RE-EXCITABLE RESPONSE REPRESENTATIONS FROM A SINGLE-WAVEFORM RADAR OBSERVATION

Anonymous authors
Paper under double-blind review

ABSTRACT

Pulsed radar measurements are intrinsically conditioned on the transmitted waveform, and the same scene can therefore produce substantially different observations under different excitation conditions. We study *radar re-excitation*: predicting how a scene observed once under a source waveform can be measured under another known waveform, without physical re-acquisition. The central challenge is that the source may not adequately capture response directions that become important under the query waveform, so source consistency alone is insufficient for reliable prediction. We address this ambiguity with a Re-Excitable Response Representation constructed from a single source observation. The representation anchors measurement-supported content to a source-evidence response, organizes unresolved content through shared propagation trajectories and compact event-aligned responses, and performs observability-conditioned closed-form completion with structural intervention restricted to source-weak directions. The completed response can subsequently be re-excited under known waveform queries without requiring target-waveform scene observations. Across comprehensive synthetic benchmarks and real-world tests, our method achieves the lowest MSE and highest SSIM among six physics-guided and learning-based baselines. Re-excited measurements further enable two frozen inversion models to approach native-waveform performance, demonstrating a reusable interface between varying sensing waveforms and downstream physical inference.

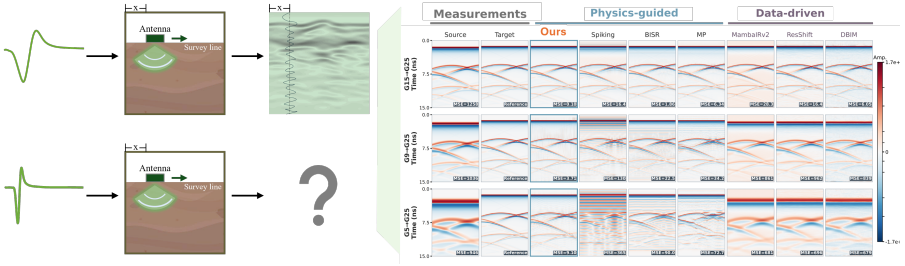

Figure 1: **Overview of the cross-waveform conversion problem and our re-excitable response representation.** The electromagnetic measurement operator encodes a fixed scene under a source waveform; our method identifies a reusable response state and re-excites it with a different known waveform to synthesize the corresponding measurement without physical reacquisition. Across seven simulated shifts, our method achieves the best conversion fidelity among three physics-guided and three learned baselines.

1 INTRODUCTION

Radar observations are ubiquitous across sensing systems, with applications ranging from autonomous navigation and subsurface inspection to planetary exploration (Li et al., 2022; Zhang et al., 2024; Paige et al., 2024; Li et al., 2024; Ort et al., 2023). Established radar systems have enabled reliable deployment in these settings, yet their measurements remain strongly conditioned on the transmitted waveform. In pulsed radar, each scan trace is shaped jointly by the scene, electromagnetic (EM) propagation, scattering, attenuation, acquisition geometry, and excitation waveform.

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.

Changing the pulse shape, bandwidth, or center frequency changes how a fixed scene response is encoded, altering echo bandwidth, temporal extent, and separability. Consequently, echoes may overlap under one excitation, waveform-specific inversion models may fail under another, and measurements collected with different settings may be difficult to compare or fuse. Under fixed geometry and a linear measurement model, source and query waveforms probe a shared scene response through different operators. A response direction that is weakly expressed by the source may nevertheless become important under the query, so source consistency alone does not guarantee query accuracy. We formulate this mismatch as *radar re-excitation*: estimating a reusable response representation from one source scan and querying it with a known waveform to synthesize the corresponding measurement.

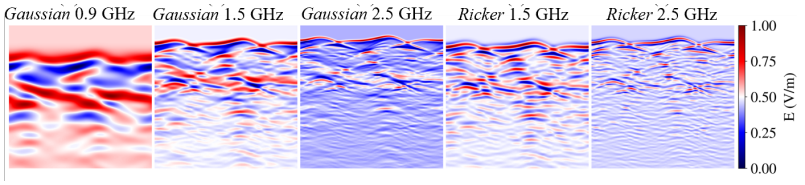

Figure 2: Measurements of the same scene under Gaussian excitations at 0.9, 1.5, and 2.5 GHz and Ricker excitations at 1.5 and 2.5 GHz (left to right). All panels share the scale shown at right.

Existing approaches address waveform variation through data-driven translation or physics-guided response reconstruction. Data-driven methods learn paired, unpaired, or waveform-conditioned mappings and can be effective when training covers the relevant waveforms and scenes (Liu et al., 2024; Zhang et al., 2025; Yang et al., 2026). Under unseen conditions, however, the learned mapping may encode the distribution of observed waveform pairs rather than a reusable response representation. A second line estimates or sharpens responses through deterministic deconvolution and calibrated multi-frequency reconstruction (Turner, 1994; Xia et al., 2003; Economou & Vafidis, 2011; Booth et al., 2009; Shen et al., 2022). These methods regularize the source inverse and select a source-consistent response, but they do not explicitly distinguish directions supported by measurement evidence from directions that require structural continuation. Neither learned translation nor a low source reconstruction error therefore resolves the cross-operator ambiguity by itself.

To address this ambiguity, we construct a re-excitable response state rather than directly learning a source-to-query mapping. A *source-evidence anchor* preserves response content supported by the realized source experiment, while an *event-aligned structured continuation* uses shared propagation trajectories and compact local kernels to organize unresolved content across scan positions. A source-observability operator combines them in closed form according to a *minimal structural intervention* principle: the continuation modifies the anchor only along source-weak response directions. Because this completion depends on source observability rather than on a particular query operator, the resulting state can be reused across multiple known waveform queries. No held-out target measurement or label under the query waveform enters its construction.

We evaluate this representation for EM super-resolution, where a lower-frequency measurement may merge nearby reflection events that become separable under a higher-frequency query. Against three physics-guided conversion methods and three frozen learned translators, our method achieves the lowest MSE and highest SSIM in all seven simulated cases spanning frequency, geometry, heterogeneous-background, and noise shifts. At the largest G5→G25 gap, its MSE is 38.8, compared with 182.9 for the strongest physics-guided baseline and 956.9 for the strongest learned baseline. On the measured TU1208 200→400 MHz limestone profile, it reduces MSE by 73.1% relative to the strongest reported baseline. Response estimation uses the source scan and calibrated source/query waveforms, while held-out target measurements and labels are used only for evaluation.

Beyond measurement fidelity, we test whether re-excitation reduces waveform dependence in downstream inversion. We train a G25 TransUNet for stratigraphic segmentation and a G15 inverse model for joint geometry and permittivity recovery, then freeze both models before applying converted measurements. For G15, G9, and G5 re-excited to G25, our method obtains foreground mIoUs of 0.909, 0.912, and 0.867, within 0.007, 0.004, and 0.049 of the native-G25 result. On the reported held-out multi-object scene, it also improves both object IoU and permittivity MAE over

the strongest physics-guided conversion baseline under unseen G9/G5→G15 shifts. These results suggest a route toward reusable cross-waveform inversion backbones in which a response-level re-excitation layer absorbs waveform variation without collecting a new labeled dataset or retraining the inverse model for every excitation condition.

Unlike direct source-to-query translation and source-only response sharpening, our work makes three contributions:

- **Cross-operator formulation.** We characterize when a source observation determines a query measurement and identify cross-operator observability transfer as the central ambiguity of single-observation re-excitation.
- **Re-excitable response representation.** We combine a source-evidence anchor with event-aligned structured continuation through an observability-gated closed-form composition, limiting structural intervention to source-weak response directions.
- **Multi-level validation.** Across seven synthetic shifts, measured TU1208 data, and cross-waveform segmentation and inversion tasks, our method achieves more accurate large-gap and high-frequency projection than physics-guided methods, while generalizing more robustly than data-driven baselines to unseen waveform, geometry, medium, noise, and sampling conditions.

2 Related Works

EM super-resolution. Learning-based methods provide a direct route to cross-frequency image translation by learning mappings between low- and high-frequency measurements. Early supervised approaches learn from paired measurements, while weakly supervised methods relax one-to-one pairing but still require substantial high-resolution target-domain data (Zhang et al., 2025; Liu et al., 2024). More recent methods use adversarial and contrastive objectives for unpaired cross-frequency translation (Yang et al., 2026). Despite this progression from paired to unpaired learning, these methods are typically tied to predefined source-to-target frequency configurations and show limited generalization to out-of-distribution frequencies, geometries, and noise levels. Deconvolution-based methods instead recover a sharpened reflection sequence directly from the measured response, avoiding the need for cross-frequency training data. Known-wavelet approaches stabilize this ill-conditioned inverse through filtering, sparse pursuit, or regularized inversion (Turner, 1994; Xia et al., 2003; Economou & Vafidis, 2011; Schmelzbach & Huber, 2015; Mallat & Zhang, 1993; Tuncer & Gürbüz, 2012; Selesnick & Bayram, 2016; Goudarzi & Dehghan-Manshadi, 2026), while cross-frequency extensions use calibrated spectral extrapolation or multi-band fusion (Xiao & Liu, 2016; Shen et al., 2022). However, stable recovery is fundamentally constrained by the ill-posed nature of the inverse and by the spectral gap between the source and query waveforms. Classical spiking deconvolution, for example, can be degraded by finite-bandwidth, mixed-phase, or spatially varying GPR responses (Bowen, 1986; Schmelzbach & Huber, 2015; Economou & Vafidis, 2012).

Cross-waveform inversion. Physics-based full-waveform inversion requires repeated forward and gradient computations, whereas learned inversion avoids this repetition through a measurement-to-property mapping. However, changing the excitation waveform changes the effective measurement operator, causing a model trained under one waveform to confound excitation effects with changes in the underlying scene. Transfer-based methods address this mismatch by fine-tuning pretrained inversion networks using labeled or semi-supervised target-frequency data (Liu et al., 2023), while domain-adaptation methods align waveform-dependent latent features through adversarial training (Imai et al., 2025). Source-independent full-waveform inversion instead modifies the inversion objective to reduce sensitivity to unknown source wavelets, but still requires iterative full-wave forward and gradient computations (Wang et al., 2025). Learned approaches have also incorporated waveform-aware constraints directly into the inversion framework (Ge et al., 2025). Existing methods therefore address waveform mismatch either by adapting network parameters, latent features, objectives, or training data. In contrast, our method estimates a reusable measurement-level response representation that can be re-excited under a specified waveform before downstream inversion.

3 METHOD

Given a source scan acquired under a known excitation, our goal is to identify a response state that can be reused under subsequent waveform queries. The key distinction is between response content supported by the realized source experiment and content that remains weakly observable but may become exposed under another sensing operator. We first characterize cross-operator query identifiability, then construct a source-evidence anchor and an event-aligned structured continuation, and finally admit that continuation according to source observability before applying a query waveform.

3.1 CROSS-OPERATOR QUERY IDENTIFIABILITY

We formulate the cross-operator ambiguity for the position-varying response, while estimating the lateral common component separately under its calibrated received-wavelet convention. Let e_s denote the realized source excitation and e_q a subsequent known query excitation, with finite-prefix convolution operators $C_s = C(e_s)$ and $C_q = C(e_q)$. For a zero-mean position-varying response $R \in \mathbb{R}^{T\times M}$ satisfying $R\mathbf{c} = 0$, the corresponding source and query measurements are

$$Y_s^{\mathrm{var}} = C_s R, \qquad Y_q^{\mathrm{var}} = C_q R. \tag{1}$$

For a single temporal response vector $r \in \mathbb{R}^T$, write $y_s = C_s r$ and $y_q = C_q r$. For a fixed source observation y_s^{obs} and discrepancy tolerance $\delta \geq 0$, define the admissible set

$$\mathcal{A}_s(\delta; y_s^{\mathrm{obs}}) = \left\{ r : \|C_s r - y_s^{\mathrm{obs}}\|_2 \leq \delta \right\}. \tag{2}$$

In the exact noiseless setting ($\delta = 0$), the query response $C_q r$ is uniquely determined by y_s^{obs} for every admissible response if and only if

$$\ker C_s \subseteq \ker C_q. \tag{3}$$

Every temporal direction invisible to the source must therefore also be invisible to the query. The result applies columnwise to the position-varying response matrix in Eq. (1); for $\delta > 0$, one must instead analyze stability or error propagation. Appendix B.1 proves Eq. (3) and gives its equivalent operator factorization.

Practical band-limited measurements are governed more often by weak than exact null directions. A perturbation may have little influence under C_s while becoming consequential under C_q. We call this change in response visibility *cross-operator observability transfer*. In frequency terms, a source waveform may weakly constrain response content in regions that a subsequent waveform excites strongly; more generally, source-weak response directions may become query-exposed. Consequently, *source consistency does not imply query sufficiency*: accurately reproducing the realized source measurement does not by itself guarantee an accurate response to another sensing operator.

3.2 EVIDENCE ANCHOR AND EVENT-ALIGNED RESPONSE CONTINUATION

We separate the lateral component shared across scan positions from position-dependent variations. With $\mathbf{c} = [1, \ldots, 1]^\top \in \mathbb{R}^M$,

$$\begin{aligned} P_{\mathrm{com}} &= \frac{\mathbf{c}\mathbf{c}^\top}{\mathbf{c}^\top\mathbf{c}}, & P_{\mathrm{var}} &= I_M - P_{\mathrm{com}}, \\ Y_s &= \bar{y}_s \mathbf{c}^\top + Y_s^{\mathrm{var}}, \quad qquad & \bar{y}_s &= \frac{Y_s \mathbf{c}}{\mathbf{c}^\top\mathbf{c}}, \quad Y_s^{\mathrm{var}} = Y_s P_{\mathrm{var}}. \end{aligned} \tag{4}$$

The common component is estimated independently from $\bar{y}_s$ by the source-regularized temporal solve in Appendix B.3; the remaining construction concerns Y_s^{var}. A thin singular-value decomposition provides compact lateral coordinates and a source-supported response:

$$\begin{aligned} Y_s^{\mathrm{var}} &= U_y \Sigma_y V_y^\top, & Q &= V_y(:, 1\!:\!r_Q), \quad B_s = Y_s^{\mathrm{var}} Q, \\ \widehat{Z}_d &= (C_s^\top C_s + \lambda_d D_d)^{-1} C_s^\top B_s, \quad & R_d &= \widehat{Z}_d Q^\top. \end{aligned} \tag{5}$$

Only nonzero (or numerically resolved) right-singular directions are retained, so $Q^\top Q = I_{r_Q}$ and $Q^\top \mathbf{c} = 0$. Here D_d is a source-derived temporal support precision. We call R_d the *source-evidence*

anchor: it preserves content supported by the realized experiment rather than inventing content that the source does not reliably constrain. The common-mode solve and the construction of D_d are given in Appendix B.3.

Source-weak content is represented in event-aligned coordinates instead of by directly extrapolating a query measurement. Suppressing the structural-candidate index for clarity, each effective event j follows a shared cross-position trajectory and carries a compact local response:

$$\begin{aligned} \tau_j(x) &= \mathcal{T}(s(x), z_j; \psi) + \mathcal{T}(z_j, r(x); \psi)\,, \\ R_g(t,x) &= \sum_{j=1}^{J} a_j(x;\psi)\,\kappa_j(t-\tau_j(x), x)\,, \\ \kappa_j(u,x) &= \sum_{p,\ell} \theta_{jp\ell}\, b_p(u)\, \chi_{j\ell}(x). \end{aligned} \tag{6}$$

Here a_j captures propagation-dependent amplitude variation, b_p are fixed compact temporal basis functions, and $\chi_{j\ell}$ are low-order spatial features. Event alignment removes the dominant cross-position timing variation before local response morphology is modeled, while the shared coefficients $\theta_{jp\ell}$ restrict each event to a finite-dimensional family. Under the ideal continuous-shift interpretation, and ignoring finite-prefix truncation, its frequency-domain form is

$$\widetilde{R}_g(f,x) \approx \sum_{j=1}^{J} a_j(x;\psi) e^{-\mathrm{i}2\pi f \tau_j(x)} \sum_{p,\ell} \theta_{jp\ell} B_p(f) \chi_{j\ell}(x), \tag{7}$$

where B_p is the Fourier transform of b_p. The source observation constrains the coefficients through the source-excited portion of this compact model; once identified, the same response family remains defined in source-weak spectral regions. This is structural continuation rather than information creation: unobserved content is restricted to responses compatible with the source-fitted event model.

Event locations, propagation parameters, and local-response coefficients are fitted exclusively from the source observation using variable projection with a regularized linear inner solve. Source-only aggregation of predefined structural hypotheses yields

$$R_g = \sum_{k=1}^{K} \pi_k R_{g,\mathrm{full}}^{(k)} P_{\mathrm{var}}, \qquad \pi_k \geq 0, \quad \sum_{k=1}^{K} \pi_k = 1. \tag{8}$$

No held-out target measurement or label under the query waveform enters this construction; Appendix B.4 gives the implementation details.

3.3 Observability-Gated Minimal-Intervention Completion

The evidence anchor should remain authoritative where the source reliably constrains the response, whereas structural continuation should intervene only in weakly observable directions. We determine this division from the source operator. Let

$$C_s = U_s \operatorname{diag}(\sigma_i) V_s^\top, \qquad \rho_i^{(s)} = \frac{\sigma_i^2}{\sigma_{\max}^2}, \qquad \sigma_{\max} = \max_i \sigma_i. \tag{9}$$

The vectors v_i are temporal response modes of the finite-prefix source operator rather than individual Fourier bins. We assign the evidence anchor a mode-specific weight

$$o_i = \frac{\sigma_i^2}{\sigma_i^2 + \lambda_s}, \qquad \gamma_i = \sqrt{\frac{\rho_i^{(s)}}{\rho_i^{(s)} + \varepsilon}}, \qquad \omega_i = o_i \gamma_i, \tag{10}$$

and assemble the source-observability operator

$$\mathcal{O}_s = V_s \operatorname{diag}(\omega_i) V_s^\top. \tag{11}$$

Because $0 \leq \omega_i \leq 1$, $\mathcal{O}_s$ and $I_T - \mathcal{O}_s$ form a soft observability partition: strongly supported modes remain under evidential control, whereas source-weak modes receive progressively greater

structural freedom. The final frozen configuration uses the source-domain regularization selected by the source-only GCV/configuration rule for λ_s; the calibration convention is fixed before this selection. The normalized-support transition uses $\varepsilon = 3.162 \times 10^{-11}$, selected on six independent development cases and then frozen.

The re-excitable variation response is

$$\begin{aligned} R^\star &= R_d + (I_T - \mathcal{O}_s)(R_g - R_d) \\ &= \mathcal{O}_s R_d + (I_T - \mathcal{O}_s) R_g. \end{aligned} \tag{12}$$

The first form exposes a *minimal structural intervention*: R_d is the evidence-supported reference, $R_g - R_d$ proposes a structured correction, and only the component admitted by $I_T - \mathcal{O}_s$ modifies the anchor. The equivalent observability-weighted variational problem, modal bounds, and uniqueness of the closed-form solution are given in Appendix B.5. Together with the independently estimated common response $\widehat{h}_{\mathrm{com}}$, $R^\star$ forms the reusable response representation.

We deliberately condition this completion on source observability rather than on a particular query operator. This keeps $R^\star$ reusable across waveform queries, while each C_q determines at readout which completed response directions affect the synthesized measurement.

3.4 Query-Time Re-Excitation

After the response state has been identified, a known query condition is applied only at the measurement projection stage. The common-mode branch uses the calibrated query received wavelet w_q, while the position-varying branch uses the known query excitation e_q:

$$\widehat{Y}_q = C(w_q)\widehat{h}_{\mathrm{com}}\mathbf{c}^\top + C_q R^\star P_{\mathrm{var}}. \tag{13}$$

The evidence anchor, structured continuation, and source-observability operator are all determined before the query is applied. Re-excitation therefore reads the same response state through a different known sensing operator rather than re-estimating it for every source–query pair. Within the response and calibration models, the same $R^\star$ can support multiple waveform queries and can place measurements acquired under different excitations into a common waveform domain for comparison or inversion.

4 Experiments

4.1 Experimental Setup

Tasks. We evaluate the proposed representation on two tasks: EM super-resolution and cross-waveform inversion. In EM super-resolution, a measurement acquired with a lower-frequency source waveform is re-excited under a higher-frequency query waveform, allowing closely spaced reflection events that overlap in the source measurement to become more separable. In cross-waveform inversion, a measurement acquired under a non-reference waveform is first re-excited into the waveform domain on which a frozen inverse model was trained, enabling the same model to recover scene properties without waveform-specific retraining or additional labeled data. Detailed acquisition, calibration, and implementation settings are provided in Appendix C.

Datasets. For EM super-resolution, we construct a seven-case synthetic benchmark. It includes three frequency shifts, G15→G25, G9→G25, and G5→G25, to test increasing source–query spectral gaps; letter-shaped and digit-shaped reflectors to evaluate irregular geometry; a spatially varying soil background to assess heterogeneous propagation; and 30-dB source noise to test robustness to corrupted measurements.

We further evaluate conversion on the TU1208 limestone Line 2 dataset at 200→400 MHz (Dérobert & Pajewski, 2018). The data were collected at a full-scale geophysical test site approximately 30 m long, with a base width of about 5 m and depths ranging from roughly 3.3 to 4.7 m. The limestone section contains buried pipes, electrical cables, artificial cavities, and rock blocks at different depths. These measurements include real antenna responses, heterogeneous propagation, acquisition effects, and interacting subsurface reflections, providing a challenging test beyond the controlled synthetic conditions.

For cross-waveform inversion, we use controlled synthetic data for two complementary downstream tasks. The first evaluates three-class stratigraphic segmentation after G15→G25, G9→G25, and G5→G25 conversion. The second recovers dense permittivity maps and object geometry in multi-object scenes after G9→G15 and G5→G15 conversion. Both tasks use inverse models trained only on their native reference-waveform measurements and then frozen.

Baselines. We compare against six strong baselines spanning physics-guided reconstruction and contemporary learning-based restoration. The physics-guided group includes Spiking deconvolution (Turner, 1994), BISR-rational (Selesnick & Bayram, 2016; Goudarzi & Dehghan-Manshadi, 2026), and Matching Pursuit (MP) (Mallat & Zhang, 1993; Tuncer & Gürbüz, 2012), covering classical inverse filtering, sparsity-enhanced deconvolution, and greedy adaptive sparse reconstruction, respectively. The learning-based group comprises three recent high-performing architectures with substantially different modeling mechanisms: MambaIRv2 (CVPR 2025), an attentive state-space restoration model (Guo et al., 2025); ResShift (NeurIPS 2023), an efficient residual-shifting diffusion model for image restoration (Yue et al., 2023); and DDBM (ICLR 2024) with DBIM sampling (ICLR 2025), which models translation directly as a diffusion bridge between paired endpoint distributions (Zhou et al., 2024; Zheng et al., 2025). Together, these baselines cover deterministic restoration, residual diffusion, diffusion-bridge translation, and three representative physics-guided reconstruction strategies. All learning-based baselines are trained on the same 3,000 paired G15→G25 scenes with direct source–query supervision, whereas our method constructs each re-excitable response from a single source scan and known waveforms without paired cross-waveform training data or held-out target measurements during inference.

4.2 EM Super-Resolution: Accuracy and Efficiency

EM super-resolution. Figure 3 and Table 1 jointly compare the proposed method with the six baselines under four benchmark groups. In the frequency-gap block (a), the G5 source produces broad, overlapping arrivals that do not reveal whether a response comes from one reflector or several nearby reflectors, whereas the G25 target separates these events more clearly; this is the super-resolution effect that the conversion is intended to recover. The learning-based baselines degrade sharply as the gap widens because they are trained only on paired G15→G25 scenes and have no response-level mechanism for completing source-weak spectral components under the new excitation. The physics-guided baselines also lose accuracy, but their wavelet and sparse-reconstruction priors make the decline less severe; their residual error exposes the source–query observability mismatch rather than eliminating it. By combining source-supported evidence with the shared event structure and observability-conditioned completion, Ours limits this degradation: its largest frequency-gap MSE is 38.8, compared with 182.9 for the best physics-guided baseline and 956.9 for the best learning-based baseline. Across the remaining groups (b)–(d), Ours preserves the reflection structure under letter- and digit-shaped geometries, nonuniform soil, and 30-dB noise, achieving the lowest MSE and highest SSIM in every case.

Time cost analysis. Across the seven cases, Ours averages 9.836 s per source realization (Table 1). The fastest learning-based baseline, MambaIRv2, takes 0.666 s per case after deployment, but requires approximately 150 h to generate the paired training data and a 26.7-min training run. ResShift and the DDBM model used with DBIM sampling require 36.2 and 35.3 min of training, respectively.

Real-world tests. TU1208 measurements provide a stringent sim-to-real evaluation because the 200→400 MHz Line 2 profile combines different antenna bands, heterogeneous limestone propagation, instrument responses, and overlapping reflections from buried targets. On this 20.64-m profile, Ours reconstructs the target-like reflection trajectories more coherently than the physics-guided baselines and achieves the lowest raw MSE, 2.13641×10^{-4}. This is lower than Spiking (7.93384×10^{-4}), BISR (7.93161×10^{-4}), and MP (7.93838×10^{-4}), corresponding to a 73.1% reduction relative to the best reported baseline, BISR. These results support the transferability of the re-excitable response representation to measured data under different waveform and acquisition conditions. Details of data preprocessing, waveform calibration, and implementation are provided in Appendix C.3.

Taken together, these results support the central claim: source-only re-excitation improves target fidelity across spectral and acquisition shifts at a practical per-scan inference cost.

Table 1: Quantitative super-resolution results on seven synthetic cases. MSE is computed on the original-amplitude target images in amplitude-squared units (lower is better); SSIM is unitless (higher is better).

	Proposed	**Physics-guided**			**Learning-based**		
Case	Ours MSE/SSIM	BISR-rat. MSE/SSIM	Spiking MSE/SSIM	MP MSE/SSIM	MambaIRv2 MSE/SSIM	ResShift MSE/SSIM	DBIM MSE/SSIM
Frequency shifts							
G15→G25	**0.1**/**1.000**	4.0/0.973	18.1/0.912	24.2/0.929	20.1/0.900	18.4/0.921	6.1/0.958
G9→G25	**6.1**/**0.981**	50.9/0.805	131.1/0.717	241.3/0.693	1069.9/0.335	1209.4/0.339	1065.2/0.359
G5→G25	**38.8**/**0.806**	182.9/0.577	375.6/0.363	687.4/0.481	969.0/0.258	956.9/0.244	968.3/0.271
Geometry shifts							
Letter-shape	**6.1**/**0.975**	55.3/0.805	130.5/0.677	142.5/0.643	945.7/0.192	1036.3/0.189	943.5/0.211
Digit-shape	**9.6**/**0.973**	80.4/0.781	151.4/0.660	202.0/0.608	1023.8/0.188	1133.7/0.182	1020.5/0.194
Background and noise							
Nonuniform soil	**2.6**/**1.000**	24.0/0.963	249.7/0.842	7.2/0.982	1766.4/0.855	2021.1/0.863	1624.8/0.873
30-dB noise	**49.6**/**0.854**	59.3/0.779	206.0/0.467	151.1/0.621	1120.9/0.213	1208.8/0.305	1081.9/0.289
Mean MSE↓	**16.1**	65.3	180.4	207.9	988.0	1083.5	958.6
Mean SSIM↑	**0.941**	0.812	0.663	0.708	0.420	0.435	0.451
Mean time (s)↓	9.836	23.732	5.223	**0.665**	0.666	5.392	0.742

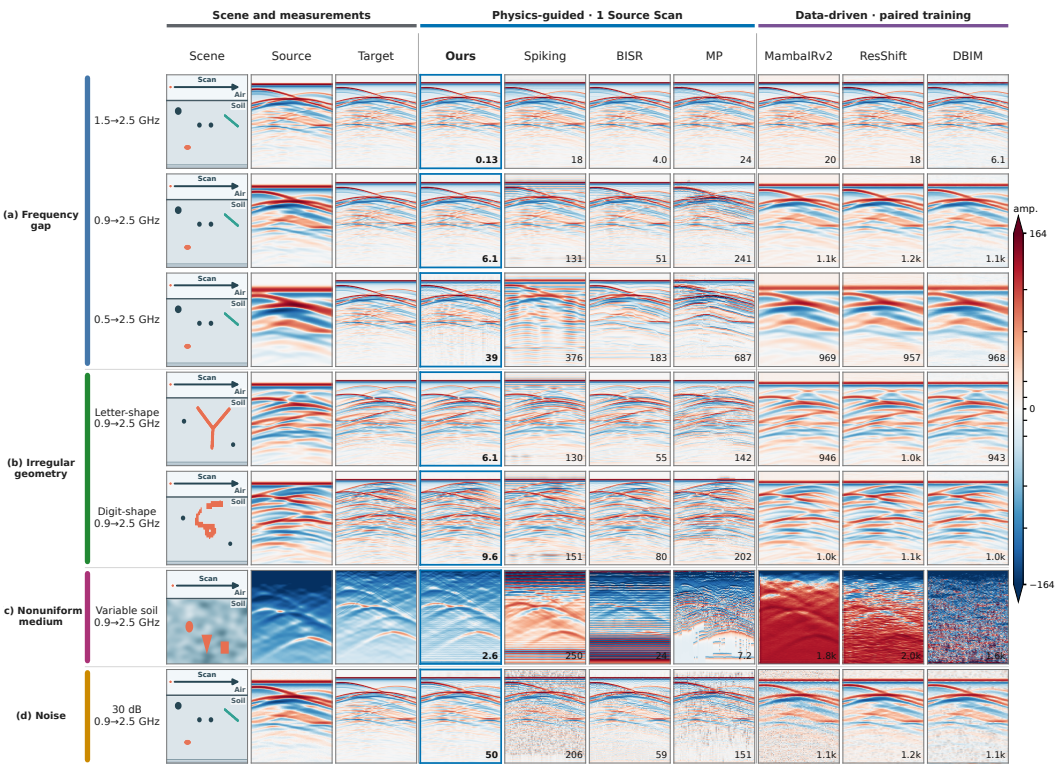

Figure 3: Comparison of our method with six baselines for EM super-resolution across seven synthetic cases: (a) frequency gaps, (b) irregular geometries, (c) a nonuniform medium, and (d) noise. The source echoes broaden and overlap as the frequency gap grows; Ours retains coherent target-like trajectories, whereas the baselines show increasing blur or artifacts under the tested shifts. Insets report full-array MSE in amplitude-squared units (lower is better).

Cross-waveform stratigraphic segmentation. We freeze a TransUNet (Chen et al., 2021) trained on labeled native G25 measurements and test direct and converted G15, G9, and G5 inputs after conversion to G25. As shown in Figure 5, Ours obtains foreground mIoUs of 0.909, 0.912, and 0.867, versus 0.710, 0.280, and 0.280 for direct inputs and 0.916 for native G25. The strongest physics-guided baselines reach 0.600 and 0.654 at G9 and G5, while learned translators degrade outside

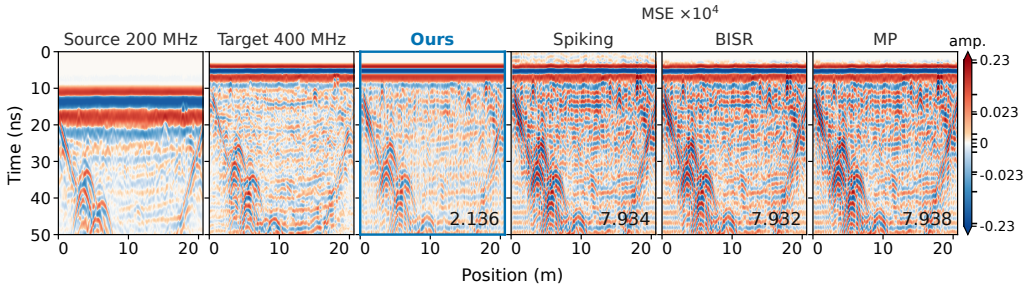

Figure 4: Comparison of our method with three baselines on the real-world TU1208 limestone Line 2 profile (20.64 m) under 200→400 MHz re-excitation. MSE is shown in each subfigure.

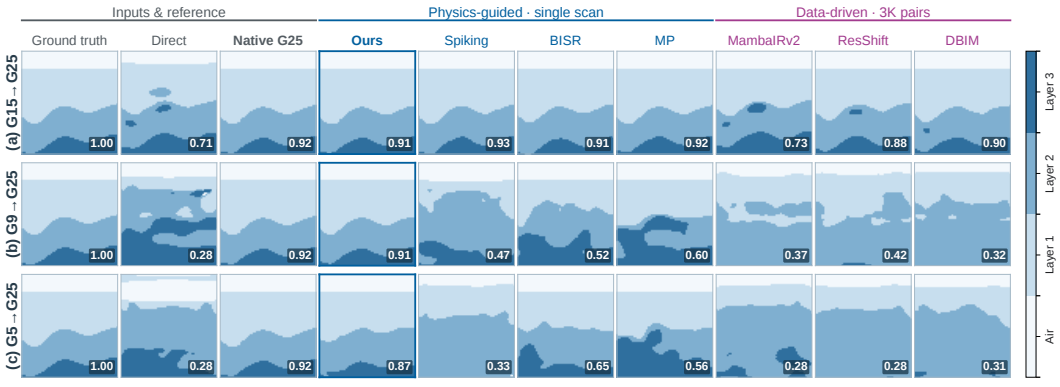

Figure 5: Cross-waveform stratigraphic inversion with a frozen G25 TransUNet. Rows show (a) G15→G25, (b) G9→G25, and (c) G5→G25. Insets report foreground mean IoU; colors denote air and three strata.

G15→G25; the gaps to native G25 are only 0.007, 0.004, and 0.049. Conversion outputs before target-domain reconstruction are shown in Appendix Figure 8; model details are in Appendix C.6.

Cross-waveform geometry and permittivity inversion. Appendix Figure 7 evaluates a frozen G15 inverse model on G9 and G5 measurements converted to G15. Ours obtains IoU/permittivity-MAE scores of 0.935/0.246 and 0.853/0.530, versus 0.841/0.478 and 0.647/1.143 for the strongest physics-guided baseline. The learned translators do not generalize to G9/G5→G15 without adaptation. Signal-space results and protocol details are in Appendix C.6 and Figure 9.

Ablation study. On G5→G25 (Appendix C.7), the full model achieves MSE 38.8/SSIM 0.806. Removing the anchor raises source NMSE to 0.117; removing continuation raises target MSE to 1488.1 despite near-zero source NMSE. Removing factorization, kinematics, or modal gating yields MSE 135.9/83.6/194.5 (scalar gating: SSIM 0.564), supporting their complementarity.

5 Conclusion and Limitations

Ours outperforms all baselines in simulated/measured EM super-resolution and downstream inversion with frozen models. It requires calibrated source/query waveforms; future work will address blind estimation and richer constraints on propagation, antenna response, boundaries, and multiple scattering.

AI USE STATEMENT

In this work, large language models (LLMs) were used to assist with the implementation of parts of the research code, the preparation of visualisations, the exploration and refinement of aspects of the methodology, and the editing of the manuscript for clarity and readability. All methodological decisions, implementations, visualisations, experimental results, analyses, and final manuscript content were reviewed and approved by the authors, who take full responsibility for the work, including material developed with LLM assistance.

REFERENCES

Adam D. Booth, Anthony L. Endres, and Tavi Murray. Spectral bandwidth enhancement of GPR profiling data using multiple-frequency compositing. *Journal of Applied Geophysics*, 2009. doi: 10.1016/j.jappgeo.2008.09.015. URL https://doi.org/10.1016/j.jappgeo.2008.09.015.

A. N. Bowen. A Comparison of Statistical and Deterministic Wiener Deconvolution of Deep-Tow Seismic Data. *Geophysical Prospecting*, 34(3):366–382, 1986. doi: 10.1111/j.1365-2478.1986.tb00472.x. URL https://doi.org/10.1111/j.1365-2478.1986.tb00472.x.

Jieneng Chen, Yongyi Lu, Qihang Yu, Xiangde Luo, Ehsan Adeli, Yan Wang, Le Lu, Alan L. Yuille, and Yuyin Zhou. TransUNet: Transformers make strong encoders for medical image segmentation. *arXiv preprint arXiv:2102.04306*, 2021. doi: 10.48550/arXiv.2102.04306. URL https://arxiv.org/abs/2102.04306.

Xavier Dérobert and Lara Pajewski. TU1208 open database of radargrams: The dataset of the IFSTTAR geophysical test site. *Remote Sensing*, 10:530, 2018. doi: 10.3390/rs10040530.

Nikos Economou and Antonis Vafidis. Deterministic deconvolution for GPR data in the t-f domain. *Near Surface Geophysics*, 9(5):427–433, 2011. doi: 10.3997/1873-0604.2011020. URL https://doi.org/10.3997/1873-0604.2011020.

Nikos Economou and Antonis Vafidis. GPR Data Time Varying Deconvolution by Kurtosis Maximization. *Journal of Applied Geophysics*, 81:117–121, 2012. doi: 10.1016/j.jappgeo.2011.09.004. URL https://doi.org/10.1016/j.jappgeo.2011.09.004.

Junkai Ge, Rui Liu, Shirong Zhang, Xiaodong Li, Huaifeng Sun, Bo Tian, and Ziqiang Zheng. Deep Learning-Based GPR Imaging for Permafrost: A Source Wavelet-Independent Inversion Approach. *IEEE Transactions on Geoscience and Remote Sensing*, 2025. doi: 10.1109/TGRS.2025.3632923. URL https://doi.org/10.1109/TGRS.2025.3632923.

Alireza Goudarzi and Seyed Hadi Dehghan-Manshadi. Enhanced Sparse Deconvolution Techniques for Ground-Penetrating Radar Data: A Comparative Study of Convex and Nonconvex Regularization Penalties. *Arabian Journal for Science and Engineering*, 51(11):14203–14216, 2026. doi: 10.1007/s13369-025-10605-z. URL https://doi.org/10.1007/s13369-025-10605-z.

Hang Guo, Yong Guo, Yaohua Zha, Yulun Zhang, Wenbo Li, Tao Dai, Shu-Tao Xia, and Yawei Li. MambaIRv2: Attentive state space restoration. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 28124–28133, 2025. doi: 10.1109/CVPR52734.2025.02619. URL https://openaccess.thecvf.com/content/CVPR2025/html/Guo_MambaIRv2_Attentive_State_Space_Restoration_CVPR_2025_paper.html.

Takanori Imai, Tsukasa Mizutani, Tatsuya Iguchi, and Toshihiro Haneda. Enhancing Deep Learning-Based GPR Data Inversion With Unsupervised Domain Adaptation: Comparison of Domain Classifiers. *IEEE Transactions on Geoscience and Remote Sensing*, 2025. doi: 10.1109/TGRS.2025.3568384. URL https://doi.org/10.1109/TGRS.2025.3568384.

Chao Li, Yikang Zheng, Xin Wang, Jinhai Zhang, Yibo Wang, Ling Chen, Lei Zhang, Pan Zhao, Yike Liu, Wenmin Lv, Yang Liu, Xu Zhao, Jinlai Hao, Weijia Sun, Xiaofeng Liu, Bojun Jia, Juan Li, Haiqiang Lan, Wenzhe Fa, Yongxin Pan, and Fuyuan Wu. Layered subsurface in utopia basin of mars revealed by zhurong rover radar. *Nature*, 610:308–312, 2022. doi: 10.1038/s41586-022-05147-5.

Haifeng Li, Jiajun Guo, and Dezhen Song. Subsurface feature-based ground robot/vehicle localization using a ground penetrating radar. In *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, pp. 1716–1722. IEEE, 2024. doi: 10.1109/ICRA57147.2024.10611579.

Dawei Liu, Mei Zhou, Xiaokai Wang, Zhensheng Shi, Mauricio D. Sacchi, Wenchao Chen, Zhaodan Liu, and Xian Wang. Enhancing Ground-Penetrating Radar (GPR) Data Resolution Through Weakly Supervised Learning. *IEEE Transactions on Geoscience and Remote Sensing*, 2024. doi: 10.1109/TGRS.2024.3410184. URL https://doi.org/10.1109/TGRS.2024.3410184.

Hanchi Liu, Jing Wang, Jiaqi Zhang, Haonan Jiang, Jing Xu, Peng Jiang, Fengkai Zhang, Qingmei Sui, and Zhengfang Wang. Semisupervised Deep Neural Network-Based Cross-Frequency Ground-Penetrating Radar Data Inversion. *IEEE Transactions on Geoscience and Remote Sensing*, 2023. doi: 10.1109/TGRS.2023.3316153. URL https://doi.org/10.1109/TGRS.2023.3316153.

Stéphane G. Mallat and Zhifeng Zhang. Matching Pursuits with Time-Frequency Dictionaries. *IEEE Transactions on Signal Processing*, 41(12):3397–3415, 1993. doi: 10.1109/78.258082. URL https://doi.org/10.1109/78.258082.

Teddy Ort, Igor Gilitschenski, and Daniela Rus. GROUNDED: A localizing ground penetrating radar evaluation dataset for learning to localize in inclement weather. *The International Journal of Robotics Research*, 42(10):901–916, 2023. doi: 10.1177/02783649231183460.

David A. Paige, Svein-Erik Hamran, Hans E. F. Amundsen, Tor Berger, Patrick Russell, Reva Kakaria, Michael T. Mellon, Sigurd Eide, Lynn M. Carter, Titus M. Casademont, Daniel C. Nunes, Emileigh S. Shoemaker Thackston, Dirk Plettemeier, Henning Dypvik, Sanna Holm-Alwmark, and Briony H. N. Horgan. Ground penetrating radar observations of the contact between the western delta and the crater floor of jezero crater, mars. *Science Advances*, 10:eadi8339, 2024. doi: 10.1126/sciadv.adi8339.

Neil R. Peplinski, Fawwaz T. Ulaby, and Myron C. Dobson. Dielectric properties of soils in the 0.3–1.3-GHz range. *IEEE Transactions on Geoscience and Remote Sensing*, 33(3):803–807, 1995. doi: 10.1109/36.387598.

Cedric Schmelzbach and Emanuel Huber. Efficient Deconvolution of Ground-Penetrating Radar Data. *IEEE Transactions on Geoscience and Remote Sensing*, 53(9):5209–5217, 2015. doi: 10.1109/TGRS.2015.2419235. URL https://doi.org/10.1109/TGRS.2015.2419235.

Ivan W. Selesnick and İlker Bayram. Enhanced Sparsity by Non-Separable Regularization. *IEEE Transactions on Signal Processing*, 64(9):2298–2313, 2016. doi: 10.1109/TSP.2016.2518989. URL https://doi.org/10.1109/TSP.2016.2518989.

Honglei Shen, Gang Tian, Chunhui Tao, Hanchuang Wang, and Jinwei Fang. Multi-frequency data fusion via joint weighted deconvolution for resolution enhancement. *Journal of Applied Geophysics*, 203:104702, 2022. doi: 10.1016/j.jappgeo.2022.104702. URL https://doi.org/10.1016/j.jappgeo.2022.104702.

Mehmet Ali Çağrı Tuncer and Ali Cafer Gürbüz. Analysis of Orthogonal Matching Pursuit Based Subsurface Imaging for Compressive Ground Penetrating Radars. *Turkish Journal of Electrical Engineering and Computer Sciences*, 20(6):979–989, 2012. doi: 10.3906/elk-1104-3. URL https://doi.org/10.3906/elk-1104-3.

Greg Turner. Subsurface radar propagation deconvolution. *Geophysics*, 59(2):215–223, 1994. doi: 10.1190/1.1443583. URL https://doi.org/10.1190/1.1443583.

Xiangyu Wang, Junhong Chen, and Hai Liu. Enhancing Full-Waveform Inversion of Field GPR Data: A Source-Independent Approach With Dynamic Reference Selection via SE-Wave-U-Net. *Geophysics*, 90(1):KS15–KS30, 2025. doi: 10.1190/geo2024-0283.1. URL https://doi.org/10.1190/geo2024-0283.1.

Craig Warren, Antonios Giannopoulos, and Iraklis Giannakis. gprMax: Open source software to simulate electromagnetic wave propagation for ground penetrating radar. *Computer Physics Communications*, 209:163–170, 2016. doi: 10.1016/j.cpc.2016.08.020.

Jianghai Xia, Evan K. Franseen, Richard D. Miller, Thomas V. Weis, and Alan P. Byrnes. Improving ground-penetrating radar data in sedimentary rocks using deterministic deconvolution. *Journal of Applied Geophysics*, 54(1-2):15–33, 2003. doi: 10.1016/S0926-9851(03)00045-4. URL https://doi.org/10.1016/S0926-9851(03)00045-4.

Jianping Xiao and Lanbo Liu. Permafrost Subgrade Condition Assessment Using Extrapolation by Deterministic Deconvolution on Multifrequency GPR Data Acquired Along the Qinghai-Tibet Railway. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 9 (1):83–90, 2016. doi: 10.1109/JSTARS.2015.2487970. URL https://doi.org/10.1109/JSTARS.2015.2487970.

Fan Yang, Shufan Hu, Huilin Zhou, Yonghui Zhao, and Kunwei Feng. High-Frequency GPR Data Reconstruction With Conditional GAN and Contrastive Unpaired Translation. *IEEE Geoscience and Remote Sensing Letters*, 2026. doi: 10.1109/LGRS.2026.3657665. URL https://doi.org/10.1109/LGRS.2026.3657665.

Zongsheng Yue, Jianyi Wang, and Chen Change Loy. ResShift: Efficient diffusion model for image super-resolution by residual shifting. In *Advances in Neural Information Processing Systems*, volume 36, pp. 13294–13307. Curran Associates, Inc., 2023. doi: 10.52202/075280-0583. URL https://proceedings.neurips.cc/paper_files/paper/2023/file/2ac2eac5098dba08208807b65c5851cc-Paper-Conference.pdf.

Hancheng Zhang, Yuanyuan Hu, Qiang Wang, Zhendong Qian, and Pengfei Liu. End-to-end frequency enhancement framework for GPR images using domain-adaptive generative adversarial networks. *Computer-Aided Civil and Infrastructure Engineering*, 2025. doi: 10.1111/mice.13525. URL https://doi.org/10.1111/mice.13525.

Lei Zhang, Chao Li, Jinhai Zhang, Bin Zhou, Yu-Yan Sara Zhao, Yang Liu, Kaichang Di, Ross N. Mitchell, Juan Li, Zhigang Zhang, Lin Chen, Xiaofeng Liang, Weijia Sun, Yike Liu, Xu Zhao, Jinlai Hao, Changyi Xu, Yikang Zheng, Yibo Wang, Xin Wang, Pan Zhao, Wenmin Lv, Yang Li, Haiqiang Lan, Yuxi Li, Wei Wang, Yang Lu, Honglei Lin, Peng Fang, Wei Lin, Yong Wei, Ling Chen, Guangyou Fang, Yangting Lin, Zhenxing Yao, and Yongxin Pan. Buried palaeo-polygonal terrain detected underneath utopia planitia on mars by the zhurong radar. *Nature Astronomy*, 8: 69–76, 2024. doi: 10.1038/s41550-023-02117-3.

Kaiwen Zheng, Guande He, Jianfei Chen, Fan Bao, and Jun Zhu. Diffusion bridge implicit models. In *The Thirteenth International Conference on Learning Representations*, 2025. URL https://openreview.net/forum?id=eghAocvqBk.

Linqi Zhou, Aaron Lou, Samar Khanna, and Stefano Ermon. Denoising diffusion bridge models. In *The Twelfth International Conference on Learning Representations*, 2024. URL https://openreview.net/forum?id=FKksTayvGo.

APPENDIX CONTENTS

Appendix	Contents	Page
Appendix A	Notation, finite-prefix operators, and waveform calibration	14
Appendix B	Query identifiability, method derivations, and implementation details	16
Appendix C	Datasets, baselines, downstream protocols, and evaluation	20
Appendix D	Operator-level identifiability analysis	28

A NOTATION AND MEASUREMENT CONVENTIONS

This appendix defines the measurement operators and distinguishes source, query, and evaluation quantities. Within the method derivation, *source* denotes the realized sensing condition and *query* denotes a known waveform applied after response-state identification. *Target* denotes a held-out measured reference used only for evaluation; when available, it is acquired under the same waveform specified by the query.

A.1 NOTATION AND DIMENSIONS

Table 2 follows the data flow from the source scan to response completion and query-time re-excitation. Unless stated otherwise, temporal operators act from the left and lateral operators act from the right.

Table 2: Response re-excitation notation and dimensions.

Symbol	Definition
Measurements and temporal operators	
$T,\ M,\ t,\ x,\ \Delta t$	Number of recorded time samples, number of scan positions, continuous time coordinate, lateral scan coordinate, and temporal sampling interval.
$Y_s, Y_q, \widehat{Y}_q \in \mathbb{R}^{T\times M}$	Realized source measurement, ideal measurement under the specified query, and synthesized query measurement.
$Y_t \in \mathbb{R}^{T\times M}$	Held-out measured target used only for evaluation; its acquisition waveform matches the declared query in a conversion experiment.
e_s, e_q	Known transmitted source and query excitations used by the position-varying branch.
w_s, w_q	Calibrated received source and query wavelets used by the common-mode branch.
$C(u) \in \mathbb{R}^{T\times T}$	Finite-prefix linear convolution operator induced by sampled waveform u; $C_s = C(e_s)$ and $C_q = C(e_q)$.
$R, \Delta R \in \mathbb{R}^{T\times M}$	Candidate finite response and response perturbation used to express cross-operator ambiguity.
$\mathcal{A}_s(\delta; y_s^{\mathrm{obs}})$	Responses consistent with a fixed source observation up to discrepancy tolerance δ.
I_T, I_M, I_{r_Q}	Identity matrices on the temporal, full lateral, and reduced lateral-coordinate spaces.
$\|\cdot\|_2, \|\cdot\|_F$	Vector Euclidean norm and matrix Frobenius norm.
Orthogonal coordinates and evidence anchor	
$\mathbf{c} \in \mathbb{R}^M$	All-ones vector defining the lateral common mode.
$P_{\mathrm{com}}, P_{\mathrm{var}} \in \mathbb{R}^{M\times M}$	Orthogonal common-mode and zero-mean variation projectors.
$\bar{y}_s, Y_s^{\mathrm{var}}$	Lateral mean trace and centered source variation, with $Y_s = \bar{y}_s\mathbf{c}^\top + Y_s^{\mathrm{var}}$.
$\widehat{h}_{\mathrm{com}}, D_{\mathrm{com}}, \lambda_{\mathrm{com}}$	Estimated common-mode response, temporal support precision, and regularization strength.
U_y, Σ_y, V_y	Thin SVD factors of $Y_s^{\mathrm{var}} = U_y\Sigma_y V_y^\top$.
r_Q, Q, B_s	Retained lateral rank, orthonormal basis $Q \in \mathbb{R}^{M\times r_Q}$, and compact source evidence $B_s = Y_s^{\mathrm{var}}Q$.

Continued on the next page

Table 2: Response re-excitation notation and dimensions (continued).

Symbol	Definition
$\widehat{Z}_d, D_d, \lambda_d, R_d$	Source-supported coordinates, temporal precision, regularization strength, and source-evidence anchor $R_d = \widehat{Z}_d Q^\top$.
$H_b^{(p)}, a_b, g_b$	Pilot response, temporal activity, and normalized smoothed activity for branch $b \in \{\mathrm{com}, d\}$.
$\rho_p, \kappa_b, \sigma_b, \nu_b$	Pilot ridge ratio, waveform energy scale, smoothing width, and support-penalty strength used to construct D_b.
Event-aligned structured continuation	
K, k, J, j	Number of structural hypotheses, hypothesis index, number of effective events, and event index.
$z_j = (\xi_j, d_j), \psi, \mathcal{T}$	Effective event location, shared propagation parameters, and one-way travel-time model.
$s(x), r(x), \tau_j(x)$	Transmitter and receiver positions and the resulting two-way delay.
$R_g^{(k)}, R_{g,\mathrm{full}}^{(k)}, R_g$	Source-fitted hypothesis, full-grid rendering, and aggregated zero-mean structured continuation.
$a_j(x;\psi_k), \kappa_j(u,x)$	Propagation-dependent amplitude and compact local response of event j.
$b_p, B_p, \chi_{j\ell}, \theta_{jp\ell}$	Compact temporal basis, its Fourier transform, low-order spatial feature, and shared kernel coefficient.
L_k, P_k	Local kernel length and number of temporal basis functions for hypothesis k.
$\mathcal{D}_k, \mathbf{D}_k, W_s, \lambda_{\mathrm{aux}}$	Source synthesis operator, vectorized dictionary, whitening operator, and ridge strength.
$\widehat{\psi}_k, \pi_k$	Fitted propagation parameters and nonnegative source-only aggregation weights, with $\sum_k \pi_k = 1$.
D_x, h_m, ν_m, ζ	Horizontal separation, layer thickness, layer velocity, and ray parameter in the layered travel-time model.
$\mathcal{K}, \eta_w$	Retained non-null singular directions and whitening floor in W_s.
ℓ_k, η	Normalized source residual and numerical floor used by the known-noise aggregation rule.
Observability-gated completion	
U_s, σ_i, V_s, v_i	SVD quantities in $C_s = U_s \operatorname{diag}(\sigma_i) V_s^\top$; v_i is a temporal response mode.
$\sigma_{\max}, \rho_i^{(s)}$	Largest source singular value and normalized modal power $\rho_i^{(s)} = \sigma_i^2/\sigma_{\max}^2$.
λ_s, o_i	Source-confidence scale and absolute confidence $o_i = \sigma_i^2/(\sigma_i^2 + \lambda_s)$.
ε, γ_i	Observability transition scale and normalized support gate $\gamma_i = \sqrt{\rho_i^{(s)}/(\rho_i^{(s)} + \varepsilon)}$.
$\omega_i, \mathcal{O}_s$	Modal evidence weight $\omega_i = o_i\gamma_i$ and source-observability operator $\mathcal{O}_s = V_s \operatorname{diag}(\omega_i) V_s^\top$.
$R^\star$	Completed response $R_d + (I_T - \mathcal{O}_s)(R_g - R_d)$ in the zero-mean variation space.
Query-time re-excitation	
$C(w_q)\widehat{h}_{\mathrm{com}}\mathbf{c}^\top$	Query-wavelet projection of the common-mode response.
$C_q R^\star P_{\mathrm{var}}$	Query-excitation projection of the completed position-varying response.

A.2 Finite-Prefix Convolution and Waveform Calibration

For a sampled waveform u and response $H \in \mathbb{R}^{T\times M}$, the recorded prefix of their linear convolution is

$$[C(u)H]_{n,i} = \sum_{\ell=0}^{T-1} u_{n-\ell} H_{\ell,i}, \qquad n = 0, \ldots, T-1, \tag{14}$$

where $u_k = 0$ outside the sampled waveform support. FFT padding implements linear convolution but does not add observations or impose zeros on an unrecorded convolution tail.

The position-varying component uses transmitted excitations e_s and e_q, with $C_s = C(e_s)$ and $C_q = C(e_q)$, whereas the common component uses calibrated received wavelets w_s and w_q. The two branches therefore follow different calibration conventions and represent different effective response quantities; they are recombined only after projection into measurement space. These operator

conventions must remain distinct. The evidence anchor, structured continuation, and their completion all remain in the zero-mean variation space:

$$R_d\mathbf{c} = R_g\mathbf{c} = R^\star\mathbf{c} = 0. \tag{15}$$

In simulation, the matching empty-scene record is subtracted before response estimation and supplies the received-wavelet calibration. The time origin, sample interval, finite-prefix convention, DC convention, and relative source-to-query gain are fixed for each waveform pair. Field-data calibration follows the direct-wave protocol in Appendix C.3.

B Method Details and Derivations

This section supplies the derivations and implementation details behind the cross-operator formulation, response construction, and observability-gated completion in the main method.

B.1 Proof of Cross-Operator Query Identifiability

Consider two temporal response vectors $r_1, r_2 \in \mathbb{R}^T$ that produce the same noiseless source measurement. With $\Delta r = r_1 - r_2$,

$$C_s r_1 = C_s r_2 \quad \Longleftrightarrow \quad C_s \Delta r = 0. \tag{16}$$

If $\ker C_s \subseteq \ker C_q$, then $C_q \Delta r = 0$, and the two responses produce the same query measurement. This proves sufficiency. Applying the argument to each temporal response column gives the corresponding matrix statement for the position-varying response.

For necessity, take any $r \in \ker C_s$ and set $r_1 = r$, $r_2 = 0$. These vectors produce the same noiseless source observation. Unique query prediction requires $C_q r = 0$, so every $r \in \ker C_s$ also belongs to $\ker C_q$, proving Eq. (3).

The same condition is equivalent, in finite dimensions, to factorization of the query operator through the source operator. Define a map on $\mathrm{range}(C_s)$ by

$$\widetilde{A}_{q\leftarrow s}(C_s r) = C_q r. \tag{17}$$

Kernel inclusion makes this definition independent of the representative r. Extending the map linearly outside $\mathrm{range}(C_s)$ yields an operator $A_{q\leftarrow s}$ satisfying

$$C_q = A_{q\leftarrow s} C_s. \tag{18}$$

Exact source-only query prediction therefore requires no additional response assumption precisely when the query introduces no sensitivity to directions invisible to the source. In finite-bandwidth and noisy measurements, exact null directions are replaced by weakly observable directions, motivating the continuous operator $\mathcal{O}_s$ used by the method.

B.2 Orthogonal Response Coordinates

The lateral projectors introduced in Eq. (4) satisfy

$$P_{\mathrm{com}}^2 = P_{\mathrm{com}}, \qquad P_{\mathrm{var}}^2 = P_{\mathrm{var}}, \qquad P_{\mathrm{com}} P_{\mathrm{var}} = 0, \qquad P_{\mathrm{com}} + P_{\mathrm{var}} = I_M. \tag{19}$$

Hence they provide a complete orthogonal partition of the lateral response space.

The compact basis Q is formed only from the nonzero, numerically resolved right-singular directions of Y_s^{var}. Thus its retained rank satisfies $r_Q \le \mathrm{rank}_\tau(Y_s^{\mathrm{var}})$ for the declared numerical threshold τ (that is, rank_τ counts singular values above τ), and the retained directions obey $Q^\top Q = I_{r_Q}$ and $Q^\top \mathbf{c} = 0$. This explicit rank restriction avoids assigning an arbitrary null-space basis to the lateral evidence coordinates.

For an arbitrary prediction error $E = \widehat{Y} - Y$,

$$\langle EP_{\mathrm{com}}, EP_{\mathrm{var}} \rangle_F = \mathrm{tr}\left(P_{\mathrm{com}} E^\top E P_{\mathrm{var}}\right) \tag{20}$$

$$= \mathrm{tr}\left(P_{\mathrm{var}} P_{\mathrm{com}} E^\top E\right) = 0, \tag{21}$$

and therefore

$$\|E\|_F^2 = \|EP_{\mathrm{com}}\|_F^2 + \|EP_{\mathrm{var}}\|_F^2. \tag{22}$$

Thus, errors in the common and position-varying branches cannot compensate for one another.

B.3 Source-Derived Temporal Support Precision

Pilot responses. The diagonal matrices D_{com} and D_d encode where the source observation supports non-negligible temporal response energy. They are computed once from a lightly regularized source-only pilot and are soft precisions rather than hard time masks. For $b \in \{\mathrm{com}, d\}$, define

$$\begin{aligned} Y_{\mathrm{com}} &= \bar{y}_s, & C_{\mathrm{com}} &= C(w_s), \\ Y_d &= Y_s^{\mathrm{var}}, & C_d &= C_s, \end{aligned} \qquad H_b^{(p)} = \left(C_b^\top C_b + \rho_p \kappa_b I_T\right)^{-1} C_b^\top Y_b, \tag{23}$$

where $\rho_p = 10^{-10}$ and $\kappa_b = \max_f |\,\mathrm{FFT}(u_b)[f]|^2$, with $u_{\mathrm{com}} = w_s$ and $u_d = e_s$. The branch activities are

$$a_{\mathrm{com}}[n] = \left|H_{\mathrm{com}}^{(p)}[n]\right|, \qquad a_d[n] = \sqrt{\frac{1}{M}\sum_{m=1}^{M}\left(H_d^{(p)}[n,m]\right)^2}. \tag{24}$$

Support precision. Let G_{σ_b} denote a unit-mass one-dimensional Gaussian kernel with zero extension at the recorded-window boundary. We normalize its smoothed activity and form the temporal precision as

$$g_b[n] = \frac{(G_{\sigma_b} * a_b)[n]}{\max_{n'}(G_{\sigma_b} * a_b)[n'] + 10^{-30}}, \qquad D_b = \mathrm{diag}_n\left(1 + \nu_b[1 - g_b[n]]^2\right). \tag{25}$$

Thus, times with strong pilot activity receive a precision near one, whereas times with little source-supported activity receive a stronger response-energy penalty. This construction only controls temporal support; source observability across operator modes is handled separately by $\mathcal{O}_s$.

Selection and final solves. For the common branch, $\sigma_{\mathrm{com}} = 0.06$ ns and $\nu_{\mathrm{com}} = 3 \times 10^5$. The clean-data experiments use the development-frozen ridge ratio $\rho_{\mathrm{com}} = 1.7783 \times 10^{-14}$, so $\lambda_{\mathrm{com}} = \rho_{\mathrm{com}}\kappa_{\mathrm{com}}$. For the variation branch, $\sigma_d \in \{0.06, 0.15, 0.30\}$ ns, $\nu_d \in \{10^2, 10^3, 10^4, 3\times 10^4, 10^5, 3\times 10^5\}$, and ρ_d is searched over 45 logarithmically spaced values from 10^{-16} to 10^{-5}. The minimum source-only generalized-cross-validation family selects (σ_d, ν_d); among solutions within 5% of its minimum GCV, the largest ρ_d is retained. With $\widetilde{Y}_{\mathrm{com}} = \bar{y}_s$ and $\widetilde{Y}_d = B_s$, the two main-paper solves are equivalently

$$\widehat{H}_b = \operatorname*{argmin}_H \left\|C_b H - \widetilde{Y}_b\right\|_F^2 + \lambda_b \left\|D_b^{1/2} H\right\|_F^2, \qquad \lambda_b = \rho_b \kappa_b, \tag{26}$$

where $\widehat{H}_{\mathrm{com}} = \widehat{h}_{\mathrm{com}}$ and $\widehat{H}_d = \widehat{Z}_d$. The known-noise branch uses the separately declared frozen values in Appendix B.6.

B.4 Shared-Event Parameterization and Optimization

B.4.1 Layered Travel Time and Source-Only Initialization

For event j at effective position $z_j = (\xi_j, d_j)$, the two-way delay is the transmit–event–receive path in Eq. (6). For a one-way horizontal separation D_x, layer thicknesses h_m, and velocities ν_m, the ray parameter ζ satisfies

$$D_x = \sum_m \frac{h_m \zeta}{\sqrt{\nu_m^{-2} - \zeta^2}}, \qquad \mathcal{T} = \zeta D_x + \sum_m h_m \sqrt{\nu_m^{-2} - \zeta^2}. \tag{27}$$

The implementation solves for ζ on the physical branch using bounded Newton iterations with bisection fallback. This is a low-dimensional propagation prior, not an assertion that every effective event is a physical interface or that the medium is exactly layered.

Candidate event locations are proposed on a fixed search grid by the reduction in the multi-trace source residual after fitting a short local kernel. A minimum spacing prevents duplicate proposals. Initialization uses only the source measurement and scan geometry; true object locations, target measurements, and target labels are not used. Event count and kernel length are capacity choices rather than estimates of the number of physical objects.

B.4.2 Compact Temporal Kernel Basis

For a temporal support of length L, the fixed basis functions used in Eq. (6) are constructed as

$$\bar{b}_p[n] = \cos\left(\frac{\pi(n+\frac{1}{2})p}{L}\right)\sin^2\left(\frac{\pi(n+\frac{1}{2})}{L}\right), \qquad n = 0, \dots, L-1, \tag{28}$$

followed by unit-ℓ_2 normalization,

$$b_p[n] = \frac{\bar{b}_p[n]}{\|\bar{b}_p\|_2}, \tag{29}$$

with $b_p[n] = 0$ outside the local support. The $\sin^2$ taper provides smooth temporal localization, while increasing p introduces progressively richer local oscillatory structure without allowing an unconstrained per-trace waveform.

Let $B_p(f)$ be the discrete-time Fourier transform of b_p. Under the ideal continuous-shift interpretation, and ignoring finite-prefix truncation, shifting the compact kernel to the event trajectory gives

$$\widetilde{R}_g(f,x) \approx \sum_{j=1}^{J} a_j(x;\psi)e^{-\mathrm{i}2\pi f\tau_j(x)}\sum_{p,\ell}\theta_{jp\ell}B_p(f)\chi_{j\ell}(x), \tag{30}$$

which is the frequency-domain form used to interpret event-aligned response continuation. The implementation itself uses the finite discrete rendering in Eq. (6).

Fractional event delays are implemented by linear interpolation.

B.4.3 Source-Domain Variable Projection

For structural hypothesis k, let $\mathcal{D}_k(\psi)[\theta] \in \mathbb{R}^{T\times r_Q}$ denote the linear source-domain synthesis operator induced by the current event geometry, compact kernel basis, source excitation, and projection into the coordinates B_s. The propagation parameters are fitted by

$$\widehat{\psi}_k = \arg\min_{\psi}\ \min_{\theta}\left[\|W_s\left(\mathcal{D}_k(\psi)[\theta] - B_s\right)\|_F^2 + \lambda_{\mathrm{aux}}\|\theta\|_2^2\right], \tag{31}$$

where W_s whitens the source-domain residual. To state the inner solve as an ordinary matrix ridge problem, define

$$\mathbf{b}_s = \mathrm{vec}(B_s), \qquad \mathbf{D}_k(\psi)\theta = \mathrm{vec}(\mathcal{D}_k(\psi)[\theta]), \qquad \mathbf{W}_s = I_{r_Q}\otimes W_s. \tag{32}$$

With $C_s = U_s\,\mathrm{diag}(\sigma_i)V_s^\top$, the implementation uses

$$W_s = \mathrm{diag}_{i\in\mathcal{K}}\left[\left(\sigma_i^2 + \eta_w\sigma_{\max}^2\right)^{-1/2}\right]U_{s,\mathcal{K}}^\top, \tag{33}$$

where $\mathcal{K}$ excludes numerical null directions and $\eta_w > 0$ is a small whitening floor. Whitening changes the weighting of source-observable directions but does not create information in a source-null direction.

For fixed propagation parameters, the kernel coefficients have the regularized linear solution

$$\widehat{\theta}_k(\psi) = \left[\mathbf{D}_k(\psi)^\top\mathbf{W}_s^\top\mathbf{W}_s\mathbf{D}_k(\psi) + \lambda_{\mathrm{aux}}I\right]^{-1}\mathbf{D}_k(\psi)^\top\mathbf{W}_s^\top\mathbf{W}_s\mathbf{b}_s. \tag{34}$$

The linear coefficients are therefore eliminated at every geometry evaluation, leaving only the nonlinear propagation parameters to the outer optimization.

In the implementation, dictionary columns are normalized before the linear solve. The absolute ridge level is obtained from the prescribed relative regularization parameter using the scale of the

whitened Gram matrix, and a small whitening floor is used for numerical stability. The outer geometry optimization uses an Adam stage followed by L-BFGS refinement. After the geometry is frozen, the final compact-kernel coefficients are refitted analytically under the prescribed source-domain capacity rule.

B.5 Observability-Conditioned Completion

B.5.1 Structural Candidate Aggregation

The predefined structural hypotheses are assigned source-only aggregation weights $\pi_k \geq 0$ with $\sum_{k=1}^{K} \pi_k = 1$. The default protocol uses equal weights,

$$\pi_k = \frac{1}{K}. \tag{35}$$

For the declared known-noise extension, let ℓ_k be the normalized source-coordinate residual of candidate k. We instead use

$$\pi_k = \frac{(\ell_k + \eta)^{-2}}{\sum_{j=1}^{K} (\ell_j + \eta)^{-2}}, \qquad \eta = 10^{-12}, \tag{36}$$

where the exponent is selected once on the independent development set and then frozen. This weighting suppresses source-inconsistent structural candidates without inspecting the held-out target response.

These weights act only when forming the single structural representation R_g in Eq. (8); they do not introduce an additional temporal modal-confidence term.

B.5.2 Modal Weight Bounds

For $\rho_i^{(s)} \geq 0$, $\lambda_s > 0$, and $\varepsilon > 0$, the two factors in Eq. (10) satisfy

$$0 \leq o_i = \frac{\sigma_i^2}{\sigma_i^2 + \lambda_s} \leq 1, \qquad 0 \leq \gamma_i = \sqrt{\frac{\rho_i^{(s)}}{\rho_i^{(s)} + \varepsilon}} \leq 1. \tag{37}$$

Consequently, $0 \leq \omega_i = o_i \gamma_i \leq 1$. Therefore $\mathcal{O}_s$ and $I_T - \mathcal{O}_s$ are positive semidefinite, so both quadratic terms in Eq. (12) are well defined. The allocation between R_d and R_g is determined directly by source observability, while the structural candidates have already been aggregated before this modal composition.

B.5.3 Closed-Form Completion

For brevity, write

$$\mathcal{O} = \mathcal{O}_s. \tag{38}$$

The objective in Eq. (12) is

$$\mathcal{J}(R) = \left\| \mathcal{O}^{1/2}(R - R_d) \right\|_F^2 + \left\| (I_T - \mathcal{O})^{1/2}(R - R_g) \right\|_F^2. \tag{39}$$

Its gradient is

$$\nabla_R \mathcal{J} = 2\mathcal{O}(R - R_d) + 2(I_T - \mathcal{O})(R - R_g) \tag{40}$$

$$= 2\left[R - \mathcal{O}R_d - (I_T - \mathcal{O})R_g\right]. \tag{41}$$

Setting the gradient to zero yields

$$\begin{aligned} R^\star &= R_d + (I_T - \mathcal{O})(R_g - R_d) \\ &= \mathcal{O}R_d + (I_T - \mathcal{O})R_g. \end{aligned} \tag{42}$$

Because

$$\nabla_R^2 \mathcal{J} = 2I_T, \tag{43}$$

the objective is strictly convex and the solution is unique. Moreover, both R_d and R_g lie in the zero-mean variation space and $\mathcal{O}$ acts only along the temporal dimension, so $R^\star \mathbf{c} = 0$ automatically. Consequently, the final multiplication by P_{var} in the query projection is theoretically redundant; it is retained as a numerical projection to suppress finite-precision mean leakage.

B.6 Known-Noise Robustness Configuration

For the 30-dB experiment, a source-only positive-part spectral shrinker is followed by a Marchenko–Pastur rank guard before response estimation. The noise variance is computed from the declared SNR with multiplier 1.0. The common mode is the frozen source-support estimate obtained with pilot ridge 3.162×10^{-8} and support ridge 3.162×10^{-9}; the calibrated G9→G25 target-wavelet gain is 1.0526. If ℓ_k denotes the normalized source-coordinate residual of structural candidate k, candidate weights follow Eq. (36). The same source-modal composition as the clean dense cases is then applied.

These settings were fixed using 15 realizations from three non-test 30-dB development scenes. The reported test uses three independent seeds (20260924–20260926) on a different scene.

C Experimental Protocol

This section records the datasets, calibration, baseline configurations, downstream protocols, and evaluation rules used in the experiments.

C.1 Synthetic Benchmarks and Acquisition Settings

The synthetic super-resolution benchmark is generated with two-dimensional FDTD simulations using gprMax (Warren et al., 2016). Except for the wider nonuniform-soil aperture, each recorded B-scan contains 637 time samples and 71 bistatic positions over 15.001 ns. The receiver moves from 0.11 to 0.81 m in 0.01-m steps and the transmitter remains 0.06 m to its right. The simulation grid spacing is 1 mm. The nonuniform-soil case contains 101 positions over a 1-m aperture. For every waveform, we subtract the matching empty-scene trace, retain the original amplitude scale, and decimate time by a factor of ten without per-case normalization. The labels G5, G9, G15, and G25 denote Gaussian excitation parameters of 0.5, 0.9, 1.5, and 2.5 GHz, respectively.

Table 3: Seven-case simulated super-resolution benchmark. Rows list the shift, waveform pair, perturbation, and evaluation domain.

Shift	Waveforms	Scene or perturbation	Evaluation
Frequency	G15→G25	Close-reflector reference scene	All 71 traces
Frequency	G9→G25	Close-reflector reference scene	All 71 traces
Frequency	G5→G25	Close-reflector reference scene	All 71 traces
Geometry	G9→G25	Letter-shape reflector	All 71 traces
Geometry	G9→G25	Digit-shape reflector	All 71 traces
Background	G9→G25	Spatially varying Peplinski soil	All 101 traces
Noise	G9→G25	30-dB Gaussian source noise; three seeds	Mean over all traces and seeds

Noise is added after direct-wave subtraction and temporal decimation, using seeds 20260924, 20260925, and 20260926; the target remains clean.

C.2 Nonuniform-Soil Benchmark

The soil-background case in Table 1 uses the Peplinski material model (Peplinski et al., 1995) with sand and clay fractions of 0.6125 and 0.1875, respectively, a bulk density of $1.4875\,\mathrm{g/cm^3}$, a particle density of $2.66\,\mathrm{g/cm^3}$, and volumetric water content in $[0.065, 0.1225]$. Spatial variation is generated with fractal dimension 1.5, 50 material bins, and random seed 2026081905 on a 1-mm grid. The background is therefore nonuniform rather than a single constant-permittivity material.

For the soil row in Fig. 3, samples before 3 ns are omitted consistently from the source, target, and all predictions to suppress the dominant direct/interface arrival and expose the attenuated reflections. This gate affects only the displayed images; the MSE shown in Fig. 3 and reported in Table 1 is computed on the full target image.

C.3 TU1208 Limestone Measurement Protocol

We use the TU1208 limestone Line 2 profiles (Dérobert & Pajewski, 2018) acquired with nominal 200- and 400-MHz antennas. The 200-MHz profile is linearly resampled from 1,035 to 512 traces to match the 400-MHz grid. A fixed rigid lateral registration aligns source columns 0–507 with target columns 4–511, corresponding to a four-trace shift.

Because no empty-site record accompanies the field profile, we estimate the effective direct wave from the laterally registered common arrival. Let $t_0 = 1.0557$ ns and let $M_r = 508$ be the number of registered trace pairs. The source and target calibration wavelets are

$$w_s[n] = \frac{1}{M_r}\sum_{i=0}^{507} Y_s(t_0 + n\Delta t, i), \qquad w_t[n] = \frac{1}{M_r}\sum_{i=4}^{511} Y_t(t_0 + n\Delta t, i), \tag{44}$$

using 422 source samples and 157 target samples, respectively. Each wavelet is normalized to unit ℓ_2 norm inside the convolution solve; the recorded norms, 983.7 and 8.64×10^4, restore the acquisition-specific amplitude scale. Figure 6 shows both estimated wavelets and their recorded amplitudes.

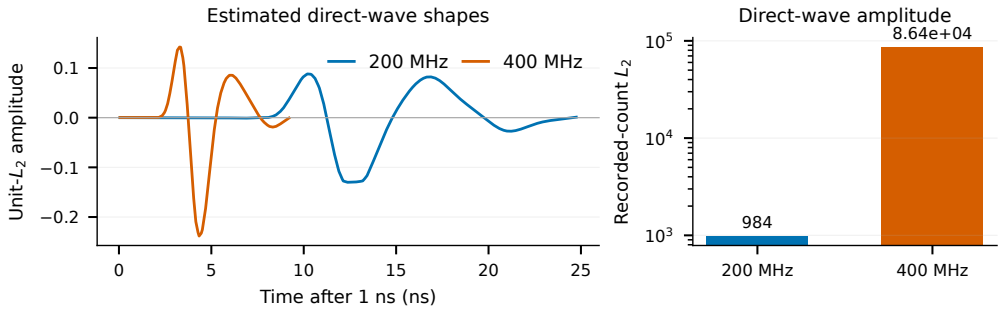

Figure 6: TU1208 direct-wave calibration from the lateral mean of 508 registered trace pairs. Left: unit-ℓ_2 waveforms used in convolution. Right: recorded norms used to restore the 200- and 400-MHz amplitude scales (log axis).

C.4 Baseline Implementation Details

All simulated baselines use the same direct-wave-subtracted source scan, original floating-point amplitude scale, temporal grid, and evaluation domain as the proposed method. The three deconvolution baselines estimate a response under the calibrated source wavelet and synthesize the target measurement by convolving that response with the known target wavelet. Their configurations are summarized in Table 4.

Table 4: Implementation of the physics-guided baselines on the simulated benchmark. Parameters are shared across the reported cases unless stated otherwise.

Method	Response estimator	Configuration
Spiking	Optimum-delay loaded FIR inverse of the calibrated source wavelet	Filter lengths 32–768; prewhitening 10^{-14}–10^{-1}; retained-wavelet energy thresholds 10^{-2}–10^{-12}
BISR-rational	Forward–backward sparse deconvolution with the adjacent-sample rational penalty	Wavelet tail 10^{-3}; $\lambda = 10^{-4}\lambda_{\max}$; 100,000 iterations
MP	Signed-correlation matching pursuit on the finite causal convolution dictionary	Wavelet tail 10^{-3}; relative source-residual threshold 0.02; at most 100,000 iterations

The learning-based baselines are signal-space adaptations of their official implementations. All three are trained on the same 3,000 paired G15–G25 training scenes and selected using 190 validation scenes. Each B-scan has shape 467×90. Source and target arrays are divided by one RMS scale computed from the training split and restored to the original amplitude scale at inference; no per-scene normalization, image resizing, or output clipping is used. Table 5 records the configurations used for the reported comparisons.

Table 5: Implementation of the learning-based baselines. Parameter counts refer to trainable network parameters.

Method	Parameters	Training	Configuration
MambaIRv2	22.65M	3,000 steps	128×90 crops; Charbonnier loss; Adam, 2×10^{-4}; FP32
ResShift	116.68M	3,000 steps	15-step residual diffusion; unweighted x_0 MSE; AdamW, 5×10^{-5}; EMA 0.999; FP32
DBIM	105.19M	10,000 steps	DDBM bridge training; RAdam, 10^{-4}; EMA 0.999; mixed precision; first-order DBIM, NFE 10, $\eta = 0$

C.5 One-Time Data and Training Cost Estimate

Data-generation and training cost. At 1.5 min per single-waveform FDTD run, the 6,000 runs required for 3,000 paired G15–G25 scenes take approximately 150 h on one worker. Available logs give estimated training times of 26.7 min for MambaIRv2, 36.2 min for ResShift, and 35.3 min for DDBM, totaling approximately 1.6 h.

Learning-based checkpoints. For compactness, *DBIM* in the main-paper tables denotes a DDBM bridge model evaluated with the first-order DBIM sampler; DBIM is the sampler rather than a separately trained network. The extended 10,000-step budget is used for this baseline to reduce undertraining, whereas MambaIRv2 and ResShift use 3,000 steps. All checkpoints are fixed before evaluation on the seven reported benchmark cases. MambaIRv2 uses one prespecified inference seed. ResShift and DBIM use two fixed draws, and their quantitative results are averaged over the draws.

All three learning baselines use batch size one, training seed 20260923, and zero output clipping. MambaIRv2 uses Adam with $(\beta_1, \beta_2) = (0.9, 0.99)$; its learning-rate milestones at 20,000 and 27,000 steps are not reached by the reported checkpoint. ResShift uses zero weight decay, an exponential 15-step schedule with minimum noise level 0.04, terminal coefficient 0.99, power 0.3, and x_0 prediction. DBIM also uses zero weight decay and a variance-preserving bridge with $\sigma_{\mathrm{data}} = 0.5$, $\sigma_{\min} = 10^{-4}$, $\sigma_{\max} = 1$, $\beta_d = 2$, $\beta_{\min} = 0.1$, and zero source–target covariance. ResShift and DBIM are evaluated with their EMA checkpoints.

TU1208 configurations. For the TU1208 experiment, the same lateral registration, amplitude convention, and reflection window are applied to every method. Because an independent query-band calibration wavelet is unavailable, the effective 400-MHz wavelet is estimated from the early direct arrival as defined in Eq. (44); later reflection samples do not enter this wavelet estimate. For each physics-guided baseline, the native parameter is selected by the minimum 200-MHz source-replay MSE over its predeclared grid on 128 source calibration traces; the measured 400-MHz reflection image is not opened until all predictions have been sealed. This gives a 512-tap Spiking filter with prewhitening 0.003162, BISR with $\lambda = 10^{-8}\lambda_{\max}$, and MP with a relative source-residual threshold of 0.005715. BISR is refitted for 400,000 iterations and MP is allowed at most one million iterations. Although Fig. 4 displays 0–50 ns, BISR and MP are fitted using the complete recorded 0–60 ns source trace and are cropped only for display. Their finite-prefix convolution dictionaries retain late response atoms whose visible energy is at least 10^{-4} of a complete atom; fitting the longer source record avoids an artificial near-zero tail at the bottom of the displayed panels.

C.6 Cross-Waveform Inversion Protocol

Each downstream inverse model is trained only on measurements from its declared reference waveform, which is used as the query waveform during re-excitation: G25 for stratigraphic segmentation

and G15 for multi-object permittivity recovery. At evaluation, a direct source measurement or a query-domain measurement produced by each conversion method is passed to the corresponding frozen inverse model; no inverse-network parameter is updated for a source waveform.

Stratigraphic segmentation. The random-layer dataset contains four semantic classes: background and three subsurface strata. Its native G25 split contains 1,875 training images and 126 validation images. We use an R50–ViT-B/16 TransUNet with three skip connections, a 512×512 input, and four output classes (Chen et al., 2021). It is trained for 150 epochs with batch size eight, SGD with initial learning rate 10^{-2}, momentum 0.9, weight decay 10^{-4}, and polynomial decay exponent 0.9. The loss is the equal-weighted sum of cross-entropy and multiclass Dice losses, and the random seed is 1234. The checkpoint with the best validation layer IoU is frozen for all reported cross-waveform tests.

The converter predicts the direct-wave-subtracted residual on a temporally decimated grid. We linearly interpolate it to the native 6361×75 grid, add the G25 direct arrival, resize the result bilinearly to 512×512, and repeat the single radar channel three times. The unconverted source input uses its corresponding source direct arrival instead. For each input, we report IoU for all four classes, pixel accuracy, and foreground mIoU averaged over the three stratum classes. For G5→G25, the displayed task input retains the task-selected five-candidate response and replaces a weight of 0.095 by the refined low-ringing capacity response shown in Fig. 8, followed by the declared 1.0-sample coarse-grid delay. Thus, the appendix panel is the audited capacity component of the Fig. 5 composition; applying that component alone is a different diagnostic.

Multi-object geometry and permittivity. The multi-object collection contains 10,000 simulated scenes with dense permittivity maps. The selected protocol uses 7,997 scenes for training, 240 for development, and 700 disjoint sealed-evaluation scenes. The inverse model is a 181.97-million-parameter `DualCoordinatePermittivityNet` implemented on a CCSeg network with a ResNet-101 encoder. Two parallel polar-coordinate and Cartesian-coordinate decoder branches reconstruct multi-scale features; a polar-to-Cartesian projection and cross-attention module exchange information between the branches before residual batch-normalized fusion. The fused features feed Cartesian and polar semantic heads, an auxiliary object-mask head, and dense permittivity regression heads. The reported `ccseg_direct_layer4` model uses the Cartesian dense head to predict a normalized log-permittivity map and the Cartesian mask logit for geometry. It is initialized from the architecture-screening checkpoint, trained on the full training split for three epochs at learning rate 10^{-4}, refined for two epochs at 3×10^{-5}, and adapted for one epoch at 10^{-5} with only the highest encoder stage unfrozen. All stages use AdamW with weight decay 10^{-4} and seed 20260925. The frozen checkpoint is selected using the development split and has no overlap with the 700-scene sealed evaluation, where it obtains object IoU 0.834, full-map permittivity MAE 0.082, and object-region MAE 0.908.

At inference, each input scan is resized bicubically to 512×512, normalized as $(Y + 4.1615 \times 10^{-6})/0.2298$, and repeated over three input channels. The predicted normalized log-permittivity is converted to relative permittivity as $\exp(\mathrm{clip}(\widehat{z}, 0, 1) \log 81)$, and the object mask is obtained from $\mathrm{sigmoid}(\ell) > 0.5$. We report pixel IoU against the true object mask and mean absolute relative-permittivity error inside the true object region. Full-map MAE is retained only as an auxiliary metric; the annotations in Appendix Figure 7 report object IoU and true-object-region permittivity MAE. The same preprocessing, frozen checkpoint, mask threshold, and display scale are used for every converter.

The displayed held-out scene is record 16378. It contains three objects with relative permittivities 11.7, 11.3, and 11.9 in a background of 7.01. The G9→G15 and G5→G15 response estimates use only the corresponding source scan and calibrated waveforms. Spiking, BISR-rational, and MP are fitted independently from each single source scan and then projected to G15. MambaIRv2, ResShift, and DBIM retain their frozen G15→G25 checkpoints and training RMS; they are applied directly to G9/G5→G15 without adaptation or per-scene re-normalization.

C.6.1 Waveform-Conversion Inputs to the Frozen Inverse Models

Figures 8 and 9 visualize the archived direct-wave-subtracted converter outputs before the target direct arrival is reinserted and the arrays are resized for the frozen inverse networks. Within each

figure, every source, target, and prediction uses one symmetric asinh color normalization whose limits are fixed by the 99.5th percentile of the measured target magnitude. The transform is used only for display; inset MSE values are computed on the original untransformed amplitudes over the complete recorded array. For the terrain conversion figure, the G5→G25 panel uses the continuous modal transition, analytic fractional delay, and source-fitted capacity bank. This refinement reduces ringing without post-hoc image smoothing.

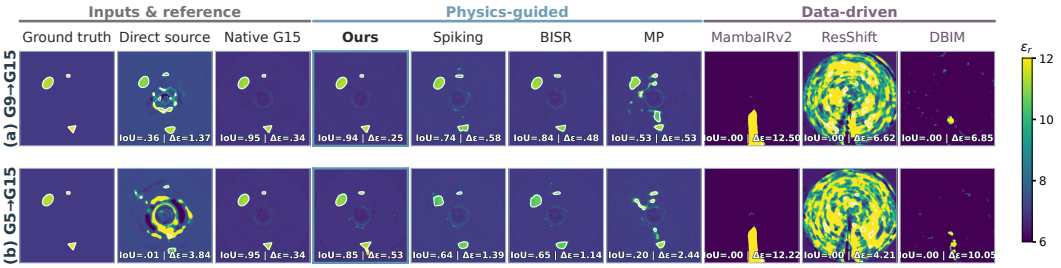

Figure 7: Multi-object geometry/permittivity inversion with a frozen G15 model: (a) G9→G15 and (b) G5→G15. The annotations report object IoU and true-object-region permittivity MAE.

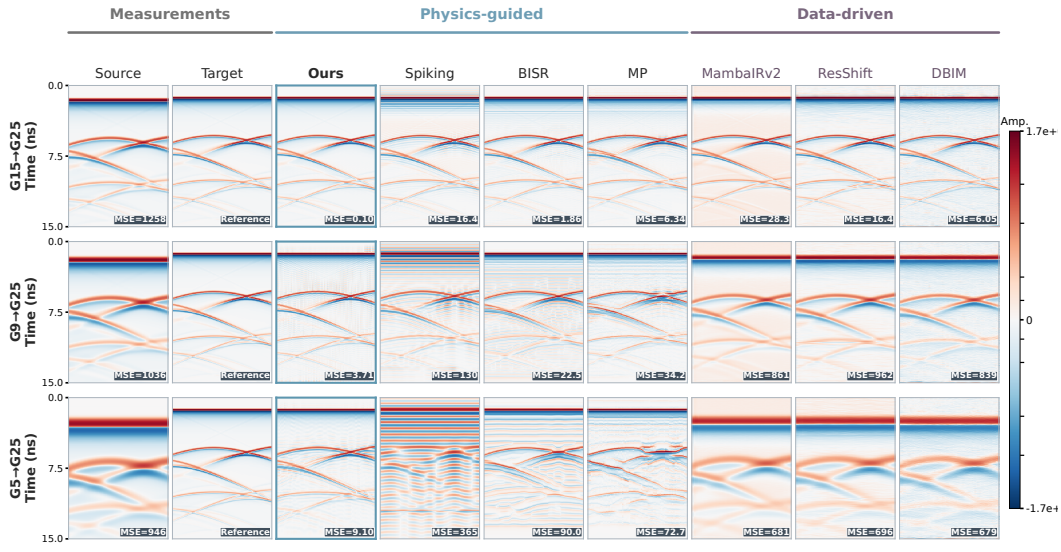

Figure 8: Waveform conversion preceding the frozen G25 stratigraphic inverse model. Rows show G15→G25, G9→G25, and G5→G25; columns show the source scan, measured G25 target, Ours, three physics-guided baselines, and three learned baselines. All panels share the target-derived amplitude scale, and insets report full-array MSE on the original amplitude. The G5→G25 Ours panel uses the refined configuration.

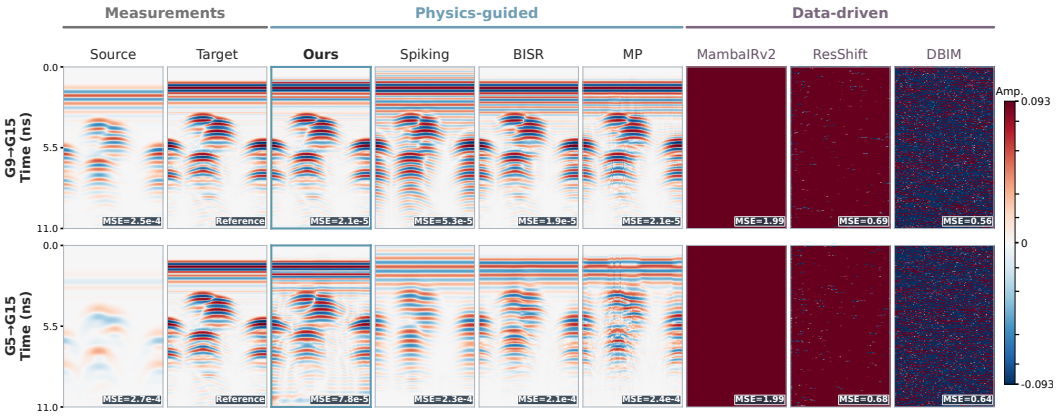

Figure 9: Waveform conversion preceding the frozen G15 multi-object inverse model. Rows show G9→G15 and G5→G15; columns follow Fig. 8. All panels share the measured-G15 amplitude scale, and insets report full-array MSE. The learned outputs are direct transfers of checkpoints trained for G15→G25 without adaptation or per-scene re-normalization.

C.7 Component Ablation

Table 6: Component ablation on the simulated G5→G25 case. Source NMSE measures consistency with the source observation, while target MSE and SSIM measure re-excitation accuracy.

Variant	Source NMSE↓	Target MSE↓	Target SSIM↑
Full model	2.53×10^{-10}	38.8	0.806
w/o evidence anchor	1.17×10^{-1}	45.9	0.756
w/o event-aligned continuation	2.23×10^{-10}	1488.1	0.245
w/o orthogonal response coordinates	3.49×10^{-10}	135.9	0.742
w/o shared-event kinematics	$(2.587 \pm 0.002) \times 10^{-10}$	83.6 ± 1.7	0.765 ± 0.003
w/o source-observability gating	5.32×10^{-2}	194.5	0.564

C.8 Evaluation Metrics and Reporting

MSE and NMSE. Let Ω be the declared set of evaluated time–position samples. We use the original measurement amplitudes and define

$$\mathrm{MSE}(\widehat{Y}, Y) = \frac{1}{|\Omega|} \sum_{q \in \Omega} (\widehat{Y}_q - Y_q)^2, \qquad \mathrm{NMSE}(\widehat{Y}, Y) = \frac{\sum_{q \in \Omega} (\widehat{Y}_q - Y_q)^2}{\sum_{q \in \Omega} Y_q^2}. \tag{45}$$

MSE is averaged over every sample of the target image for the simulated cases. For the 30-dB case, each noise seed is scored separately and the table reports the arithmetic mean; uncertainty is the population standard deviation over the three seeds. The source NMSE in the ablation table follows the same definition with the source measurement as Y.

SSIM. SSIM is computed on the raw-amplitude arrays without per-image normalization. We use a 7×7 uniform window, sample covariance, $K_1 = 0.01$, $K_2 = 0.03$, and

$$L_Y = \max_{q \in \Omega} Y_q - \min_{q \in \Omega} Y_q \tag{46}$$

as the data range. Scores are averaged over valid window centers, matching the scikit-image implementation with a uniform rather than Gaussian window and sample covariance enabled.

C.9 Computing Environment

Hardware. Experiments are run on one workstation with an Intel Core i9-14900KF CPU, 64 GB system memory, and one NVIDIA GeForce RTX 4090 with 24 GB memory.

D Operator-Level Identifiability Study

This study tests whether a single source excitation uniquely constrains the response components exposed by a query excitation. It is distinct from the component ablation in Table 6: the ablation measures the contribution of each proposed component, whereas this study measures ambiguity in the inverse problem before choosing a completion method. We use the same finite-prefix operators and frozen $\varepsilon = 3.162 \times 10^{-11}$ as in the main method, with the pair-gap scene fixed while varying the source excitation from G15 to G9 and G5; the query is G25 in all three cases.

D.1 Query Energy in the Source-Weak Subspace

Let v_i be the ith right singular vector of C_s and retain $\rho_i^{(s)}$ from Eq. (9). We define the normalized query modal power, the source-weak index set, and the fraction of query-operator energy contained in that subspace as

$$\begin{aligned} \tau_i^{(q)} &= \frac{\|C_q v_i\|_2^2}{\max_j \|C_q v_j\|_2^2}, \qquad & \mathcal{W}_\varepsilon &= \{i : \rho_i^{(s)} < \varepsilon\}, \\ \eta_{q|\mathcal{W}} &= \frac{\sum_{i \in \mathcal{W}_\varepsilon} \|C_q v_i\|_2^2}{\|C_q\|_F^2}. \end{aligned} \tag{47}$$

Because V_s is orthogonal, $\eta_{q|\mathcal{W}}$ is exactly the fraction of query-operator Frobenius energy acting on source-weak temporal directions. Figure 10(a–b) shows that it rises from 0.0536% for G15→G25 to 2.967% for G9→G25 and 21.326% for G5→G25. Thus, as the frequency separation increases, a larger portion of the query depends on response directions that the source operator constrains weakly.

D.2 Random Source-Indistinguishable Response Families

Perturbation construction. We next test whether the modal mismatch produces a family of materially different query predictions, rather than only one extreme singular vector. After removing the lateral mean, we construct a source-supported reference response by a truncated source pseudoinverse,

$$R_{\mathrm{ref}} = \sum_{i \notin \mathcal{W}_\varepsilon} v_i \sigma_i^{-1} u_i^\top Y_s^{\mathrm{var}}. \tag{48}$$

Its relative source residual is 0.00063%, 0.00096%, and 0.00168% for G15, G9, and G5, respectively. We then draw 256 response perturbations from the 32 source-weak temporal modes with the largest known-query operator energy and the leading eight lateral modes of Y_s^{var}. Gaussian coefficients use seed 20260925, and every draw is rescaled to $\|\Delta R\|_F = 0.1 \|R_{\mathrm{ref}}\|_F$. For each draw we measure

$$\delta_s = \frac{\|C_s \Delta R\|_F}{\|Y_s^{\mathrm{var}}\|_F}, \qquad \delta_q = \frac{\|C_q \Delta R\|_F}{\|Y_t^{\mathrm{var}}\|_F}. \tag{49}$$

The known query operator is used to identify query-relevant directions; the held-out target measurement is used only as the denominator of the reported δ_q evaluation scale and does not determine a perturbation direction. Because both operators are linear, the ratio δ_q/δ_s and the cross-waveform ordering are invariant to the chosen 0.1 perturbation scale.

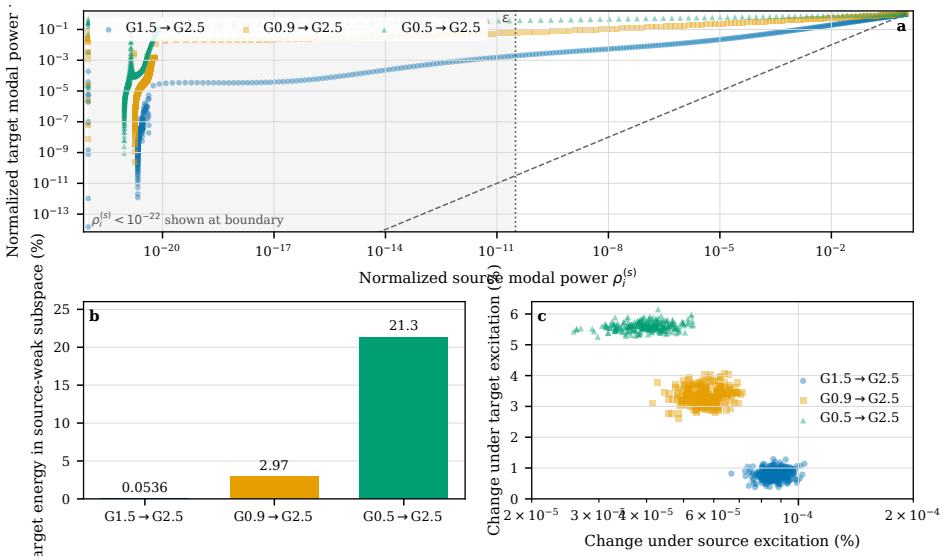

Figure 10: Operator-level identifiability study. (a) Source and query modal powers in the source right-singular basis; values of $\rho_i^{(s)} < 10^{-22}$ are displayed at the left boundary. The dotted line marks the frozen ε, and the dashed diagonal denotes equal normalized modal power. (b) Fraction of query-operator energy in the source-weak subspace. (c) Source- and query-measurement changes for 256 random response perturbations of fixed relative response norm.

Table 7: Identifiability statistics. Measurement changes and amplification are medians over 256 perturbations; Figure 10(c) shows their full distributions.

Source–query	$\lvert\mathcal{W}_\varepsilon\rvert$	$\eta_{q\lvert\mathcal{W}}$ (%)	δ_s (%)	δ_q (%)	δ_q/δ_s
G15→G25	421	0.0536	8.66×10^{-5}	0.789	9.18×10^{3}
G9→G25	509	2.967	5.70×10^{-5}	3.348	5.78×10^{4}
G5→G25	568	21.326	4.02×10^{-5}	5.599	1.41×10^{5}

Although each perturbation has 10% of the reference-response norm, its median source-measurement effect remains below $10^{-4}\%$ in every case. The corresponding median query effect increases from 0.789% to 5.599% as the frequency gap widens. This supplies constructive evidence for multiple practically source-indistinguishable responses with distinct query predictions. The ordering is insensitive to the exact transition threshold: varying ε from 10^{-2} to 10^{2} times its frozen value gives source-weak query-energy intervals of 0.0150–0.1596%, 2.026–4.491%, and 17.487–25.847% for G15, G9, and G5, respectively.

This study establishes the need for additional structure in source-weak directions, but it does not claim that any structural prior uniquely recovers the physical response. Whether the proposed shared-event completion improves the query prediction is evaluated separately by the main benchmarks and component ablations.