

Parameter compensation reveals weak physical effects in multi-parameter inverse learning

Yifan Yang

Chongqing Jiaotong University <https://orcid.org/0000-0001-8907-8135>

Zengwei Guo

zengweiguo@cqjtu.edu.cn

Chongqing Jiaotong University

Article

Keywords: Scientific machine learning, Inverse problems, Physics-informed neural networks, Parameter identifiability, Parameter compensation

Posted Date: October 15th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10934849/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: There is **NO** Competing Interest.

Yifan Yang, Chongqing Jiaotong University, No. 66 Xuefu Avenue, Nan'an District, Chongqing, P. R. China.

Email: yyf_cqjtu@163.com; ORCID: <https://orcid.org/0000-0001-8907-8135>

(*Corresponding author) Zengwei Guo, Chongqing Jiaotong University, No. 66 Xuefu Avenue, Nan'an District, Chongqing, P. R. China.

Email: zengweiguo@cqjtu.edu.cn; ORCID: <https://orcid.org/0000-0001-9972-4014>

Abstract

Inverse learning from sparse observations can reproduce measured responses while assigning them to an incorrect combination of physical parameters. This failure can remain hidden at low response error or optimized loss, especially when weak effects are absorbed by compensatory changes in another parameter. We formalize this as parameter compensation and establish local conditions for the existence, persistence and recoverability of compensatory representations. We demonstrate the principle in the minimal nontrivial two-parameter setting, providing a basis for hierarchical extensions. We then develop a bias-compensated neural network (BcNN) with two single-parameter anchors in adjacent physical representations that expose structured pseudo-solutions. Their ordered difference defines a physical gauge for complementary bias reparameterization. Across controlled theoretical benchmarks and sparse structural measurements, BcNN recovers weak physical effects more reliably than baseline PINN. These results show accurate response reconstruction is not sufficient evidence of physical recovery and that representation-dependent pseudo-solutions can provide information for inverse learning.

Keywords: Scientific machine learning; Inverse problems; Physics-informed neural networks; Parameter identifiability; Parameter compensation

Inverse learning is increasingly used not only to reproduce observable fields, but also to infer latent quantities that are interpreted as physical properties, mechanisms or system states [1]. Physics-informed neural networks (PINN) and related physics-informed operator methods have expanded this capability by combining sparse observations with governing equations [2,3,4]. Yet an important difficulty remains. A model may reproduce the measured response accurately while assigning that response to an incorrect combination of physical parameters. This is especially problematic in multi-parameter inverse problems, where the contribution of a weak physical mechanism can be absorbed by changes in a more influential parameter. An inverse solution may therefore appear successful according to response error or optimized loss while still giving the wrong physical explanation. When the inferred parameters themselves carry scientific meaning, this distinction is not secondary. It determines whether the inverse analysis has actually recovered the underlying physics.

Non-uniqueness of physical parameters has long been studied through structural and practical identifiability [5,6,7,8], equifinality [9,10] and sloppy or weakly constrained parameter directions [11,12,13,14]. Related behaviour has also been described through compensatory parameters and parameter trade-offs [15,16]. Similar ambiguity has been observed in neural inverse rendering, where different latent physical explanations can remain compatible with the same data [17,18,19]. Recent work on PINN inverse transfer learning has likewise reported parameter compensation when low field error coexists with inaccurate recovered parameters [20]. These studies show that ambiguity can persist even when the observable response is fitted well. What remains less clear is how that ambiguity can be organized so that it reveals which parameter is being masked, which parameter carries the compensatory change, and whether that relation changes when the physical representation is altered.

Scientific machine learning has recently broadened physics-informed inverse analysis through physics encoding, operator learning and improved treatment of difficult inverse problems [21,22,23,24,25,26,27]. Structural identifiability has also been extended to PDE settings [28], and non-unique inverse solution families have begun to be represented explicitly [29]. These advances have improved how ambiguity is detected and described, but they do not explicitly organize parameter compensation as a directed relation across related physical representations. As a result, a pseudo-solution that reproduces the observations is usually viewed as an undesirable consequence of non-uniqueness. It may, however, contain information about how the inverse problem redistributes physical contributions among parameters.

We use changes in physical representation to probe this possibility. Consider two related governing models that retain the same shared physical structure. The enriched representation contains an additional physical mechanism, while the reduced representation removes that mechanism. If both models can reproduce the same sparse observations, the retained carrier parameter may need to shift in order to compensate for what has been removed. That shift provides information about how the observable contribution of the missing mechanism is redistributed. Representation reduction is therefore used here as a controlled probe of the inverse problem rather than simply as a way to simplify the governing model. By removing one mechanism while keeping the observations and shared physical structure unchanged, compensation that is difficult to see within a single representation can become exposed. In this sense, an observationally admissible pseudo-solution need not be only a failed estimate. It can also contain structured information about the inverse problem.

Here we establish a parameter compensation principle for multi-parameter inverse learning and formulate compensation as a directed and representation-dependent inverse structure. Existence describes when a locally unidentifiable or near-null direction involves both a compensation target and at least one distinct carrier parameter. Persistence describes whether the observable change caused by representation reduction can be reproduced locally by changing the designated carrier parameter under the conditions given in Methods. Recoverability describes when an admissible parameter representative can be locally isolated from the observational equivalence class. We introduce a remained-bias representation that records the target and carrier parameters together with their physical representation orders, and we use it to develop a bias-compensated neural network (BcNN). The present study examines the simplest nontrivial realization with two unknown physical parameters and two adjacent physical representations. The same notation leaves room for more parameters and deeper representation hierarchies, although constructing compatible anchors and admissible regions in those settings remains an open problem. Extension to nonlinear governing systems may also be possible when a meaningful hierarchy of physical representations and sufficiently regular observable response maps can be defined.

BcNN uses two independently trained anchors with one trainable physical parameter each. These anchors expose compensatory pseudo-solutions in adjacent physical representations. Their ordered difference defines a physical gauge that provides the scale for complementary bias reparameterization in the dual-parameter target network. BcNN does not create observational information and does not increase the rank of the local parameter sensitivity Jacobian. Instead, it organizes compensation that is already present in the inverse problem so that weak but distinguishable physical information is less likely to remain absorbed in an unstructured compensatory state. We evaluate this mechanism using controlled Sturm-Liouville inverse problems and sparse measurements from composite beams exhibiting compression softening. Response reconstruction and physical parameter recovery are evaluated separately throughout.

Results

A parameter compensation principle for inverse learning

Accurate response reconstruction does not necessarily imply unique physical parameter recovery. Even when the corresponding forward problem is well-posed, the response map can be non-injective or locally ill-conditioned, allowing distinct parameter configurations to remain observationally indistinguishable. We define parameter compensation as coordinated variation of distinct physical parameters within an observational equivalence class or, for approximate observations, within the corresponding admissible set. This differs from parameter coupling. Coupling is prescribed by the governing equations, whereas compensation describes indistinguishability created by the inverse problem.

The parameter compensation principle is formalized through three properties whose mathematical conditions are given in Methods. Existence describes when a locally unidentifiable or near-null direction involves both a compensation target and at least one distinct carrier parameter. Persistence describes whether the response change induced by removing the target mechanism can be reproduced locally by variation of the designated carrier parameter in the reduced representation. BcNN does not evaluate this local first-order response condition directly. Instead, the reduced anchor provides a finite computational probe by testing whether variation of the same carrier can restore observational admissibility after representation reduction. Recoverability describes when an admissible parameter representative can be locally isolated from the observational equivalence class.

To represent compensation explicitly, the remained-bias notation records the target parameter r , carrier parameter i , carrier representation order j and target representation order c . Coordinates with $i=r$ are parameter-diagonal and describe self-retention. Coordinates with $i \neq r$ are off-diagonal and describe cross-parameter compensation, in which a carrier parameter reproduces an observable contribution associated with a distinct target parameter. The complete notation and normalization are defined in Methods.

The present BcNN realization considers the simplest nontrivial setting with two unknown physical parameters and two representation orders. Parameter 2 in the enriched representation is selected as the compensation target, and parameter 1 is retained as the carrier. The reduced and enriched anchors are optimized independently against the same sparse observations, with one trainable physical parameter in each anchor. The enriched anchor exposes an observationally admissible compensatory pseudo-solution. The reduced anchor probes whether the same relation between the target and carrier remains observationally admissible after the target mechanism is removed. The frozen off-diagonal remained-bias estimates from the reduced and enriched anchors are denoted by $\hat{b}_L$ and $\hat{b}_H$. Their ordered difference defines the physical gauge,

$$\hat{g} = \hat{b}_H - \hat{b}_L.$$

The gauge quantifies the representation-dependent change in the compensatory carrier coordinate. It is not itself a recovered physical parameter and does not establish structural identifiability. Once the two anchor estimates have been frozen, the dual-parameter target network optimizes complementary auxiliary coordinates and reconstructs the fully diagonal remained-bias coordinates that describe parameter self-retention. The corresponding transformation can change parameter scale, orientation and optimization conditioning while preserving the rank of the local parameter sensitivity Jacobian, as shown in Methods.

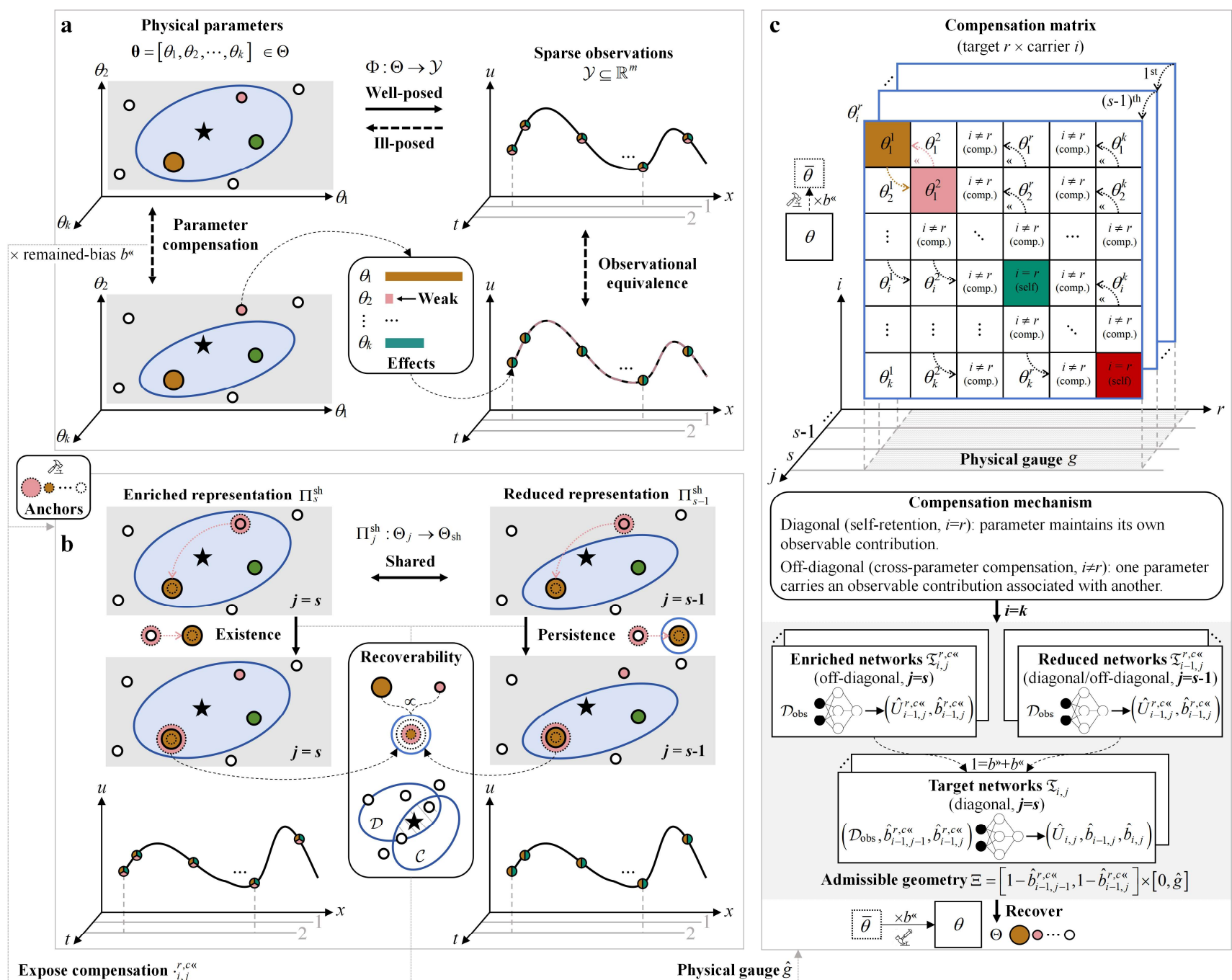


Fig. 1 | Parameter compensation principle for multi-parameter inverse learning. **a**, Observational equivalence and parameter compensation. Even when the forward problem is well-posed, distinct physical parameter vectors can generate identical or nearly identical responses under sparse observations, giving rise to observationally admissible pseudo-solutions rather than a uniquely determined physical parameter vector. Coordinated changes in distinct parameters

can preserve the observable response, allowing a weak physical effect to be masked by compensatory changes in another parameter. **b**, Parameter compensation across hierarchical physical representations. The enriched representation contains the additional physical mechanism, whereas the reduced representation removes this mechanism while retaining the shared physical coordinates. The enriched anchor exposes an observationally admissible compensatory pseudo-solution, providing computational evidence consistent with Existence. The reduced anchor probes whether the same relation between the target and carrier remains observationally admissible after representation reduction, as motivated by Persistence. Recoverability requires an admissible parameter representative to be locally isolated within the relevant observational equivalence class and admissible set. **c**, General compensation representation and its dual-parameter realization in BcNN. For a fixed target representation order c , remained-bias coordinates are organized by target parameter r , carrier parameter i and carrier representation order j . Parameter-diagonal coordinates ($i=r$) describe self-retention, whereas off-diagonal coordinates ($i \neq r$) describe cross-parameter compensation. Off-diagonal anchor coordinates at adjacent representation orders expose the representation-dependent compensatory change, whose ordered difference defines the physical gauge. In the present dual-parameter realization, the frozen anchor estimates define the target search geometry and the gauge scales complementary bias reparameterization used to reconstruct the parameter-diagonal coordinates.

BcNN therefore does not remove the dual-parameter inverse problem or add information that is absent from the observations. It replaces an unstructured joint search with two independently optimized anchor problems followed by a dual-parameter target problem conditioned on the frozen anchor estimates. Fig. 1 summarizes the parameter compensation principle and Fig. 2 shows its computational realization in BcNN.

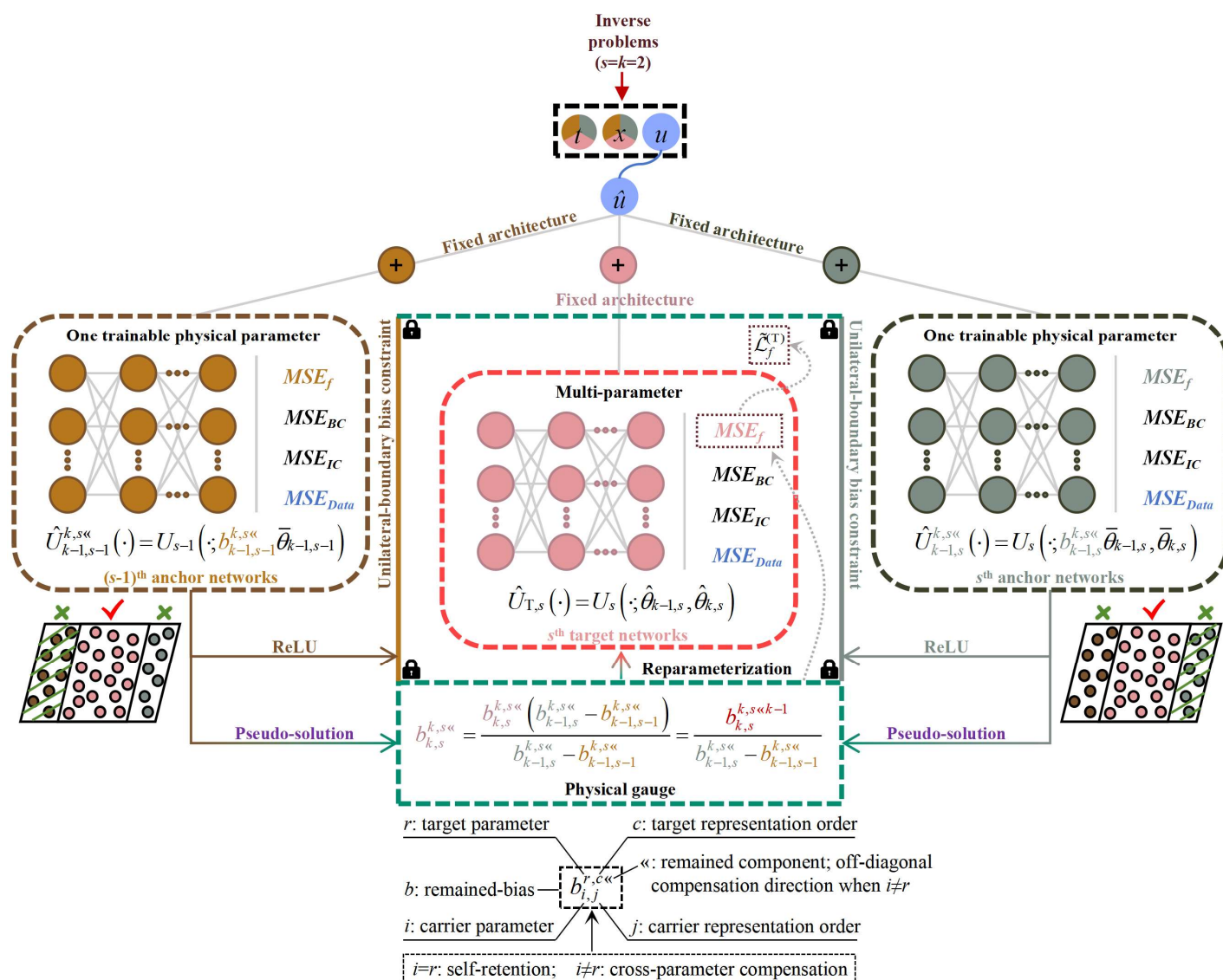


Fig. 2 | Architecture and compensation notation of the dual-parameter BcNN. The reduced and enriched anchor networks are trained separately using the same sparse observations. In both anchors, parameter 2 at representation order 2 is the compensation target and parameter 1 is the carrier. The two anchors estimate off-diagonal remained-bias coordinates in adjacent physical representations, and their ordered difference defines the physical gauge. After the anchor estimates are frozen, the target network optimizes two auxiliary coordinates derived from the evolved-bias forms and reconstructs the fully diagonal remained-bias coordinates through complementary bias reparameterization. The frozen anchor estimates also define the admissible intervals used in target optimization. The soft unilateral-boundary constraint is inactive inside these intervals and modifies the governing equation residual after a boundary violation.

Benchmark design and evaluation

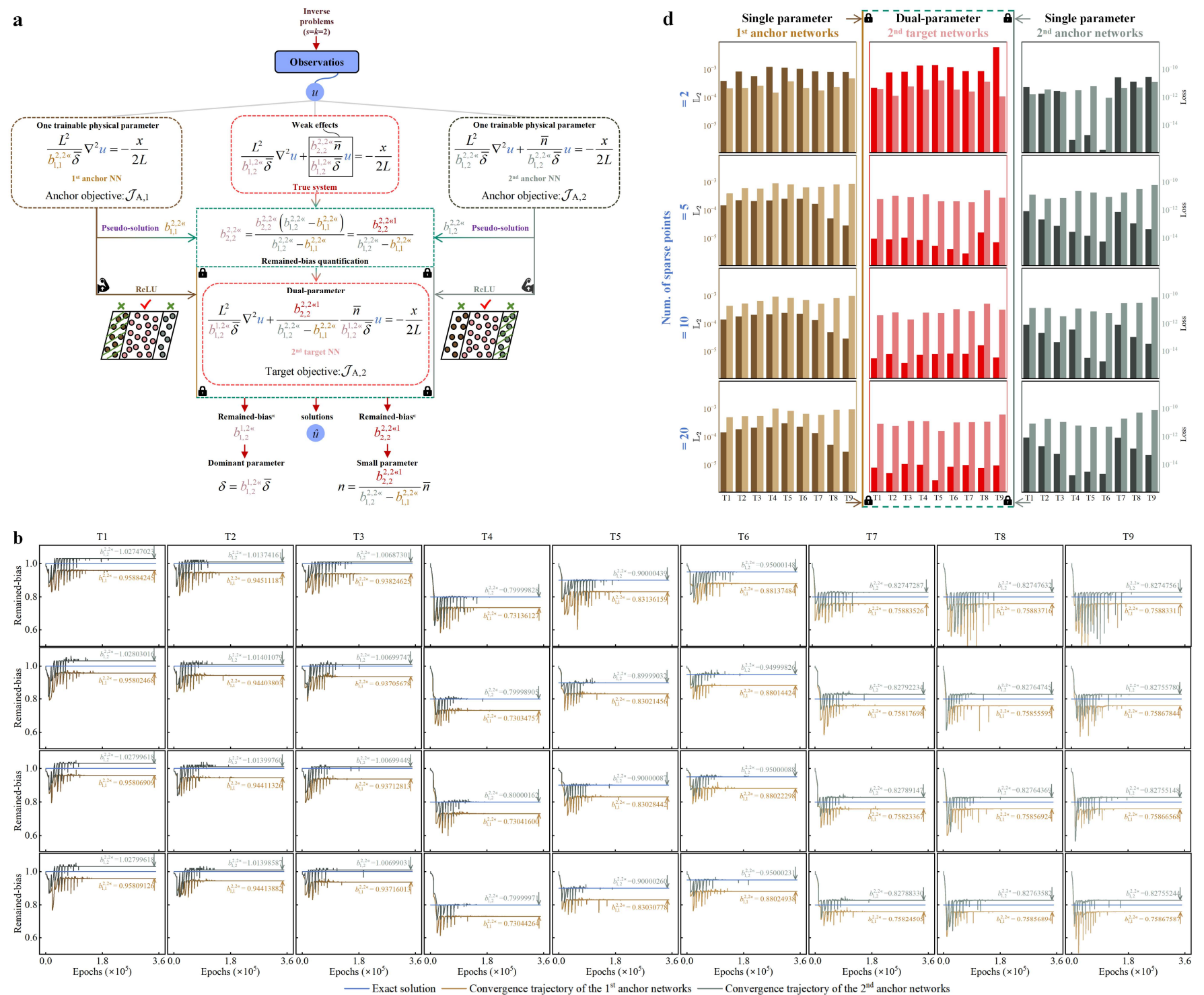
We evaluated BcNN in two complementary inverse settings. The controlled Sturm-Liouville benchmark used sparse samples from exact solutions and separately examined observation sparsity, component ablation, response scale and architecture sensitivity. The structural benchmark used published composite beam measurements for which the source study reported sensor-related experimental uncertainty but did not provide quantitative bounds. Each inverse problem was trained independently. Throughout the benchmarks, dominant and small parameters are distinguished by their relative influence on the observable response under the prescribed physical condition rather than by their numerical magnitude. Evaluation therefore separates response reconstruction from physical parameter recovery. We report response error, parameter error and parameter trajectories, together with convergence behaviour where applicable. Full benchmark construction, configuration counts, data provenance and training controls are provided in Methods and Extended Data Table 1.

BcNN reveals weak physical effects from sparse samples of exact solutions

We first tested BcNN in a controlled setting designed to separate the intrinsic difficulty of multi-parameter inverse learning from experimental noise and measurement uncertainty. The observations were sparse samples of exact solutions, so failures of parameter recovery could not be attributed to an imposed noise model or experimental uncertainty. The benchmark contained a dominant parameter with a comparatively strong influence on the observable response and a small parameter whose contribution was weaker and therefore more easily absorbed by compensatory changes in the dominant parameter. Nine physical conditions were considered by varying the dominant parameter, the small parameter and the loading level, and each response was sampled using 2, 5, 10 and 20 observation points. The exact parameter values were retained only for post hoc evaluation and were not supplied to the inverse optimization. This design therefore provides a controlled test of whether accurate response reconstruction can be separated from genuine recovery of a weak physical effect.

Fig. 3a shows how BcNN reorganizes this dual-parameter inverse problem. The reduced and enriched anchors are trained independently against the same sparse observations, with one trainable physical parameter in each. Both use the same compensation target and carrier, while the target mechanism is removed from the reduced representation and fixed at its prescribed normalization reference in the enriched representation. The resulting anchor solutions are not expected to reproduce the underlying dual-parameter state directly. Instead, they expose observationally admissible compensatory states in which the carrier absorbs different amounts of the target contribution. Their ordered difference therefore measures how the compensatory carrier state changes when the physical representation changes, and this quantity defines the physical gauge used by the target network. The anchor trajectories in Fig. 3b show that this representation-dependent compensation pattern is consistently exposed across the nine physical conditions and four observation densities, although the sparsest observations produce larger fluctuations. The target trajectories in Fig. 3c reveal the more difficult part of the problem. The dominant parameter stabilizes comparatively easily, whereas the small parameter exhibits larger transients and slower convergence because its observable contribution is weaker and more readily exchanged with the dominant parameter. Increasing the number of observations generally improves both response reconstruction and parameter recovery (Fig. 3d), but the difference between the two parameters persists. This is precisely the regime in which response fitting alone can be misleading. A network may reproduce the sparse observations while leaving the weak physical contribution embedded in a compensatory parameter configuration.

The controlled benchmark therefore tests more than numerical accuracy. The enriched anchor exposes an observationally admissible compensatory pseudo-solution, the reduced anchor tests whether the same target-carrier relation remains admissible after the target mechanism is removed, and the target stage asks whether the weak physical contribution can be recovered within the region defined by the two anchor states. The resulting behaviour is consistent with the computational roles associated with Existence, Persistence and Recoverability. Because these effects are observed with exact data and known reference parameters, they provide a controlled basis for the subsequent structural experiments, where the same distinction between response fitting and physical recovery must be examined in the presence of experimental uncertainty.



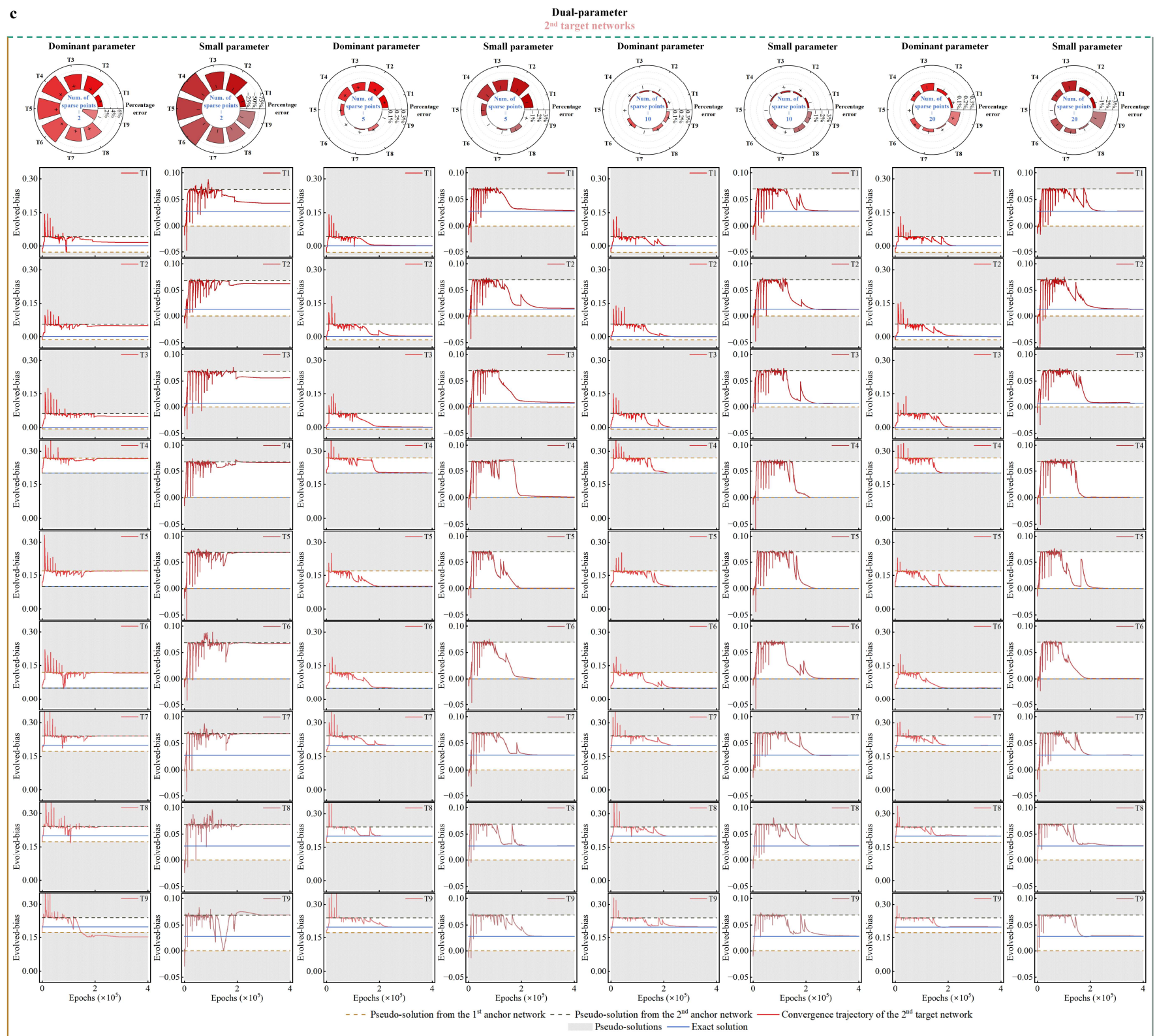


Fig. 3 | BcNN reveals weak physical effects from sparse samples of exact solutions. **a**, Bias-compensated reformulation of the dual-parameter Sturm-Liouville inverse problem. Two anchor networks, each with one trainable physical parameter, estimate the off-diagonal remained-bias coordinates in the reduced and enriched physical representations. Their signed difference defines the physical gauge that conditions subsequent complementary bias reparameterization and dual-parameter target recovery. **b**, Convergence trajectories of the remained-bias coordinates inferred by the two anchor networks for nine benchmark conditions and four observation counts. **c**, Trajectories of the dominant and small parameters in the dual-parameter target network. Horizontal reference lines indicate the exact parameter values, and pseudo-solutions inferred from the two anchor networks are shown for comparison. **d**, Summary response errors, parameter errors and optimized losses for the two anchor networks and the target network using 2, 5, 10 and 20 observations.

Response accuracy can conceal incorrect physical parameters

We next evaluated BcNN using sparse measurements from independent composite beam experiments. Each experiment provided five observation points. The source study reports sensor-related experimental uncertainty but does not provide quantitative uncertainty bounds for individual cases. We therefore used the published measurements without assigning an artificial noise model and treated the experimental uncertainty as unquantified.

Across the structural experiments, BcNN recovered the dominant parameter with errors ranging from -1.77% to 2.87% and the small parameter with errors ranging from -10.22% to 8.83%. Baseline PINN was markedly less reliable for the small parameter, with errors exceeding 20% in magnitude in five of the nine experiments. Response reconstruction showed a different pattern. Notably, BcNN reduced the absolute error of the small parameter in every structural experiment. Whenever baseline PINN achieved a lower response error than BcNN, its estimate of the small parameter was nevertheless less accurate. A better response fit therefore frequently coincided with a less accurate physical parameter estimate. Judged by response error or optimized loss alone, these runs would often appear successful despite their inaccurate physical parameters. The complete values for individual experiments are provided in Extended Data Table 2.

Fig. 4 shows that this discrepancy is also visible in the optimization dynamics. The BcNN target stage reaches stable estimates of both physical parameters after the anchor networks have been optimized and frozen, whereas baseline PINN continues to exhibit parameter drift over the common comparison horizon. A low response error can therefore correspond to an observationally admissible compensatory state rather than to recovery of the underlying physical parameters. The distinction is most evident in the trajectories of the small parameter, where baseline PINN can remain near an apparently plausible endpoint while continuing to drift or oscillate, whereas the BcNN trajectories settle into stable plateaus after anchor construction.

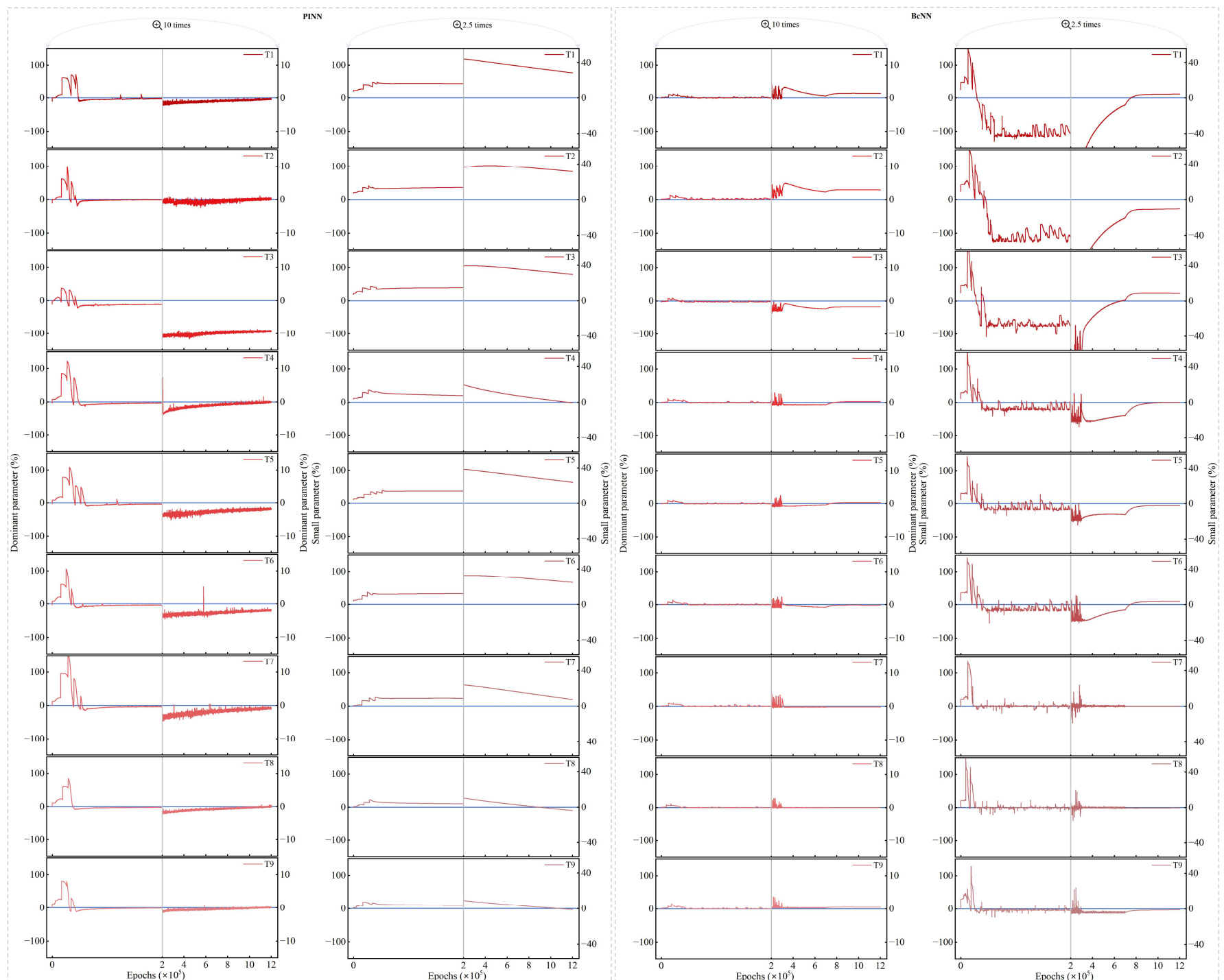


Fig. 4 | BcNN recovers weak physical effects from sparse structural measurements. Parameter trajectories of baseline PINN and the BcNN target stage are shown across the independent structural experiments. Each experiment uses five sparse measurements subject to unquantified sensor-related experimental uncertainty. Baseline PINN optimizes both physical parameters directly. BcNN first optimizes and freezes the two anchor estimates before training the dual-parameter target network. The horizontal axis for BcNN therefore reports target-stage epochs after anchor construction. Horizontal reference lines denote the reported reference parameters. Baseline PINN continues to exhibit parameter drift or oscillation within the common comparison horizon, whereas the BcNN target stage approaches stable and more accurate parameter estimates across the structural experiments.

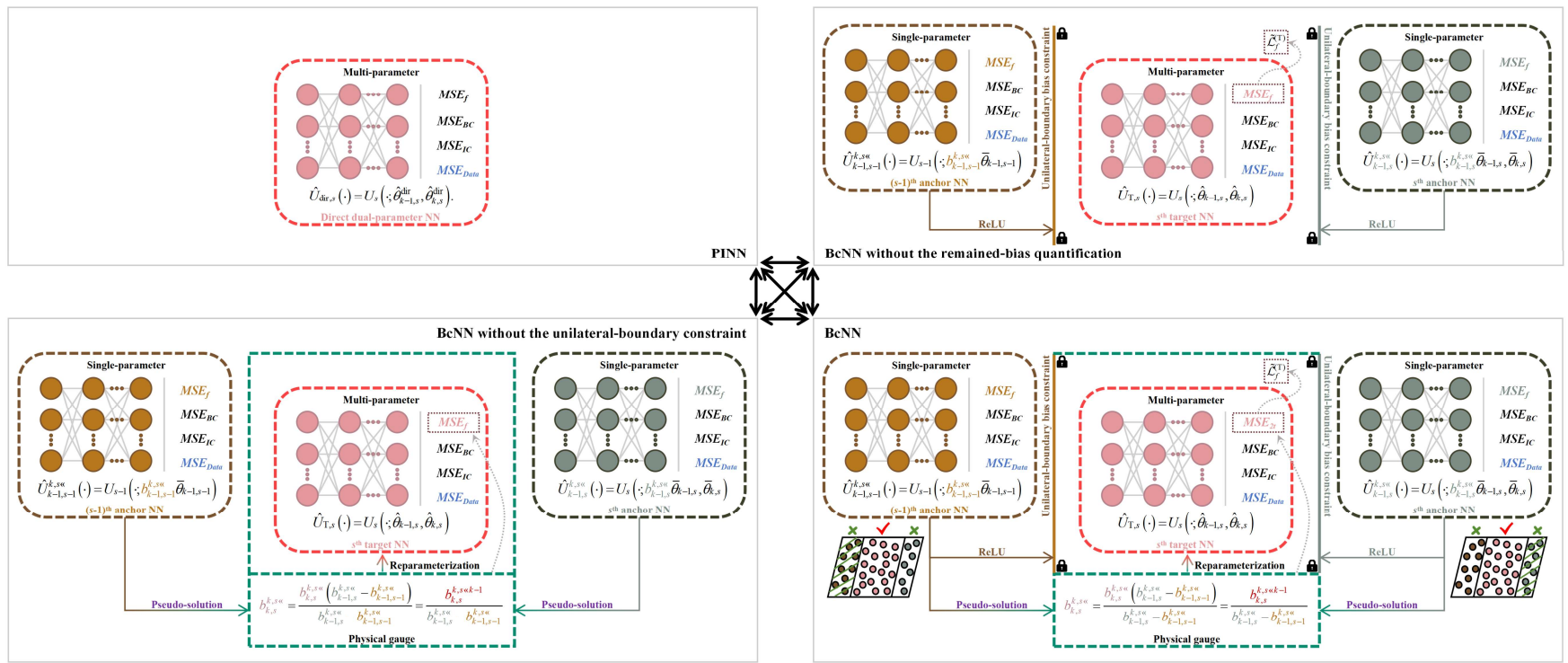
Remained-bias quantification and the soft unilateral-boundary constraint jointly support discovery of weak effects

We next examined the model comparison and ablation branch of the benchmark based on exact solutions (Extended Data Table 1). Four variants were compared. These were baseline PINN, BcNN without remained-bias quantification, BcNN without the soft unilateral-boundary constraint and full BcNN. Their computational structures are shown in Fig. 5a. Architecture sensitivity was evaluated separately and is discussed with Fig. 6.

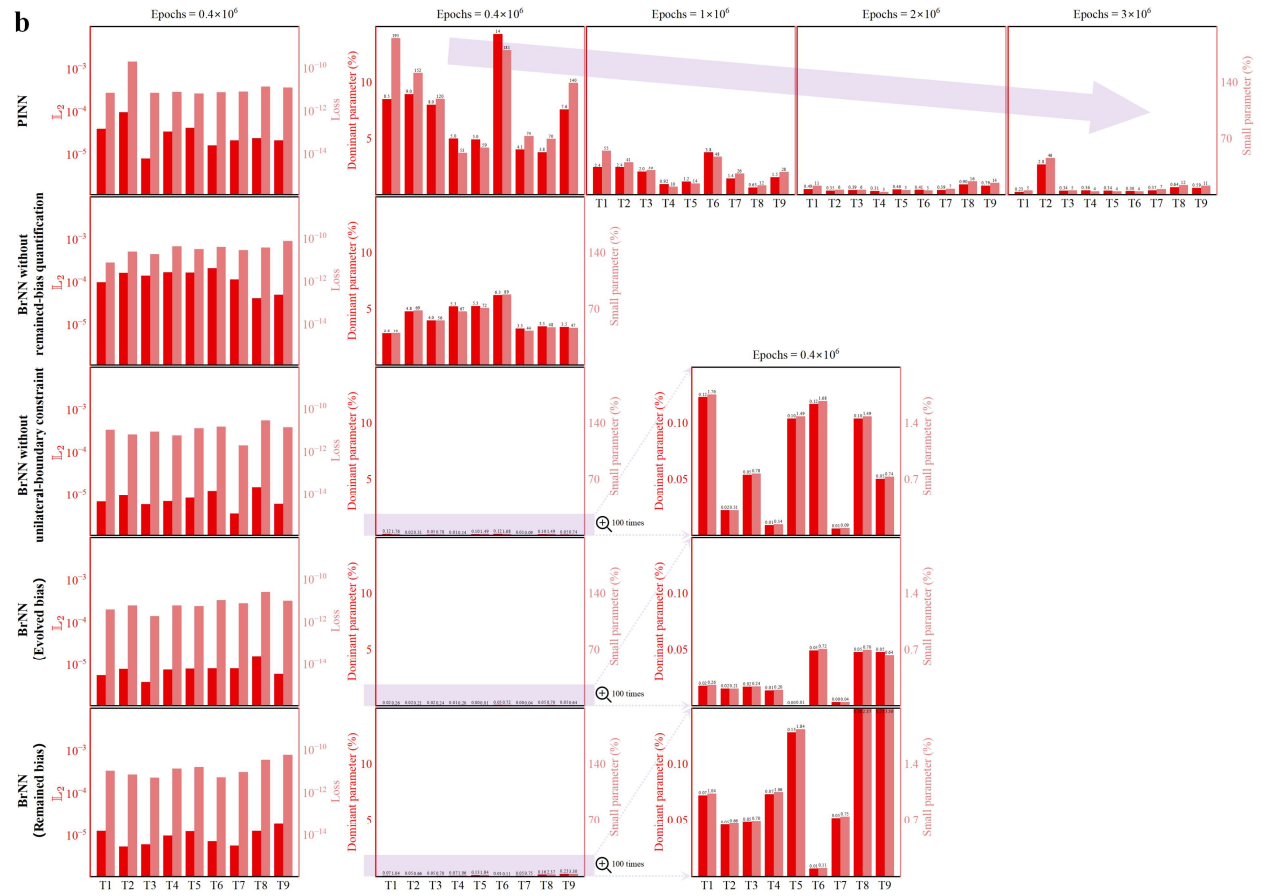
Baseline PINN optimizes both physical parameters directly and contains neither an explicit compensation representation nor an anchor-defined search region. Its trajectories show prolonged transients, oscillations and, in several cases, continued drift after accurate response reconstruction. Removing remained-bias quantification retains the anchors but prevents the representation-dependent change between their compensatory states from being transferred explicitly to the target parameterization. Removing the soft unilateral-boundary constraint retains the compensation representation but allows the auxiliary coordinates to leave the anchor-defined intervals, producing larger overshoot and oscillations. Full BcNN combines both components. The signed difference between the off-diagonal remained-bias coordinates inferred by the adjacent anchors defines the physical gauge, while the ReLU channels modify the governing equation residual only after a boundary violation. In the reported experiments, this combination stabilizes reconstruction of the diagonal remained-bias coordinates. Fig. 5b,c shows that response error and optimized loss alone do not reliably rank the variants: response fitting can stabilize while the physical parameters remain inaccurate or unstable. The difference is not simply attributable to a longer optimization horizon. BcNN attains more accurate physical parameter estimates within the target stage, whereas baseline PINN remains less accurate even after substantially longer training. The ablations expose two qualitatively different failure modes. Removing remained-bias quantification leaves parameter recovery in a compensatory state, whereas removing the unilateral-boundary mechanism primarily destabilizes the optimization trajectory.

Recovering a weak physical effect also requires greater parameter accuracy than may be needed for forward reconstruction. A response error that is adequate for predicting the observable field can still be too large to distinguish parameter combinations within a narrow equivalence class induced by compensation. The accuracy required for inverse recovery must therefore be judged against parameter identifiability, not response reconstruction alone.

a



b



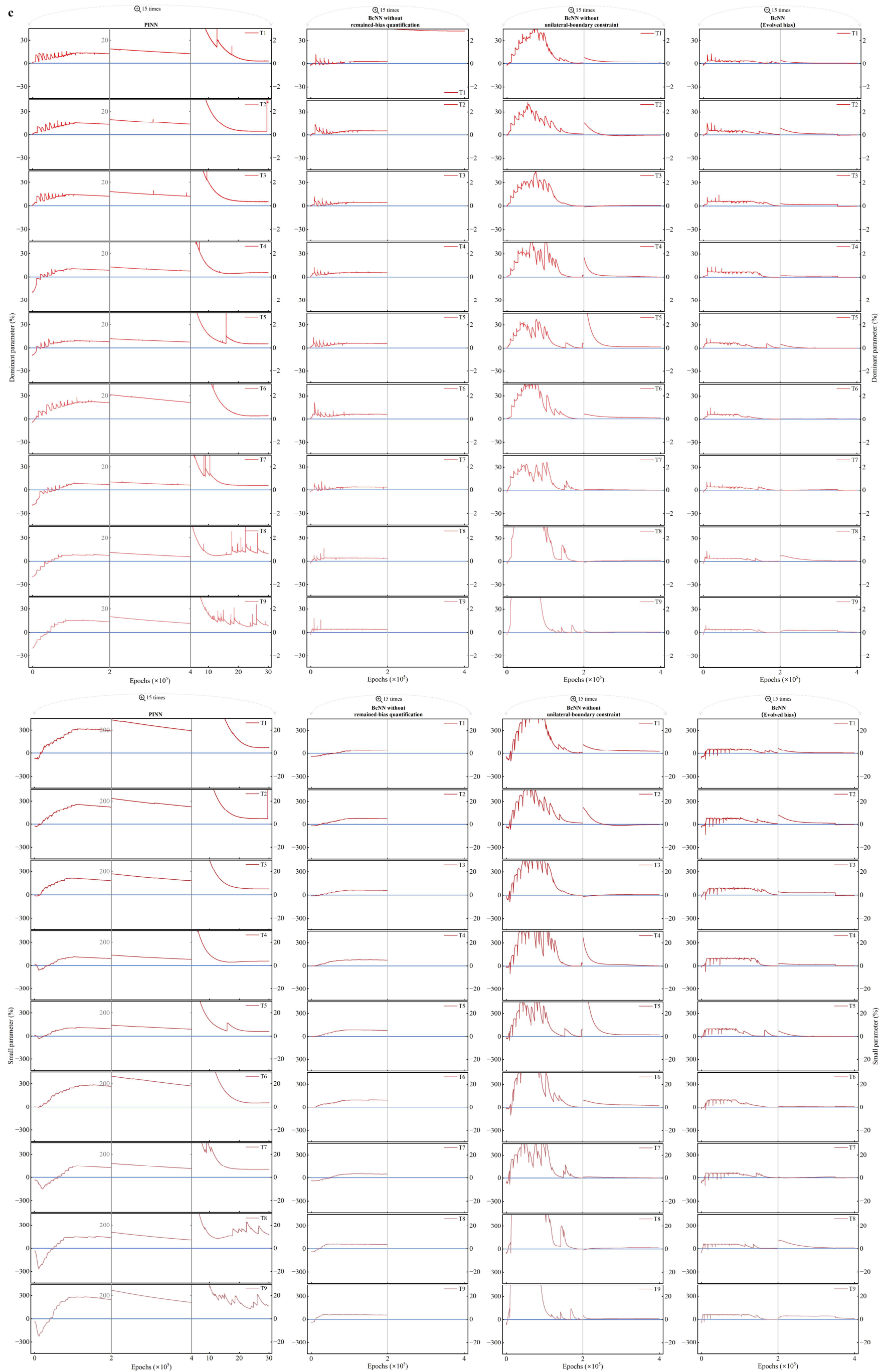


Fig. 5 | Remained-bias quantification and the soft unilateral-boundary constraint jointly support reliable discovery of weak effects. a, Computational

structures of baseline PINN, BcNN without remained-bias quantification, BcNN without the unilateral-boundary constraint, and full BcNN. **b**, Comparison of relative L_2 response error, comparable loss, error in the dominant parameter and error in the small parameter across the nine theoretical benchmark conditions. Baseline PINN is additionally shown at extended training horizons up to 3×10^6 epochs to assess whether continued optimization resolves the parameter instability. Parameter estimates reconstructed from the complementary remained-bias and evolved-bias forms are shown separately as a check of implementation consistency for full BcNN. **c**, Trajectories of the dominant and small parameters for the four model variants. Removing remained-bias quantification prevents the representation-dependent compensation change from being transferred explicitly to the target parameterization, whereas removing the soft unilateral-boundary constraint produces oscillatory or unstable parameter evolution. Full BcNN provides the most consistent parameter recovery.

Weak observability and architecture sensitivity define the operating limits of BcNN

BcNN remains limited by the information contained in the observations. As the normalized loading amplitude decreases, errors in both physical parameters increase, with the largest deterioration occurring in the small parameter at the lowest response levels (Fig. 6a). Response error and optimized loss can nevertheless remain small. The deterioration at low loading amplitudes is therefore interpreted primarily as a practical observability limitation within the tested response range rather than as direct evidence of floating-point failure. Neural representation also affects parameter recovery more strongly than response fitting. Several activation functions and combinations of network width and depth produce similar response errors while yielding substantially different physical parameter estimates (Fig. 6b,c). Sine and tanh give the smallest parameter errors among the tested activations in the present Sturm-Liouville benchmark, but this comparison does not establish a universal architecture ranking. Increasing network size likewise does not produce monotonic improvement in parameter recovery.

These tests define the present operating regime of BcNN. Reliable recovery requires observations that resolve the weak physical contribution and a neural representation that preserves this information sufficiently for stable inverse optimization.

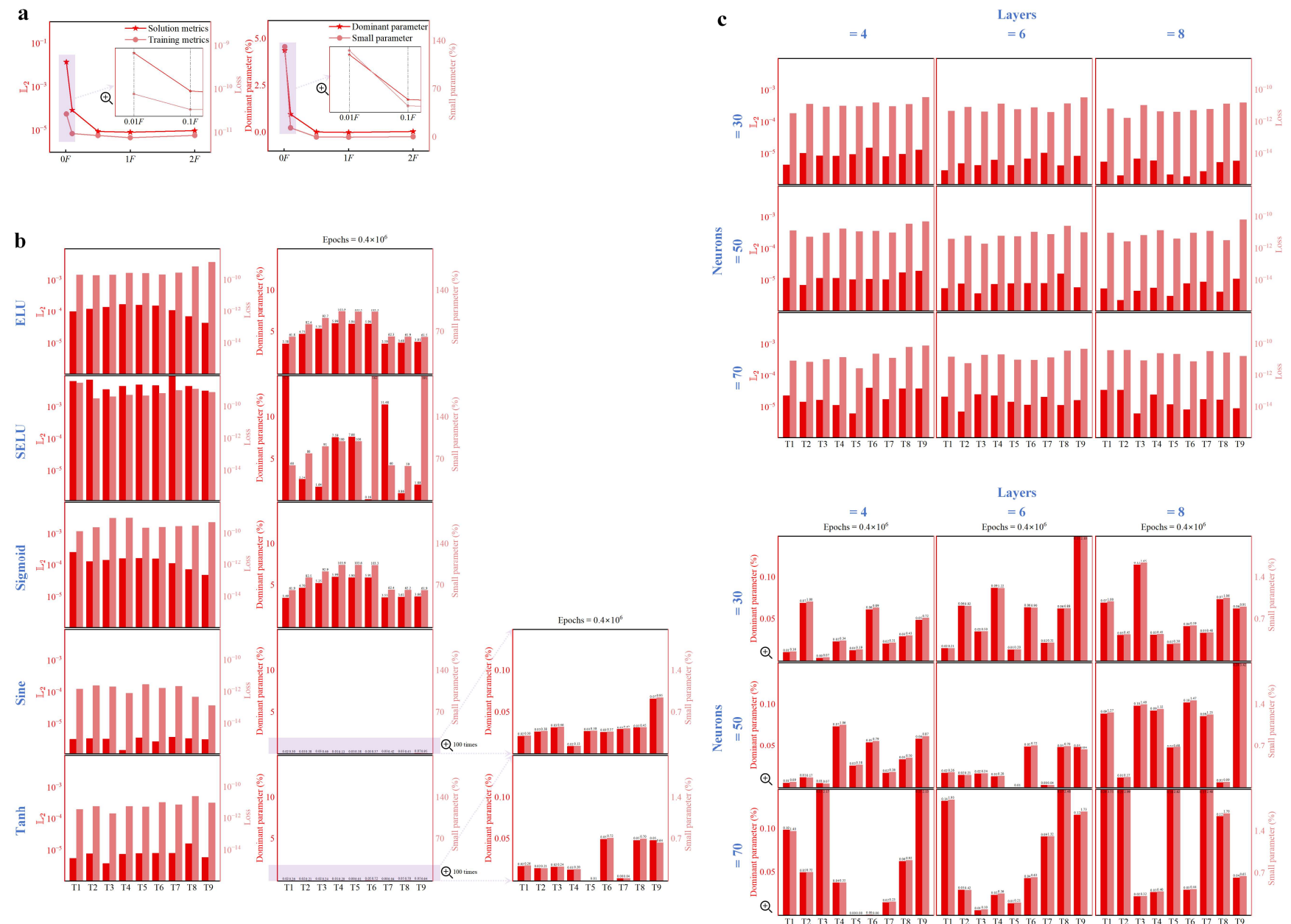


Fig. 6 | Weak observability and architecture sensitivity define the operating limits of BcNN. **a**, Stress test of response scale. Relative L2 response error, comparable loss, errors in the dominant parameter and the small parameter are evaluated at normalized loading amplitudes of $0.01F$, $0.1F$, $0.5F$, $1F$ and $2F$. The magnified view shows parameter recovery at the two lowest loading amplitudes. **b**, Stress test of the activation function across the nine theoretical benchmark conditions. ELU, SELU, sigmoid, sine and tanh are evaluated using identical network widths, depths and training settings. The magnified views show the smaller parameter errors obtained with sine and tanh. **c**, Stress test of network width and depth using 30, 50 and 70 neurons per hidden layer and 4, 6 and 8 hidden layers. Response metrics remain low for many configurations, whereas parameter errors vary substantially, particularly for the small parameter. All parameter errors are evaluated at the common training epoch indicated above each panel.

Discussion

One of the main difficulties exposed by these results is that the usual signs of a successful inverse model can be misleading. A low response error and a small optimized loss do not necessarily mean that the underlying physical parameters have been recovered. In our experiments, response reconstruction could remain accurate while the weaker physical parameter was inaccurate or continued to drift. When the parameters themselves are the quantities of scientific interest, this is not simply an optimization imperfection. The model has reproduced what was observed but has assigned that response to the wrong physical explanation.

Parameter compensation provides a way to understand this failure. It is related to structural and practical non-identifiability [5,6,7,8], equifinality [9,10], sloppy sensitivity directions [11,12,13,14] and compensatory parameter trade-offs [15,16]. Similar ambiguity has also been studied in neural inverse problems [17,18,19,20]. The contribution of the present work is not the observation that compensation can occur. It is the explicit description of compensation as a directed relation that also depends on the physical representation. The framework identifies the parameter whose contribution is masked, the carrier that absorbs that contribution, and the representation in which the relation appears.

This view also changes the role of pseudo-solutions. An observationally admissible pseudo-solution is usually treated as something that should be rejected because its parameters are not the underlying physical state. Here it is used differently. The off-diagonal remained-bias coordinates of the two anchors describe compensatory carrier states in adjacent physical representations. Their ordered difference records how that state changes when the physical mechanism is added or removed, and this difference defines the physical gauge used by the target parameterization. The ablation results support the separate roles of remained-bias quantification and the soft unilateral-boundary constraint. The first transfers the representation-dependent compensation change into

the target problem, while the second discourages the auxiliary coordinates from leaving the region defined by the anchors. Neither operation creates new observational information. Unique recovery still depends on sufficient observability and local identifiability.

The structural experiments make the practical importance of this distinction particularly clear. With sparse measurements and unquantified sensor-related uncertainty, BcNN recovered the weak parameter more accurately than baseline PINN in every experiment. The baseline error in the small parameter exceeded 20% in magnitude in five of the nine cases. Yet these failures were not reliably revealed by response fitting. Baseline PINN could produce a comparable or even smaller response error while returning a less accurate weak parameter. Similar inverse identification problems with sparse, noisy or experimental measurements have been considered in elasticity and engineering systems [30-33,36], together with regularization for inverse PDE problems [34] and simultaneous identification of multiple constitutive parameters [35]. The present results highlight a different concern. Agreement with measured responses alone can hide a compensatory parameter state. Low response error or optimized loss should therefore not be treated as sufficient evidence that the inferred physics is correct, which is consistent with broader concerns about incomplete optimization objectives in physics-informed learning [37].

This has consequences for how inverse learning should be evaluated. Response accuracy, parameter accuracy and trajectory stability are not interchangeable indicators. A response error that is acceptable for forward prediction may still be too large to distinguish nearby parameter combinations along a compensatory direction. The architecture tests reinforce the same point. Different neural representations can fit the observable response similarly and still converge to noticeably different parameter states. Increasing network size does not remove this ambiguity automatically. These results suggest that validation of scientific inverse models should examine the recovered physical quantities directly rather than relying mainly on how well the observable field is fitted.

The remained-bias representation may also help separate parameter compensation from other sources of discrepancy. In the present work, remained-bias coordinates describe how observable contributions are redistributed among physical parameters. They are not estimates of sensor bias or statistical bias. With additional independent information and an explicit observation model, future formulations could represent compensation-induced ambiguity separately from random observational uncertainty, systematic measurement bias, numerical approximation error and model discrepancy. The present experiments do not attempt this decomposition because quantitative information about the individual discrepancy sources is not available.

The current realization has clear limits. It considers two unknown physical parameters and two physical representation orders. Extending the construction to higher-dimensional parameter spaces will require compatible sets of anchors and admissible regions that remain physically meaningful. The structural experiments also do not establish how recovery changes with a prescribed noise amplitude. Recovery deteriorates when the observable response becomes too weak to distinguish the parameter contributions, and broader sensitivity to learning rate and initialization remains to be examined. The normalization references are coordinate scales fixed before optimization rather than benchmark reference values, and their choice remains application dependent. More fundamentally, BcNN cannot recover information that is absent from the observations. Its role is narrower. It helps prevent weak but still observable physical information from being absorbed into an unstructured compensatory state.

The broader implication is that reproducing observations is not the same as recovering the physics that produced them. As physics-informed models are increasingly used to infer latent mechanisms and parameters, this distinction becomes increasingly important. Pseudo-solutions that would normally be discarded can sometimes reveal how one parameter masks another when they are examined across related physical representations. BcNN provides one interpretable way to use that information while keeping the limits imposed by identifiability and observability explicit. The results also suggest that neural architecture can influence which admissible parameter state is selected, even when different networks reproduce the observations almost equally well [23,26].

Methods

Problem formulation, observational equivalence and admissibility

Let $\Omega \subset \mathbb{R}^d$ denote the spatial domain and let $\Omega_t \subset \mathbb{R}$ denote the temporal domain when time dependence is present. At representation order $j \in \{1, \dots, s\}$, the physical field U_j satisfies

$$\mathcal{G}_j[U_j; \boldsymbol{\theta}_j](\mathbf{x}, t) = 0, \quad (\mathbf{x}, t) \in \Omega \times \Omega_t, \quad (1)$$

together with the boundary and initial operators

$$\mathcal{B}_j[U_j; \boldsymbol{\theta}_j] = 0, \quad \mathcal{I}_j[U_j; \boldsymbol{\theta}_j] = 0. \quad (2)$$

The active physical parameter vector at representation order j is

$$\boldsymbol{\theta}_j = (\theta_{i,j})_{i \in \mathcal{K}_j} \in \Theta_j, \quad \mathcal{K}_j \subseteq \{1, \dots, k\}, \quad k_j = |\mathcal{K}_j|. \quad (3)$$

Here, $\mathcal{K}_j$ is the index set of physical mechanisms active at representation order j , and k_j is the number of active parameters in that representation. We write $k := k_s$ for the number of unknown physical parameters in the inverse problem at the highest representation order. Parameter identities are preserved across the hierarchy even when a mechanism is absent from a reduced representation. The quantities k and s are independent: k counts unknown physical parameters, whereas s denotes the highest representation order.

It is assumed that, for every parameter vector in the neighbourhood considered, the corresponding forward problem admits a locally unique solution field $U_j(\boldsymbol{\theta}_j)$. Let $\mathcal{H}$ denote the observation operator, and let the observable response belong to a finite-dimensional space $\mathcal{Y} \subseteq \mathbb{R}^m$. The response map at representation order j is

$$\Phi_j : \Theta_j \rightarrow \mathcal{Y}, \quad \Phi_j(\boldsymbol{\theta}_j) = \mathcal{H}U_j(\boldsymbol{\theta}_j). \quad (4)$$

The observations are represented as

$$\mathbf{y}_{\text{obs}} = \Phi_s(\boldsymbol{\theta}_s^*) + \boldsymbol{\epsilon}, \quad \boldsymbol{\epsilon} \in \mathbb{R}^m \quad (5)$$

where $\boldsymbol{\theta}_s^*$ denotes the underlying physical parameter vector associated with the observations, and $\boldsymbol{\epsilon}$ collects observational uncertainty, numerical approximation error and model discrepancy. Two parameter vectors are exactly observationally equivalent at representation order j when they generate identical observable responses:

$$\boldsymbol{\theta}_j \sim_j \tilde{\boldsymbol{\theta}}_j \Leftrightarrow \Phi_j(\boldsymbol{\theta}_j) = \Phi_j(\tilde{\boldsymbol{\theta}}_j). \quad (6)$$

The corresponding exact observational equivalence class is

$$\mathcal{E}_j(\boldsymbol{\theta}_j) = \left\{ \tilde{\boldsymbol{\theta}}_j \in \Theta_j : \Phi_j(\tilde{\boldsymbol{\theta}}_j) = \Phi_j(\boldsymbol{\theta}_j) \right\}. \quad (7)$$

For noisy or numerically approximate observations, exact equivalence is replaced by an observationally admissible set centred on the measured data:

$$\mathcal{A}_{\tau,j}^{\text{obs}} = \left\{ \tilde{\boldsymbol{\theta}}_j \in \Theta_j : \left\| \Phi_j(\tilde{\boldsymbol{\theta}}_j) - \mathbf{y}_{\text{obs}} \right\|_{\mathbf{W}_{\text{obs}}} \leq \tau \right\}. \quad (8)$$

Here, $\tau \geq 0$ is the prescribed observational tolerance. The weighted observation norm is

$$\left\| \mathbf{z} \right\|_{\mathbf{W}_{\text{obs}}} = \left(\mathbf{z}^T \mathbf{W}_{\text{obs}} \mathbf{z} \right)^{1/2}, \quad \mathbf{W}_{\text{obs}} \in \mathbb{R}^{m \times m}, \quad \mathbf{W}_{\text{obs}} \succ 0. \quad (9)$$

Parameter compensation refers to coordinated variations of multiple physical parameters that remain within an exact observational equivalence class or, under noisy and approximate observations, within the corresponding observationally admissible set. It is distinct from parameter coupling. Parameter coupling is prescribed by the governing equations, whereas parameter compensation describes observational indistinguishability among coordinated parameter changes.

Local sufficient conditions for parameter compensation

Let

$$\mathbf{J}_j(\boldsymbol{\theta}_j) = D\Phi_j(\boldsymbol{\theta}_j) = \frac{\partial \Phi_j}{\partial \boldsymbol{\theta}_j}(\boldsymbol{\theta}_j) \quad (10)$$

denote the parameter sensitivity Jacobian, where D denotes the Fréchet derivative. In the finite-dimensional observation space considered here, it is represented by the corresponding Jacobian matrix.

Property 1 (Existence). Suppose that $\boldsymbol{\theta}_s^*$ is an interior point of Θ_s , that Φ_s is continuously differentiable, and that $\mathbf{J}_s$ has locally constant rank $\varrho < k$ near $\boldsymbol{\theta}_s^*$:

$$\text{rank } \mathbf{J}_s(\boldsymbol{\theta}_s^*) = \varrho < k. \quad (11)$$

Under the constant-rank condition, the exact observational level set is locally a manifold of dimension

$$\dim \mathcal{E}_s(\boldsymbol{\theta}_s^*) = k - \varrho. \quad (12)$$

Its infinitesimal compensation directions belong to

$$\ker \mathbf{J}_s(\boldsymbol{\theta}_s^*). \quad (13)$$

Rank deficiency establishes the existence of locally unidentifiable directions, but it does not show that a particular target parameter is compensated by a distinct carrier parameter. For parameter r to act as the compensation target, the null direction must also satisfy

$$\exists \mathbf{v} \in \ker \mathbf{J}_s(\boldsymbol{\theta}_s^*), \quad \text{such that } v_r \neq 0, \quad \sum_{i \neq r} |v_i| > 0. \quad (14)$$

The first condition ensures that the target parameter varies along the locally unidentifiable direction. The second ensures that at least one distinct carrier parameter also varies and can compensate for the target contribution. The exact rank-deficient case describes local non-identifiability of the chosen response

map. Sparse observations and finite numerical resolution can also produce practical non-identifiability when the sensitivity Jacobian has full rank but is ill-conditioned. For a prescribed near-null sensitivity threshold $\tau_J \geq 0$, define the approximate null-direction set

$$\mathcal{Z}_{\tau_J, j}(\boldsymbol{\theta}_j) = \left\{ \mathbf{v} \in \mathbb{R}^{k_j} : \|\mathbf{v}\|_2 = 1, \left\| \mathbf{W}_{\text{obs}}^{1/2} \mathbf{J}_j(\boldsymbol{\theta}_j) \mathbf{v} \right\|_2 \leq \tau_J \right\}. \quad (15)$$

For a sufficiently small scalar a and a direction $\mathbf{v} \in \mathcal{Z}_{\tau_J, j}$,

$$\left\| \Phi_j(\boldsymbol{\theta}_j + a\mathbf{v}) - \Phi_j(\boldsymbol{\theta}_j) \right\|_{\mathbf{W}_{\text{obs}}} \leq |a| \tau_J + o(|a|). \quad (16)$$

Accordingly, the experiments are interpreted in terms of observational admissibility under the stated tolerance unless exact rank deficiency is established independently.

Property 2 (Persistence). Let $\boldsymbol{\theta}_s^*$ and $\boldsymbol{\theta}_{s-1}^*$ denote corresponding states in the enriched and reduced representations. For $j \in \{s-1, s\}$, let

$$\Pi_j^{\text{sh}} : \Theta_j \rightarrow \Theta_{\text{sh}} \quad (17)$$

denote the projection from representation j onto the common space of shared parameters Θ_{sh} . The corresponding states are required to have the same shared coordinates:

$$\Pi_s^{\text{sh}}(\boldsymbol{\theta}_s^*) = \Pi_{s-1}^{\text{sh}}(\boldsymbol{\theta}_{s-1}^*) = \boldsymbol{\eta}^*. \quad (18)$$

Choose local coordinates $\boldsymbol{\theta}_{s-1} = (\boldsymbol{\eta}, \boldsymbol{\zeta}_{s-1})$ compatible with Π_{s-1}^{sh} , where $\boldsymbol{\eta}$ contains the shared coordinates and $\boldsymbol{\zeta}_{s-1}$ contains the coordinates specific to

the reduced representation. At the corresponding reduced state $\boldsymbol{\theta}_{s-1}^* = (\boldsymbol{\eta}^*, \boldsymbol{\zeta}_{s-1}^*)$, define the sensitivity Jacobian with respect to the shared coordinates as

$$\mathbf{J}_{s-1}^{\text{sh}}(\boldsymbol{\theta}_{s-1}^*) = \left. \frac{\partial \Phi_{s-1}(\boldsymbol{\eta}, \boldsymbol{\zeta}_{s-1})}{\partial \boldsymbol{\eta}} \right|_{(\boldsymbol{\eta}^*, \boldsymbol{\zeta}_{s-1}^*)}. \quad (19)$$

Representation reduction removes the target mechanism while retaining the shared physical coordinates. Define the resulting response change at the matched states as $\mathbf{y}_{\text{rep}} := \Phi_s(\boldsymbol{\theta}_s^*) - \Phi_{s-1}(\boldsymbol{\theta}_{s-1}^*)$. Assume that $\mathbf{y}_{\text{rep}} \neq \mathbf{0}$, let $\mathbf{e}_i$ denote the coordinate direction associated with the designated carrier parameter i in the shared parameter space. A sufficient first-order local condition for the directed relation from target r to carrier i to persist after representation reduction is that there exists a sufficiently small nonzero carrier increment a_i , remaining within the local neighbourhood considered, such that

$$\mathbf{J}_{s-1}^{\text{sh}}(\boldsymbol{\theta}_{s-1}^*) \mathbf{e}_i a_i = \mathbf{y}_{\text{rep}}, \quad a_i \neq 0. \quad (20)$$

Equivalently, the representation response change must lie in the local response direction generated by the designated carrier parameter:

$$\mathbf{y}_{\text{rep}} \in \text{span} \left\{ \mathbf{J}_{s-1}^{\text{sh}}(\boldsymbol{\theta}_{s-1}^*) \mathbf{e}_i \right\}. \quad (21)$$

Under this condition, a sufficiently small change in the retained carrier parameter reproduces, to first order, the observable contribution associated with the change in physical representation. Persistence therefore does not require the carrier direction to be locally unobservable in the reduced representation. The carrier may remain locally identifiable. Persistence instead requires the carrier to be capable of reproducing the representation-induced response change. This condition is local and first order. The BcNN anchors do not evaluate the local response condition in Eq. (20), its equivalent span condition in Eq. (21), or the matched states directly. Instead, they provide a finite computational probe motivated by Persistence. With the observations, target parameter, carrier parameter and normalization convention held fixed, the reduced anchor tests whether variation of the same carrier can restore observational admissibility after the target mechanism is removed. The finite pseudo-solutions learned by the two anchors can therefore provide computational evidence consistent with Persistence, but they do not establish Eqs. (20) and (21).

Property 3 (Recoverability). Let $\mathcal{C}_A \subseteq \Theta_s$ denote an admissible parameter set supplied either by the anchor construction or by independent physical information. When $\mathcal{C}_A$ is constructed from the BcNN anchors, it is an anchor-defined admissible region conditioned on the inverse problem rather than an independent source of physical information. A parameter vector $\boldsymbol{\theta}_s^\dagger \in \mathcal{C}_A$ is locally recoverable within $\mathcal{C}_A$ if there exists a neighbourhood $\mathcal{U}$ of $\boldsymbol{\theta}_s^\dagger$ such that

$$\mathcal{E}_s(\boldsymbol{\theta}_s^\dagger) \cap \mathcal{C}_A \cap \mathcal{U} = \{\boldsymbol{\theta}_s^\dagger\}. \quad (22)$$

Let $\varrho^\dagger := \text{rank } \mathbf{J}_s(\boldsymbol{\theta}_s^\dagger)$. When the active part of the admissible set can be represented locally by an admissibility map defined by equalities

$$h_A : \Theta_s \rightarrow \mathbb{R}^{k-\varrho^\dagger}, \quad h_A(\boldsymbol{\theta}_s^\dagger) = \boldsymbol{\gamma}, \quad (23)$$

a sufficient local transversality condition is

$$\text{rank} \begin{bmatrix} \mathbf{J}_s(\boldsymbol{\theta}_s^\dagger) \\ Dh_A(\boldsymbol{\theta}_s^\dagger) \end{bmatrix} = k. \quad (24)$$

Equivalently,

$$\ker \mathbf{J}_s(\boldsymbol{\theta}_s^\dagger) \cap \ker Dh_A(\boldsymbol{\theta}_s^\dagger) = \{\mathbf{0}\}. \quad (25)$$

For an admissible set defined by inequalities, recoverability instead depends on whether the observational level set intersects the active admissible region locally at an isolated point. If the recovered point lies in the interior of an admissible region with full dimension, the inequalities do not reduce the local dimension. Exact local recovery then requires the response map restricted to that region to be locally injective. The scalar physical gauge introduced below is a coordinate scale derived from the anchors and used by the present BcNN implementation. It should not be confused with the general admissibility map $\mathbf{h}_A$, which is defined by equalities. A nonzero scalar gauge guarantees that the target reparameterization is well defined, but does not by itself establish structural identifiability or unique recovery.

These three properties characterize the local compensation structure of the inverse problem. The BcNN stages introduced below are computational

constructions motivated by these properties. The learned anchor and target states illustrate their computational counterparts but do not replace the formal rank and intersection conditions.

Compensation representation

For a carrier parameter i at representation order j and target label (r, c) , the remained-bias coordinate is written as

$$b_{i,j}^{r,c\leftarrow}. \quad (26)$$

The subscript pair (i, j) identifies the carrier parameter and the representation order in which it acts. The superscript pair (r, c) identifies the target parameter and its representation order. The left arrow $(\cdot^{\leftarrow})$ denotes the component retained in the selected normalized parameter coordinate. For an off-diagonal coordinate, it also records the directed relation from the target parameter to the carrier parameter. Let $\bar{\theta}_{i,j} \neq 0$ denote the prescribed reference value used to normalize parameter i at representation order j . The normalization reference $\theta_{i,j}$ forms part of the anchor and target parameterization. It should be distinguished from the underlying state θ_s^* associated with the observations and from the benchmark reference values used only for post hoc evaluation, even when some of these quantities coincide numerically in a controlled benchmark. The normalization references are prescribed physical scales fixed before BcNN optimization. Their selection is application dependent and may use physical information available for the system or a common scale selected from reported specimen parameters. They define the scale of the normalized anchor and target coordinates and remain fixed across the corresponding benchmark rather than being adjusted to the case specific reference values. The reference selection used here is described in Supplementary Methods.

For an off-diagonal anchor coordinate, let $\theta_{i,j}^{r,c}$ denote the value of the carrier parameter inferred by the anchor system in which parameter r at representation order c is selected as the compensation target. The remained-bias coordinate is defined as

$$b_{i,j}^{r,c\leftarrow} = \begin{cases} \frac{\theta_{i,j}}{\bar{\theta}_{i,j}}, & i = r, \\ \frac{\theta_{i,j}^{r,c}}{\bar{\theta}_{i,j}}, & i \neq r. \end{cases} \quad (27)$$

When $i=r$, the target and carrier parameters are identical. The coordinate describes parameter self-retention and does not encode compensation between distinct parameters. When $i \neq r$, the target and carrier parameters are distinct. The coordinate describes cross-parameter compensation carried by parameter i . A remained-bias coordinate is termed parameter-diagonal when $i=r$. It is termed fully diagonal when both parameter identity and representation order coincide with the target label:

$$(i, j) = (r, c). \quad (28)$$

Cross-parameter compensation is represented only by off-diagonal coordinates satisfying

$$i \neq r. \quad (29)$$

The complementary evolved-bias coordinate is denoted by

$$b_{i,j}^{r,c\gg} \quad (30)$$

and is defined through

$$b_{i,j}^{r,c\gg} + b_{i,j}^{r,c\leftarrow} = 1. \quad (31)$$

The remained-bias and evolved-bias forms are complementary descriptions of the same normalized physical coordinate. They are not independent physical degrees of freedom and should not both be optimized freely. For an off-diagonal anchor system, define the inverse training operator

$$\mathfrak{T}_{i,j}^{r,c\leftarrow} : \mathcal{D}_{\text{obs}} \mapsto \left(\hat{U}_{i,j}^{r,c\leftarrow}, \hat{b}_{i,j}^{r,c\leftarrow} \right), \quad i \neq r, \quad (32)$$

where the observation dataset is

$$\mathcal{D}_{\text{obs}} = \left\{ \left(\mathbf{x}_n, t_n, \mathbf{U}_n^{\text{obs}} \right) \right\}_{n=1}^{N_{\text{obs}}}. \quad (33)$$

The observation vector used in Eqs. (5), (8) and (9) is obtained by stacking the measured responses contained in $\mathcal{D}_{\text{obs}}$:

$$\mathbf{y}^{\text{obs}} = \text{col}(\mathbf{U}_1^{\text{obs}}, \dots, \mathbf{U}_{N_{\text{obs}}}^{\text{obs}}) \in \mathcal{Y} \subseteq \mathbb{R}^m. \quad (34)$$

Here, $\text{col}(\cdot)$ stacks the measured response vectors into a single column. When every observation contains d_U response components, $m = N_{\text{obs}} d_U$. The trained anchor neural field receives only the physical coordinates:

$$\hat{U}_{i,j}^{r,c\leftarrow}(\mathbf{x}, t) = \mathcal{N}_{i,j}^{r,c\leftarrow}(\mathbf{x}, t; \hat{\mathbf{w}}_{i,j}^{r,c\leftarrow}). \quad (35)$$

The observations condition optimization through the data loss rather than serving as direct neural network features.

Decomposition into anchor and target problems

The simplest nontrivial case contains two unknown physical parameters and two representation orders ($k=s=2$). The direct dual-parameter inverse problem optimizes the two physical parameters simultaneously with the neural field weights:

$$\left(\hat{\theta}_{1,2}^{\text{dir}}, \hat{\theta}_{2,2}^{\text{dir}}, \hat{\mathbf{w}}_{\text{dir}} \right) = \arg \min_{\theta_{1,2}, \theta_{2,2}, \mathbf{w}_{\text{dir}}} \mathcal{J}_{\text{dir}}(\theta_{1,2}, \theta_{2,2}, \mathbf{w}_{\text{dir}}; \mathcal{D}_{\text{obs}}). \quad (36)$$

This joint optimization leaves both physical parameters free to move along strongly correlated or approximately compensatory directions.

Design condition 1. Reduction to a single physical parameter

Parameter 2 at representation order 2 is selected as the compensation target. Parameter 1 is retained as the carrier parameter in both anchor representations. For $j \in \{1, 2\}$, the anchor parameter space is

$$\Theta_{\Lambda,j}^{2,2} = \left\{ \mathcal{P}_{1,j}^{2,2}(b) : b \in \mathcal{D}_{1,j}^{2,2} \right\}, \quad \dim \Theta_{\Lambda,j}^{2,2} = 1. \quad (37)$$

$\mathcal{D}_{1,j}^{2,2} \subseteq \mathbb{R}$ denotes the prescribed domain of the normalized anchor coordinate b at representation order j . For the present dual-parameter hierarchy, the anchor parameterization is

$$\mathcal{P}_{1,j}^{2,2}(b) = \begin{cases} b\bar{\theta}_{1,1}, & j=1, \\ (b\bar{\theta}_{1,2}, \bar{\theta}_{2,2})^\top, & j=2. \end{cases} \quad (38)$$

For $j=1$, the target mechanism is absent from the reduced representation and only the carrier parameter is active. For $j=2$, the target parameter remains fixed at its prescribed normalization reference while the normalized carrier coordinate b is the only trainable physical parameter. Neural network weights remain trainable in both cases. The idealized anchor response map is

$$\Phi_{A,j}^{2,2}(b) = \Phi_j(\mathcal{P}_{1,j}^{2,2}(b)). \quad (39)$$

Assume that b_j^* is an interior point of the prescribed scalar anchor domain and that $\Phi_{A,j}^{2,2}$ is continuously differentiable in a neighbourhood of b_j^* . A sufficient local identifiability condition for the single physical anchor coordinate is

$$\text{rank} \left[\frac{\partial \Phi_{A,j}^{2,2}}{\partial b}(b_j^*) \right] = 1, \quad j \in \{1, 2\}. \quad (40)$$

Here, b_j^* denotes the locally identifiable anchor coordinate associated with representation order j . Because the physical anchor coordinate is scalar, the rank-one condition means that at least one observable component has a nonzero local derivative with respect to b , which is sufficient for local identification of the anchor coordinate. This condition distinguishes a meaningful anchor problem with one trainable physical parameter from an anchor representation that is locally insensitive to the retained carrier coordinate.

Design condition 2. Conditioning on common observations

The two anchor systems use the same observation coordinates, measured responses and observation operator. Their implemented estimators are

$$(\hat{b}_j, \hat{\mathbf{w}}_{A,j}) = \arg \min_{b \in \mathcal{D}_{1,j}^{2,2}, \mathbf{w}_{A,j}} \mathcal{J}_{A,j}(b, \mathbf{w}_{A,j}; \mathcal{D}_{\text{obs}}), \quad j \in \{1, 2\}. \quad (41)$$

The corresponding fitted anchor observations are

$$\hat{\mathbf{y}}_{A,j} = \mathcal{H} \hat{\mathbf{U}}_{A,j}. \quad (42)$$

An anchor pseudo-solution is observationally admissible when

$$\|\hat{\mathbf{y}}_{A,j} - \mathbf{y}^{\text{obs}}\|_{\mathbf{w}_{\text{obs}}} \leq \tau_A, \quad \tau_A \geq 0. \quad (43)$$

τ_A is the prescribed observational tolerance for each anchor system. The two anchors use the same observations, carrier parameter, compensation target and normalization convention, together with matched network architectures and training settings. These controls limit differences due to data conditioning and network design. The remaining difference in the inferred remained-bias coordinates is therefore interpreted as arising from the change in physical representation, subject to neural approximation and optimization error.

Design condition 3. Hierarchical carrier compatibility

The two anchors use the same compensation target, the same carrier parameter, adjacent physical representations and a common normalization convention:

$$(r, c) = (2, 2), \quad i = 1, \quad j \in \{1, 2\}. \quad (44)$$

The reduced and enriched anchor estimates are

$$\hat{b}_L = \hat{b}_{1,1}^{2,2\kappa}, \quad \hat{b}_H = \hat{b}_{1,2}^{2,2\kappa}. \quad (45)$$

The enriched anchor exposes an observationally admissible compensatory pseudo-solution of finite amplitude, consistent with Existence. The reduced anchor tests whether the same relation between the target and carrier remains observationally admissible after the physical representation is reduced, thereby providing a computational probe motivated by Persistence.

Design condition 4. Frozen transfer from the anchors to the target network

The two anchor systems are optimized independently before the target network is trained. Once both anchors have been optimized, their neural field weights and remained-bias estimates ($\hat{b}_L$ and $\hat{b}_H$) are held fixed and used only to condition the target objective. The target stage therefore optimizes $\mathbf{w}_T$, μ and ν without updating either anchor. This prevents optimization in the target stage from retrospectively changing the pseudo-solutions used to construct the physical gauge and the intervals for the auxiliary coordinates. The complete computational sequence is

$$\begin{aligned} (\hat{b}_L, \hat{\mathbf{w}}_{A,1}) &= \arg \min_{b, \mathbf{w}_{A,1}} \mathcal{J}_{A,1}, \\ (\hat{b}_H, \hat{\mathbf{w}}_{A,2}) &= \arg \min_{b, \mathbf{w}_{A,2}} \mathcal{J}_{A,2}, \end{aligned} \quad (46)$$

followed by

$$(\hat{\mu}_{1,2}, \hat{\nu}_{2,2}, \hat{\mathbf{w}}_T) = \arg \min_{\mu, \nu, \mathbf{w}_T} \mathcal{J}_T(\mu, \nu, \mathbf{w}_T | \hat{b}_L, \hat{b}_H). \quad (47)$$

Each anchor problem contains one trainable physical coordinate. Its neural field weights remain trainable. The target stage still recovers two physical parameters. BcNN therefore does not eliminate the dual-parameter recovery task. It replaces an unstructured joint search with two independently optimized anchor problems, each with one trainable physical parameter, followed by a dual-parameter target problem.

Physical gauge and complementary bias reparameterization

The ordered difference between the two off-diagonal anchor remained-bias coordinates defines the scalar physical gauge:

$$\hat{g}_{2,2} = \hat{b}_H - \hat{b}_L = \hat{b}_{1,2}^{2,2\kappa} - \hat{b}_{1,1}^{2,2\kappa}. \quad (48)$$

The ordering is fixed as the estimate from the enriched representation minus that from the reduced representation. The present implementation adopts the orientation

$$\hat{g}_{2,2} > 0, \quad (49)$$

equivalently, $\hat{b}_H > \hat{b}_L$, and all reported experiments satisfied this precondition. Equal anchor estimates yield $\hat{g} = 0$, which is a degenerate case. Oppositely

ordered estimates yield a negative but nonzero gauge. Such a gauge is not mathematically singular, but it falls outside the interval orientation adopted in Eqs. (62)-(64). The gauge is not replaced by its absolute value because its sign records the direction of the change from the reduced to the enriched representation. The target network does not directly optimize the two fully diagonal remained-bias coordinates. Instead, it optimizes two auxiliary coordinates defined from the corresponding evolved-bias forms. For the dominant parameter,

$$\hat{\mu}_{1,2} = \hat{b}_{1,2}^{1,2\gg}, \quad (50)$$

and for the small parameter,

$$\hat{\nu}_{2,2} = \hat{g}_{2,2} \hat{b}_{2,2}^{2,2\gg}. \quad (51)$$

The corresponding diagonal remained-bias coordinates are reconstructed deterministically:

$$\hat{b}_{1,2}^{1,2\ll} = 1 - \hat{\mu}_{1,2}, \quad (52)$$

and

$$\hat{b}_{2,2}^{2,2\ll} = 1 - \frac{\hat{\nu}_{2,2}}{\hat{g}_{2,2}}. \quad (53)$$

The recovered dominant and small parameters are

$$\hat{\theta}_{1,2} = \hat{b}_{1,2}^{1,2\ll} \bar{\theta}_{1,2} = (1 - \hat{\mu}_{1,2}) \bar{\theta}_{1,2}, \quad (54)$$

and

$$\hat{\theta}_{2,2} = \hat{b}_{2,2}^{2,2\ll} \bar{\theta}_{2,2} = \left(1 - \frac{\hat{\nu}_{2,2}}{\hat{g}_{2,2}}\right) \bar{\theta}_{2,2}. \quad (55)$$

The target neural field is represented as

$$\hat{U}_{2,2}(\mathbf{x}, t) = \mathcal{N}_{2,2}^T(\mathbf{x}, t; \hat{\mathbf{w}}_T). \quad (56)$$

The target training operator is

$$\mathfrak{T}_{2,2}^T : (\mathcal{D}_{\text{obs}}, \hat{b}_L, \hat{b}_H) \mapsto (\hat{U}_{2,2}, \hat{b}_{1,2}^{1,2\ll}, \hat{b}_{2,2}^{2,2\ll}). \quad (57)$$

The observations and frozen anchor estimates condition the optimization problem. The neural field receives only the physical coordinates as direct inputs, whereas the reconstructed physical parameters enter the governing equation residual. The physical parameter reparameterization can be written as

$$\Psi_{\hat{g}}(\mu, \nu) = \begin{pmatrix} (1 - \mu) \bar{\theta}_{1,2} \\ \left(1 - \frac{\nu}{\hat{g}_{2,2}}\right) \bar{\theta}_{2,2} \end{pmatrix}. \quad (58)$$

Its Jacobian is

$$D\Psi_{\hat{g}} = \begin{pmatrix} -\bar{\theta}_{1,2} & 0 \\ 0 & -\frac{\bar{\theta}_{2,2}}{\hat{g}_{2,2}} \end{pmatrix}. \quad (59)$$

When the normalization references are nonzero and $\hat{g}_{2,2} > 0$, the transformation is locally invertible. Consequently,

$$D(\Phi_2 \circ \Psi_{\hat{g}})(\mu, \nu) = \mathbf{J}_2(\Psi_{\hat{g}}(\mu, \nu)) D\Psi_{\hat{g}}(\mu, \nu), \quad (60)$$

and

$$\text{rank } D(\Phi_2 \circ \Psi_{\hat{g}})(\mu, \nu) = \text{rank } \mathbf{J}_2(\Psi_{\hat{g}}(\mu, \nu)). \quad (61)$$

The complementary bias reparameterization can therefore change parameter scale, orientation and optimization conditioning, but it does not by itself increase the rank of the local parameter sensitivity Jacobian. Unique physical recovery remains conditional on local identifiability within the admissible region defined by the anchors.

Admissible region defined by the anchors

The auxiliary coordinate for the dominant parameter is assigned the interval

$$1 - \hat{b}_H \leq \hat{\mu}_{1,2} \leq 1 - \hat{b}_L, \quad (62)$$

whereas the auxiliary coordinate for the small parameter is assigned

$$0 \leq \hat{\nu}_{2,2} \leq \hat{g}_{2,2}. \quad (63)$$

The admissible region defined by the anchors in the auxiliary coordinate space is

$$\Xi_A = [1 - \hat{b}_H, 1 - \hat{b}_L] \times [0, \hat{g}_{2,2}]. \quad (64)$$

The corresponding admissible set in the physical parameter coordinates is

$$\mathcal{C}_A(\hat{b}_L, \hat{b}_H) = \Psi_{\hat{g}}(\Xi_A). \quad (65)$$

Equivalently,

$$\mathcal{C}_A(\hat{b}_L, \hat{b}_H) = \left\{ \boldsymbol{\theta}_2 : \hat{b}_L \leq \frac{\theta_{1,2}}{\theta_{1,2}} \leq \hat{b}_H, \quad 0 \leq \frac{\theta_{2,2}}{\theta_{2,2}} \leq 1 \right\}. \quad (66)$$

For an exact response, local recovery within this region requires

$$\mathcal{E}_2(\boldsymbol{\theta}_2^\dagger) \cap \mathcal{C}_A(\hat{b}_L, \hat{b}_H) \cap \mathcal{U} = \{\boldsymbol{\theta}_2^\dagger\}. \quad (67)$$

For noisy observations, the corresponding requirement is an isolated minimum, or a sufficiently well-conditioned local minimum, of the combined observation and physics objective within the region defined by the anchors.

Equivalence of complementary coordinates and invariance of the anchor difference

For the dominant parameter,

$$\hat{b}_{1,2}^{1,2\ll} = 1 - \hat{\mu}_{1,2} = 1 - \hat{b}_{1,2}^{1,2\gg}. \quad (68)$$

For the small parameter,

$$\hat{b}_{2,2}^{2,2\ll} = 1 - \frac{\hat{\nu}_{2,2}}{\hat{g}_{2,2}} = 1 - \hat{b}_{2,2}^{2,2\gg}. \quad (69)$$

The remained-bias and evolved-bias forms therefore reconstruct algebraically identical physical parameters. Any difference between the two reconstructions arises only from numerical implementation or postprocessing error. The physical gauge remains invariant under a common translation a of the two anchor coordinates:

$$\hat{b}_L' = \hat{b}_L + a, \quad \hat{b}_H' = \hat{b}_H + a, \quad (70)$$

because

$$\hat{g}_{2,2}' = \hat{b}_H' - \hat{b}_L' = \hat{g}_{2,2}. \quad (71)$$

This result applies only to the signed difference between the two anchor coordinates. It does not imply invariance of the individual anchor coordinates, the admissible intervals, the target coordinates, the recovered physical parameters or the complete complementary bias reparameterization. All anchor and target coordinates must use a consistent normalization convention.

Soft unilateral-boundary constraint defined by the anchors

For a scalar coordinate z with prescribed interval $[z_{\min}, z_{\max}]$, define the unilateral-boundary violation as

$$C_{\text{UB}}(z; z_{\min}, z_{\max}) = \text{ReLU}(z_{\min} - z) + \text{ReLU}(z - z_{\max}). \quad (72)$$

The two coordinate violations are

$$C_{\mu} = C_{\text{UB}}(\hat{\mu}_{1,2}; 1 - \hat{b}_H, 1 - \hat{b}_L), \quad (73)$$

and

$$C_{\nu} = C_{\text{UB}}(\hat{\nu}_{2,2}; 0, \hat{g}_{2,2}). \quad (74)$$

The theoretical and structural benchmarks considered here have scalar governing equation residuals, and the following unilateral-boundary construction is therefore written for a scalar residual. Let $j(q)$ denote the corresponding representation order, with $j(q)=1$ for the reduced anchor and $j(q)=2$ for both the enriched anchor and the target network. Let N_f be the number of collocation points used for the governing equation, and denote these points by

$\{(\mathbf{x}_{f,n}, t_{f,n})\}_{n=1}^{N_f}$. The unconstrained equation residual at collocation point n is

$$\mathcal{R}_{f,n}^{(q)} = \mathcal{G}_{j(q)}[\hat{U}_q; \hat{\boldsymbol{\theta}}_q](\mathbf{x}_{f,n}, t_{f,n}). \quad (75)$$

For the target network, the implemented residual is

$$\tilde{\mathcal{R}}_{f,n}^{(\text{T})} = \mathcal{R}_{f,n}^{(\text{T})} + \alpha C_{\mu} + \beta C_{\nu}, \quad (76)$$

The governing equation residual was nondimensionalized before the unilateral-boundary terms were introduced. C_{μ} and C_{ν} are dimensionless. $\alpha > 0$ and $\beta > 0$ control their numerical contribution to the normalized residual. Their numerical values are specified in Supplementary Methods. The modified target physics loss is

$$\tilde{\mathcal{L}}_f^{(\text{T})} = \frac{1}{N_f} \sum_{n=1}^{N_f} |\tilde{\mathcal{R}}_{f,n}^{(\text{T})}|^2. \quad (77)$$

The unconstrained physics loss is

$$\mathcal{L}_{f,0}^{(\text{T})} = \frac{1}{N_f} \sum_{n=1}^{N_f} |\mathcal{R}_{f,n}^{(\text{T})}|^2. \quad (78)$$

When both auxiliary coordinates lie inside their intervals,

$$C_{\mu} = C_{\nu} = 0, \quad (79)$$

the original residual and physics loss are recovered exactly:

$$\tilde{\mathcal{R}}_{f,n}^{(\text{T})} = \mathcal{R}_{f,n}^{(\text{T})}, \quad \tilde{\mathcal{L}}_f^{(\text{T})} = \mathcal{L}_{f,0}^{(\text{T})}. \quad (80)$$

When either coordinate leaves its prescribed interval, the corresponding ReLU channel becomes positive and modifies the target residual. The ReLU violation channels are therefore inactive inside the region defined by the anchors and active outside it. Because the violation terms are added to the signed residual before squaring, the present construction is a soft residual bias rather than a hard projection or barrier. Let

$$C = \alpha C_{\mu} + \beta C_{\nu}. \quad (81)$$

Then

$$\tilde{\mathcal{L}}_f^{(\text{T})} - \mathcal{L}_{f,0}^{(\text{T})} = 2C \bar{\mathcal{R}}_f^{(\text{T})} + C^2, \quad (82)$$

where

$$\bar{\mathcal{R}}_f^{(\text{T})} = \frac{1}{N_f} \sum_{n=1}^{N_f} \mathcal{R}_{f,n}^{(\text{T})}. \quad (83)$$

The cross term is not guaranteed to be positive. The soft unilateral-boundary constraint therefore acts as a residual bias that discourages excursions outside the prescribed intervals and favours the region defined by the anchors. It does not guarantee a monotonic increase in the physics loss for every interval

violation or ensure that every iterate remains inside that region.

Training objectives and sequential optimization

For component network q , the observation loss is

$$\mathcal{L}_{\text{data}}^{(q)} = \frac{1}{N_{\text{obs}}} \sum_{n=1}^{N_{\text{obs}}} \left\| \hat{U}_q(\mathbf{x}_n^{\text{obs}}, t_n^{\text{obs}}) - U_n^{\text{obs}} \right\|_2^2. \quad (84)$$

The unconstrained governing equation loss is

$$\mathcal{L}_{f,0}^{(q)} = \frac{1}{N_f} \sum_{n=1}^{N_f} \left| \mathcal{R}_{f,n}^{(q)} \right|^2. \quad (85)$$

The benchmarks reported here are stationary and therefore require no initial condition loss. Let $N_{B,0}$ and $N_{B,1}$ denote the numbers of samples used for the

Dirichlet condition and the condition of zero slope, respectively, and let $\mathbf{x}_{B,0}^{(n)}$ and $\mathbf{x}_{B,1}^{(n)}$ denote the corresponding boundary locations. Their losses are

$$\mathcal{L}_{B,0}^{(q)} = \frac{1}{N_{B,0}} \sum_{n=1}^{N_{B,0}} \left| \hat{U}_q(\mathbf{x}_{B,0}^{(n)}) \right|^2, \quad (86)$$

and

$$\mathcal{L}_{B,1}^{(q)} = \frac{1}{N_{B,1}} \sum_{n=1}^{N_{B,1}} \left| \frac{\partial \hat{U}_q}{\partial x}(\mathbf{x}_{B,1}^{(n)}) \right|^2. \quad (87)$$

The two anchor objectives are

$$\mathcal{J}_{A,j} = \lambda_f \mathcal{L}_{f,0}^{(j)} + \lambda_{B,0} \mathcal{L}_{B,0}^{(j)} + \lambda_{B,1} \mathcal{L}_{B,1}^{(j)} + \lambda_{\text{data}} \mathcal{L}_{\text{data}}^{(j)}, \quad j \in \{1, 2\}. \quad (88)$$

Here, $\lambda_f, \lambda_{B,0}, \lambda_{B,1}, \lambda_{\text{data}} > 0$ are the weights assigned to the governing equation, boundary conditions and observation loss. Their numerical values are specified in Supplementary Methods. The set of trainable physical coordinates in each anchor is

$$\mathcal{V}_A^{(j)} = \{b_{1,j}^{2,2\kappa}\}, \quad |\mathcal{V}_A^{(j)}| = 1. \quad (89)$$

The target objective is

$$\mathcal{J}_T = \lambda_f \tilde{\mathcal{L}}_f^{(T)} + \lambda_{B,0} \mathcal{L}_{B,0}^{(T)} + \lambda_{B,1} \mathcal{L}_{B,1}^{(T)} + \lambda_{\text{data}} \mathcal{L}_{\text{data}}^{(T)}. \quad (90)$$

The vector of trainable target auxiliary coordinates is

$$\xi_T = (\mu_{1,2}, \nu_{2,2})^\top \in \mathbb{R}^2. \quad (91)$$

The BcNN components are trained in stages. The reduced and enriched anchors are first optimized independently. Their neural field weights and remained-bias estimates are then frozen. The frozen estimates define the physical gauge and the intervals for the target auxiliary coordinates. The dual-parameter target network is optimized only after the anchors have been frozen.

Experimental design, evaluation metrics and data provenance

Unlike the evaluation of a forward problem, assessment of multi-parameter inverse learning must examine parameter recovery as well as response reconstruction. The benchmark design uses three controls. First, each target problem contains two unknown physical parameters. BcNN addresses this target through two anchor subproblems, each with one trainable physical parameter, followed by a dual-parameter target stage. Second, except in the designated architecture stress tests, the compared models use the same default network width, depth and activation function, together with the same optimizer family and initialization protocol. The loss weights and optimization schedules used for each model are reported in Supplementary Methods and were kept fixed across parameter and loading conditions within each benchmark. This avoids case-specific network tuning and limits architectural confounding. Third, the observations are deliberately sparse. The theoretical benchmarks use sparse samples from exact solutions, while the structural experiments use sparse measurements with unquantified sensor-related experimental uncertainty.

The unknown parameters differ in scale and response sensitivity, which can lead to strong ill-conditioning or practical non-identifiability. Dominant and small are relative labels defined within each prescribed physical condition. The dominant parameter is the one whose prescribed variation produces the larger change in the observable response over the benchmark range, while the small parameter produces the weaker change. The labels describe relative observable influence rather than numerical magnitude. Evaluation therefore includes relative L_2 response error, optimized objective value, relative errors in both parameters, convergence epochs and full parameter trajectories.

The experimental program combines theoretical benchmarks without added noise with sparse structural measurements. Component ablation tests assess remained-bias quantification and the soft unilateral-boundary constraint. Separate architecture stress tests vary the activation function, network width and depth. The structural experiments evaluate parameter recovery in composite beams exhibiting compression softening under sparse measurements. For the theoretical benchmark, observations are generated from exact solutions of the full Sturm-Liouville model. The highest physical representation order is $s=2$. The full representation is indexed by $j=2$, while the reduced representation is indexed by $j=1$. For the structural experiments, the compression softening model and its reduced Euler-Bernoulli counterpart provide the corresponding physical hierarchy. The benchmark tests whether the anchors expose the off-diagonal remained-bias coordinates, whether the same relation between the target and carrier remains admissible in the reduced and enriched representations, whether the signed anchor difference defines a nondegenerate physical gauge, whether reparameterization and the soft unilateral-boundary constraint support recovery of both physical parameters, and whether the same construction remains useful for sparse structural measurements.

The controlled benchmark uses the analytical cases, parameter settings and exact responses reported by Yang et al. [40], with the governing beam-column formulation recast here in the Sturm-Liouville representation adopted for the BcNN hierarchy. The baseline conditions vary the dominant parameter, the small parameter and the loading level. We then evaluated BcNN using sparse measurements from nine independent composite beam experiments reported by Bonopera et al. [41]. These experiments span variations in axial compressive force, bending rigidity and loading. The benchmark separates response reconstruction, recovery of the dominant parameter and recovery of the weak physical effect associated with the small parameter because success in the first does not imply success in the latter two.

The controlled theoretical benchmark comprises the same nine analytical cases, denoted T1-T9. Exact response fields were generated analytically from the recast full Sturm-Liouville representation used in this study and sampled at 2, 5, 10 and 20 observation points. The physical parameters used to generate each response served as exact reference values for evaluating parameter recovery. The structural benchmark comprises nine independent experimental cases reported by Bonopera et al. [41], denoted T1^r-T9^r. Each experiment provides five measurements and a pair of reference physical parameters. The same

measurements were supplied separately to baseline PINN and BcNN.

The controlled benchmark contains 203 configurations spanning observation sparsity, component ablation, response scale and architecture sensitivity. The structural benchmark contains nine independent experiments, each evaluated separately using baseline PINN and BcNN. The complete design is summarized in Extended Data Table 1.

Training horizons and epoch accounting

Training horizons were selected separately for the controlled benchmark and the structural experiments. Unless otherwise stated, controlled runs were evaluated at a common horizon of 0.4×10^6 epochs for the optimization stage being compared. In the analysis of Fig. 5b,c, baseline PINN was additionally continued to 1×10^6 , 2×10^6 and 3×10^6 epochs to examine whether longer training resolved the observed parameter instability. For BcNN, the reported epoch count refers to the target stage after the two anchors have been trained and frozen. For each structural experiment, baseline PINN and the BcNN target stage were evaluated over the same maximum horizon of 1.2×10^6 epochs. The reported BcNN epochs therefore exclude the epochs used to train the anchors and characterize optimization after construction of the physical gauge rather than the total computational cost of BcNN. The maximum training horizon was fixed in advance and was not used as an early stopping criterion. Parameter accuracy and trajectory behaviour were assessed afterwards from the saved parameter trajectories.

Methodological rationale

BcNN does not rely on a more expressive neural network. Its main methodological change is the organization of the inverse problem. Two anchor problems, each with one trainable physical parameter, are solved independently before the dual-parameter target problem.

The two anchors serve different roles despite sharing the same observations, carrier parameter, compensation target and normalization convention because they operate at different physical representation orders. Their signed difference is interpreted as the compensatory change associated with the shift in physical representation, subject to neural approximation and optimization error. Because the anchor estimates are frozen before the target is trained, optimization in the target stage cannot retrospectively alter the pseudo-solutions used to construct the physical gauge and the intervals for the auxiliary coordinates.

The target network does not optimize the original physical parameters directly. Instead, it optimizes two auxiliary coordinates defined from the evolved-bias forms and reconstructs the fully diagonal remained-bias coordinates and physical parameters through a reparameterization conditioned on the anchor estimates. In this sequence, the enriched anchor exposes an observationally admissible compensatory state, the reduced anchor evaluates whether the same relation between the target and carrier remains observationally admissible after the representation is reduced, and the target network searches for a stable parameter representative within the region defined by the anchors.

This staged construction makes parameter compensation explicit without creating information that is absent from the observations or converting a locally non-injective response map into a locally injective one.

Extended Data Table 1 | Dual-parameter benchmark matrix ($k=2, s=2$). Here, k denotes the number of unknown physical parameters and s denotes the highest representation order. The controlled theoretical benchmark comprises nine baseline Sturm-Liouville conditions generated by varying the dominant parameter, the small parameter and the loading level one factor at a time, with three levels assigned to each factor. These conditions are used to evaluate observation sparsity, component ablation and architecture sensitivity, yielding 198 configurations. Five additional configurations evaluate response scale at normalized loading amplitudes of $0.01F$, $0.1F$, $0.5F$, $1F$ and $2F$, bringing the theoretical total to 203 configurations. The structural benchmark comprises nine independent structural experiments, each containing five observation points and subject to unquantified sensor-related experimental uncertainty. Baseline PINN and full BcNN are evaluated separately for each experiment, yielding 18 nominal configurations. Except in the designated architecture stress tests, the compared models use the same default network width, depth and activation function, together with the same optimizer family and initialization protocol. The loss weights and optimization schedules used for each model are reported in Supplementary Methods. Dominant and small parameters are distinguished by their relative influence on the observable response under the considered physical and observational conditions, rather than by their numerical magnitude. Representation order refers to the physical model hierarchy and is distinct from the differential order of the governing operator. Comparable loss denotes the aggregate of the loss terms shared by the compared models. For the controlled theoretical benchmark, parameter errors are reported relative to the exact values. For the structural experiments, parameter errors are reported relative to the reference values for each experiment. The equations shown above the benchmark matrix summarize the hierarchy of physical representations used for the theoretical Sturm-Liouville benchmark. For the theoretical Sturm-Liouville benchmark, the observed response is u , the dominant parameter is δ , and the small parameter is n . The structural experiments use the compression softening model and its reduced Euler-Bernoulli representation, as detailed in Methods.

Full Sturm-Liouville benchmark model: $\frac{L^2}{\delta} \nabla^2 u + \frac{n}{\delta} u = -\frac{x}{2L} \rightarrow$ Compensation representation at $j=s=2$: $\frac{L^2}{b_{1,2}^{1,2\kappa} \delta} \nabla^2 u + \frac{b_{2,2}^{2,2\kappa} \bar{n}}{b_{1,2}^{1,2\kappa} \delta} u = -\frac{x}{2L}$

Reduced anchor representation at $j=s-1=1$: $\frac{L^2}{b_{1,1}^{2,2\kappa} \delta} \nabla^2 u = -\frac{x}{2L}$ & Enriched anchor representation at $j=s=2$: $\frac{L^2}{b_{1,2}^{2,2\kappa} \delta} \nabla^2 u + \frac{\bar{n}}{b_{1,2}^{2,2\kappa} \delta} u = -\frac{x}{2L}$

Benchmark type	Varied physical condition	Number of observations	Model comparison and ablation	Stress tests	Evaluation metrics
Theoretical benchmark (exact and sparse)	Dominant parameter: 3 levels; Small parameter: 3 levels; Loading: 3 levels	2, 5, 10, 20	Baseline PINN;	Activations: ELU, SELU, sigmoid, sine and tanh;	Relative L ₂ response error; Comparable loss; Epochs to convergence; Small parameter error; Dominant parameter error; Parameter trajectories
			BcNN without remained-bias	Widths: 30, 50 and 70 neurons per hidden layer;	
			quantification;	Depths: 4, 6 and 8 hidden layers	
			BcNN without the unilateral-boundary constraint;	Width-depth combinations: 3×3=9	
			Full BcNN	Response scale: 0.01 <i>F</i> , 0.1 <i>F</i> , 0.5 <i>F</i> , 1 <i>F</i> and 2 <i>F</i>	

Structural experiments (sparse observations with unquantified uncertainty)	Dominant parameter: 3 levels; Small parameter: 3 levels; Loading: 3 levels	5	Baseline PINN; Full BcNN	Not performed	Relative L ₂ response error; Comparable loss; Epochs to convergence; Small parameter error; Dominant parameter error; Parameter trajectories
Evaluation count	Theoretical benchmark: 203; Structural experiments: 18				

Extended Data Table 2 | Parameter recovery from independent structural experiments. T1^r-T9^r denote independent structural experiments, each containing five observation points. Values outside parentheses are obtained using BcNN, whereas values in parentheses are obtained using baseline PINN. Parameter errors are reported relative to the reference values for each experiment. The relative L₂ response error is evaluated across the five observation points of each experiment. Comparable loss denotes the aggregate of the loss terms shared by BcNN and baseline PINN. The maximum training horizon was 1.2×10⁶ epochs for both methods in every structural experiment. For BcNN, the reported convergence epoch refers to the dual-parameter target stage after the two anchor networks have been optimized and frozen. The epochs used to train the anchors are not included because the comparison concerns only the target stage. The reported BcNN convergence epoch was determined post hoc from the complete parameter trajectories. “Failure” indicates that baseline PINN did not simultaneously satisfy the criteria for parameter accuracy and trajectory stability for both physical parameters within the maximum training horizon.

Metrics	T1 ^r	T2 ^r	T3 ^r	T4 ^r	T5 ^r	T6 ^r	T7 ^r	T8 ^r	T9 ^r
Dominant parameter error	1.29% (-0.36%)	2.87% (0.37%)	-1.77% (-9.51%)	0.28% (-0.09%)	0.33% (-1.88%)	-0.17% (-1.87%)	-0.21% (-0.84%)	-0.04% (0.22%)	0.50% (0.07%)
Small parameter error	4.11% (28.62%)	-10.22% (31.49%)	8.83% (29.60%)	0.12% (-0.54%)	-2.22% (24.13%)	3.58% (24.69%)	0.00% (7.47%)	-0.30% (-3.64%)	-1.13% (-1.24%)
Relative L ₂ response error	1.90E-02 (1.39E-02)	2.61E-02 (2.42E-02)	1.34E-02 (7.54E-02)	3.14E-03 (1.26E-03)	1.89E-03 (8.66E-04)	8.63E-04 (1.19E-03)	1.08E-02 (4.04E-03)	3.11E-03 (4.26E-03)	4.56E-03 (1.10E-03)
Comparable loss	4.50E-13 (1.73E-12)	5.97E-13 (1.83E-12)	5.60E-13 (2.56E-12)	1.30E-12 (2.49E-12)	1.29E-12 (1.34E-12)	9.97E-13 (1.86E-12)	1.61E-12 (1.64E-12)	1.59E-12 (1.84E-12)	5.97E-13 (2.45E-12)
Epochs to convergence (maximum budget: 1.2×10 ⁶)	0.8×10 ⁶ (Failure)	0.8×10 ⁶ (Failure)	0.8×10 ⁶ (Failure)	0.7×10 ⁶ (Failure)	0.7×10 ⁶ (Failure)	0.7×10 ⁶ (Failure)	0.6×10 ⁶ (Failure)	0.6×10 ⁶ (Failure)	0.6×10 ⁶ (Failure)

516 **Data availability**

517 The controlled theoretical data are generated from the exact solutions and parameter settings described in Methods and Supplementary Methods. The
518 structural measurements analysed in this study were obtained from Bonopera et al. [41]. Processed data underlying the reported results and the structural
519 input data used in the analyses will be available to editors and reviewers during peer review and will be made publicly available in Zenodo upon publication
520 under DOI 10.5281/zenodo.22019417.

521 **Code availability**

522 The code used to implement BeNN, reproduce the controlled benchmarks and generate the reported analyses will be available to editors and reviewers
523 during peer review and will be made publicly available in Zenodo upon publication under DOI 10.5281/zenodo.22019493.

524 **Funding**

525 This work was supported by the National Natural Science Foundation of China (grant nos. 52678187), the Joint Training Base Construction Project for
526 Graduate Students in Chongqing (JDLHPYJD2020023) and the Team Building Project for Graduate Tutors in Chongqing (JDDSTD2022003).

527 **Author contributions**

528 Y.Y. conceived the study, developed the theoretical framework and methodology, implemented BeNN, performed the analytical derivations, numerical
529 experiments and data analyses, prepared the figures and tables, and wrote the manuscript. Z.G. supervised the study, acquired funding, provided project
530 administration, and reviewed and edited the manuscript. Both authors reviewed and approved the submitted version.

531 **Competing interests**

532 The authors declare no competing interests.

533 **Supplementary information**

534 See Supplementary Methods.

536 Network architecture and numerical implementation

537 All neural networks were implemented using DeepXDE with the TensorFlow backend. Unless otherwise stated in the architecture sensitivity tests, the
 538 reduced anchor, enriched anchor and target network used the same fully connected architecture, consisting of six hidden layers with 50 neurons in each
 539 hidden layer. The default activation function was tanh, and the network weights were initialized using Glorot uniform initialization. The baseline PINN used
 540 the same default network architecture unless otherwise stated. Each network takes the spatial coordinate x as input and predicts the response field $u(x)$.
 541 Observations enter the model through the data loss rather than as additional network inputs. Spatial derivatives required by the governing equations and
 542 boundary conditions are evaluated using automatic differentiation.

543 The problems considered here are stationary. Owing to symmetry, only one half of the physical domain is modelled. The displacement condition is
 544 imposed at the support, whereas the zero slope condition is imposed at midspan. The detailed beam-column formulation, its transformation to the
 545 Sturm-Liouville representation and the analytical solution used to generate the theoretical responses follow Yang et al. [40] and are not repeated here.

546 For the default calculations, 10,000 collocation points were sampled uniformly inside the computational domain, together with 10 boundary samples. For
 547 the theoretical benchmark, an additional set of 10,000 uniformly distributed points was used to evaluate the predicted response against the analytical solution.
 548 Trainable physical parameters or auxiliary coordinates were recorded every 100 epochs so that their optimization trajectories could be examined afterwards.
 549 No artificial noise was added to the theoretical observations. For the structural experiments, the published measurements were used directly. Because
 550 quantitative uncertainty bounds were not available for the individual measurements, no assumed probability distribution or synthetic noise level was
 551 introduced. Architecture sensitivity was examined separately from the default calculations. Five activation functions were considered: ELU, SELU, sigmoid,
 552 sine and tanh. Network width was varied among 30, 50 and 70 neurons per hidden layer, whereas network depth was varied among 4, 6 and 8 hidden layers.
 553 Apart from the architecture quantity being tested, the remaining physical and numerical settings were kept unchanged.

554 Procedure for compensation representation

555 The present BcNN implementation considers two unknown parameters and two physical representation orders. In the controlled Sturm-Liouville
 556 benchmark, δ is the carrier coordinate and n is the compensation target. These coefficients are related to the underlying beam quantities EI and N through the
 557 transformation given below. In the structural benchmark, bending rigidity EI and prestress force N serve as the corresponding carrier and compensation target,
 558 respectively. The detailed transformation between the beam-column formulation and the Sturm-Liouville representation follows Yang et al. [40].

559 The coefficients used in the Sturm-Liouville formulation can be written as

$$\delta = \frac{FL^3}{EI}, \quad n = \frac{NL^2}{EI}, \quad (1)$$

560 where F is the applied lateral load and L is the span length. The derivation and analytical solution are given in Yang et al. [40] and are therefore not repeated
 561 here. The normalization references used in BcNN require a separate clarification. They provide physical scales for the anchor and target coordinates and
 562 should not be confused with the benchmark reference values used later to calculate parameter errors. This distinction is also made explicitly in Methods. For
 563 clarity, $\overline{EI}$ and $\overline{N}$ denote the normalization references below, whereas EI^{ref} and N^{ref} denote the physical reference values used only for post hoc
 564 evaluation. For the theoretical benchmark, the common normalization references were

$$\overline{EI} = 7.00 \times 10^8 \text{ N} \cdot \text{m}^2, \quad \overline{N} = 1.20 \times 10^6 \text{ N}, \quad L = 20 \text{ m}. \quad (2)$$

565 These values define a common physical scale for the theoretical calculations. The same normalization references were retained when the dominant parameter,
 566 the small parameter, the loading level and the number of observations were varied. For the structural experiments, the common normalization references
 567 were

$$\overline{EI} = 5.17213 \times 10^7 \text{ N} \cdot \text{m}^2, \quad \overline{N} = 8.20 \times 10^5 \text{ N}, \quad L = 6.62 \text{ m}. \quad (3)$$

568 The structural normalization references were selected from the physical values reported by Bonopera et al. [41]. The largest reported values of bending
 569 rigidity and prestress force were adopted as common reference scales. The same pair was then kept fixed for all structural experiments, while the case
 570 specific values were treated as variations relative to these common scales. The normalization references therefore provide a common coordinate scale rather
 571 than case specific information supplied to the inverse solver. They may coincide numerically with the reference values of one experimental condition, but
 572 they are not adjusted to match the individual cases. Choosing the largest reported values was a convenient normalization convention rather than a
 573 requirement of BcNN. The important point is that these normalization references provide approximate physical scale information; they do not specify the
 574 parameter values that should be recovered from a particular set of observations. The benchmark reference values used to calculate errors are introduced only
 575 after optimization. The reduced and enriched anchors are trained independently against the same observations. In both anchors, only the carrier parameter is
 576 trainable. The target mechanism is absent from the reduced representation and retained at its prescribed normalization reference in the enriched
 577 representation. The two anchors therefore probe the same carrier parameter under adjacent physical representations.

578 Following the notation established in Methods, the off-diagonal remained-bias coordinates obtained from the reduced and enriched anchors are denoted
 579 here in the shorter form $\hat{b}_L$ and $\hat{b}_H$, respectively. Once both anchors have been optimized, their estimates are frozen. Their ordered difference defines the
 580 physical gauge,

$$\hat{g} = \hat{b}_H - \hat{b}_L. \quad (4)$$

581 The ordering is always enriched minus reduced. The physical gauge therefore measures the representation dependent change in the carrier remained-bias
 582 coordinate. As defined in Methods, it provides a scale for the subsequent complementary bias reparameterization and should not be interpreted as an
 583 independently recovered physical parameter.

584 The target network is trained only after the two anchors have been completed and frozen. The frozen anchor estimates determine the intervals of the two
 585 target coordinates, μ and ν . For the coordinate associated with the carrier parameter,

$$1 - \hat{b}_H \leq \mu \leq 1 - \hat{b}_L. \quad (5)$$

586 For the coordinate associated with the compensation target,

$$0 \leq \nu \leq \hat{g}. \quad (6)$$

The corresponding underlying beam parameters are reconstructed as

$$\widehat{EI} = \overline{EI}(1 - \mu). \quad (7)$$

and

$$\hat{N} = \bar{N} \left(1 - \frac{\nu}{\hat{g}} \right). \quad (8)$$

The remained-bias and evolved-bias descriptions are complementary representations of the same normalized physical coordinates. They are therefore not optimized as two independent copies of the physical parameters. Once the anchors have been frozen, they remain unchanged throughout target training. No gradient from the target network is propagated back to either anchor. This separation prevents the target optimization from retrospectively modifying the pseudo-solutions used to construct the physical gauge.

Benchmark settings and observations

The controlled theoretical benchmark follows the analytical cases reported by Yang et al. [40], recast in the Sturm-Liouville hierarchy adopted in the present study. Nine baseline conditions, T1-T9, were considered. For T1-T3, the normalized bending rigidity was fixed at

$$\frac{EI}{\overline{EI}} = 1.0. \quad (9)$$

while the normalized prestress force was set to

$$\frac{N}{\bar{N}} = 0.6, 0.8, 0.9. \quad (10)$$

respectively. The applied lateral load was $F=100$ kN. For T4-T6, the normalized prestress force was fixed at

$$\frac{N}{\bar{N}} = 1.0. \quad (11)$$

while the normalized bending rigidity was set to

$$\frac{EI}{\overline{EI}} = 0.8, 0.9, 0.95. \quad (12)$$

respectively. The applied lateral load was again $F=100$ kN. For T7-T9, the physical state was fixed at

$$\frac{EI}{\overline{EI}} = 0.8, \quad \frac{N}{\bar{N}} = 0.6 \quad (13)$$

and the applied lateral load was 100, 200 and 300 kN, respectively. For the observation sparsity analysis, each analytical response was sampled using 2, 5, 10 and 20 uniformly distributed observation points over the half span. No artificial noise was added. The exact parameter values used to generate the analytical responses were retained only as references for evaluating the recovered parameters. They were not supplied to the BcNN optimization.

The response scale test used

$$\frac{EI}{\overline{EI}} = 0.8, \quad \frac{N}{\bar{N}} = 0.6, \quad (14)$$

with a nominal load of $F=100$ kN. Five loading amplitudes were considered: $0.01F$, $0.1F$, $0.5F$, F and $2F$. The purpose of this test was simply to vary the observable response scale while retaining the same underlying parameter relation.

The structural experiments use the sparse measurements reported by Bonopera et al. [41]. Consistent with Methods, the structural hierarchy comprises the enriched physical representation and its reduced Euler-Bernoulli counterpart. The same compensation construction is then applied to the two representations. Five response locations over the half span were supplied to the inverse calculation:

$$x = 0, 0.8275, 1.655, 2.4825, 3.31 \text{ m} \quad (15)$$

The value at $x=0$ represents the prescribed zero displacement at the support. The remaining four locations correspond to the experimental deflection measurements. For convenience, these five response inputs are referred to collectively as observations elsewhere in the manuscript. The same observations were supplied separately to BcNN and baseline PINN. The applied lateral loads for T1-T9 were 20.2, 22.6, 25.0, 20.1, 22.6, 25.1, 20.2, 22.9 and 25.1 kN, respectively. Their corresponding physical reference parameters were retained for post hoc evaluation and were not used to guide BcNN optimization. No additional random noise was added to the experimental responses. The aim here was to use the available measurements as reported rather than impose an artificial uncertainty model on top of them.

Engineering relevance and infrastructure scale

The identification of bending rigidity EI and prestress force N is relevant well beyond the beam experiments considered here. The two quantities describe different aspects of structural condition. Bending rigidity governs the relation between load, curvature and displacement, and changes in its effective value can reflect cracking, deterioration or changes in the effective load carrying section. Prestress force describes the internal precompression introduced into prestressed concrete members. Its residual value is important for both serviceability and structural capacity, yet it is difficult to determine reliably once a structure is in service. Field studies of existing prestressed concrete bridges have therefore treated residual prestress as an important quantity for structural assessment, and monitoring of prestress force remains a recognized challenge in structural health monitoring [42,43]. The practical difficulty is that changes in EI and N can influence the same measured structural response. In systems whose global behaviour is dominated by beam action, a change in stiffness can partly mask the comparatively weak response associated with prestress force, especially when only a few displacement measurements are available. Separating these effects is useful because they have different physical meanings and may point to different forms of structural deterioration or maintenance need. In this sense, the present benchmark is not simply a parameter fitting exercise for an idealized beam. It represents a broader assessment problem in which sparse measurements must be translated into physically meaningful quantities. Recent studies have likewise emphasized the need for scalable structural health monitoring using inexpensive or limited sensing, as well as practical methods for measuring bridge displacement in the field [44,45].

The structural scope considered here is deliberately narrower than the full bridge population. It is most directly relevant to prestressed concrete girders and composite girder deck systems whose global response can reasonably be represented by the beam hierarchy used in this study. The experiments of Bonopera et al. [41], for example, provide a controlled setting in which bending rigidity and prestress force jointly influence the measured deflection. The present

formulation is not intended to represent cable stayed or suspension bridges, arches, strongly torsional systems or structures dominated by distributed nonlinear damage without extending the physical model. This limitation matters because the response and residual capacity of a bridge after deterioration or component failure depend strongly on the structural form and on the resistance mechanisms that become available. Recent experimental testing and validated simulations have shown that secondary resistance mechanisms can substantially change how damaged bridge systems redistribute load after component failure [46].

Even within this restricted scope, the potential engineering population is substantial. A harmonized global database that classifies bridges by prestressing system and beam type is not available, so national inventories should not simply be added together. They nevertheless indicate the scale of the broader infrastructure in which beam and girder systems are widely used. China reported 1.1441 million highway bridges at the end of 2025, with a total length of about 108.7 million linear metres [47]. The 2025 US National Bridge Inventory contained 624,193 bridges, including 41,685 classified as being in poor condition [48]. Japan reports approximately 730,000 road bridges [49]. European Commission sources refer to approximately 1.1 million bridges in the European Union and estimate annual maintenance and repair expenditure at around €4–6 billion [50]. These inventories use different definitions and include many structural forms outside the present model, but together they show that bridge assessment is an infrastructure problem involving millions of structures rather than a collection of isolated laboratory cases.

This scale also explains why identification from sparse response measurements is of practical interest. Dense permanent monitoring remains uncommon even for globally important bridges. A recent global analysis of bridges with long spans reported that fewer than 20% of the bridges in the database had installed structural health monitoring systems [51]. The ability to infer physically interpretable quantities from a small number of measurements therefore addresses a different need from simply increasing sensor density. In the present setting, EI and N provide two such quantities: EI characterizes the bending rigidity of the member, whereas N describes the internal prestress force, whose effect can be comparatively weak in the measured response and can therefore remain hidden within a compensatory inverse solution.

Numerical values for the soft unilateral-boundary constraint defined by the anchors

The mathematical construction of the soft unilateral-boundary constraint is given in Methods. Here we provide the numerical settings used in the calculations. For a scalar coordinate z with interval $[z_{\min}, z_{\max}]$, the unilateral-boundary violation is

$$C_{\text{UB}} = \text{ReLU}(z_{\min} - z) + \text{ReLU}(z - z_{\max}). \quad (16)$$

The two violation terms associated with μ and ν are introduced into the normalized target residual as

$$\tilde{\mathcal{R}} = \mathcal{R} + \alpha C_{\mu} + \beta C_{\nu}. \quad (17)$$

The numerical coefficients used in the reported BcNN calculations were $\alpha=1$, $\beta=1$. The interval tolerance was set to zero. When both auxiliary coordinates remain inside the intervals defined by the anchors,

$$C_{\mu} = C_{\nu} = 0 \quad (18)$$

and the target residual reduces to the original governing equation residual. The constraint is intentionally soft. Once an auxiliary coordinate leaves its prescribed interval, the corresponding ReLU term modifies the governing residual. It does not project the coordinate back to the boundary or impose a hard feasibility condition. In this sense, the constraint favours the region defined by the anchors rather than forcing the optimization to remain there.

This interpretation is the same as that used in Methods: because the violation term is added to the signed residual before squaring, the construction acts as a residual bias and does not imply that every boundary violation must monotonically increase the physics loss. In the ablation without the soft unilateral-boundary constraint, the ReLU terms were removed while the anchor estimates, physical gauge and complementary bias reparameterization were retained.

Numerical values for training objectives and sequential optimization

For each component network, the objective contains the governing equation loss, displacement boundary loss, slope boundary loss and observation loss:

$$\mathcal{L} = \lambda_f \mathcal{L}_f + \lambda_{B,0} \mathcal{L}_{B,0} + \lambda_{B,1} \mathcal{L}_{B,1} + \lambda_{\text{data}} \mathcal{L}_{\text{data}}. \quad (19)$$

For the controlled theoretical benchmark, the loss weights were

$$\lambda_f = 10^{-4}, \quad \lambda_{B,0} = 1, \quad \lambda_{B,1} = 1, \quad \lambda_{\text{data}} = 1. \quad (20)$$

The same loss weights were used for the reduced anchor, enriched anchor and target network in the theoretical benchmark. Training was carried out sequentially. The reduced and enriched anchors were first optimized independently. Each theoretical anchor was trained with Adam for 2.0×10^5 epochs at a learning rate of 1.0×10^{-3} , followed by 1.0×10^5 epochs at 5×10^{-4} and 0.5×10^5 epochs at 10^{-4} . Each anchor therefore used 3.5×10^5 epochs in total. After both anchors had been optimized and frozen, the target network was initialized and trained separately. The theoretical target was optimized for 2.0×10^5 epochs at 10^{-3} , followed by 1.0×10^5 epochs at 5×10^{-4} and 1.0×10^5 epochs at 10^{-4} . Its total training horizon was therefore 4.0×10^5 epochs.

For the structural experiments, the BcNN loss weights were

$$\lambda_f = 10^{-4}, \quad \lambda_{B,0} = 10, \quad \lambda_{B,1} = 1, \quad \lambda_{\text{data}} = 10^{-3}. \quad (21)$$

Each structural anchor and target network was trained with Adam for 3.0×10^5 epochs at a learning rate of 10^{-3} , followed by 4.0×10^5 epochs at 5×10^{-4} and 5.0×10^5 epochs at 10^{-4} . The maximum training horizon of each stage was therefore 1.2×10^6 epochs. The reduced and enriched anchors were always completed before target optimization began. Their remained-bias estimates were then frozen and used to construct the physical gauge and target coordinate ranges. Neural field weights were not transferred from either anchor to the target network. The baseline PINN used the same observations and governing formulation but optimized the two inverse parameters simultaneously, without the anchor construction or complementary bias reparameterization.

For the theoretical benchmark, baseline PINN was optimized with Adam using a learning rate of 10^{-3} . In the analysis associated with Fig. 5, training was extended to as many as 3.0×10^6 epochs to examine whether prolonged optimization resolved the observed parameter instability. For the structural experiments, baseline PINN was trained for at most 1.2×10^6 epochs using Adam with a learning rate of 10^{-3} . The corresponding loss weights were

$$\lambda_f = 10^{-4}, \quad \lambda_{B,0} = 1, \quad \lambda_{B,1} = 1, \quad \lambda_{\text{data}} = 10^{-3}. \quad (22)$$

These numerical settings are reported to make the implementation transparent. The baseline comparison is intended to contrast direct dual-parameter inversion with the staged BcNN construction rather than imply that the two optimization paths are identical at every stage.

Evaluation of training horizons and epoch accounting

Reference parameter values were not used to terminate or guide the optimization. Training was carried out over the prescribed horizon, and the stored parameter trajectories were analysed afterwards. For the controlled benchmark, the principal BcNN target stage contained 4.0×10^5 epochs. Baseline PINN trajectories were additionally inspected at 1.0×10^6 , 2.0×10^6 and 3.0×10^6 epochs to determine whether continued direct optimization changed the parameter behaviour. For the structural experiments, both baseline PINN and the BcNN target stage were evaluated over a maximum horizon of 1.2×10^6 epochs. For BcNN, the target stage begins only after both anchors have been trained and frozen. The epoch coordinate reported for the target network therefore excludes anchor training. It describes optimization after construction of the physical gauge and should not be interpreted as the total computational cost of BcNN. This convention is also used in the main text. Parameter accuracy and trajectory behaviour were assessed after training from the saved parameter trajectories. No early stopping criterion based on parameter values was used. In other words, knowledge of the benchmark reference parameters did not alter the optimization path.

For a physical parameter p , the signed relative error is defined as

$$\varepsilon_p = \frac{\hat{p} - p^{\text{ref}}}{p^{\text{ref}}} \times 100\%. \quad (23)$$

For the theoretical benchmark, p^{ref} is the exact parameter value used to generate the analytical response. For the structural experiments, it is the corresponding reference value for that experiment. In both cases, these values are used only for evaluation. The relative L_2 response error is

$$\varepsilon_u = \frac{\|\hat{\mathbf{u}} - \mathbf{u}^{\text{ref}}\|_2}{\|\mathbf{u}^{\text{ref}}\|_2}. \quad (24)$$

For the theoretical benchmark, $\mathbf{u}^{\text{ref}}$ denotes the exact analytical response. For the structural experiments, the corresponding measured response is used.

Response reconstruction, parameter accuracy and trajectory stability were evaluated separately. This distinction is important in the present problem because the observable response can already be reproduced accurately while the physical parameters continue to move within an observationally admissible compensatory region. The complete parameter trajectory therefore provides information that cannot be inferred from the final response error alone.

Reproducibility and scope

The theoretical observations were generated deterministically from the analytical model reported in Yang et al. [40], with no synthetic observational noise. The structural calculations used the measurements reported by Bonopera et al. [41] without imposing an additional artificial noise model.

Within each inverse problem, the reduced and enriched anchors were trained using the same observations. What changes between them is the physical representation, not the data. After training, the two anchor estimates were frozen before optimization of the target network. The normalization references were also fixed before optimization. They provide coordinate scales for the anchor and target systems and remain distinct from the benchmark reference values used afterwards to quantify parameter errors. The present implementation is limited to two unknown physical parameters and two physical representation orders. The remained-bias notation introduced in Methods is written more generally, but systematic construction of compatible anchors and admissible regions for higher dimensional parameter spaces remains outside the scope of this study.

Finally, a nonzero physical gauge should not be interpreted as evidence of structural identifiability. Its role is more limited: it makes the complementary bias reparameterization well defined and provides the scale used by the target coordinates. BcNN does not increase the rank of the local parameter sensitivity Jacobian or create information that is absent from the observations. Instead, it uses the compensation exposed by the anchors across different physical representations to organize the inverse search when the weak physical contribution remains distinguishable from the available observations.

References

1. Tarantola, Albert. Inverse problem theory and methods for model parameter estimation. Society for industrial and applied mathematics, 2005.
2. Raissi, Maziar, Paris Perdikaris, and George E. Karniadakis. "Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations." *Journal of Computational physics* 378 (2019): 686-707.
3. Karniadakis, George Em, et al. "Physics-informed machine learning." *Nature Reviews Physics* 3.6 (2021): 422-440.
4. Jiao, Anran, et al. "Solving Forward and Inverse Partial Differential Equation Problems on Unknown Manifolds via Physics-Informed Neural Operators." *SIAM Journal on Scientific Computing* 48.1 (2026): C136-C163.
5. Bellman, R. & Åström, K. J. "On structural identifiability." *Mathematical biosciences* 7.3-4 (1970): 329-339.
6. Raue, Andreas, et al. "Structural and practical identifiability analysis of partially observed dynamical models by exploiting the profile likelihood." *Bioinformatics* 25.15 (2009): 1923-1929.
7. Wieland, Franz-Georg, et al. "On structural and practical identifiability." *Current opinion in systems biology* 25 (2021): 60-69.
8. Miao, Hongyu, et al. "On identifiability of nonlinear ODE models and applications in viral dynamics." *SIAM review* 53.1 (2011): 3-39.
9. Beven, Keith, and Jim Freer. "Equifinality, data assimilation, and uncertainty estimation in mechanistic modelling of complex environmental systems using the GLUE methodology." *Journal of hydrology* 249.1-4 (2001): 11-29.
10. Beven, Keith. "A manifesto for the equifinality thesis." *Journal of hydrology* 320.1-2 (2006): 18-36.
11. Gutenkunst, Ryan N., et al. "Universally sloppy parameter sensitivities in systems biology models." *PLoS computational biology* 3.10 (2007): e189.
12. Brown, Kevin S., and James P. Sethna. "Statistical mechanical approaches to models with many poorly known parameters." *Physical review E* 68.2 (2003): 021904.
13. Cavazzoni, Simone, Paolo Bordone, and Matteo GA Paris. "Quantum metrology of electric and magnetic dipole moments: ultimate limits and optimal regimes." *arXiv preprint arXiv:2606.25510* (2026).
14. Transtrum, Mark K., et al. "Perspective: Sloppiness and emergent theories in physics, biology, and beyond." *The Journal of chemical physics* 143.1 (2015).
15. Clark, Martyn P., and Jasper A. Vrugt. "Unraveling uncertainties in hydrologic model calibration: Addressing the problem of compensatory parameters." *Geophysical Research Letters* 33.6 (2006).
16. Daly, Aidan C., et al. "Inference-based assessment of parameter identifiability in nonlinear biological models." *Journal of The Royal Society Interface* 15.144 (2018): 20180318.
17. Kouros, Georgios, et al. "Unveiling the ambiguity in neural inverse rendering: A parameter compensation analysis." *2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*. IEEE, 2024.

18. Lyu, Linjie, et al. "Diffusion posterior illumination for ambiguity-aware inverse rendering." *ACM Transactions on Graphics (ToG)* 42.6 (2023): 1-14.
19. Ost, Julian, et al. "Towards generalizable and interpretable three-dimensional tracking with inverse neural rendering." *Nature Machine Intelligence* 7.8 (2025): 1322-1330.
20. Hu, Qian, et al. "Target-Guided Selective Reweighting for Physics-Informed Neural Network Inverse Problems: A Transfer Learning Approach." *arXiv preprint arXiv:2607.05271* (2026).
21. Rao, Chengping, et al. "Encoding physics to learn reaction – diffusion processes." *Nature Machine Intelligence* 5.7 (2023): 765-779.
22. Cao, Qianying, Somdatta Goswami, and George Em Karniadakis. "Laplace neural operator for solving differential equations." *Nature Machine Intelligence* 6.6 (2024): 631-640.
23. Berner, J., Liu-Schiaffini, M., Kossaifi, J. et al. Principled approaches for extending neural architectures to function spaces for operator learning. *Nat. Mach. Intell.* 8, 1173 – 1181 (2026).
24. Brunton, S. L. & Kutz, J. N. Promising directions of machine learning for partial differential equations. *Nat. Comput. Sci.* 4, 483 – 494 (2024).
25. Peng, M. & Tang, H. Information-distilled physics informed deep learning for high order differential inverse problems with extreme discontinuities. *Commun. Eng.* 4, 150 (2025).
26. Liu, Z. et al. Automatic network structure discovery of physics informed neural networks via knowledge distillation. *Nat. Commun.* 16, 9558 (2025).
27. Aarts, G. et al. Physics-driven learning for inverse problems in quantum chromodynamics. *Nat. Rev. Phys.* 7, 154 – 163 (2025).
28. Salmani, Y. & Browning, A. P. Structural identifiability of linear-in-parameter parabolic PDEs through auxiliary elliptic operators. *J. Math. Biol.* 91, 4 (2025).
29. Wu, Z., Li, X., Wan, J.-H. & Xu, T. A physics-informed machine learning framework for inverse multi-solution problems. *Appl. Math. Model.* 152, 116618 (2026).
30. Srikitrungruang, Tatthapong, et al. "Robust Physics - Informed Neural Network Approach for Estimating Heterogeneous Elastic Properties from Noisy Displacement Data." *Advanced Science* 12.48 (2025): e08445.
31. Ni, F., Dai, Y., Zheng, Q. & Jia, W. Physics-informed machine learning for robust inverse problem solving in maglev train levitation systems under noisy measurement. *Nonlinear Dyn.* 113, 26981 – 26999 (2025).
32. Mahar, N., Sen, S. & Mevel, L. A parallel framework of physics-informed neural networks for model identification of linear and nonlinear systems. *Structures* 79, 109454 (2025).
33. Mahar, Nikhil, Subhamoy Sen, and Laurent Mevel. "Implicit Runge Kutta physics informed neural network for parameter identification of structural systems." *Computers & Structures* 323 (2026): 108157.
34. Shan, S., Zhu, M., Lin, Y. et al. Regularization by denoising diffusion models for solving inverse PDE problems with application to full waveform inversion. *Commun. Phys.* 9, 242 (2026).
35. Zhang, Yi, et al. "Inverse identification of multiple constitutive parameters of composites utilizing physics-informed neural networks." *Composite Structures* (2026): 120318.
36. Tang, Hong, et al. "Physics informed neural networks for solving inverse thermal wave coupled boundary-value problems." *International Journal of Heat and Mass Transfer* 245 (2025): 126985.
37. Chen, Nanxi, et al. "Enforcing hidden physics in physics-informed neural networks." *Communications Physics* (2026).
38. McGreivy, Nick, and Ammar Hakim. "Weak baselines and reporting biases lead to overoptimism in machine learning for fluid-related partial differential equations." *Nature machine intelligence* 6.10 (2024): 1256-1269.
39. Fabiani, G., Vandecasteele, H., Goswami, S. et al. Enabling local neural operators to perform equation-free system-level analysis. *Nat. Mach. Intell.* 8, 1127 – 1141 (2026).
40. Yang, Yifan, Zengwei Guo, and Zhiyuan Liu. "Parameter identification in prestressed concrete beams by incremental beam – column equation and physics - informed neural networks." *Computer - Aided Civil and Infrastructure Engineering* 40.19 (2025): 2876-2899.
41. Bonopera, Marco, et al. "Feasibility study of prestress force prediction for concrete beams using second-order deflections." *International Journal of Structural Stability and Dynamics* 18.10 (2018): 1850124.
42. Bagge, Niklas, Jonny Nilimaa, and Lennart Elfgren. "In-situ methods to determine residual prestress forces in concrete bridges." *Engineering Structures* 135 (2017): 41-52.
43. Abdel-Jaber, Hiba, and Branko Glisic. "Monitoring of prestressing forces in prestressed concrete structures—An overview." *Structural Control and Health Monitoring* 26.8 (2019): e2374.
44. Li, X. et al. Mechanics-informed autoencoder enables automated detection and localization of unforeseen structural damage. *Nat. Commun.* 15, 9229 (2024).
45. Ri, S., Ye, J., Toyama, N. & Ogura, N. Drone-based displacement measurement of infrastructures utilizing phase information. *Nat. Commun.* 15, 395 (2024).
46. Reyes-Suárez, J. C. et al. Latent resistance mechanisms of steel truss bridges after critical failures. *Nature* 645, 101–107 (2025).
47. Ministry of Transport of the People's Republic of China. Statistical Bulletin on Transport Industry Development 2025 (Ministry of Transport, 2026).
48. Federal Highway Administration. Bridge Condition by Highway System 2025, National Bridge Inventory (U.S. Department of Transportation, 2025).
49. Ministry of Land, Infrastructure, Transport and Tourism. Roads in Japan 2026 (MLIT, Tokyo, 2026).
50. European Commission. Buildings, tunnels and bridges could soon repair themselves. *Horizon: The EU Research & Innovation Magazine* (2021).
51. Malinowska, D. et al. Global geo-hazard risk assessment of long-span bridges enhanced with InSAR availability. *Nat. Commun.* 16, 9048 (2025).