform simulation in pandax-4t. *Chinese Physics C*, 48(7):073001, 2024. doi: 10.1088/1674-1137/ad380f.

Shixiao Liang, Aaron Higuera, Christina Peters, Venkat Roy, Waheed U. Bajwa, Hagit Shatkay, and Christopher D. Tunnell. Domain-informed neural networks for interaction localization within astroparticle experiments. *Frontiers in Artificial Intelligence*, Volume 5 - 2022, 2022. ISSN 2624-8212. doi: 10.3389/frai.2022.832909. URL https://www.frontiersin.org/journals/artificial-intelligence/articles/10.3389/frai.2022.832909.

David B. Lindell, Dave Van Veen, Jeong Joon Park, and Gordon Wetzstein. Bacon: Band-limited coordinate networks for multiscale scene representation. In *CVPR*, 2022.

Jianglai Liu, Xun Chen, and Xiangdong Ji. Current status of direct dark matter detection experiments. *Nature Physics*, 13(3):212–216, 2017. doi: 10.1038/nphys4039. URL https://doi.org/10.1038/nphys4039.

Zhen Liu, Hao Zhu, Qi Zhang, Jingde Fu, Weibing Deng, Zhan Ma, Yanwen Guo, and Xun Cao. Finer: Flexible spectral-bias tuning in implicit neural representation by variableperiodic activation functions. In *2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 2713–2722, 2024. doi: 10.1109/CVPR52733.2024.00262.

Lingyin Luo, Deqing Fang, Ke Han, Di Huang, Xiaofeng Shang, Anqing Wang, Qiuhong Wang, Shaobo Wang, Siguang Wang, Xiang Xiao, Binbin Yan, and Xiyu Yan. Improvement on the linearity response of pandax-4t with new photomultiplier tube bases. *Journal of Instrumentation*, 19(05):P05021, may 2024. doi: 10.1088/1748-0221/19/05/P05021. URL https://doi.org/10.1088/1748-0221/19/05/P05021.

LUX Collaboration. Position reconstruction in lux. *Journal of Instrumentation*, 13(02):P02001, feb 2018. doi: 10.1088/1748-0221/13/02/P02001. URL https://doi.org/10.1088/1748-0221/13/02/P02001.

LUX-ZEPLIN Collaboration. First dark matter search results from the lux-zeplin (lz) experiment. *Phys. Rev. Lett.*, 131:041002, Jul 2023. doi: 10.1103/PhysRevLett.131.041002. URL https://link.aps.org/doi/10.1103/PhysRevLett.131.041002.

PandaX-4T Collaboration. Dark matter search results from the pandax-4t commissioning run. *Physical Review Letters*, 127(26), December 2021. ISSN 1079-7114. doi: 10.1103/physrevlett.127.261802. URL http://dx.doi.org/10.1103/PhysRevLett.127.261802.

PandaX-4T Collaboration. Low radioactive material screening and background control for the pandax-4t experiment. *Journal of High Energy Physics*, 2022(6):147, June 2022. doi: 10.1007/JHEP06(2022)147.

PandaX Collaboration. Horizontal position reconstruction in pandax-ii. *Journal of Instrumentation*, 16(11):P11040, nov 2021. doi: 10.1088/1748-0221/16/11/P11040. URL https://doi.org/10.1088/1748-0221/16/11/P11040.

PandaX Collaboration. Limits on the luminance of dark matter from xenon recoil data. *Nature*, 618(7963):47–50, 2023a. doi: 10.1038/s41586-023-05982-0. URL https://doi.org/10.1038/s41586-023-05982-0.

PandaX Collaboration. Search for light dark matter with ionization signals in the pandax-4t experiment. *Physical Review Letters*, 130:261001, 2023b. doi: 10.1103/PhysRevLett.130.261001.

PandaX Collaboration. Signal response model in PandaX-4T. *Physical Review D*, 110:023029, 2024. doi: 10.1103/PhysRevD.110.023029.

Planck Collaboration. Planck 2018 results: Vi. cosmological parameters. *Astronomy & Astrophysics*, 641:A6, 2020. ISSN 1432-0746. doi: 10.1051/0004-6361/201833910. URL http://dx.doi.org/10.1051/0004-6361/201833910.

Zhicheng Qian et al. Position reconstruction and surface background model for the pandax-4t detector. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 1077:170548, August 2025. ISSN 0168-9002. doi: 10.1016/j.nima.2025.170548. URL http://dx.doi.org/10.1016/j.nima.2025.170548.

Vishwanath Saragadam, Daniel LeJeune, Jasper Tan, Guha Balakrishnan, Ashok Veeraraghavan, and Richard G. Baraniuk. Wire: Wavelet implicit neural representations. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 18507–18516, June 2023.

Liming Shen, Liang Deng, Chongke Bi, Yu Wang, Xinhai Chen, Yueqing Wang, and Jie Liu. PEINR: A physics-enhanced implicit neural representation for high-fidelity flow field reconstruction. In Aarti Singh, Maryam Fazel, Daniel Hsu, Simon Lacoste-Julien, Felix Berkenkamp, Tegan Maharaj, Kiri Wagstaff, and Jerry Zhu (eds.), *Proceedings of the 42nd International Conference on Machine Learning*, volume 267 of *Proceedings of Machine Learning Research*, pp. 54364–54380. PMLR, 13–19 Jul 2025. URL https://proceedings.mlr.press/v267/shen25a.html.

Donald Shepard. A two-dimensional interpolation function for irregularly-spaced data. In *Proceedings of the 1968 23rd ACM National Conference*, ACM '68, pp. 517–524, New York, NY, USA, 1968. Association for Computing Machinery. ISBN 9781450374866. doi: 10.1145/800186.810616. URL https://doi.org/10.1145/800186.810616.

Kexuan Shi, Hai Chen, Leheng Zhang, and Shuhang Gu. Inductive gradient adjustment for spectral bias in implicit neural representations. In Aarti Singh, Maryam Fazel, Daniel Hsu, Simon Lacoste-Julien, Felix Berkenkamp, Tegan Maharaj, Kiri Wagstaff, and Jerry Zhu (eds.), *Proceedings of the 42nd International Conference on Machine Learning*, volume 267 of *Proceedings of Machine Learning Research*, pp. 54864–54891. PMLR, 13–19 Jul 2025. URL https://proceedings.mlr.press/v267/shi25a.html.

Vincent Sitzmann, Julien Martel, Alexander Bergman, David Lindell, and Gordon Wetzstein. Implicit neural representations with periodic activation functions. In H. Larochelle, M. Ranzato, R. Hadsell, M.F. Balcan, and H. Lin (eds.), *Advances in Neural Information Processing Systems*, volume 33, pp. 7462–7473. Curran Associates, Inc., 2020. URL https://proceedings.neurips.cc/paper_files/paper/2020/file/53c04118df112c13a8c34b38343b9c10-Paper.pdf.

Matthew Tancik, Pratul Srinivasan, Ben Mildenhall, Sara Fridovich-Keil, Nithin Raghavan, Utkarsh Singhal, Ravi Ramamoorthi, Jonathan Barron, and Ren Ng. Fourier features let networks learn high frequency functions in low dimensional domains. In H. Larochelle, M. Ranzato, R. Hadsell, M.F. Balcan, and H. Lin (eds.), *Advances in Neural Information Processing Systems*, volume 33, pp. 7537–7547. Curran Associates, Inc., 2020. URL https://proceedings.neurips.cc/paper_files/paper/2020/file/55053683268957697aa39fba6f231c68-Paper.pdf.

Kaiyuan Tang and Chaoli Wang. Stsr-inr: Spatiotemporal super-resolution for multivariate time-varying volumetric data via implicit neural representation. *Computers Graphics*, 119:103874, 2024. ISSN 0097-8493. doi: https://doi.org/10.1016/j.cag.2024.01.001. URL https://www.sciencedirect.com/science/article/pii/S0097849324000013.

Qing Wu, Yuwei Li, Yawen Sun, Yan Zhou, Hongjiang Wei, Jingyi Yu, and Yuyao Zhang. An arbitrary scale super-resolution approach for 3d mr images via implicit neural representation. *IEEE Journal of Biomedical and Health Informatics*, 27(2):1004–1015, 2023. doi: 10.1109/JBHI.2022.3223106.

XENON Collaboration. Observation of two-neutrino double electron capture in ^{124}Xe with xenon1t. *Nature*, 568(7753):532–535, 2019a. doi: 10.1038/s41586-019-1124-4. URL https://doi.org/10.1038/s41586-019-1124-4.

XENON Collaboration. Xenon1t dark matter data analysis: Signal reconstruction, calibration, and event selection. *Phys. Rev. D*, 100:052014, Sep 2019b. doi: 10.1103/PhysRevD.100.052014. URL https://link.aps.org/doi/10.1103/PhysRevD.100.052014.

XENON Collaboration. Xenonnt analysis: Signal reconstruction, calibration, and event selection. *Phys. Rev. D*, 111:062006, Mar 2025a. doi: 10.1103/PhysRevD.111.062006. URL https://link.aps.org/doi/10.1103/PhysRevD.111.062006.

XENON Collaboration. First search for light dark matter in the neutrino fog with xenonnt. *Phys. Rev. Lett.*, 134:111802, Mar 2025b. doi: 10.1103/PhysRevLett.134.111802. URL https://link.aps.org/doi/10.1103/PhysRevLett.134.111802.

Junshen Xu, Daniel Moyer, Borjan Gagoski, Juan Eugenio Iglesias, P. Ellen Grant, Polina Golland, and Elfar Adalsteinsson. Nesvor: Implicit neural representation for slice-to-volume reconstruction in mri. *IEEE Transactions on Medical Imaging*, 42(6):1707–1719, 2023. doi: 10.1109/TMI.2023.3236216.

Binbin Yan, Di Huang, Jianglai Liu, Xiaoying Lu, Anqing Wang, and Mengjiao Xiao. Characterization of Hamamatsu R11410-23 photomultiplier tubes and performance in the PandaX-4T detector. *Radiation Detection Technology and Methods*, Aug 2026. doi: 10.1007/s41605-026-00709-x. URL https://doi.org/10.1007/s41605-026-00709-x.

Siqi Ye, Liyue Shen, Md Tauhidul Islam, and Lei Xing. Super-resolution biomedical imaging via reference-free statistical implicit neural representation. *Physics in Medicine Biology*, 68(20): 205020, oct 2023. doi: 10.1088/1361-6560/acfdf1. URL https://doi.org/10.1088/1361-6560/acfdf1.

Felix J. Yu, Nicholas Kamp, and Carlos A. Argüelles. Enhancing events in neutrino telescopes through deep-learning-driven superresolution. *Phys. Rev. D*, 111:L041301, Feb 2025. doi: 10.1103/PhysRevD.111.L041301. URL https://link.aps.org/doi/10.1103/PhysRevD.111.L041301.

Qingtian Zhu, Yumin Zheng, Yuling Sang, Yifan Zhan, Ziyan Zhu, Jun Ding, and Yinqiang Zheng. SUICA: Learning super-high dimensional sparse implicit neural representations for spatial transcriptomics. In Aarti Singh, Maryam Fazel, Daniel Hsu, Simon Lacoste-Julien, Felix Berkenkamp, Tegan Maharaj, Kiri Wagstaff, and Jerry Zhu (eds.), *Proceedings of the 42nd International Conference on Machine Learning*, volume 267 of *Proceedings of Machine Learning Research*, pp. 80448–80462. PMLR, 13–19 Jul 2025. URL https://proceedings.mlr.press/v267/zhu25af.html.

APPENDIX

This appendix supplements the benchmark construction in Section 3, the method in Section 4, and the evaluation in Sections 5 and 6.

A DETECTOR AND BENCHMARK DETAILS

A.1 SIGNAL FORMATION AND DETECTOR READOUT

PandaX-4T is a dual-phase liquid-xenon time-projection chamber (TPC) designed for rare-event searches, with weakly interacting massive particles (WIMPs) among its primary physics targets. The detector response itself, however, is not specific to dark matter: energy depositions in liquid xenon can arise from candidate dark-matter interactions as well as ordinary radioactive backgrounds, neutrons, calibration sources, and neutrino-induced interactions (Aalbers et al., 2022). BambooSR therefore operates on the detector signal representation rather than on a particular particle identity.

An interaction in liquid xenon produces both prompt scintillation photons and ionization electrons. The prompt scintillation forms the primary signal, S1, whereas the ionization electrons drift upward under the electric field and are extracted into the gaseous xenon, where proportional electroluminescence produces the delayed secondary signal, S2 (Aalbers et al., 2022). The S1–S2 time separation provides information along the drift direction, while the S2 light distribution is used to reconstruct the transverse interaction position (PandaX-4T Collaboration, 2021; Qian et al., 2025).

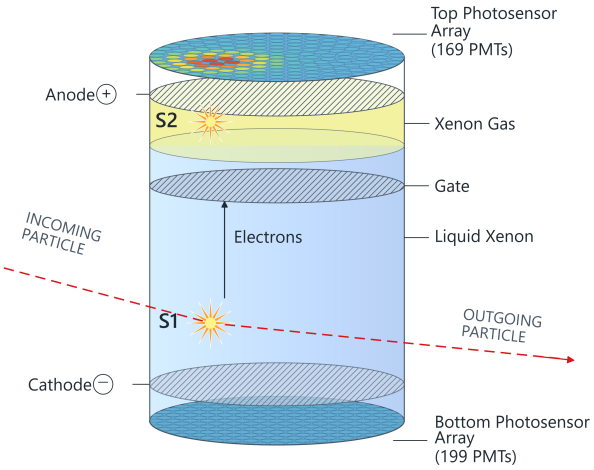


Figure 6: S1 and S2 signal formation in the PandaX-4T TPC. Adapted from Aalbers et al. (2022), Fig. 4, under the CC BY 4.0 license.

For transverse position reconstruction, S2 provides the relevant spatial signature for this work. The ionization electrons drift approximately along the vertical electric field while retaining the transverse coordinates of the original interaction. After extraction into the gas phase, the resulting electroluminescence is generated near the liquid–gas interface above the interaction position, producing a localized light pattern across the top photosensor array. PandaX-4T consequently reconstructs the horizontal position from the top-array S2 charge pattern (PandaX-4T Collaboration, 2021; Qian et al., 2025). Fig. 6 summarizes the signal-formation process underlying the input representation used by BambooSR.

A.2 PandaX-4T Detector and Sensing Geometry

The PandaX-4T TPC contains approximately 3.7 tonnes of liquid xenon in its sensitive volume and is bounded by reflective PTFE panels and electrodes that establish the drift and extraction fields (Luo et al., 2024). Three-inch Hamamatsu R11410-23 PMTs are installed above and below the active xenon volume, with 169 PMTs in the top array and 199 PMTs in the bottom array (Luo et al., 2024). An overview of the detector structure is shown in Fig. 7.

Although both PMT arrays collect scintillation light, their roles in event reconstruction are different. The S2 light pattern is more concentrated on the top array and varies strongly with the transverse interaction position, making it the primary input for horizontal position reconstruction. The bottom array receives a more spatially uniform S2 response and is used primarily for energy reconstruction (Luo et al., 2024). Accordingly, the present study uses the 169-channel top-array S2 readout as its observed detector input.

The physical LR and virtual HR configurations considered in this work form *co-centered heterogeneous sensing layouts* over the same circular readout region. They share the detector center and sensing domain but differ in sensor footprint, channel count, sensor-center locations, and spatial sampling density. In particular, the HR configuration is not a nested pixel grid or a fixed $2\times/4\times$ magnification of the LR layout. This distinction motivates representing the underlying S2 response as a continuous field that can be integrated over different finite-area sensor geometries. The construction of the paired LR–HR measurements is detailed in Section A.3.

A.3 Photon-Level LR–HR Pair Construction

The physical PandaX-4T top array provides only one fixed sensing configuration, so a denser sensor readout cannot be acquired simultaneously for the same real interaction. We therefore construct paired LR–HR supervision from photon-level BambooMC simulation rather than from image-space interpolation.

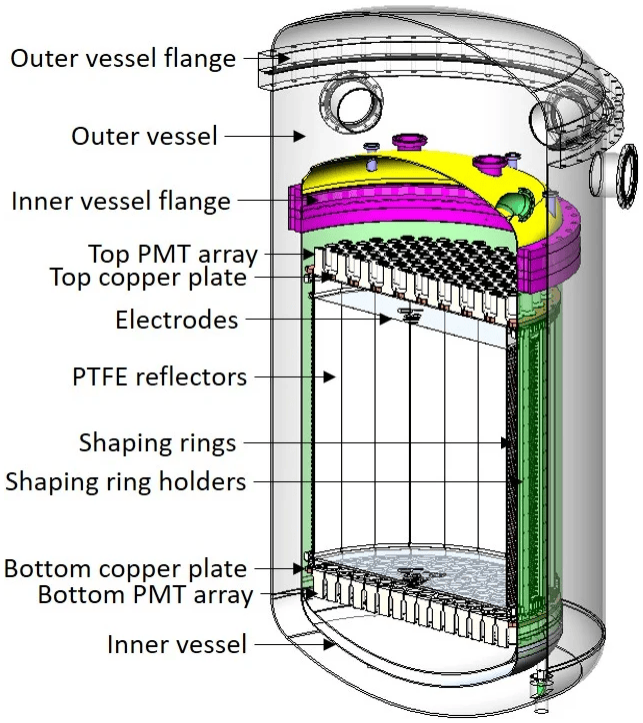


Figure 7: Layout of the PandaX-4T detector. Reproduced from PandaX-4T Collaboration (2022), Fig. 1, under the CC BY 4.0 license.

For each simulated event, photon landing positions on the top readout plane are retained before channel aggregation. The same propagated photon realization is then read out under two sensor geometries. Photons falling within the physical 169-PMT photocathodes form the LR observation, while the same stored photon coordinates are independently aggregated over a virtual 1,261-channel geometry to form the HR target. LR and HR are therefore two geometry-derived finite-area readouts of the same propagated photon realization, rather than a low-resolution image and an interpolated or synthetically resized counterpart.

The virtual HR geometry is anchored to a physically available smaller photosensor rather than to an arbitrary image-scale factor. We use the Hamamatsu R7057 as a geometric reference. Its manufacturer specification gives a 28-mm tube diameter and a 25-mm active photocathode diameter, substantially smaller than the 76-mm tube and approximately 64-mm active photocathode of the R11410 family used in PandaX-4T (Hamamatsu Photonics; Yan et al., 2026). Packing this smaller active footprint over the same circular readout region yields 1,261 virtual channels. The resulting HR configuration is therefore determined by a physically realizable sensor scale rather than by choosing a conventional $2\times$ or $4\times$ image magnification factor.

Importantly, the R7057 is used only to define the virtual sensor footprint and layout. The HR labels do not simulate a separate R7057-specific photo-detection chain. Instead, they are obtained by counting the same propagated photon coordinates within the virtual photocathode regions. R7057-specific quantum efficiency, collection efficiency, gain, and electronics response are therefore not modeled in the HR target. The HR labels should consequently be interpreted as geometry-derived virtual sensor counts rather than as a complete simulation of a small-PMT detector readout.

Because the physical LR and virtual HR layouts have different photocathode footprints and spatial coverage, their total sensor counts are not constrained to be identical for the same photon realization. Together, the two arrays form co-centered heterogeneous sensing layouts that share the detector center and readout domain but differ in sensor footprint, channel count, and sampling locations.

This construction motivates the continuous representation adopted by BambooSR. Rather than predicting a fixed HR vector directly, the model reconstructs an event-conditioned signal field and obtains each sensor response by integrating the field over its finite photocathode region.

A.4 Masking, Signal-Yield Variation, and Dataset Splits

A fixed channel ordering is used throughout data preparation, training, and evaluation so that PMT counts, sensor coordinates, and detector identities remain aligned. The non-operational physical PMT is masked only in the observed LR input, before the event total and charge centroid used by BambooSR are computed. The corresponding spatial region remains fully supervised in the virtual HR target, and all 1,261 HR channels contribute to training. The masked region therefore provides a direct test of whether the model can infer signal structure that is not locally observed by the physical LR array.

The simulated benchmark contains 11,289 source positions. At each position, we generate ten independent events with stochastic photon realizations and varying signal strength. For each event, a nominal PE-equivalent light-yield scale is sampled uniformly from 100 to 5,000 PE. This nominal scale is converted to the number of emitted optical photons using the empirical 6.5% conversion adopted in the simulation before photon propagation.

The 100–5,000 PE interval therefore specifies the simulation input scale rather than a hard bound on the final LR or virtual HR sensor counts. After propagation, the realized readouts are determined by the stochastic photon landing coordinates together with the corresponding photocathode geometry. Consequently, the final LR or HR total can fall outside the nominal generation interval, and the LR and HR totals need not match for the same propagated photon realization.

The lower end of this nominal light-yield range is deliberately included to expose the model to weak-signal conditions, where limited photon statistics make spatial reconstruction more difficult. This regime is scientifically relevant to low-energy rare-event searches. For context, the PandaX-4T light-dark-matter analysis selected ionization-only events with S2 signals of only 60–200 PE (PandaX Collaboration, 2023b). The nominal generation scale used in our simulation is not defined as the same detector-level observable interval. Rather, its lower end is intended to provide substantial exposure to weak-signal realizations, while the broader range prevents the learned representation from being restricted to a single signal scale.

Generating multiple realizations at each spatial location also prevents the benchmark from reducing to a deterministic mapping between source position and sensor pattern. Events generated at the same coordinate can differ in signal scale, photon propagation, and resulting sensor counts. However, because these events share the same underlying spatial condition, randomly splitting individual events would allow nearly identical source locations to appear in both training and evaluation. We therefore partition the benchmark by source position rather than by event. The resulting split contains 90,330 training events, 11,280 validation events, and 11,280 held-out test events, with no source position shared across partitions.

Simulated interaction coordinates are retained only as evaluation metadata and are never supplied to the SR model as an inference input or position-supervision target. The real PandaX-4T Run 2 wall events are likewise excluded from supervised SR training and are reserved for downstream real-data position-reconstruction evaluation.

B Model Details and Design Insights

B.1 Event Conditioning and INR Backbone

BambooSR is a shared conditional implicit neural representation (INR) across events, rather than a separate network fitted independently to each event. Let $\widetilde{\mathbf{q}}_e \in \mathbb{R}^{169}_{\geq 0}$ denote the masked LR readout of event e, and define the observed event total as

$$Q_e = \sum_{i=1}^{169} \widetilde{q}_{e,i}. \tag{9}$$

The event encoder receives the normalized channel fractions together with a compressed representation of the total signal,

$$\mathbf{z}_e = E_\phi\left(\left[\frac{\widetilde{\mathbf{q}}_e}{\max(Q_e,\epsilon)}, \frac{\log(1+Q_e)}{10}\right]\right), \tag{10}$$

where $\mathbf{z}_e \in \mathbb{R}^{32}$. The encoder consists of a $170 \to 128$ linear layer, a ReLU activation, and a $128 \to 32$ linear layer.

Spatial coordinates are represented using fixed Gaussian Fourier features following Tancik et al. (2020). For a query coordinate $\mathbf{u} = (x, y)$ in millimeters, the coordinate is normalized by $S = 650$ mm and projected using 64 fixed Gaussian frequencies with standard deviation 3. The sine and cosine responses are concatenated to form a 128-dimensional coordinate representation $\gamma(\mathbf{u})$.

The base conditional INR takes $[\gamma(\mathbf{u}), \mathbf{z}_e]$ as input and uses three width-128 ReLU hidden layers followed by a scalar Softplus output. The network parameters are shared across all events, while $\mathbf{z}_e$ conditions the reconstructed field on the observed S2 pattern.

B.2 Event-Dependent Physics Context

The physics branch first predicts an event-dependent context from the latent representation $\mathbf{z}_e$. A linear head

$$P_\xi : \mathbb{R}^{32} \to \mathbb{R}^6 \tag{11}$$

produces six values $\mathbf{p}_e$. These six head outputs should be distinguished from the six query-dependent physics features introduced below.

The first two head outputs define a bounded correction to the charge-weighted center of gravity (COG), yielding an effective source location

$$\mathbf{s}_e = \sum_i \frac{\widetilde{q}_{e,i}}{\max(Q_e,\epsilon)}\mathbf{c}_i + 120\,\tanh\left(\mathbf{p}_e^{(1:2)}\right), \tag{12}$$

where $\mathbf{c}_i$ denotes the center of LR sensor i. The simulated truth interaction coordinate is not used in this construction.

The remaining four head outputs are mapped to positive or bounded event-dependent response parameters,

$$h_e = 10 + 100\,\mathrm{softplus}(p_e^{(3)}), \tag{13}$$

$$\tau_e = 10 + 100\,\mathrm{softplus}(p_e^{(4)}), \tag{14}$$

$$\sigma_e = 10 + 100\,\mathrm{softplus}(p_e^{(5)}), \tag{15}$$

$$\rho_e = \mathrm{sigmoid}(p_e^{(6)}). \tag{16}$$

Here h_e controls the effective direct-response height, τ_e controls the source-to-boundary mixing scale, σ_e controls the spatial spread of the boundary response, and ρ_e controls the mixture between localized and diffuse boundary components.

The physics head is initialized so that the learned source correction is initially zero and the transformed response parameters are approximately

$$h_e = 50\ \mathrm{mm}, \qquad \tau_e = 200\ \mathrm{mm}, \qquad \sigma_e = 230\ \mathrm{mm}, \qquad \rho_e \simeq 0.10. \tag{17}$$

These quantities are learnable effective parameters and are not interpreted as independently calibrated detector constants.

B.3 Query-Dependent Physics Features

Given the effective source $\mathbf{s}_e$ and a query location $\mathbf{u}$, BambooSR constructs six query-dependent physics features. The direct-response component is

$$D_e(\mathbf{u}) = \frac{h_e}{2\pi\left(\|\mathbf{u}-\mathbf{s}_e\|^2 + h_e^2\right)^{3/2}}. \tag{18}$$

To describe boundary-dependent behavior, we place 24 fixed reference anchors uniformly on a circle of radius $R = 592.5$ mm. The source-to-anchor mixing weights are

$$\pi_{e,k} = \frac{\exp\left(-\|\mathbf{b}_k - \mathbf{s}_e\|^2/(2\tau_e^2)\right)}{\sum_\ell \exp\left(-\|\mathbf{b}_\ell - \mathbf{s}_e\|^2/(2\tau_e^2)\right)}, \tag{19}$$

where $\mathbf{b}_k$ denotes boundary anchor k. Each anchor contributes a Gaussian spatial kernel,

$$K_{e,k}(\mathbf{u}) = \frac{\exp\left(-\|\mathbf{u} - \mathbf{b}_k\|^2/(2\sigma_e^2)\right)}{2\pi\sigma_e^2}. \tag{20}$$

The localized boundary response is

$$H_e^{\text{local}}(\mathbf{u}) = \sum_k \pi_{e,k} K_{e,k}(\mathbf{u}), \tag{21}$$

and the full phenomenological boundary response is

$$H_e(\mathbf{u}) = (1 - \rho_e) H_e^{\text{local}}(\mathbf{u}) + \frac{\rho_e}{\pi R^2}. \tag{22}$$

The final six-dimensional physics feature vector is

$$\boldsymbol{\psi}_e(\mathbf{u}) = \left[\frac{u_x - s_{e,x}}{S}, \frac{u_y - s_{e,y}}{S}, \log\big(1 + D_e(\mathbf{u})S^2\big), \log\big(1 + H_e(\mathbf{u})S^2\big), \frac{R - \|\mathbf{u}\|}{S}, \frac{R - \|\mathbf{s}_e\|}{S}\right], \tag{23}$$

with $S = 650$ mm.

The six components therefore correspond to: (i) source-relative displacement in x, (ii) source-relative displacement in y, (iii) compressed direct response, (iv) compressed boundary response, (v) the signed radial gap of the query location to the reference circle, and (vi) the signed radial gap of the effective source to the same reference circle.

These features are phenomenological inductive biases rather than a calibrated photon-acceptance function or a complete optical transport model. In particular, R defines a reference geometry for constructing the features; it is not imposed as a hard boundary on the reconstructed continuous field.

B.4 Feature-Level Physics Adaptation

The physics information is injected into the INR through an additive feature adapter rather than through a multiplicative output template or a hard physical constraint. For each event and query point, the adapter receives

$$[\mathbf{z}_e, \boldsymbol{\psi}_e(\mathbf{u})] \in \mathbb{R}^{38} \tag{24}$$

and applies

$$A_\eta : 38 \to 32 \xrightarrow{\text{SiLU}} 128. \tag{25}$$

Let the first linear transformation of the conditional INR be

$$\mathbf{h}_e^{(0)}(\mathbf{u}) = W_0[\gamma(\mathbf{u}), \mathbf{z}_e] + \mathbf{b}_0. \tag{26}$$

BambooSR adds the adapter output directly to this first hidden preactivation,

$$\widetilde{\mathbf{h}}_e^{(0)}(\mathbf{u}) = \mathbf{h}_e^{(0)}(\mathbf{u}) + A_\eta\left([\mathbf{z}_e, \boldsymbol{\psi}_e(\mathbf{u})]\right), \tag{27}$$

before the first ReLU activation.

Only the final linear layer of the adapter is initialized to zero. Consequently, at initialization the adapter contributes zero and the network reproduces the behavior of the underlying conditional INR. During joint optimization, the model can learn how strongly the physics-derived features should alter the hidden representation.

This construction treats physics as a soft inductive bias: the analytical features provide structured information about source-relative geometry, direct response, and boundary response, while the learned INR remains free to compensate for discrepancies between the phenomenological features and the simulated detector response.

B.5 Finite-Area Sensor Readout

The INR produces a nonnegative continuous response $f_\theta(\mathbf{u} \mid \widetilde{\mathbf{q}}_e)$. For continuous-field visualization in physical units, we define

$$\lambda_\theta(\mathbf{u} \mid \widetilde{\mathbf{q}}_e) = \frac{Q_e}{A_{\mathrm{ref}}} f_\theta(\mathbf{u} \mid \widetilde{\mathbf{q}}_e), \tag{28}$$

where

$$A_{\mathrm{ref}} = 169\pi(32\ \mathrm{mm})^2. \tag{29}$$

Thus λ_θ has units of PE/mm^2, whereas the internal network response f_θ is a normalized nonnegative field.

A PMT measurement is not treated as a point sample at the sensor center. Instead, each virtual HR photocathode is represented by 128 precomputed deterministic quadrature points $\mathbf{u}_{j,s}$ with physical area weights $\alpha_{j,s}$. Defining

$$w_{j,s} = \frac{\alpha_{j,s}}{A_{\mathrm{ref}}}, \tag{30}$$

the predicted count of HR channel j is

$$\mu_{e,j} \approx \sum_s \alpha_{j,s} \lambda_\theta(\mathbf{u}_{j,s} \mid \widetilde{\mathbf{q}}_e) \tag{31}$$

$$= Q_e \sum_s w_{j,s} f_\theta(\mathbf{u}_{j,s} \mid \widetilde{\mathbf{q}}_e). \tag{32}$$

The same finite-area readout operator is used for BambooSR and the continuous neural baselines. Multiplication by Q_e restores event-level signal scale, but does not constrain the sum of predicted HR counts to equal the observed LR total. Likewise, the continuous response is nonnegative but is not normalized to integrate to unity.

B.6 Training Loss and Optimization

Training uses the paired simulated HR channel counts. For a sampled set of event–channel pairs $\mathcal{S}$, the loss is the Poisson negative log-likelihood

$$\mathcal{L}_{\mathrm{HR}} = \frac{1}{|\mathcal{S}|} \sum_{(e,j)\in\mathcal{S}} \left[\mu_{e,j} - y_{e,j} \log\left(\max(\mu_{e,j}, 10^{-8})\right)\right], \tag{33}$$

where the target-only $\log(y_{e,j}!)$ term is omitted because it does not affect optimization. No LR-consistency loss, residual penalty, or auxiliary physics loss is used in the final model; the encoder, INR backbone, physics head, and adapter are optimized jointly through the HR count loss.

BambooSR is trained with Adam at a learning rate of 10^{-4} and a batch size of 8. For each training minibatch, 32 of the 1,261 HR channels are sampled per event to reduce computation and memory use. Validation evaluates the complete HR layout, processed in chunks of 64 channels. Gradients are clipped to a maximum norm of 1.

Training uses a maximum budget of 160 epochs. A ReduceLROnPlateau scheduler reduces the learning rate by a factor of 0.5 after five validation epochs without improvement. Early stopping uses a patience of 15 epochs, and the best checkpoint is updated only when the validation Poisson NLL improves by more than 10^{-4}. The reported BambooSR model is selected by validation NLL. The experiments use random seed 42.

C Baseline Implementations and Controlled Variants

We compare BambooSR with two complementary groups of baselines. The first consists of deterministic geometry-based mappings that do not learn from paired simulation. These methods test how much of the LR–HR reconstruction problem can be explained by sensor geometry alone. The second group consists of neural implicit representations with different coordinate parameterizations, allowing us to compare BambooSR against established INR architectures under the same event-conditioned continuous-field task.

Table 5: Main architecture and optimization settings of BambooSR.

Component	Setting
Event latent dimension	32
INR hidden width	128
INR hidden layers	3
Fourier frequencies	64
Fourier feature dimension	128
Fourier scale σ	3
Coordinate scale S	650 mm
Physics-head outputs	6
Physics-feature dimension	6
Boundary anchors	24
Reference radius R	592.5 mm
Adapter	$38 \to 32 \to 128$
HR quadrature points / channel	128
Optimizer	Adam
Learning rate	10^{-4}
Batch size	8
Sampled HR channels / event	32
Maximum epochs	160
Early-stopping patience	15
LR-scheduler patience	5
LR-scheduler factor	0.5
Gradient clipping	1
Random seed	42
Checkpoint criterion	Validation Poisson NLL

We additionally construct two controlled variants of BambooSR to separate the contribution of the proposed physics features from that of additional learnable capacity. Unless otherwise stated, all learned methods use the same aligned LR inputs, source-position-grouped data partitions, masked physical channel, virtual HR geometry, and HR supervision. Continuous neural models are converted to HR sensor counts through the same finite-area observation operator described in Appendix B.5.

C.1 Geometry-Based Baselines

We include three training-free geometry-based baselines: Nearest, inverse-distance weighting (IDW), and Area. These methods operate directly on the observed LR measurements and the known LR and HR sensor geometries, without fitting to paired HR simulation. They represent two distinct assumptions about the detector readout: Nearest and IDW treat LR measurements as samples associated with sensor centers, whereas Area explicitly models the finite spatial support of the photocathodes.

Let $\mathbf{q}^{\mathrm{LR}} \in \mathbb{R}^{169}$ denote the aligned LR readout, let $\mathbf{c}_i^{\mathrm{LR}}$ and $\mathbf{c}_j^{\mathrm{HR}}$ denote LR and HR sensor centers, and let $\mathcal{V}$ denote the set of valid LR channels after excluding the non-operational PMT. The LR and HR photocathode radii are

$$r_{\mathrm{LR}} = 32\ \mathrm{mm}, \qquad r_{\mathrm{HR}} = 12.5\ \mathrm{mm}. \tag{34}$$

Nearest. For each HR sensor j, Nearest selects the closest valid LR sensor

$$i_j^* = \arg\min_{i \in \mathcal{V}} \left\| \mathbf{c}_j^{\mathrm{HR}} - \mathbf{c}_i^{\mathrm{LR}} \right\|_2 . \tag{35}$$

The corresponding LR measurement is then rescaled by the photocathode area ratio,

$$\widehat{q}_j^{\mathrm{HR}} = q_{i_j^*}^{\mathrm{LR}} \left(\frac{r_{\mathrm{HR}}}{r_{\mathrm{LR}}} \right)^2 . \tag{36}$$

This produces a piecewise-constant spatial reconstruction based only on the nearest sensor center.

Inverse-distance weighting. IDW uses the four nearest valid LR sensors for each HR sensor. Following the inverse-distance interpolation principle of Shepard (1968), we use inverse-squared center-to-center distances. Let $\mathcal{N}_j^{(4)}$ denote the four nearest LR sensors to HR sensor j and

$$d_{ji} = \left\| \mathbf{c}_j^{\mathrm{HR}} - \mathbf{c}_i^{\mathrm{LR}} \right\|_2 . \tag{37}$$

The normalized interpolation weights are

$$\alpha_{ji} = \frac{\max(d_{ji}, \epsilon_d)^{-2}}{\sum_{k \in \mathcal{N}_j^{(4)}} \max(d_{jk}, \epsilon_d)^{-2}}, \qquad \epsilon_d = 10^{-6}. \tag{38}$$

The HR prediction is

$$\widehat{q}_j^{\mathrm{HR}} = \left(\frac{r_{\mathrm{HR}}}{r_{\mathrm{LR}}} \right)^2 \sum_{i \in \mathcal{N}_j^{(4)}} \alpha_{ji} q_i^{\mathrm{LR}}. \tag{39}$$

Thus, Nearest and IDW both reconstruct the HR readout from sensor-center proximity and account for the different sensor sizes through a global photocathode-area scaling factor.

Area. The Area baseline instead treats both LR and HR photocathodes as finite circular sensing regions. For every LR–HR pair, we compute the geometric intersection area

$$A_{ji} = \mathrm{Area}\left[B\left(\mathbf{c}_j^{\mathrm{HR}}, r_{\mathrm{HR}}\right) \cap B\left(\mathbf{c}_i^{\mathrm{LR}}, r_{\mathrm{LR}}\right) \right], \tag{40}$$

where $B(\mathbf{c}, r)$ denotes a disk centered at $\mathbf{c}$ with radius r. The corresponding transfer coefficient is

$$M_{ji}^{\mathrm{area}} = \frac{A_{ji}}{\pi r_{\mathrm{LR}}^2}. \tag{41}$$

Coefficients associated with the masked LR channel are set to zero, and the predicted HR count is

$$\widehat{q}_j^{\mathrm{HR}} = \sum_{i \in \mathcal{V}} M_{ji}^{\mathrm{area}} q_i^{\mathrm{LR}}. \tag{42}$$

This construction corresponds to assuming that the measured signal within each LR photocathode is spatially uniform and redistributing that signal according to the fractional photocathode overlap with each HR sensor. Unlike Nearest and IDW, Area does not use center-to-center interpolation and does not require an additional global area-rescaling factor, because the finite sensor supports enter the mapping directly.

Together, the three baselines provide complementary non-learning references. Nearest and IDW test whether point-based spatial interpolation is sufficient, whereas Area tests whether finite-support geometry alone can explain the change in sensing layout. None of the three methods learns event-dependent signal structure or models event-dependent direct and boundary responses.

C.2 Neural INR Baselines

We further compare BambooSR with four neural implicit representations: FFN, SIREN, WIRE, and SASNet. Each model is adapted to the same event-conditioned detector reconstruction task. The masked LR event is encoded into an event representation, while the spatial decoder defines a continuous field over the detector plane. Predicted HR counts are then obtained through the same finite-area readout geometry used by BambooSR.

Fourier Feature Network (FFN). We use a Fourier Feature Network (FFN) baseline based on the positional encoding of Tancik et al. (2020). FFN is the closest generic INR baseline to BambooSR. It uses the same conditional event-encoding formulation and fixed Gaussian Fourier coordinate representation, but contains no physics head, physics-derived features, or feature adapter.

The event latent and Fourier-encoded query coordinate are concatenated and decoded by a ReLU MLP into a nonnegative continuous field. Because BambooSR retains this conditional FFN backbone and augments it with the physics-aware feature pathway, the FFN comparison directly tests whether the proposed physics adaptation provides information beyond the underlying continuous Fourier representation.

SIREN. SIREN (Sitzmann et al., 2020) represents continuous signals using sinusoidal activation functions instead of fixed Fourier encoding followed by a ReLU decoder. We adapt SIREN to the same conditional setting by coupling its coordinate representation with the LR-derived event latent. The resulting model predicts an event-conditioned continuous field, which is integrated over the same HR photocathodes used for the other neural baselines.

This comparison therefore changes the coordinate representation and decoder family while preserving the detector input, HR supervision, and finite-area observation task.

WIRE. WIRE (Saragadam et al., 2023) represents continuous signals using complex Gabor wavelets, combining oscillatory structure with spatially localized envelopes. We adapt WIRE to the same event-conditioned detector reconstruction setting.

The final implementation uses three complex-Gabor layers with complex width 90, frequency parameter $\omega = 10$, and scale parameter 5. To prevent large event-conditioning values from excessively attenuating the Gabor response, the event latent supplied to the WIRE decoder is bounded as

$$\widetilde{\mathbf{z}}_e = 0.25 \tanh\left(\frac{\mathbf{z}_e}{0.25}\right). \tag{43}$$

The encoder and decoder are optimized with learning rates of 10^{-4} and 5×10^{-4}, respectively. These settings are architecture-specific stabilization choices; the dataset, masking, HR geometry, supervision, and evaluation protocol remain unchanged.

SASNet. We additionally evaluate SASNet (Feng et al., 2026), a recent INR architecture with adaptive spatial representation components. We port the released SASNet design to our event-conditioned continuous field interface while retaining its backbone and grid/mask components. Because our task requires conditioning on a sparse detector event and prediction through finite-area HR sensors, this implementation is a task adaptation rather than a bitwise reproduction of the original code path.

The SASNet backbone is optimized with a learning rate of 10^{-3}, while its grid/mask components use a learning rate of 5×10^{-3}. As with the other neural baselines, the output is evaluated on the same HR sensing geometry and with the same reconstruction metrics.

C.3 Controlled Physics-Prior Variants

In addition to external baselines, we construct two controlled variants of BambooSR to distinguish the contribution of the proposed physics feature content from that of additional model capacity. Both variants are trained independently from initialization rather than obtained by fine-tuning a pretrained FFN or BambooSR model.

Direct-only. The Direct-only variant removes the explicit boundary-response feature while retaining the remaining BambooSR architecture. It still uses the learned effective source, the two source-relative coordinate features, the direct-response feature, and the query and source radial-gap features. In terms of the feature vector in Eq. 23, the boundary-response component is set to zero while the remaining components are preserved.

This is therefore not a geometry-free or physics-free model. Instead, it specifically tests whether the explicit boundary-dependent response contributes information beyond source-relative geometry and the direct-response prior. Because some boundary-related physics-head outputs become inactive in this variant, we do not interpret it as having exactly the same active degrees of freedom as the complete BambooSR model.

Neural control. The Neural control preserves the same allocated six-output head and the same adapter dimensions as BambooSR, but replaces the six physics-derived features with six generic event- and coordinate-dependent learned features. Let $\gamma(\mathbf{u})_{1:6}$ denote the first six components of the coordinate encoding and let $P_\xi(\mathbf{z}_e) \in \mathbb{R}^6$ denote the output of the six-dimensional control head. The replacement features are

$$\boldsymbol{\psi}_e^{\mathrm{ctrl}}(\mathbf{u}) = \tanh\left[\gamma(\mathbf{u})_{1:6} + P_\xi(\mathbf{z}_e)\right]. \tag{44}$$

They are then passed through the same $38 \to 32 \to 128$ adapter used by the full model.

This control preserves an additional event- and coordinate-dependent feature pathway without providing the structured direct, boundary, and radial features of BambooSR. Its purpose is therefore to test whether improvements can be explained simply by the added head and adapter capacity. Differences between BambooSR and this control provide evidence, under the present controlled architecture, for the usefulness of the physics-derived feature content rather than additional parameters alone.

C.4 Comparison Protocol

All learned models use the same paired simulation dataset, source-position-grouped training, validation, and test partitions, aligned LR channel ordering, masked non-operational PMT, and 1,261-channel HR target geometry. The simulated interaction coordinate is never supplied as an inference input to any neural baseline.

For continuous neural models, predictions are converted to HR sensor counts using the same finite-area photocathode observation operator described in Appendix B.5. Model checkpoints are selected using validation Poisson negative log-likelihood, and all methods are evaluated using the same signal-reconstruction metrics described in Appendix D.

FFN and SIREN use the same Adam learning rate of 10^{-4} as BambooSR. WIRE and SASNet retain the architecture-specific optimizer settings described above because their parameterizations require different optimization scales for stable training. We therefore control the dataset, sensing geometry, supervision, checkpoint criterion, and evaluation operator, rather than forcing architecturally different models to use identical optimizer hyperparameters.

The geometric baselines contain no trainable parameters and are applied deterministically to the same masked LR observations. Consequently, the complete comparison spans three levels of prior structure: point- or area-based geometric transfer, learned generic continuous representations, and the physics-aware continuous representation introduced by BambooSR.

D Evaluation Protocol and Metrics

We evaluate signal super-resolution from several complementary perspectives rather than relying on a single channel-level error. For a test event e and virtual HR channel j, let $y_{e,j}$ denote the photon-level simulated HR target and $\mu_{e,j}$ the corresponding model prediction. Unless otherwise specified, the reported simulation metrics are computed over all $N = 11{,}280$ test events and all $M = 1{,}261$ HR channels.

The metrics characterize three aspects of reconstruction quality: local channel-count accuracy, event-level signal magnitude, and the normalized spatial distribution of the reconstructed signal. We additionally evaluate two challenging detector regions: simulated interactions near the detector boundary and the HR sensing region surrounding the masked physical PMT.

D.1 Signal-Reconstruction Metrics

Channel-level RMSE and MAE. The primary reconstruction metric is the root mean squared error (RMSE),

$$\mathrm{RMSE} = \sqrt{\frac{1}{NM} \sum_{e=1}^{N} \sum_{j=1}^{M} (\mu_{e,j} - y_{e,j})^2}. \tag{45}$$

We additionally report the mean absolute error (MAE),

$$\mathrm{MAE} = \frac{1}{NM} \sum_{e=1}^{N} \sum_{j=1}^{M} |\mu_{e,j} - y_{e,j}| \,. \tag{46}$$

Both quantities are reported in PE/channel. RMSE places greater emphasis on large local reconstruction errors, whereas MAE provides a less outlier-sensitive measure of the typical channel-wise discrepancy. These metrics quantify HR sensor-count reconstruction and should not be interpreted as interaction-position errors.

Poisson deviance. Because the HR targets are photon counts, we also report the mean Poisson deviance,

$$D_{\mathrm{Pois}} = \frac{2}{NM} \sum_{e,j} \left[\mu_{e,j} - y_{e,j} + y_{e,j} \log \left(\frac{y_{e,j}}{\mu_{e,j}} \right) \right]. \tag{47}$$

The $y_{e,j} = 0$ term is evaluated by its limiting value. For numerical stability, arguments of the logarithm are floored at $\epsilon = 10^{-8}$. Poisson deviance is dimensionless and complements RMSE and MAE by measuring discrepancy under a count-statistical formulation.

This evaluation metric is distinct from the Poisson negative log-likelihood used during training. The training loss omits target-only terms that do not affect optimization, whereas the deviance is used only for evaluation.

Event-total signal accuracy. Channel-wise reconstruction quality does not by itself determine whether the total signal of an event is preserved. We therefore define the target and predicted HR totals as

$$Y_e = \sum_{j=1}^{M} y_{e,j}, \qquad \widehat{Y}_e = \sum_{j=1}^{M} \mu_{e,j}. \tag{48}$$

The event-total mean absolute error is

$$\text{Total MAE} = \frac{1}{N} \sum_{e=1}^{N} \left| \widehat{Y}_e - Y_e \right|, \tag{49}$$

and is reported in PE.

To compare events spanning different signal strengths, we additionally use the event-wise relative total error,

$$\text{Total Relative MAE} = \frac{100}{|\mathcal{E}_+|} \sum_{e \in \mathcal{E}_+} \frac{\left| \widehat{Y}_e - Y_e \right|}{Y_e}, \tag{50}$$

where $\mathcal{E}_+ = \{e : Y_e > 0\}$. This metric is reported as a percentage and gives each event a relative contribution rather than allowing high-yield events to dominate the error.

We further report the signed global total bias,

$$\text{Total Bias} = 100 \frac{\sum_e \widehat{Y}_e - \sum_e Y_e}{\sum_e Y_e}. \tag{51}$$

Unlike the event-wise absolute metrics, positive and negative errors can cancel in this quantity. It should therefore be interpreted as a measure of global signal calibration rather than event-wise accuracy. When $|\text{Total Bias}|$ is shown in a comparison table, only the absolute value of this signed quantity is used for presentation.

Normalized spatial-shape error. To evaluate the spatial distribution independently of total signal magnitude, we normalize the predicted and target HR patterns event-by-event:

$$p_{e,j} = \frac{\mu_{e,j}}{\max(\widehat{Y}_e, \epsilon)}, \qquad q_{e,j} = \frac{y_{e,j}}{\max(Y_e, \epsilon)}. \tag{52}$$

For each event, we compute their total-variation distance,

$$d_{\mathrm{TV},e} = \frac{1}{2} \sum_{j=1}^{M} \left| p_{e,j} - q_{e,j} \right|, \tag{53}$$

and report

$$\text{Shape TV} = \frac{1}{|\mathcal{E}_+|} \sum_{e \in \mathcal{E}_+} d_{\mathrm{TV},e}. \tag{54}$$

Shape TV is dimensionless and isolates the reconstructed spatial pattern from its total PE. A value of zero indicates identical normalized HR channel distributions. Together, the event-total metrics and Shape TV distinguish errors in *how much* signal is reconstructed from errors in *where* that signal is distributed.

D.2 Region-Specific and Spatial Metrics

Global metrics can obscure reconstruction failures in detector regions where the LR sensing geometry is particularly sparse or incomplete. We therefore report dedicated metrics for the detector boundary and the region surrounding the masked physical PMT.

Edge RMSE. We define the near-wall subset using the simulated source coordinate $\mathbf{x}_e^{\mathrm{true}}$:

$$\mathcal{E}_{\mathrm{edge}} = \left\{ e : \left\| \mathbf{x}_e^{\mathrm{true}} \right\|_2 \geq 520 \text{ mm} \right\}. \tag{55}$$

The current test partition contains 2,530 such events. Edge RMSE is computed over all HR channels of this subset,

$$\text{Edge RMSE} = \sqrt{\frac{1}{|\mathcal{E}_{\mathrm{edge}}| M} \sum_{e \in \mathcal{E}_{\mathrm{edge}}} \sum_{j=1}^{M} (\mu_{e,j} - y_{e,j})^2}. \tag{56}$$

The truth coordinate is used only to define this evaluation subset and is never provided to the reconstruction model as an inference input.

Masked-PMT-region RMSE. Let $\mathbf{c}_{\mathrm{mask}}$ denote the center of the masked physical LR PMT. We define the corresponding HR neighborhood as

$$\mathcal{J}_{\mathrm{mask}} = \left\{ j : \left\| \mathbf{c}_j^{\mathrm{HR}} - \mathbf{c}_{\mathrm{mask}} \right\|_2 \leq 32 \text{ mm} \right\}. \tag{57}$$

The masked-region metric is then

$$\text{PMT3-region RMSE} = \sqrt{\frac{1}{N |\mathcal{J}_{\mathrm{mask}}|} \sum_{e=1}^{N} \sum_{j \in \mathcal{J}_{\mathrm{mask}}} (\mu_{e,j} - y_{e,j})^2}. \tag{58}$$

This metric selects HR *sensor locations* surrounding the missing physical channel across all test events. It does not select events according to their interaction position. It therefore directly evaluates reconstruction in a region where the corresponding LR sensor measurement is unavailable.

Channel-wise spatial diagnostics. For additional spatial analysis, we compute an RMSE independently for each HR channel,

$$\mathrm{RMSE}_j = \sqrt{\frac{1}{N} \sum_{e=1}^{N} (\mu_{e,j} - y_{e,j})^2}. \tag{59}$$

Plotting RMSE_j at the corresponding HR sensor coordinates reveals detector regions that are systematically difficult to reconstruct.

We may additionally visualize the signed mean channel error,

$$\mathrm{Bias}_j = \frac{1}{N} \sum_{e=1}^{N} (\mu_{e,j} - y_{e,j}), \tag{60}$$

to distinguish systematic over-prediction from under-prediction. Unlike channel RMSE, this signed bias can take both positive and negative values.

Single-event residual maps are reported separately as $\mu_{e,j} - y_{e,j}$ or $|\mu_{e,j} - y_{e,j}|$ and are not referred to as RMSE maps.

D.3 Evaluation and Reporting Protocol

All methods are evaluated on the same source-position-disjoint test partition, with identical HR targets and channel ordering. For learned models, checkpoints are selected using validation Poisson negative log-likelihood rather than test performance. The geometry-based baselines contain no trainable parameters and therefore require no checkpoint selection.

The reported neural-model metrics are point estimates from the selected training runs. They should not be interpreted as estimates of variability across independent training seeds, and metric differences alone are not reported as formal statistical significance across retraining. Because the dataset contains ten stochastic realizations at each simulated source position, events sharing the same position are also not independent samples of detector geometry. For future uncertainty estimates, source position therefore provides the natural grouping unit for a cluster-level bootstrap rather than resampling individual events independently.

The fixed test partition was used only after validation-based checkpoint selection for each reported model. However, aggregate test diagnostics were consulted during method development; we therefore describe the partition as held out under the fixed source-position split rather than claiming that it remained completely unseen throughout the entire development process.

Finally, visualization transforms are used only for presentation. For example, nonlinear intensity scaling may be applied to make weak signals visible in qualitative figures, and signed-log scaling may be used for residual maps. All quantitative metrics reported above are computed from the original, untransformed PE predictions and targets.

E Additional Signal-Reconstruction Analysis

Beyond the aggregate metrics reported in the main text, we further analyze how reconstruction quality varies with event signal strength, source position, and event-level signal magnitude. These analyses probe failure modes that can be obscured by a single test-set average and provide additional context for interpreting the physical reliability of the reconstructed detector signals.

E.1 Dependence on Signal Strength

We first stratify the simulated test events by their observed LR total PE and compute the HR channel RMSE within each signal-strength range. Fig. 8 shows a clear separation between learned continuous representations and the training-free geometric mappings. The geometric baselines exhibit substantially larger absolute errors and a strong increase in RMSE as the event signal grows, whereas the neural methods remain considerably more stable across the evaluated signal range.

Among the learned models, BambooSR maintains the lowest or near-lowest RMSE across the signal range and exhibits a smooth increase in absolute error as the total signal grows. FFN and WIRE show similarly stable trends, while SIREN and SASNet produce larger reconstruction errors over most of the range. Because higher-PE events also contain larger channel counts, an increase in absolute RMSE at high PE should not by itself be interpreted as evidence that high-energy events are intrinsically more difficult in relative terms.

Conversely, weak-signal reliability cannot be fully characterized by absolute channel RMSE alone. At low PE, even a modest absolute error can represent a substantial fraction of the available signal, and event-level failure modes can be partially hidden by an average channel metric. We therefore examine event-total calibration separately in Section E.3.

E.2 Dependence on Source Radius

We next evaluate reconstruction error as a function of the true simulated source radius. This analysis probes the transition from the better-covered detector interior toward the boundary, where the relation between the interaction location, finite sensor supports, and detector geometry becomes more challenging.

As shown in Fig. 9, BambooSR, FFN, and WIRE remain relatively stable through most of the active region. The differences between methods become more pronounced in the outermost radial region, where several learned baselines show a sharper increase in reconstruction error. BambooSR also becomes less accurate near the wall, but its degradation remains comparatively limited.

The geometry-only mappings exhibit substantially larger RMSE throughout the detector and therefore do not recover the underlying HR signal structure through sensor geometry alone. The radial analysis complements the single-number Edge RMSE defined in Appendix D.2 by showing how reconstruction quality changes continuously from the detector interior toward the wall.

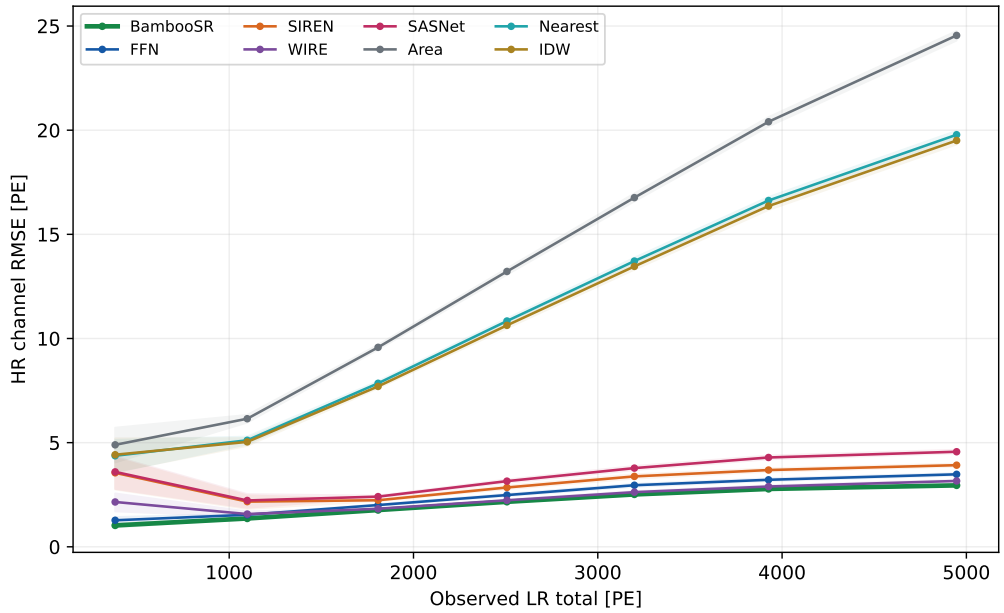


Figure 8: HR channel RMSE as a function of observed LR total PE. The analysis characterizes reconstruction accuracy across event signal strengths, including the weak-signal regime.

This behavior is also consistent with the motivation for the boundary-related features in BambooSR. Near the wall, reconstruction is affected by both sparse sensing and the changing spatial relationship between the signal source and the detector boundary, motivating a representation that includes more than generic coordinate interpolation.

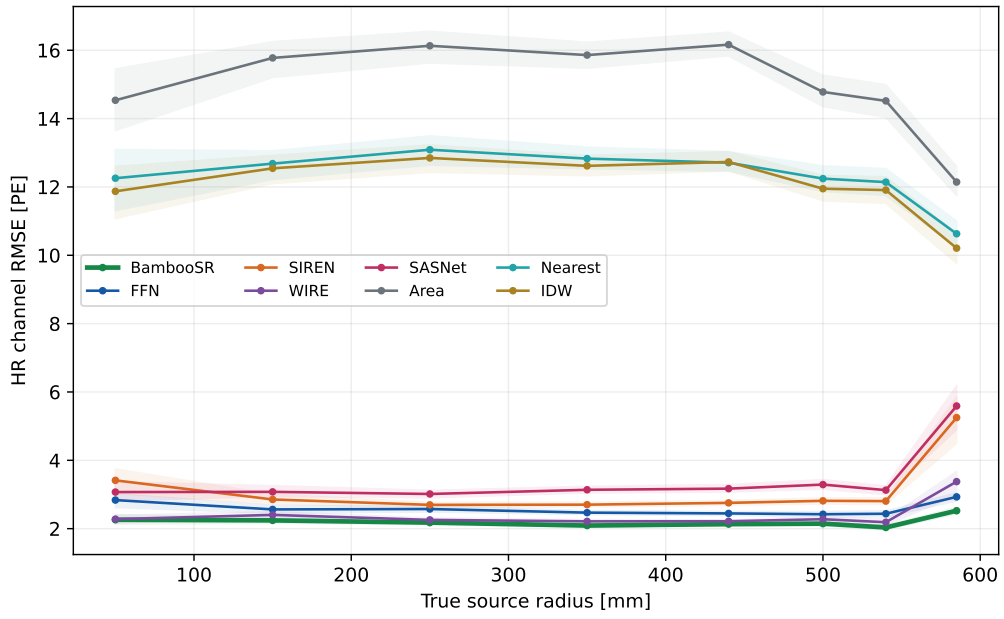


Figure 9: HR channel RMSE as a function of the true simulated source radius. The outer radial region highlights the increased reconstruction difficulty near the detector boundary.

E.3 EVENT-LEVEL TOTAL-PE CALIBRATION

Channel-wise accuracy does not necessarily imply reliable preservation of an event's overall signal magnitude. We therefore compare the predicted and ground-truth HR total PE for every simulated test event. In Fig. 10, the diagonal corresponds to perfect preservation of the event-level HR signal total.

BambooSR follows a compact, nearly single-valued relation along the diagonal over the full signal range, with a global signed total bias of approximately $+0.05\%$. FFN similarly retains a coherent calibration relation with a bias of approximately $+1.14\%$, while WIRE also preserves a comparatively well-defined diagonal trend with a bias of approximately $+1.75\%$.

The two-dimensional calibration plots reveal failure modes that are not captured by the aggregate signed bias alone. SIREN, SASNet, Nearest, and IDW exhibit an additional low-prediction population in the low-PE regime, in which a subset of events separates from the principal calibration relation. Area exhibits a different failure mode, substantially underestimating the event total across the signal range and producing a large negative global bias.

Importantly, the existence of a secondary low-PE population can remain hidden behind a small aggregate signed bias. For example, over-predicted and under-predicted events can cancel when summed over the complete test set. The calibration relation therefore provides information complementary to both the channel-level RMSE and the scalar Total Bias defined in Appendix D.1.

Total S2 PE is a physically meaningful measure of the overall detector signal magnitude. Although this study does not evaluate energy reconstruction, a super-resolution model intended for detector analysis should avoid introducing artificial event populations or strong signal-dependent distortions in this quantity, particularly in the weak-signal regime. We therefore use event-level calibration as an additional indicator of upstream reconstruction reliability rather than relying solely on average RMSE.

This consideration is also relevant when selecting representative upstream reconstruction methods for the subsequent real-data position-reconstruction study.

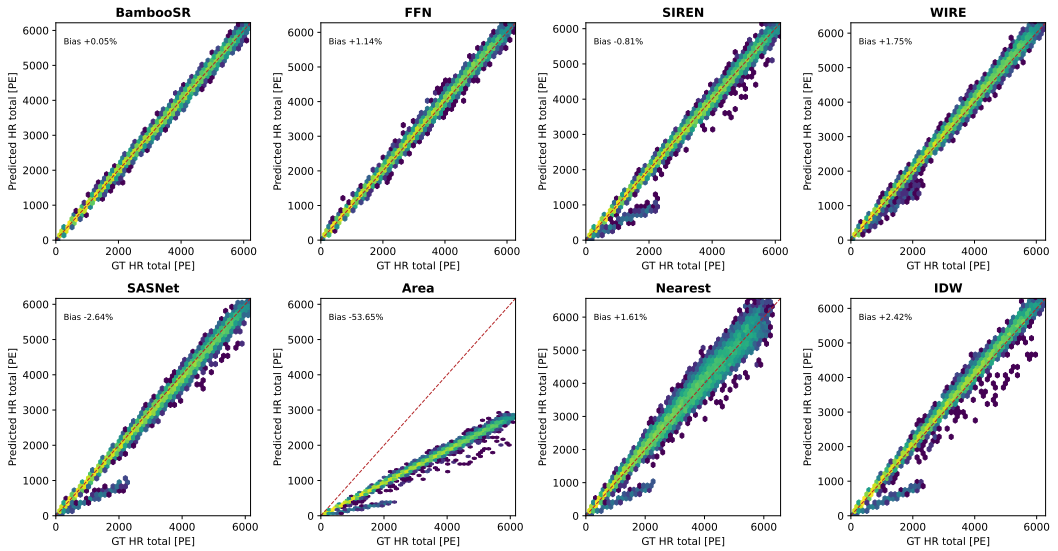


Figure 10: Event-level HR total-PE calibration on the simulated test set. Each panel compares predicted and ground-truth HR total PE for one reconstruction method; the diagonal indicates perfect total-signal preservation. The two-dimensional distributions expose low-PE failure modes that can be obscured by a single aggregate bias value.

E.4 Representative Reconstructions under Challenging Conditions

We further visualize four representative events spanning two experimentally relevant sources of difficulty: signal strength and detector geometry. The selected examples form a 2×2 set of conditions: low- and high-PE events near the masked physical PMT, and low- and high-PE events near the detector boundary.

For each example, the visualization contains the physical 169-channel LR observation, the photon-level 1,261-channel HR target, and the reconstructed fields from all compared methods. The enlarged local views focus on the region of interest, while the corresponding HR residual panels display signed prediction errors. Within each event, the methods share the same visualization scale so that spatial differences can be compared directly.

Masked-PMT neighborhood. Figures 11 and 12 examine events whose signal pattern lies near the non-operational physical PMT. Because the corresponding LR measurement is unavailable, successful reconstruction must infer the missing local structure from the surrounding sensor pattern. The low-PE example combines this missing spatial measurement with weak photon statistics, whereas the high-PE example provides a stronger signal under a similar local sensing challenge.

The learned continuous representations recover a localized field in both cases, while the geometry-only mappings retain visible structures associated with their point-interpolation or overlap-transfer assumptions. These examples qualitatively complement the PMT3-region RMSE reported in the main evaluation.

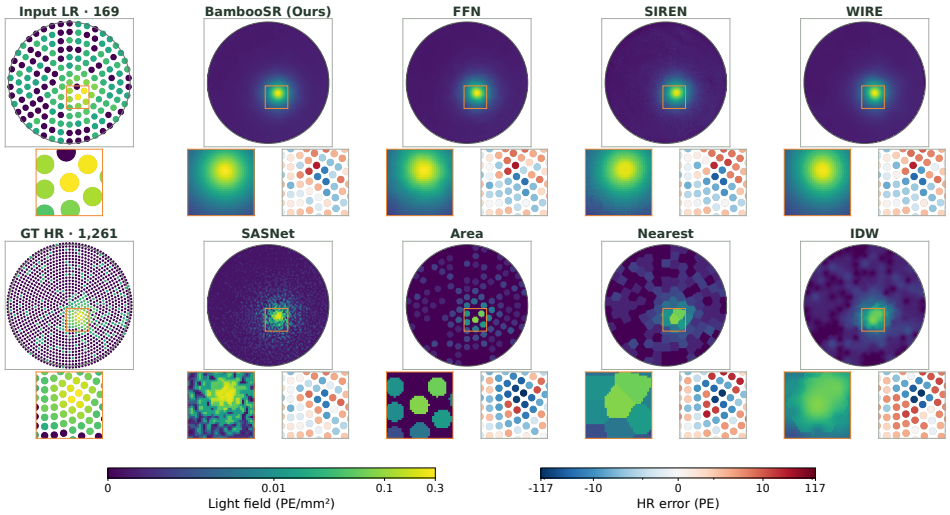


Figure 11: Representative low-PE event near the masked physical PMT. The panels compare the LR observation, photon-level HR target, reconstructed continuous fields, and local HR residuals.

Detector boundary. Figures 13 and 14 show low- and high-PE interactions near the detector boundary. In this regime the observed spatial pattern becomes asymmetric and is partially truncated by the available sensing region, making simple center-based interpolation particularly limited.

The local crops and signed residuals make the differences between methods more apparent than the global field alone. The geometry-based baselines retain coarse or piecewise spatial structures inherited from the sparse LR layout, whereas the learned continuous representations recover substantially smoother localized fields. BambooSR remains consistent with the localized HR structure in both signal regimes, qualitatively supporting the radial behavior discussed in Section E.2.

The four representative examples are intended as qualitative illustrations rather than independent statistical evidence. Quantitative conclusions are based on the complete test partition and the evaluation protocol described in Appendix D.

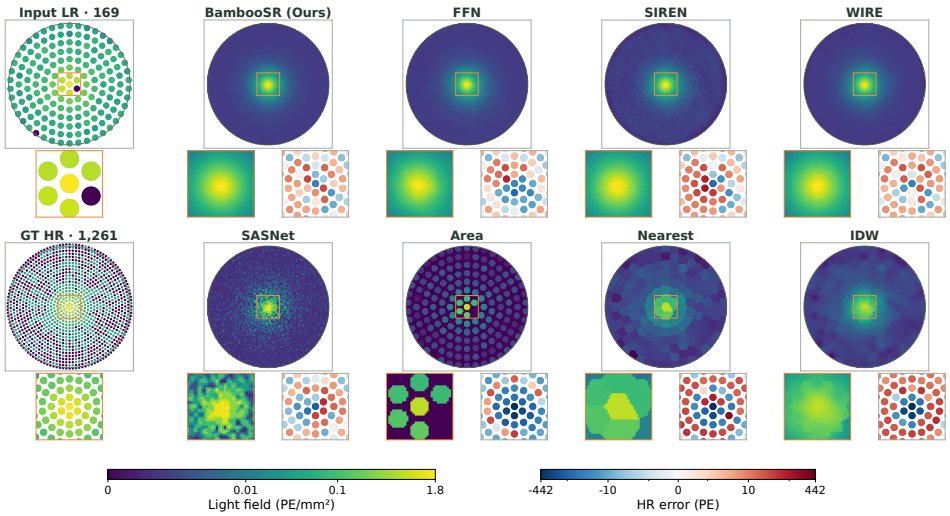


Figure 12: Representative high-PE event near the masked physical PMT. This example separates the challenge of missing local LR coverage from the additional difficulty associated with a weak signal.

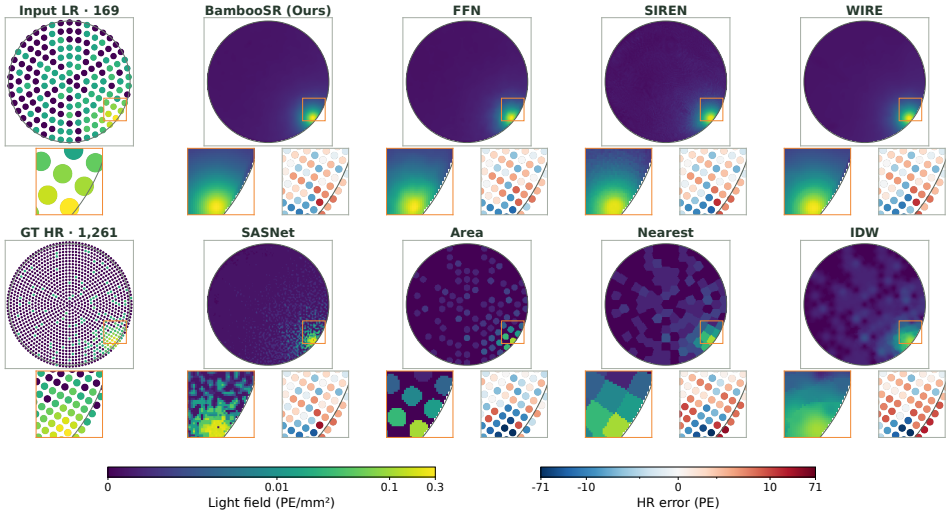


Figure 13: Representative low-PE event near the detector boundary. The boundary-focused crops emphasize reconstruction behavior where sensing coverage is incomplete and the signal pattern is asymmetric.

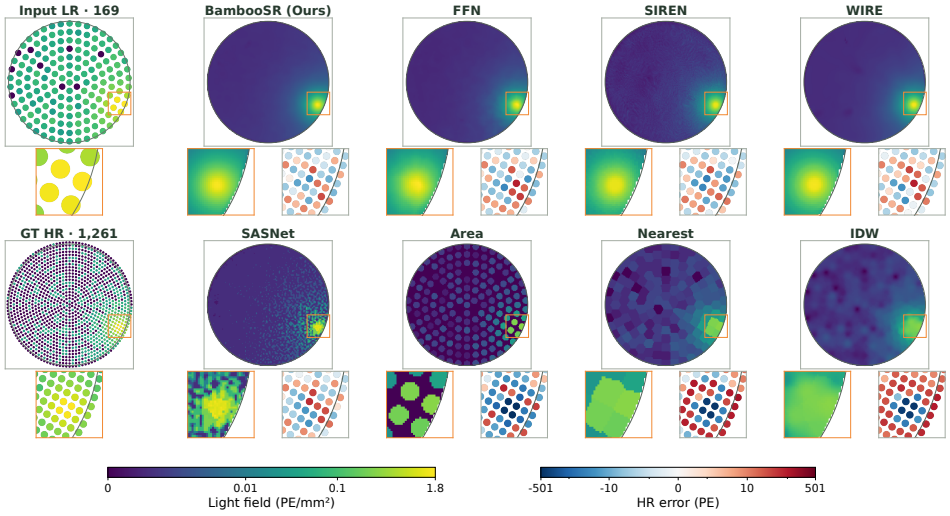


Figure 14: Representative high-PE event near the detector boundary. The example illustrates reconstruction of a stronger signal under the same boundary-related geometric challenge.

F Real-Data Wall-Event Position Reconstruction

F.1 Physical Motivation and Wall-Event Sample

Accurate transverse position reconstruction is particularly important near the PandaX-4T side boundary, where radioactive contaminants on the PTFE reflector give rise to surface backgrounds. Previous PandaX-4T studies identify radon-progeny isotopes including ^{210}Pb, ^{210}Bi, and ^{210}Po as relevant contributors and use the reconstructed radial distribution of surface events to model their leakage toward the detector fiducial volume (Qian et al., 2025). The spatial resolution of such events therefore provides a physics-motivated probe of reconstruction quality near the detector boundary.

We use a fixed evaluation sample of 5,583 selected PandaX-4T Run 2 wall events. The event list is defined before applying any super-resolution method and is kept identical for the original 169-channel LR readout and all reconstructed 1,261-channel HR readouts. The purpose of this sample is not to construct a new surface-background model, but to test whether a denser reconstructed S2 representation improves transverse position reconstruction under real detector conditions.

Event-level interaction-position ground truth is unavailable for these data. The real-data study therefore does not use supervised two-dimensional position error. Instead, we characterize the radial spread of the reconstructed wall population relative to the known detector boundary.

F.2 Common COG Reconstruction and Cross-Calibration

To isolate the effect of the upstream readout representation, every readout is passed to the same downstream position estimator. For either the original LR readout or a reconstructed HR readout, the charge-centroid (COG) position is

$$\mathbf{r}_{\mathrm{COG}} = \frac{\sum_j q_j \mathbf{c}_j}{\sum_j q_j}, \tag{61}$$

where q_j is the signal associated with sensor j and $\mathbf{c}_j$ is its transverse coordinate. No method-specific position network or supervised downstream retraining is introduced.

The PandaX-4T side boundary is represented by a regular 24-sided polygon with an apothem

$$R_0 = 592.5\ \mathrm{mm}. \tag{62}$$

For an event at azimuth φ, the corresponding radial wall position is denoted by $R_{\mathrm{wall}}(\varphi)$. Within each polygon sector,

$$R_{\mathrm{wall}}(\varphi) = \frac{R_0}{\cos\delta(\varphi)}, \tag{63}$$

where $\delta(\varphi)$ is the angular offset from the corresponding sector normal.

The wall-event sample is randomly divided into two folds using a fixed seed. For each readout method and each of the 24 angular sectors, the radial scale correction for one fold is estimated exclusively from the opposite fold. For a calibration fold, we define

$$a = \left[\mathrm{median}\left(\frac{r}{R_{\mathrm{wall}}(\varphi)}\right)\right]^{-1}, \tag{64}$$

and apply

$$r_{\mathrm{corr}} = a\, r \tag{65}$$

to events in the held-out fold. The procedure is then reversed so that every event is corrected using calibration information obtained from disjoint events.

Fig. 15 illustrates the role of this procedure. Before calibration, the reconstructed COG radius exhibits both a method-dependent radial offset and an azimuth-dependent structure. After sector-wise cross-calibration, the reconstructed wall population follows the known polygonal boundary for all compared readouts. The calibration therefore removes systematic radial and angular distortions before the residual widths are compared; it is not itself used as the performance metric.

F.3 Wall Residual and Calibrated Position Resolution

After cross-calibration, the radial residual of event e is defined as

$$\Delta r_e = r_{e,\mathrm{corr}} - R_{\mathrm{wall}}(\varphi_e). \tag{66}$$

Negative values correspond to positions reconstructed inside the nominal detector boundary, whereas positive values correspond to positions reconstructed outside it. Because the resulting wall distribution can be asymmetric, we model the residuals using a split-normal distribution with independent inward and outward widths,

$$p(\Delta r) \propto \begin{cases} \exp\left[-\dfrac{(\Delta r-\mu)^2}{2\sigma_{\mathrm{L}}^2}\right], & \Delta r < \mu, \\ \exp\left[-\dfrac{(\Delta r-\mu)^2}{2\sigma_{\mathrm{R}}^2}\right], & \Delta r \geq \mu. \end{cases} \tag{67}$$

The distribution is fit over $\Delta r \in [-80, 80]$ mm using 1-mm bins and a truncated, normalized likelihood. The free parameters are the distribution center μ, the inward width σ_{L}, and the outward width σ_{R}. The center μ remains a free fit parameter because the preceding median-based radial calibration does not require the mode or fitted center of the residual distribution to be exactly zero. Although σ_{R} is retained as a free parameter to model the asymmetric residual distribution, we do not use it as a resolution metric. The outward side is strongly affected by the hard detector boundary and therefore does not provide the same interpretable measure of inward migration as σ_{L}. We consequently report σ_{L} as the primary fitted wall-resolution metric.

We report the inward width σ_{L} as the calibrated wall-event position resolution. It is the fitted width of the inward side of the wall distribution, rather than the ordinary standard deviation of only events

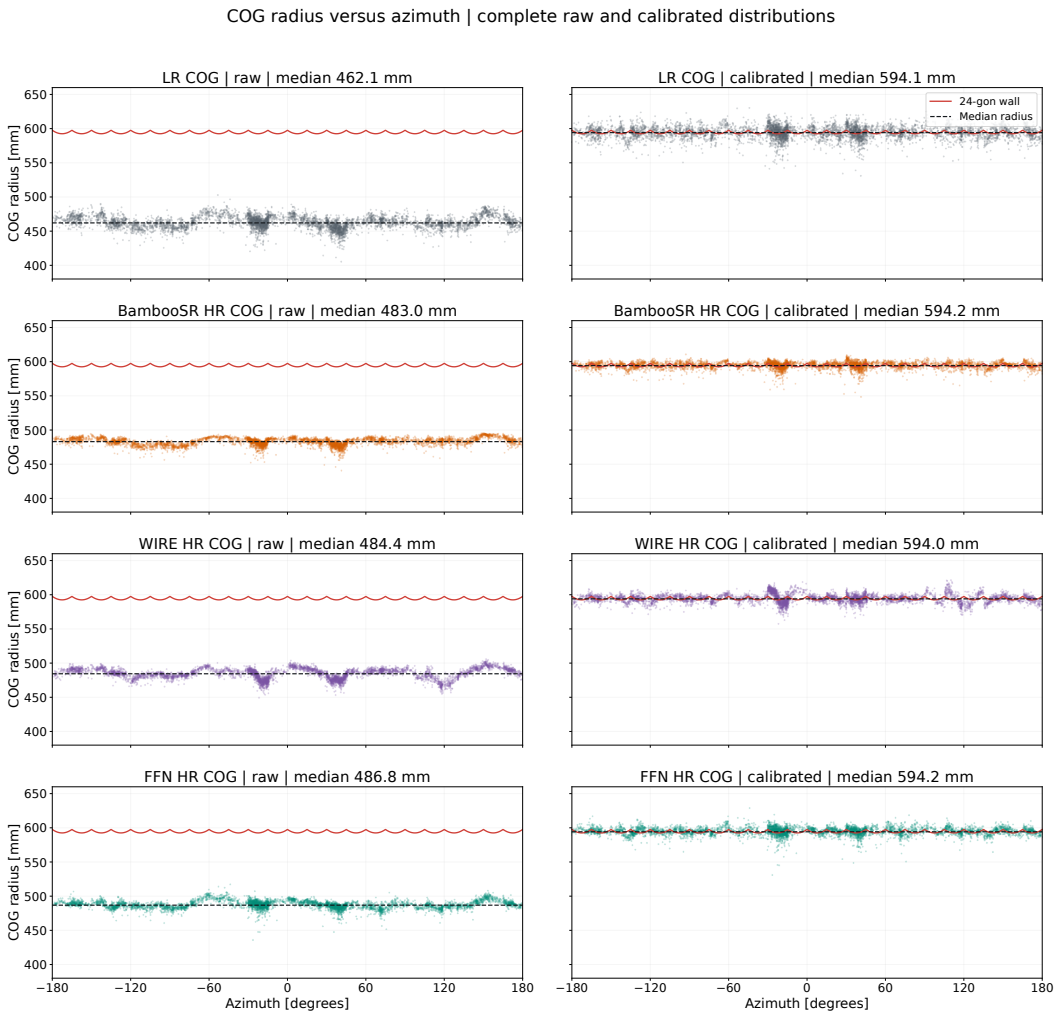


Figure 15: COG radius as a function of azimuth before and after the two-fold, 24-sector radial calibration. The red curve denotes the known 24-sided detector boundary. The right panels show that the calibration removes systematic radial and azimuthal offsets before wall-width comparison.

with negative residuals. This inward width is particularly relevant to a surface-event population because poorer radial reconstruction broadens the apparent migration of boundary events toward the detector interior. To verify that the comparison does not depend only on the parametric split-normal fit, we additionally report two fit-free statistics of the same calibrated residual distribution. The 68% containment half-width is defined as

$$w_{68} = \frac{Q_{84}(\Delta r) - Q_{16}(\Delta r)}{2}, \tag{68}$$

where $Q_p(\Delta r)$ denotes the pth percentile of the calibrated radial residuals. For a Gaussian distribution, this quantity corresponds approximately to one standard deviation, while remaining well defined without assuming a Gaussian residual shape.

We also report the mean absolute error

$$\mathrm{MAE} = \frac{1}{N}\sum_{e=1}^{N} |\Delta r_e|, \tag{69}$$

which summarizes the average absolute deviation from the calibrated detector boundary.

The calibrated residual distributions are shown in Fig. 5, and the corresponding metrics are reported in Table 3. Under the common protocol, the original LR COG gives $\sigma_{\mathrm{L}} = 11.01$ mm, whereas BambooSR reduces the width to 6.65 mm. FFN and WIRE yield 6.96 mm and 7.04 mm, respectively. The fit-free metrics show the same ordering: BambooSR obtains $w_{68} = 4.19$ mm and an MAE of 3.57 mm, compared with 8.16 mm and 6.88 mm, respectively, for LR COG.

F.4 Interpretation and Scope

The wall-event study is intended as a real-data downstream validation of the reconstructed S2 representation rather than an independent measurement of absolute two-dimensional position accuracy. Because true event-level interaction coordinates are unavailable, the reported metrics characterize the calibrated radial spread of a known boundary population rather than the Euclidean error between a reconstructed and true interaction position.

The comparison is deliberately controlled across readouts. The same 5,583 events, the same COG estimator, the same polygonal wall geometry, and the same two-fold sector calibration are applied to the original LR readout and to every reconstructed HR readout. Consequently, differences in the resulting wall widths measure how the upstream signal representation affects a fixed downstream estimator, rather than differences in event selection or downstream model capacity.

This evaluation is particularly relevant to signal super-resolution because the objective is not only to reduce simulated channel-level errors, but also to recover spatial structure that remains useful for real detector analysis. A narrower calibrated wall distribution therefore provides evidence that improvements observed under paired simulation transfer to position reconstruction on real detector data.

At the same time, this measure is boundary-specific and does not provide event-level two-dimensional ground-truth errors. It should also not be interpreted as evidence of improved energy reconstruction or dark-matter sensitivity. Its role here is specifically to evaluate the downstream utility of the reconstructed S2 representation under a simple and common position estimator.

G Reproducibility and Planned Release

G.1 Reproducibility Protocol

The experiments in this work are defined by fixed data splits, sensor geometries, model configurations, and evaluation protocols. The paired simulation benchmark contains 112,890 events generated at 11,289 source positions, with ten stochastic realizations per position. The split is performed by

source position, yielding 90,330 training events, 11,280 validation events, and 11,280 test events, with no source position shared across partitions.

All learned methods use the same paired LR–HR benchmark, physical PMT mask, virtual HR geometry, and evaluation metrics. For BambooSR, checkpoint selection is based on validation Poisson NLL. The main training configuration uses Adam with a learning rate of 10^{-4}, batch size 8, 32 sampled HR channels per event during training, a maximum of 160 epochs, and early stopping with patience 15. Additional implementation and optimization details are provided in Appendix B and Appendix C.

The real-data downstream evaluation uses a fixed sample of 5,583 PandaX-4T Run 2 wall events. All compared readouts are evaluated using the same event sample, charge-centroid estimator, 24-sector detector-boundary model, and two-fold radial cross-calibration described in Appendix F. No event-level true interaction coordinates are used in this evaluation.

G.2 Planned Code and Data Release

Following completion of the ICLR review process, we plan to release the BambooSR implementation, experiment configurations, evaluation code, and scripts required to reproduce the main simulation and wall-position analyses.

We also plan to release the paired simulation benchmark used in this work. The dataset will include the aligned 169-channel LR observations, 1,261-channel virtual HR targets, fixed train/validation/test splits, sensor geometries, and finite-area quadrature information required by the model. As described in Appendix A.3, the HR targets are geometry-derived virtual sensor counts constructed from the same propagated photon realization as the LR observation, rather than a complete simulation of a separate small-PMT electronics chain.

The nominal 100–5,000 PE-equivalent light-yield range will be documented as a simulation input scale rather than a hard bound on the realized LR or HR sensor counts. Because the same propagated photon realization is aggregated under different photocathode geometries, the final LR and HR totals are not required to be identical and may fall outside this nominal interval.

Subject to the corresponding collaboration data-release policy, we also plan to provide the fixed 5,583-event Run 2 wall-event evaluation sample used for the real-data downstream study. This sample will be released as the specific evaluation population used in this work rather than as the complete PandaX-4T Run 2 dataset.

H Scope and Limitations

BambooSR is developed as a detector-reconstruction method rather than as a complete dark-matter analysis pipeline. The scope of this work is to test whether spatial information lost under sparse and incomplete S2 readout can be recovered through a continuous signal representation, and whether the recovered structure remains useful for downstream detector reconstruction. We directly evaluate the first step using paired LR–HR simulation and the second using real PandaX-4T Run 2 wall events. In particular, the real-data study shows that the improved S2 representation can sharpen wall-event position reconstruction under a common downstream estimator.

The scientific impact of improved position reconstruction can extend beyond the position metric itself. More accurate transverse reconstruction can improve position-dependent signal corrections, sharpen fiducial-volume boundaries, and reduce the apparent inward migration of surface-background events (Qian et al., 2025). These effects can propagate to reconstructed energy quantities and background rejection and may therefore improve the sensitivity of rare-event searches. Quantifying the resulting change in the final dark-matter sensitivity, however, requires the complete downstream physics-analysis chain and is beyond the scope of the present study.

The HR supervision used in this work is simulation-derived. The virtual 1,261-channel targets are constructed by aggregating the same propagated photon realization over a denser photocathode geometry, rather than by operating a second physical sensor array. As detailed in Appendix A.3, the virtual HR labels therefore represent geometry-derived finite-area photon counts and do not include a complete small-PMT-specific photo-detection and electronics response. The present results establish

the utility of this physically motivated virtual sensing configuration, while experimental validation with an actual denser readout remains a separate step.

Although the implicit representation can be queried continuously in the detector plane, the experiments in this work use a fixed pair of co-centered sensing layouts: the physical 169-channel PandaX-4T top array and the 1,261-channel virtual HR geometry. We therefore do not claim experimentally demonstrated cross-layout generalization to arbitrary sensor configurations. Evaluating whether a trained field representation can transfer across sensor densities, photocathode sizes, or detector geometries is an important direction for future work.

Finally, the real wall-event dataset does not provide event-level true interaction coordinates. The reported real-data resolution therefore measures the calibrated radial spread of a boundary population rather than supervised two-dimensional position error. This boundary-based evaluation is directly relevant to surface-event reconstruction, but extending the validation to real events with independent position ground truth would provide a stronger assessment of absolute two-dimensional reconstruction accuracy. The neural-model comparisons reported in this work are also based on single training runs under a fixed random seed; repeated-seed studies would further characterize optimization variability.