FROST: LEARNING FREE-BOUNDARY EQUILIBRIA FOR SIMULATION AND INVERSE DESIGN

Anonymous authors
Paper under double-blind review

ABSTRACT

Free-boundary problems are partial differential equations (PDEs) whose domain evolves and is not known in advance; they govern solidification, growth and transport across the physical sciences. Classical sharp-interface solvers remesh as the boundary moves, lose robustness when domains merge or split, and are recomputed from scratch for every new instance, while existing neural operators act on a fixed, known domain or track the boundary by a single smooth deformation and so cannot represent changes in topology. We introduce FROST, a free-boundary equilibrium operator that represents the solution as a joint fixed point of the physical field and a level set on one fixed background grid, in which the physical free-boundary condition is the fixed-point residual. Because the geometry is a level set rather than a deformation, a single operator represents merging, splitting and nucleation without remeshing. In its equilibrium realisation the fixed point is solved and trained by implicit differentiation in the manner of a deep equilibrium model, so training memory is constant in the number of solver steps and convergence is monitored; a conditioned realisation of the same operator predicts time-dependent and parametric families in a single pass. The trained operator is then differentiated with respect to a design variable, so forward simulation and inverse design reduce to one problem and the physics residual serves as a trust signal for the design. Across steady and time-dependent free-boundary benchmarks FROST recovers boundaries and the changes in topology that a single deformation cannot represent, learns from only a handful of simulations, and steers a merging front to a prescribed target.

1 INTRODUCTION

Free-boundary problems (FBPs) are among the most demanding PDEs in computational science: the domain on which they are posed is not known in advance, and it changes as the solution evolves. They describe solidification and melting (Stefan, 1891; Chen, 2008), the growth of biological tissue (De Angelis & Preziosi, 2000; Droniou et al., 2020), and transport through reacting or porous media, among many other moving-interface phenomena (Friedman & Spruck, 2012; Hanzawa, 1981). The difficulty is that the interface where the boundary conditions apply must be found together with the solution, and this interface can move, merge, split or nucleate during the evolution. Classical methods either track the interface with a moving mesh (Baines et al., 2005; Peskin, 2002) or reformulate the problem on a fixed domain through enthalpy or phase-field variables (Chen, 2008); either way the solve is accurate but tied to a single instance, so a family of solutions, as needed for design or control, becomes prohibitively expensive.

Operator learning offers a route around repeated solves: a neural operator learns a map between function spaces (Chen & Chen, 1995; Lu et al., 2021b; Li et al., 2021b; Kovachki et al., 2023), so one trained model predicts the solution for many inputs at a fraction of a solver's cost. Yet operator learning rests on an assumption that FBPs violate. Existing neural operators map functions on a fixed, known domain (Li et al., 2021b; Lu et al., 2021b), and even geometry-aware variants assume a domain of fixed topology (Li et al., 2023a;b; Yin et al., 2024), so they cannot represent a domain that is itself unknown. One line of work removes this restriction with a single smooth, invertible deformation of a fixed reference shape (Long et al., 2026); but a diffeomorphism preserves connectedness (Furuya et al., 2023), so it follows the interface only while the topology stays fixed and cannot describe merging, splitting or nucleation. A second line trains a physics-informed network for each

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

instance (Raissi et al., 2019; Wang & Perdikaris, 2021; Zhu et al., 2024), which returns one solution rather than a reusable operator; such surrogates are also data-hungry (Jiao et al., 2025; Boullé et al., 2023), and inverse design is typically handled by separate search rather than by differentiating the operator across a change of topology (Kou et al., 2025; Lu et al., 2021a).

Here we introduce FROST, a free-boundary residual operator that is self-consistent in topology. The operator places the physical field and a level set on one fixed background grid, and represents the solution as their joint fixed point, the interface given by the zero level set $\{\varphi = 0\}$, in which the physical free-boundary condition is the fixed-point residual (Fig. 1). Because the geometry is a level set rather than a deformation, merging, splitting and nucleation appear on the grid without remeshing and without any prior description of the domain. We solve and train the fixed point by implicit differentiation, as in a deep equilibrium model (DEQ) (Bai et al., 2019; Marwah et al., 2023), so inference is amortised, training memory is constant in the number of solver steps, and convergence is monitored. Because the operator is differentiable, the same equilibrium is differentiated with respect to a design variable, so forward simulation and inverse design reduce to one problem, and the physics residual, readable at inference without a reference solution, marks where the operator can be trusted.

Contributions.

- **A free-boundary operator on a fixed grid.** We represent the solution of a free-boundary PDE as a joint fixed point of a physical field u and a level set φ on one fixed background grid, so that topology change (merging, splitting, nucleation) is intrinsic to the representation and requires no remeshing (Section 3).
- **The free-boundary condition as a fixed-point residual and trust signal.** The residual that the equilibrium solver drives to zero *is* the physical free-boundary condition; read at inference without a reference solution, it serves as a trust signal that marks where the operator has reached a physical equilibrium.
- **One differentiable object for simulation and design.** The same differentiation that trains the operator returns its design gradient, so forward simulation and inverse design are the same computation: in the equilibrium realisation it is a single implicit-function-theorem adjoint (one linear solve), and in the conditioned realisation it reduces to ordinary backpropagation.
- **Guarantees and efficiency.** We give a local-contraction theorem for the equilibrium (Section 4), and show empirically that FROST recovers boundaries and their changing topology from only a handful of simulations where prior operators need thousands.

2 Related work

Neural operators and geometry. Operator learning maps between function spaces with DeepONets (Lu et al., 2021b), Fourier neural operators (FNOs) (Li et al., 2021b; Kovachki et al., 2023) and their geometry-aware descendants, which handle varying but fixed-topology shapes by meshing per instance or by a signed-distance input (Li et al., 2023a;b; Yin et al., 2024; Wu et al., 2024). FROST differs in that the geometry is not a fixed input but is *co-solved* as part of the fixed point, and its topology is free to change.

Free-boundary operators. The closest prior work learns reusable operators on evolving domains by mapping each instance to a fixed reference domain through a diffeomorphism or homeomorphism (Long et al., 2026). Such a map preserves the number of connected components and so cannot represent a boundary that merges or splits; FROST's level-set state removes this limitation and, unlike these methods, differentiates the operator itself for inverse design.

Deep equilibrium and implicit operators. A DEQ writes a network's output as a fixed point and differentiates through it with the implicit function theorem (Bai et al., 2019); this has been used to solve steady PDEs as the fixed point of a weight-tied FNO (Marwah et al., 2023). FROST generalises this from a single field on a fixed domain to a *joint* field-and-geometry equilibrium on free domains, and reuses the same adjoint for design.

Level sets, topology optimisation and residual signals. Level-set methods represent moving interfaces implicitly and admit topology change in classical simulation (Osher & Sethian, 1988), and differentiable level sets drive structural topology optimisation (Bendsoe & Sigmund, 2013; Li et al., 2021a; Kou et al., 2025), but as an optimisation over a conventional solver rather than a learned operator over a family. Physics-informed operators add the PDE residual as a training signal (Karniadakis et al., 2021; Li et al., 2024), and sensitivity- or design-oriented operators use differentiability for inversion (Behroozi et al., 2025); FROST reinterprets the residual as the *fixed-point* residual and elevates it to a runtime trust signal.

3 THE FREE-BOUNDARY EQUILIBRIUM OPERATOR

State on a fixed discrete grid. We pose every problem on a fixed background box $D \subset \mathbb{R}^d$, discretised once and for all by a fixed Cartesian grid of $N = m \times n$ nodes that never moves. The physical field u lives on a domain $\Omega_t \subseteq D$ that is itself unknown and varies with a time or pseudo-time t, with moving interface $\Gamma_t = \partial\Omega_t$, and we encode the geometry by a signed-distance level set φ sampled on the same grid (Fig. 1a):

$$\Omega_t = \{x : \varphi(x,t) < 0\}, \qquad \Gamma_t = \{x : \varphi(x,t) = 0\}, \qquad V_n = -\frac{\varphi_t}{|\nabla\varphi|}, \tag{1}$$

with V_n the normal interface velocity. The state is the pair (u, φ), held as grid functions: throughout, $u, \varphi \in \mathbb{R}^{m\times n} \cong \mathbb{R}^N$ are arrays of nodal values, and the operator, its fixed point and every norm below act on this discrete state, while continuous relations such as (1) are read at the grid nodes. Because Ω_t is a sublevel set of φ on a fixed grid, its components merge and split as φ changes sign, with no remeshing and no map between domains, so the topology of the free boundary is free to change.

The joint fixed point. FROST represents the solution as a joint fixed point of a learned map (Fig. 1b). Given the problem data e (the initial geometry, the physical parameters and any forcing), the operator $\mathcal{G}_\theta$ returns $(u^\star, \varphi^\star) = \mathcal{G}_\theta(e)$ with

$$(u^\star, \varphi^\star) = T_\theta(u^\star, \varphi^\star;\, e)\,, \tag{2}$$

where T_θ couples a field update (enforcing the interior equation given the current geometry) with a level-set update (advancing φ by the interface velocity read from the field). At the fixed point the interior equation holds in $\{\varphi^\star < 0\}$, the boundary conditions hold on the grid, and the physical free-boundary condition

$$\mathcal{B}(u^\star, \nabla u^\star, V_n) = 0 \qquad \text{on } \Gamma^\star \tag{3}$$

holds on the interface $\Gamma^\star = \{\varphi^\star = 0\}$. Here $\mathcal{B}$ is the free-boundary condition of the specific problem, the trace or flux law the solution must satisfy on the moving interface (for the Stefan problem the latent-heat balance $L\,V_n = -\partial_n T$, with ∂_n the outward normal derivative on the interface; for the obstacle problem the contact complementarity); its concrete forms are in Appendix C. Equation (3) is the residual of the fixed point: the free-boundary condition is what the equilibrium drives to zero, so the geometry is found together with the field rather than supplied in advance.

Solving and differentiating the equilibrium. The equilibrium is produced in either of two ways. In the *equilibrium realisation*, an explicit solve treats (2) as a fixed-point problem and reaches it with an accelerated (Anderson) iteration; the backward pass obtains gradients from one linear solve with $I - \partial T_\theta/\partial(u,\varphi)$ at the fixed point, so training memory does not grow with the number of iterations. Writing $z = (u, \varphi)$ and $J_z = \partial_z T_\theta(z^\star; e)$ for the Jacobian of the state map at the fixed point, differentiating $z^\star = T_\theta(z^\star; e)$ gives the implicit-function-theorem (IFT) sensitivity

$$\big(I - J_z\big)\,\frac{\mathrm{d}z^\star}{\mathrm{d}e} = \frac{\partial T_\theta}{\partial e}, \tag{4}$$

a single linear solve that supplies the gradient for both training and inverse design; the concrete latent-cell realisation f_θ and the Anderson solver are given in Appendix A. In the *conditioned realisation*, used for the time-dependent and parametric families, a fixed stack of blocks is applied once and conditioned on the query time or the design; it involves no fixed-point solve at inference, and its design gradient is obtained by ordinary back-propagation. To keep the interface sharp, φ is

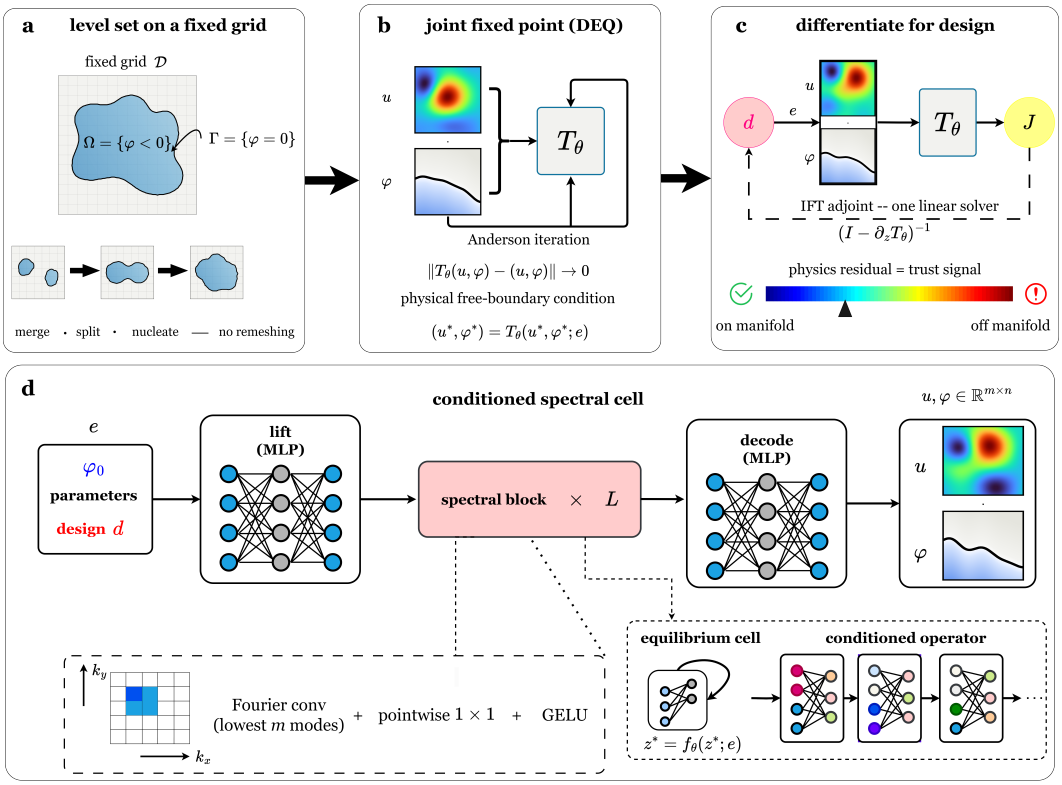

Figure 1: **The free-boundary equilibrium operator.** **(a)** A free boundary on a fixed background grid: the domain is the sublevel set $\Omega = \{\varphi < 0\}$ and the interface is $\Gamma = \{\varphi = 0\}$, so separate regions merge and split without remeshing. **(b)** The field u and the level set φ are a joint fixed point, with the physical free-boundary condition as the fixed-point residual r, solved as a deep equilibrium model. **(c)** The same equilibrium is differentiated for inverse design through the implicit function theorem, with the physics residual as a trust signal. **(d)** Network architecture: a conditioned spectral cell whose discrete output $u, \varphi \in \mathbb{R}^{m \times n}$ is trained to a relative-L^2 loss $\mathcal{L}$, distinct from the solve residual r.

periodically re-initialised to the signed distance to Γ (Appendix A). Two consequences frame the rest of the paper: the physical residual (3), evaluated on the prediction, is a trust signal readable without a reference solution (Appendix A distinguishes it from the solve residual of the equilibrium iteration and from the exact certificate available where a cheap solver exists), and, because field and geometry form one differentiable equilibrium, a design variable enters through e and the *same* fixed point is differentiated for design (Fig. 1c).

4 Convergence of the Equilibrium

We state the conditions under which the fixed point of (2) exists, is unique, and is reached by the forward iteration, and under which the implicit gradient used in training and design is well defined. Write $z = (u, \varphi) \in \mathbb{R}^{2N}$ for the discrete state and $\|\cdot\|$ for the normalised grid ℓ^2 norm, $\|z\|^2 = \frac{1}{2N} \sum_{i=1}^{2N} z_i^2$.

Assumption 1 (Local contraction)**.** *In the equilibrium realisation, for each instance e there is a non-empty closed set $\mathcal{N}_e$ and a constant $L \in [0, 1)$ such that $T_\theta(\cdot\,; e)$ maps $\mathcal{N}_e$ into itself and $\|T_\theta(z; e) - T_\theta(z'; e)\| \le L \|z - z'\|$ for all $z, z' \in \mathcal{N}_e$.*

Theorem 1 (Convergence and differentiability)**.** *Under Assumption 1, for each e the map $T_\theta(\cdot\,; e)$ has a unique fixed point $z^\star(e) \in \mathcal{N}_e$, and the iteration $z_{k+1} = T_\theta(z_k; e)$ from any $z_0 \in \mathcal{N}_e$ converges to it with $\|z_k - z^\star(e)\| \le L^k \|z_0 - z^\star(e)\|$. If in addition T_θ is continuously differentiable in (z, e),*

then $I - \partial_z T_\theta(z^\star(e); e)$ is invertible and $z^\star$ is continuously differentiable in e, with derivative given by (4).

We prove Theorem 1 in Appendix B. The geometric bound is the residual decay the solver shows in practice and sets the stopping tolerance; the invertibility of $I - \partial_z T_\theta$ is the same condition that makes the implicit gradient well posed, so one assumption supports both the forward solve and the backward pass. The contraction is what training *through* the fixed point promotes: a map fit only to reproduce a single iteration can be accurate at every step yet admit a fixed point away from the true solution, so we differentiate the converged equilibrium rather than one application of T_θ (Appendix E).

5 Experiments

We evaluate FROST first where the free boundary is steady, then where it changes topology in time, and finally differentiate the trained operator for inverse design. Metrics are the relative L^2 error of each field, the contact intersection-over-union (IoU) of the predicted and reference sublevel sets $\{\varphi < 0\}$, and the topology accuracy (the fraction of held-out instances or frames whose predicted number of connected components matches the reference, components counted after removing spurious sub-grid-scale islands). Datasets, architectures and training are in Appendix C.

5.1 Steady free boundaries and contact topology

The obstacle problem is a steady, elliptic FBP: a field u on $D = (-1, 1)^2$ is constrained to lie above an obstacle χ and is harmonic where it does not touch it, $u \geq \chi,\ -\Delta u \geq 0,\ (u - \chi)(-\Delta u) = 0$ in D, $u = 0$ on ∂D. Here the problem data is the obstacle itself, $e = \chi$. The contact set $\Omega_c = \{u = \chi\}$ has one or two connected components depending on χ, so a single family already spans a change in the topology of $\Gamma = \partial\Omega_c$.

On a held-out set containing both topologies, the direct realisation recovers the field to 0.91% and the level set to 0.85% in relative L^2, and the operator assigns the correct number of contact components on every instance (Fig. 2a,d). Reading the contact set from the explicit level set raises the contact IoU above thresholding the predicted field, most of all on the merged one-component cases (Fig. 2b). *Ablation.* The equilibrium (DEQ) realisation reaches accuracy within about twice the direct map's with one-seventh of the parameters and constant-memory training, and its fixed-point solve converges in fewer than thirty iterations; a locally one-step-accurate map instead admits a spurious fixed point that the iterated solution drifts to, which training through the equilibrium removes (Appendix E). A single smooth deformation of one reference cannot reproduce this behaviour: fitting, per held-out instance, the best fold-free diffeomorphic warp of a fixed reference (smallest relative Jacobian determinant $+0.04$, so the limit is structural) matches every instance of its own topology class and none of the other, while the equilibrium operator is correct on both (Fig. 2c).

5.2 Topology change in time

The tumour-growth problem is a time-dependent FBP: a nutrient field u on a growing domain Ω_t solves a screened-Poisson balance with linear uptake at rate k, the front advancing with the local nutrient flux (Appendix C). Two initially separate tumours grow toward each other and coalesce, so Γ_t changes from two components to one within a single trajectory. The operator maps the initial geometry, the rate k and the time t to $u(t)$ and the level set $\varphi(t)$.

Across held-out trajectories the operator tracks the moving boundary and its topology: the level set is within 1.1% in relative L^2, the tumour region has an IoU of 0.97, and the predicted boundary carries the correct number of components on 98% of frames, with errors confined to the coalescence instant (Fig. 3a,b). The nutrient field is a near-binary indicator with a sharp boundary rim, so its raw relative error is larger, about 12%, set by a Gibbs-type overshoot of the truncated spectral representation at the sharp boundary rim. A time-continuous deformation cannot cross this coalescence, because it is a homeomorphism at every instant: the best fold-free warp of the initial two-component domain stays at two components throughout and is wrong on every frame after the merge, at a trajectory topology accuracy of 73%, while the operator follows the two-to-one transition (Fig. 3c,d).

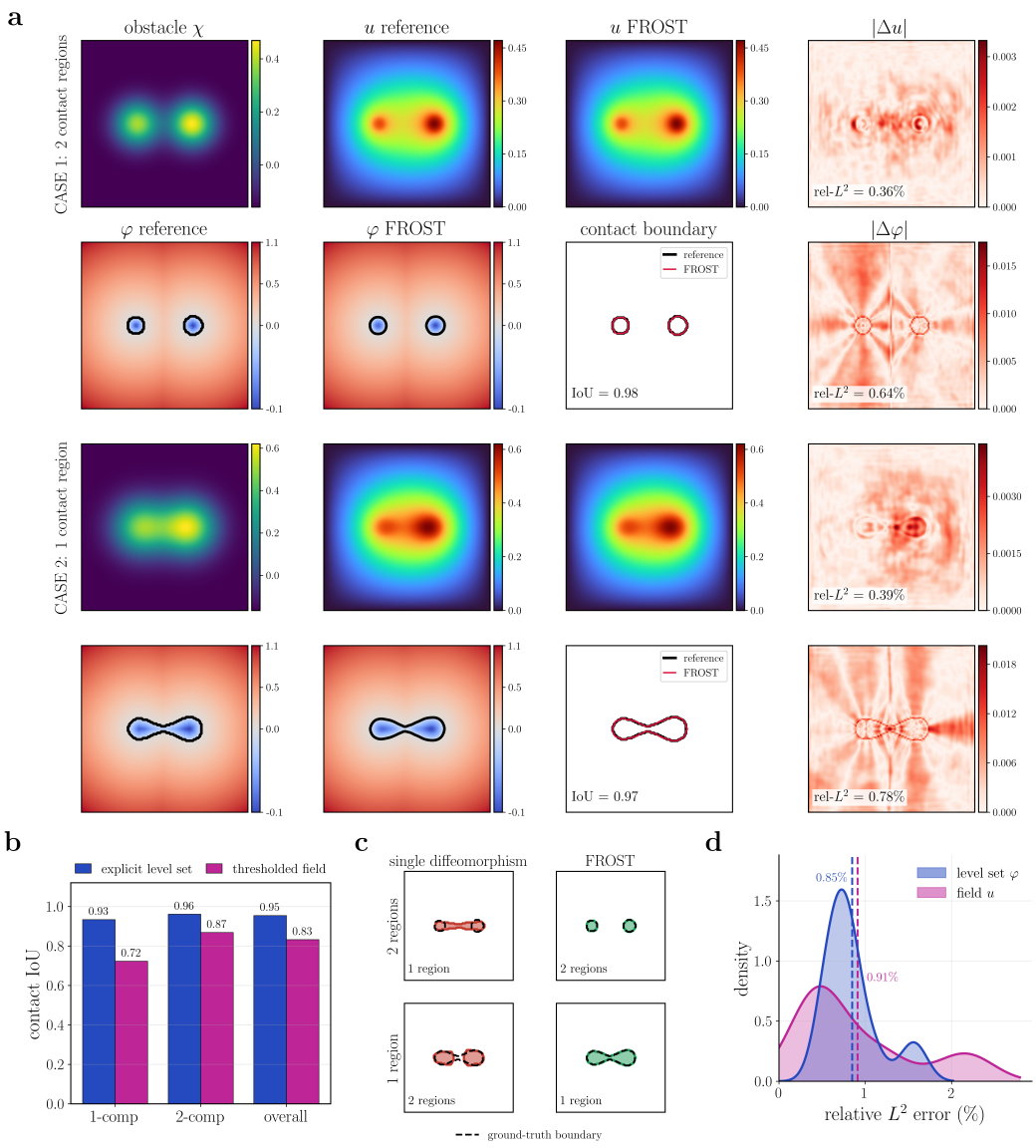

Figure 2: **Steady free boundaries and contact topology on the obstacle problem.** **(a)** A two-component and a one-component held-out instance (field row and level-set row; relative L^2 and contact IoU boxed). **(b)** Contact IoU of the explicit level-set head against a thresholded field, by topology class. **(c)** The single-deformation barrier: an oracle diffeomorphic warp cannot change the number of components, while the equilibrium operator recovers both. **(d)** Distribution of field and level-set errors over the held-out set.

5.3 Inverse design of a topology-changing front

Because the operator is differentiable, forward simulation and inverse design are the same computation: the trained operator is held fixed and its output is differentiated with respect to a design variable, here through the conditioned realisation by ordinary back-propagation (the equilibrium counterpart, differentiated through the IFT adjoint and certified against the exact solver, is demonstrated on the obstacle design task in Appendix D). We demonstrate this on the hardest case: controlling a free boundary that changes topology. The setting is solidification, a one-phase Stefan problem, in which a liquid cooled below freezing nucleates solid seeds whose interface Γ_t grows and merges over time (Appendix C). A time-conditioned operator maps the seed layout φ_0, the latent heat L, the cold-boundary temperature T_{cold} and the time t to $(T(t), \varphi(t))$. Holding it fixed, we seek the con-

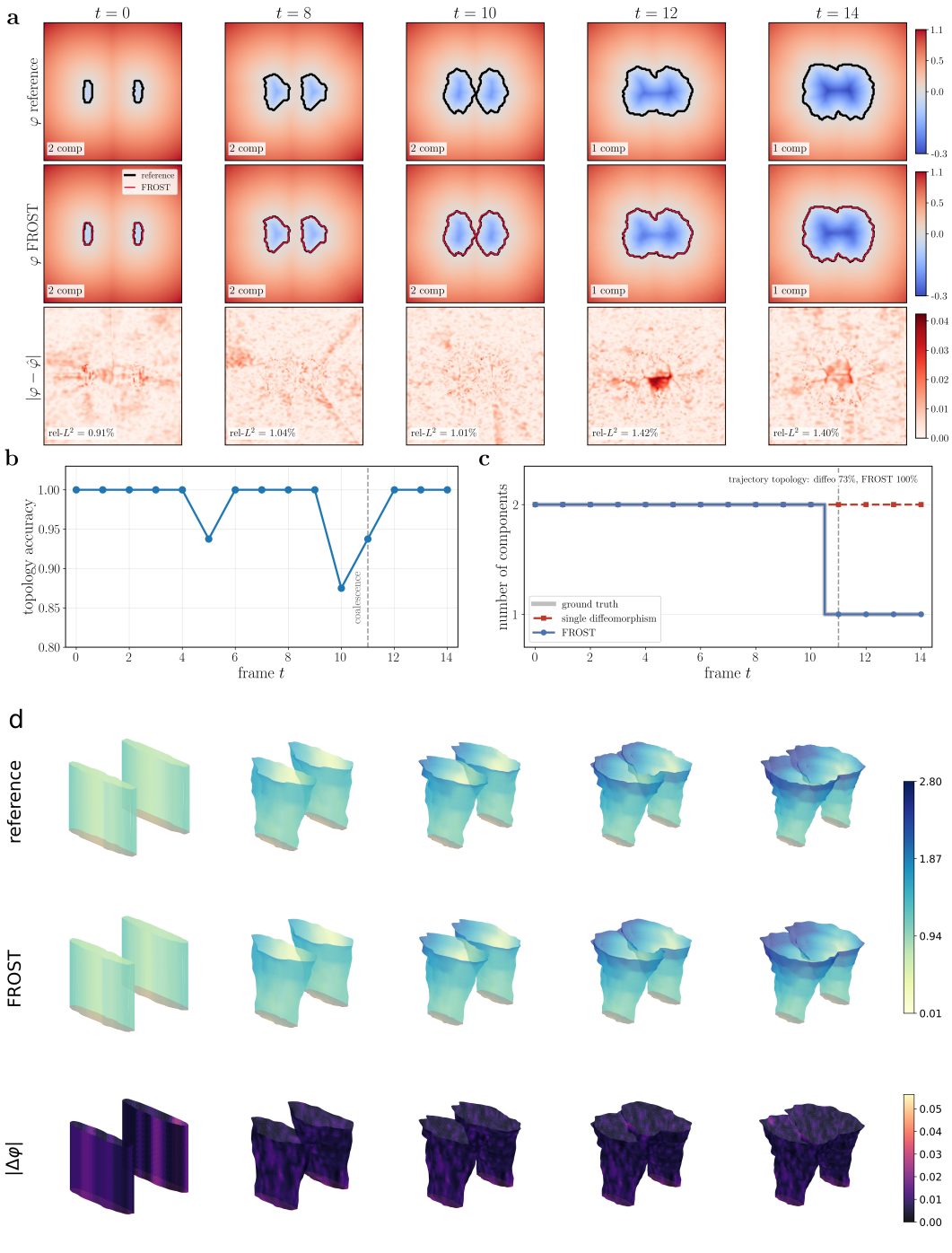

Figure 3: **Topology change in time on the tumour-growth problem.** **(a)** Space-time evolution of a held-out trajectory through the coalescence (reference and predicted level set, error, number of components boxed). **(b)** Topology accuracy through time, dipping only at the coalescence instant. **(c)** The time-continuous-deformation barrier. **(d)** Space-time surfaces of the two bodies fusing.

trols driving the final solid region to a target shape $M^\star$, $\min_{\varphi_0, T_{\text{cold}}} \|H_\tau(-\varphi(\cdot, t{=}1)) - M^\star\|_2^2$, by gradient descent through the operator (H_τ a smooth indicator).

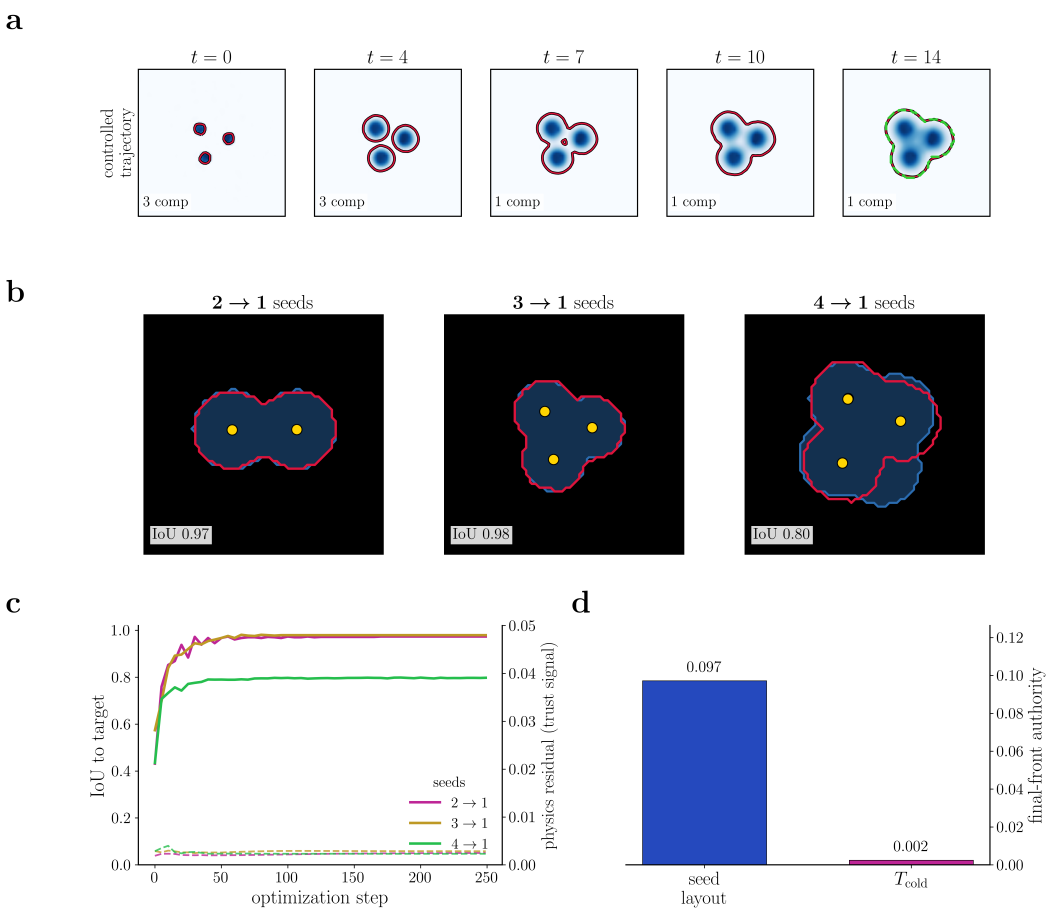

Figure 4: **Inverse design: controlling a topology-changing solidification front.** **(a)** Controlled trajectory: the optimised seed layout grows and merges three seeds into one to reach the target (green dashed); number of components boxed. **(b)** Diverse targets merging two, three and four seeds into one (contact IoU boxed). **(c)** Optimisation: the contact IoU rises while the physics-consistency residual stays low. **(d)** Control authority: the seed layout dominates T_{cold}.

The optimised controls steer the front through a change in topology: separate seeds grow and coalesce into a single grain that matches the target (Fig. 4a); a single-deformation operator cannot represent this merge, so the design problem is only well posed for a topology-capable operator. Across targets that merge two, three and four seeds into one, the controlled final front reaches the target at a contact IoU of 0.97, 0.98 and 0.80, with the correct topology in every case (Fig. 4b), and the optimisation converges within a few tens of steps while the physics-consistency residual (the solid-cold/liquid-warm consistency check of Appendix A) stays near 0.003 (Fig. 4c): the proposed design is consistent with the physics the operator learned, read from the same forward pass. Sweeping each control alone, the seed layout carries about $40\times$ the authority of T_{cold} over the final extent (Fig. 4d). Forward simulation and inverse design are thus one differentiation of one trained operator, and the physics-consistency residual, read from the same forward pass that produces the design, doubles as a trust signal.

6 Conclusion

FROST represents the solution of a free-boundary problem as a joint fixed point of the physical field and a level set on a fixed grid, with the physical free-boundary condition enforced as the residual of that fixed point. Because the geometry is a level set rather than a smooth deformation, a single operator represents interfaces that merge, split and nucleate without remeshing, and a local-contraction

argument shows the equilibrium realisation to be well posed. Across steady and time-dependent benchmarks the operator recovers free boundaries and their changing topology from only a handful of simulations, and the same operator, differentiated with respect to a design variable, steers a moving boundary to a prescribed target, so simulation and design become one computation. *Limitations.* Accuracy is set by the sharpness of the field: a near-discontinuous field carries a Gibbs-type error floor from resolving a sharp interface with truncated Fourier modes on a fixed grid, which a locally adapted representation could lower. The physics residual is a necessary rather than sufficient signal. The equilibrium realisation is exercised on the steady benchmark, while the time-dependent families use the conditioned realisation; the topology changes demonstrated empirically are merging events, with splitting and nucleation shown as representational capacity rather than learned behaviour; and the Stefan designs are evaluated through the operator, with solver certification demonstrated for the obstacle design. Finally, the contraction guarantee is local, with the equilibrium solved here in two dimensions for one problem family at a time. Extending the construction to three dimensions, coupled multi-physics equilibria, and a convergence theory for time-dependent fixed points would carry the same single differentiable object to the broad class of moving-boundary problems that pervade the physical sciences.

Reproducibility statement

All datasets are generated by the procedures in Appendix C (with sampling distributions, grids, splits and instance counts in Table 1); the operator, the accelerated forward solve and the implicit-differentiation backward pass are specified in Appendix A and Algorithms 1–2, with architectures and hyperparameters in Table 2; the proof of Theorem 1 is in Appendix B. Source code implementing the operator, its accelerated forward solver and implicit-differentiation backward pass, and the data generators is available at the anonymised repository `https://anonymous.4open.science/r/FROST-free-boundary-residual-operator-DBEC/`.

Ethics statement

This work develops a numerical operator-learning method evaluated on simulated physical benchmarks (an obstacle problem, tumour-growth and solidification models). It involves no human subjects, personal data, or sensitive information, and we foresee no direct ethical concerns; as with any surrogate for physical design, deployment should retain the physics-residual trust signal and a fallback to a conventional solver before acting on a proposed design.

AI use statement

A generative AI assistant (Claude Fable 5) was used only for minor language editing and for assistance with code and LaTeX formatting. It was not used to generate research ideas, experimental results, or scientific claims. All AI-assisted text and code were reviewed and verified by the authors, who take responsibility for the final content of this work.

References

Shaojie Bai, J Zico Kolter, and Vladlen Koltun. Deep equilibrium models. *Advances in neural information processing systems*, 32, 2019.

M. J. Baines, M. E. Hubbard, and P. K. Jimack. A moving mesh finite element algorithm for the adaptive solution of time-dependent partial differential equations with moving boundaries. *Applied Numerical Mathematics*, 54:450–469, 2005.

Abdolmehdi Behroozi et al. Sensitivity-constrained Fourier neural operators for forward and inverse problems in parametric differential equations. In *International Conference on Learning Representations (ICLR)*, 2025.

Martin Philip Bendsoe and Ole Sigmund. *Topology optimization: theory, methods, and applications*. Springer Science & Business Media, 2013.

N. Boullé, D. Halikias, and A. Townsend. Elliptic PDE learning is provably data-efficient. *Proceedings of the National Academy of Sciences*, 120:e2303904120, 2023.

L.-Q. Chen. Phase-field method of phase transitions/domain structures in ferroelectric thin films: a review. *Journal of the American Ceramic Society*, 91:1835–1844, 2008.

Tianping Chen and Hong Chen. Universal approximation to nonlinear operators by neural networks with arbitrary activation functions and its application to dynamical systems. *IEEE transactions on neural networks*, 6(4):911–917, 1995.

E. De Angelis and L. Preziosi. Advection-diffusion models for solid tumour evolution in vivo and related free boundary problem. *Mathematical Models and Methods in Applied Sciences*, 10:379–407, 2000.

J. Droniou, J. A. Flegg, and G. C. Remesan. Numerical solution of a two dimensional tumour growth model with moving boundary. *Journal of Scientific Computing*, 85:22, 2020.

Avner Friedman and Joel Spruck. *Variational and free boundary problems*, volume 53. Springer Science & Business Media, 2012.

T. Furuya, M. Puthawala, M. Lassas, and M. V. de Hoop. Globally injective and bijective neural operators. *Advances in Neural Information Processing Systems*, 36:57713–57753, 2023.

E.-I. Hanzawa. Classical solutions of the Stefan problem. *Tohoku Mathematical Journal*, 33:297–335, 1981.

Anran Jiao, Haiyang He, Rishikesh Ranade, Jay Pathak, and Lu Lu. One-shot learning for solution operators of partial differential equations. *Nature Communications*, 16:8386, 2025.

G. E. Karniadakis, I. G. Kevrekidis, L. Lu, P. Perdikaris, S. Wang, and L. Yang. Physics-informed machine learning. *Nature Reviews Physics*, 3:422–440, 2021.

Chenhui Kou, Yuhui Yin, Min Zhu, Shengkun Jia, Yiqing Luo, Xigang Yuan, and Lu Lu. Neural topology optimization via active learning for efficient channel design in turbulent mass transfer. *Advanced Science*, 12:e08386, 2025.

N. Kovachki, Z. Li, B. Liu, K. Azizzadenesheli, K. Bhattacharya, A. Stuart, and A. Anandkumar. Neural operator: learning maps between function spaces with applications to PDEs. *Journal of Machine Learning Research*, 24:1–97, 2023.

Q. Li, O. Sigmund, J. S. Jensen, and N. Aage. Reduced-order methods for dynamic problems in topology optimization: a comparative study. *Computer Methods in Applied Mechanics and Engineering*, 387:114149, 2021a.

Z. Li, D. Z. Huang, B. Liu, and A. Anandkumar. Fourier neural operator with learned deformations for PDEs on general geometries. *Journal of Machine Learning Research*, 24:1–26, 2023a.

Z. Li, N. Kovachki, C. Choy, B. Li, J. Kossaifi, S. Otta, M. A. Nabian, M. Stadler, C. Hundt, K. Azizzadenesheli, and A. Anandkumar. Geometry-informed neural operator for large-scale 3D PDEs. In *Advances in Neural Information Processing Systems*, volume 36, 2023b.

Zongyi Li, Nikola Kovachki, Kamyar Azizzadenesheli, Burigede Liu, Kaushik Bhattacharya, Andrew Stuart, and Anima Anandkumar. Fourier neural operator for parametric partial differential equations. In *International Conference on Learning Representations (ICLR)*, 2021b.

Zongyi Li, Hongkai Zheng, Nikola Kovachki, David Jin, Haoxuan Chen, Burigede Liu, Kamyar Azizzadenesheli, and Anima Anandkumar. Physics-informed neural operator for learning partial differential equations. *ACM/IMS Journal of Data Science*, 2024.

Zongjia Long, Qi Zhou, Aiqing Zhu, Dong Dai, and Yiqun Liu. Deep neural operator for free boundary problems. *Nature Machine Intelligence*, pp. 1–12, 2026.

L. Lu, R. Pestourie, W. Yao, Z. Wang, F. Verdugo, and S. G. Johnson. Physics-informed neural networks with hard constraints for inverse design. *SIAM Journal on Scientific Computing*, 43: B1105–B1132, 2021a.

Lu Lu, Pengzhan Jin, Guofei Pang, Zhongqiang Zhang, and George Em Karniadakis. Learning nonlinear operators via DeepONet based on the universal approximation theorem of operators. *Nature Machine Intelligence*, 3(3):218–229, 2021b.

Tanya Marwah, Ashwini Pokle, J Zico Kolter, Zachary Lipton, Jianfeng Lu, and Andrej Risteski. Deep equilibrium based neural operators for steady-state pdes. *Advances in Neural Information Processing Systems*, 36:15716–15737, 2023.

Stanley Osher and James A. Sethian. Fronts propagating with curvature-dependent speed: algorithms based on Hamilton–Jacobi formulations. *Journal of Computational Physics*, 79(1):12–49, 1988.

C. S. Peskin. The immersed boundary method. *Acta Numerica*, 11:479–517, 2002.

Maziar Raissi, Paris Perdikaris, and George Em Karniadakis. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational Physics*, 378:686–707, 2019.

J. Stefan. Über die Theorie der Eisbildung, insbesondere über die Eisbildung im Polarmeere. *Annalen der Physik*, 278:269–286, 1891.

S. Wang and P. Perdikaris. Deep learning of free boundary and Stefan problems. *Journal of Computational Physics*, 428:109914, 2021.

H. Wu, H. Luo, H. Wang, J. Wang, and M. Long. Transolver: a fast transformer solver for PDEs on general geometries. In *Proceedings of the International Conference on Machine Learning*, pp. 53681–53705, 2024.

M. Yin, N. Charon, R. Brody, L. Lu, N. Trayanova, and M. Maggioni. A scalable framework for learning the geometry-dependent solution operators of partial differential equations. *Nature Computational Science*, 4:928–940, 2024.

Y. Zhu, W. Kong, J. Deng, and X. Bian. Physics-informed neural networks for incompressible flows with moving boundaries. *Physics of Fluids*, 36:013617, 2024.

APPENDIX CONTENTS

This appendix is organised as follows.

- **Appendix A**: extended method and implementation (level-set representation, the equilibrium cell, the Anderson solve and implicit differentiation, architecture, and training objective), with Algorithms 1–2.
- **Appendix B**: proof of Theorem 1.
- **Appendix C**: benchmark configurations and data generation, with Tables 1–2.
- **Appendix D**: extended forward and inverse-design results.
- **Appendix E**: operator properties and ablations (data efficiency, convergence, and the spurious fixed point).
- **Appendix F**: relation to prior methods.

A EXTENDED METHOD AND IMPLEMENTATION

FROST realises $\mathcal{G}_\theta$ as the fixed point of a learned cell $z^\star = f_\theta(z^\star; e)$ on a latent feature field z, decoded to the state by $(u, \varphi) = \mathcal{D}(z^\star)$. The phase, interface and interface geometry are read from φ: $\Omega = \{\varphi < 0\}$, $\Gamma = \{\varphi = 0\}$, $\hat{n} = \nabla\varphi/|\nabla\varphi|$, so a field supported inside the domain is the masking $u \mathbb{1}[\varphi < 0]$, and the topology is the number of connected components of $\{\varphi < 0\}$, obtained from the sign pattern of a single grid function. To keep the interface sharp, φ is periodically reinitialised to the signed distance to Γ, the steady state in a pseudo-time τ of $\partial_\tau\varphi + \mathrm{sgn}(\varphi_0)(|\nabla\varphi| - 1) = 0$, where φ_0 denotes the level set before the reinitialisation (not the initial seed layout of Section 5.3).

The equilibrium cell. The solution of a free-boundary problem is itself a fixed point; the steady obstacle problem is the clearest instance, its solution being the unique fixed point of the projected iteration

$$u = \Pi_\chi\big(\mathcal{A}u\big), \qquad \Pi_\chi(w) = \max(\chi, w), \qquad (\mathcal{A}u)_{ij} = \tfrac{1}{4} \sum_{(i',j')\in\mathcal{N}(i,j)} u_{i'j'} + \tfrac{h^2}{4} f_{ij}, \tag{5}$$

with $\mathcal{N}(i,j) = \{(i \pm 1, j), (i, j \pm 1)\}$ the four axial grid neighbours of node (i,j), h the grid spacing, $u = 0$ on ∂D, and f the source ($f = 0$ for the harmonic obstacle problem). Here $\mathcal{A}$ is the Jacobi sweep of the homogeneous-Dirichlet Laplacian (spectral radius below one) and Π_χ is non-expansive, so $\Pi_\chi \circ \mathcal{A}$ is a contraction; moreover the discrete complementarity conditions $u \geq \chi$, $-\Delta u \geq f$, $(u - \chi)(-\Delta u - f) = 0$ hold exactly when (5) leaves u unchanged, so for this instance the fixed-point residual literally is the free-boundary condition. FROST generalises this construction: a learned cell f_θ, into which e is injected additively at each application, has a fixed point that a pointwise decoder maps to (u, φ), and the solve is monitored through the residual $r(z; e) = \|f_\theta(z; e) - z\|$, driven below tolerance and reported as a self-consistency diagnostic; for the learned cell this is a numerical convergence check on the solve, distinct from the physical residuals defined below, which are evaluated on the decoded state.

Anderson solve and implicit differentiation. The forward pass solves $z = f_\theta(z; e)$ by Anderson acceleration: with residual $g(z) = f_\theta(z; e) - z$, each step keeps the last m_A residuals in a history matrix G_k and combines the corresponding iterates with mixing weights α chosen to minimise $\|G_k\alpha\|_2$ subject to $\mathbf{1}^\top\alpha = 1$. Gradients are taken by the IFT rather than through the iterates. With $J_z = \partial f_\theta/\partial z|_{z^\star}$, differentiating the fixed-point relation gives $(I - J_z)\,\mathrm{d}z^\star/\mathrm{d}e = \partial f_\theta/\partial e$; the contraction $\rho(J_z) < 1$, with ρ the spectral radius, makes $I - J_z$ invertible with a convergent Neumann series, so the adjoint $v = J_z^\top v + (\partial\mathcal{L}/\partial z^\star)^\top$ is solved by the same Anderson routine, each step a single vector–Jacobian product (VJP). Only $z^\star$ enters, so training memory is independent of the number of solver steps. The same adjoint serves training and, with the design objective in place of the data loss, inverse design. Two scope notes: the geometric guarantee of Theorem 1 covers the plain iteration $z_{k+1} = f_\theta(z_k; e)$, with Anderson used as an accelerator on top of it and the plain iteration retained as a fallback; and Theorem 1 is stated for the state-level map T_θ, which the latent cell together with the decoder realises, the contraction being assumed and monitored at the latent level, where in practice we track the linearised indicator $\rho(\partial f_\theta/\partial z|_{z^\star}) < 1$.

Architecture. The cell and the conditioned operator share a spectral backbone (Li et al., 2021b): an input layer lifts e with the grid coordinates to a width-w feature field; each spectral block adds a Fourier convolution on the lowest m retained modes to a pointwise 1×1 convolution followed by a GELU nonlinearity, and a two-layer pointwise head decodes to (u, φ). In the equilibrium realisation a single block is the cell f_θ, applied repeatedly with the lifted input re-injected (used for the steady obstacle problem); for the time-dependent problems a fixed stack of blocks is applied once and conditioned on the query time and the design.

Training objective. The operator minimises the relative L^2 error of the field and the level set against reference solutions, $\mathcal{L} = \|\hat{u} - u\|_2/\|u\|_2 + \lambda \, \|\hat{\varphi} - \varphi\|_2/\|\varphi\|_2$, with $\hat{u}, \hat{\varphi}$ the predictions and the weight λ set so the two terms enter at a comparable scale; the free-boundary condition itself carries no label and is enforced through the residual of the equilibrium solve. Independently of training, FROST reports from the prediction alone a *physics-consistency residual* (the trust signal of the main text), evaluated on the decoded state and distinct both from the solve residual r above and from the certificate gap $|\mathcal{J}(\mathcal{G}_\theta) - \mathcal{J}(\text{solver})|$ of Appendix C, which requires the exact solver: for the obstacle, $\rho_{\text{obs}} = \| \min(u - \chi, -\Delta u - f) \|$; for the Stefan problem, the inconsistency between predicted solid ($\varphi < 0$) and warm temperature ($T > 0$), a necessary sign check on the phase–temperature pairing rather than the full latent-heat balance.

Algorithm 1 FROST forward solve and implicit backward

Require: instance e, cell f_θ, history depth m_A, relaxation β_A, tolerance τ_f
1: initialise z_0, $k \leftarrow 0$
2: **while** $\|f_\theta(z_k; e) - z_k\|/\|z_k\| > \tau_f$ **do**
3: form the residual history G_k (last m_A residuals); solve the constrained least squares for the mixing weights α; Anderson update z_{k+1}; $k \leftarrow k + 1$
4: **end while**
5: $z^\star \leftarrow z_k$; **return** $(u, \varphi) = \mathcal{D}(z^\star)$
6: **Backward** (cotangent $y = \partial\mathcal{L}/\partial z^\star$): solve $v = J_z^\top v + y$ by Anderson using only the VJP $J_z^\top v$ at $z^\star$; **return** $(\partial f_\theta/\partial\theta)^\top v$, $(\partial f_\theta/\partial e)^\top v$

Algorithm 2 Differentiable inverse design with an active trust region

Require: frozen $\mathcal{G}_\theta$, objective $\mathcal{J}$, parameterisation $e(c)$, step η, threshold τ_d, budget S
1: **for** $s = 1, \ldots, S$ **do**
2: forward solve $z^\star$; $J \leftarrow \mathcal{J}(\mathcal{D}(z^\star))$; design gradient $\partial J/\partial\xi$ via the equilibrium adjoint (Alg. 1)
3: $\xi \leftarrow \xi - \eta \, \partial J/\partial\xi$; on-manifold map $c \leftarrow c_{\text{lo}} + (c_{\text{hi}} - c_{\text{lo}}) \, \sigma(\xi)$, with ξ the unconstrained optimisation variable and σ the logistic sigmoid
4: re-solve $z^\star$ at $e(c)$; trust signal $\rho_d \leftarrow r(z^\star; e(c))$
5: **if** $\rho_d > \tau_d$ and a reference solver is available **then** solve the true problem at c; add to training set; warm-restart $\mathcal{G}_\theta$
6: **end if**
7: **end for**
8: **return** design c with $\rho_d \le \tau_d$ (trusted: residual below the threshold)

B Proof of Theorem 1

Throughout, $z = (u, \varphi)$ denotes the joint state on the background grid and ranges over a finite-dimensional vector space V that is therefore complete in the norm $\|\cdot\|$ (the normalised grid ℓ^2 norm of Section 4); $T_\theta(\cdot\,; e) : V \to V$ is the equilibrium map, and $\mathcal{N}_e \subset V$ is the non-empty closed set of Assumption 1, on which $T_\theta(\cdot\,; e)$ is an L-contraction with $L \in [0, 1)$ and which it maps into itself. As T_θ is assembled from smooth operations, it is continuously differentiable on $\mathcal{N}_e$.

Proof. Existence, uniqueness and geometric convergence. Since V is finite-dimensional it is complete, and the closed set $\mathcal{N}_e$ is complete in the induced metric. A contraction of a non-empty complete metric space into itself has, by the Banach fixed-point theorem, a unique fixed

point $z^\star(e) \in \mathcal{N}_e$. For any $z_0 \in \mathcal{N}_e$ the iterates remain in $\mathcal{N}_e$ and satisfy $\|z_k - z^\star(e)\| = \|T_\theta(z_{k-1}; e) - T_\theta(z^\star; e)\| \le L\,\|z_{k-1} - z^\star(e)\|$, so by induction $\|z_k - z^\star(e)\| \le L^k\,\|z_0 - z^\star(e)\|$, which converges as $L < 1$.

Invertibility of $I - \partial_z T_\theta$ and differentiability. Set $J_z = \partial_z T_\theta(z^\star(e); e)$. The L-Lipschitz bound gives $\|J_z v\| \le L$ for any unit vector v, hence $\|J_z\| \le L < 1$; the Neumann series therefore converges in operator norm, $(I - J_z)^{-1} = \sum_{k\ge 0} J_z^k$ with $\|(I - J_z)^{-1}\| \le 1/(1 - L)$, so $I - J_z$ is invertible. Writing $F(z, e) = z - T_\theta(z; e)$, we have $F(z^\star(e), e) = 0$ and $\partial_z F = I - J_z$ invertible; as $F \in C^1$, the implicit function theorem furnishes a unique continuously differentiable $e \mapsto z^\star(e)$ whose derivative solves $(I - J_z)\,\mathrm{d}z^\star/\mathrm{d}e = \partial_e T_\theta(z^\star(e); e)$, which is (4). □

Remark 1. *A single constant governs both passes: the plain forward iteration reduces its initial error by the factor τ_f in at most $\lceil \log \tau_f / \log L \rceil$ steps (Anderson acceleration is used on top of this guaranteed baseline), and the adjoint iteration contracts at the same rate $\rho(J_z^\top) = \rho(J_z) \le L$. The hypothesis $L < 1$ is also what separates the learned equilibrium from a map fitted only to reproduce a single iteration, which may be pointwise accurate yet settle on a spurious fixed point away from the true solution (Appendix E).*

C Benchmark configurations and data generation

Reference solutions are produced with a dedicated solver for each problem and rasterised to the fixed grid; the operator is trained on a held-in split and every error is reported on held-out instances. Table 1 lists the datasets and Table 2 the architectures.

Obstacle. The steady contact problem is solved by projected successive over-relaxation on a 128×128 grid, over a family of smooth obstacles whose contact set has one or two connected components, split by topology. For design, the obstacle χ is optimised so the equilibrium contact set matches a target footprint $M^\star$, $J_{\mathrm{obs}}(\chi) = \|H_\tau(-\varphi^\star) - M^\star\|_2^2$, differentiated through the equilibrium by the IFT adjoint; because the exact solver is cheap, the optimised design is re-solved and validated by the certificate gap $|\mathcal{J}(\mathcal{G}_\theta) - \mathcal{J}(\mathrm{solver})|$.

Tumour. On the growing domain Ω_t the nutrient field solves $-\Delta u + k\,u = 0$ in Ω_t, $u = 1$ on Γ_t, with the level-set front advancing at $V_n \propto \partial_n u$; fronts are recorded on a 128×128 grid over time frames, split by instance and stratified by uptake-rate quartile. This problem is used only in the forward direction.

Stefan. Solidification is generated by a one-phase enthalpy discretisation on a 128×128 grid: with Ω_t the solid region and L the latent heat, $\partial_t T = \nabla^2 T$ in $D \setminus \Omega_t$, $T = 0$ on Γ_t, $L\,V_n = -\partial_n T$ on Γ_t; cold seeds at $T_{\mathrm{cold}} \in [-1.6, -0.7]$ nucleate solid that grows and merges. The control task optimises the seed layout, T_{cold} and seed radius to drive the final front to an achievable target (the final front of a held-out run), with controls held in their trained ranges by bounded maps so the operator is never queried off-manifold.

Single-deformation oracle. The barrier baseline of Figs 2c and 3c fits, for each held-out instance (or frame), a smooth invertible displacement of a fixed reference domain directly to the reference level set by minimising their mismatch under a fold-free constraint (positive Jacobian determinant), and is reported as the best such fit. A fold-free warp is a homeomorphism and therefore preserves the number of connected components, so its failure across a topology change is structural rather than a matter of underfitting.

Table 1. Per-benchmark datasets. Stefan is downsampled by a factor of two from the 128^2 generation grid.

Benchmark	grid	instances	frames	split (tr/te)
Obstacle (steady)	128^2	80	–	$64/16$
sampling: uniform geometry parameters; topology-stratified				
Tumour (time)	128^2	80	15	$64/16$
sampling: uniform uptake k, separation, scale; k-quartile-stratified				
Stefan (time)	64^2	80	15	$65/15$
sampling: uniform seed count, L, T_{cold}; seed-count-stratified				

Table 2. Per-benchmark architecture and training. All models use AdamW (lr 10^{-3}, weight decay 10^{-4}, cosine-annealed) and the relative-L^2 objective. The DEQ obstacle cell uses Anderson acceleration (history depth $m_A = 4$, ≤ 30 forward / 14 backward iterations).

Model	modes m	width w	depth L	parameters	epochs	batch
Obstacle (FNO)	16	32	4	2.11 M	400	8
Obstacle (DEQ)	12	32	1 (iterated)	302 k	160	16
Tumour	20	32	4	3.29 M	100	32
Stefan	16	32	4	2.11 M	100	32

D Extended results

Stefan forward operator. The same operator, evaluated forward on held-out solidification trajectories, tracks the field, the front and the $N \to 1$ change of topology; the temperature carries a relative-L^2 error of 14.6% (a near-binary field with sharp grain cores sets a Gibbs-type floor for the truncated spectral representation), while the level set is recovered to 3.8% and the solid region to an IoU of 0.909, with topology exact away from the coalescence instant (Fig. 5).

Tumour nutrient field. Slicing the complementary nutrient field over full trajectories, the field error stays near its test-set mean of 12.1% while the level set is recovered to 1.1% and the tumour region to an IoU of 0.974 (Fig. 6).

Obstacle, three-dimensional view and inverse design. Rendering the membrane as a height surface makes the contact set and its topology explicit (Fig. 7). Inverse design runs through the equilibrium itself: differentiating the frozen DEQ through the IFT adjoint matches single- and two-component target footprints at IoU 0.87–0.96, and re-solving with the true solver certifies the design (certificate gap 0.0002–0.003). The operator can be exploited off-manifold, where an ensemble disagreement fails to flag the error (correlation -0.36 with the true gap) but the FROST physics fixed-point residual succeeds (correlation $+0.97$), giving a self-contained trust signal (Fig. 8).

E Operator properties and ablations

The operator reaches useful accuracy from few simulations. On the steady obstacle problem, topology accuracy reaches 0.92 at $K = 4$ training instances and 1.0 by $K = 32$, and the level set is within a few percent by $K = 8$, between roughly one hundred and seven hundred times fewer simulations than comparable data-driven operators require (Jiao et al., 2025; Boullé et al., 2023) (Fig. 9a,b). The fixed-point solve converges geometrically to its tolerance, and the equilibrium form matches the direct map to within about twice its field error (2.1% against 0.9%) at roughly seven times fewer parameters (Fig. 9c). The spurious fixed point of a one-step-accurate map is visible directly: such a map iterates to a solution far from the truth, whereas training through the fixed point does not (Fig. 9d), confirming that the contraction of Assumption 1 is what training through the equilibrium promotes.

F Relation to prior methods

FROST combines four properties that prior families supply only separately: a solution operator learned once and amortised over a family; an output domain that is itself unknown, bounded by a

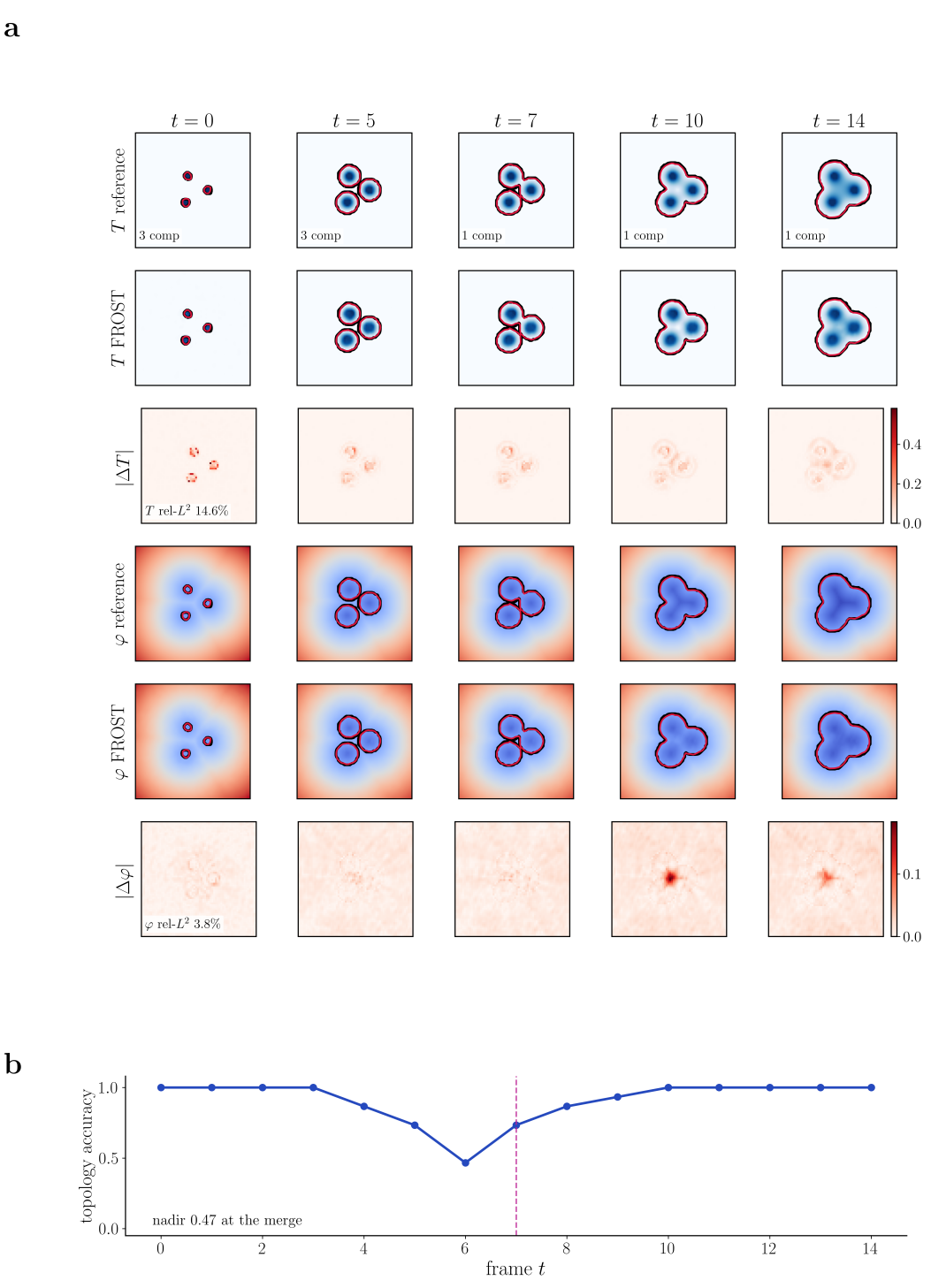

Figure 5: **Stefan forward operator.** A held-out three-seed solidification through its $3 \to 1$ merge; rows: temperature T and its error, level set φ and its error, with reference (black) and FROST (crimson) fronts overlaid and the component count boxed. Right: topology accuracy against frame.

free boundary; a topology that may change within the family; and end-to-end differentiability for design. Implicit and equilibrium neural operators learn a steady PDE's fixed point (Marwah et al., 2023; Bai et al., 2019) but on a fixed domain and over the field alone, whereas FROST's fixed point

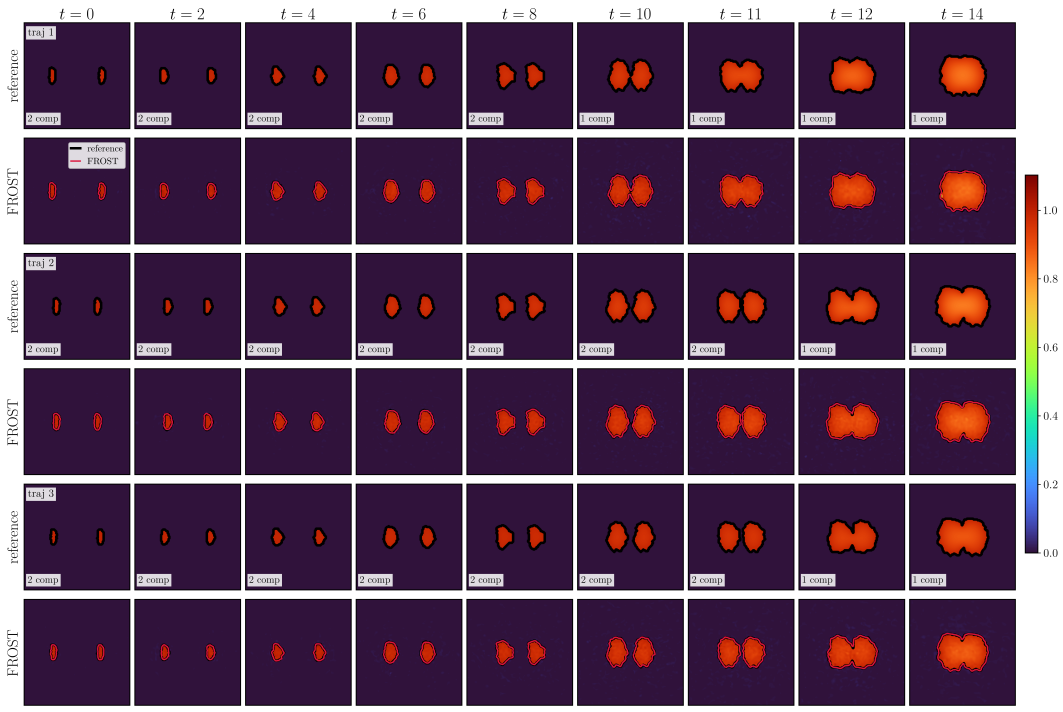

Figure 6: **Tumour forward operator: the nutrient field across trajectories.** Field u of three held-out two-tumour coalescences at nine times, reference against FROST, fronts overlaid and component count boxed.

also determines the geometry. Topology optimisation, classical and neural (Bendsoe & Sigmund, 2013; Li et al., 2021a; Kou et al., 2025), optimises a single design against a conventional solver rather than learning an operator over the family. Geometry-aware operators (Li et al., 2023a;b; Yin et al., 2024; Wu et al., 2024) interpolate over shapes meshed per instance at fixed topology. Reusable free-boundary operators (Long et al., 2026) track the boundary by a single smooth deformation of a fixed reference, so they are confined to a fixed topology and their inverse design is not through the operator. The level-set state on a fixed grid and the differentiable equilibrium supply all four properties at once.

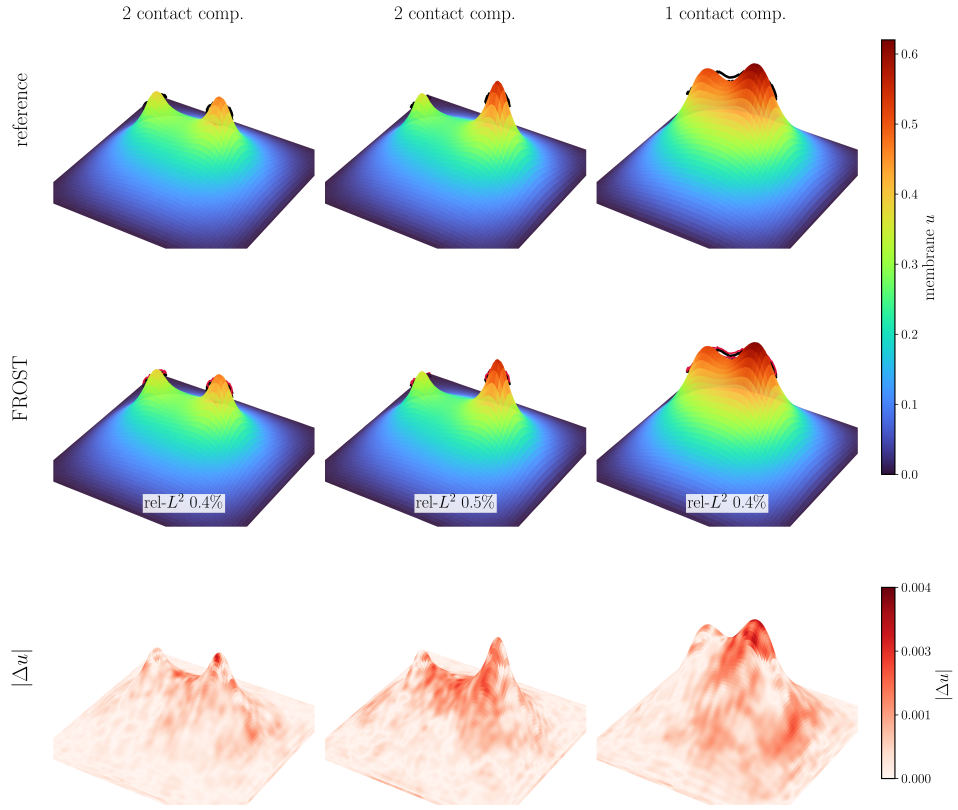

Figure 7: **Obstacle forward operator, three-dimensional view.** The membrane $u(x, y)$ as a height surface for three held-out obstacles; rows: reference, FROST, absolute error, with the free boundary lifted onto each surface (black reference, crimson FROST).

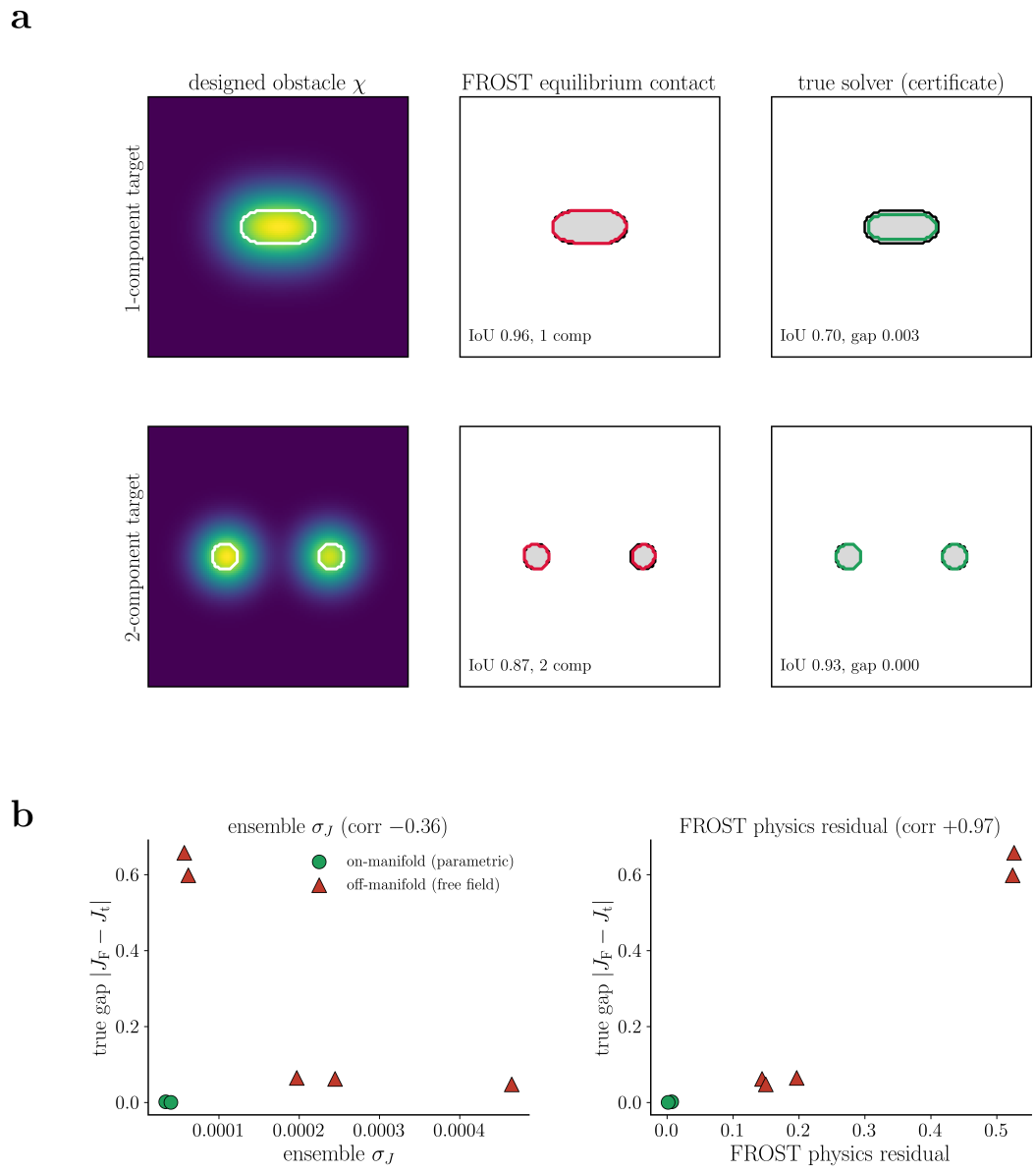

Figure 8: **Obstacle inverse design through the equilibrium.** **(a)** Parametric design for single- and two-component targets: designed obstacle, FROST equilibrium contact against the target, and the same design re-solved by the true solver (certificate). **(b)** Trust-signal sweep: the ensemble disagreement does not track the true certificate gap (corr -0.36), whereas the FROST physics fixed-point residual does (corr $+0.97$).

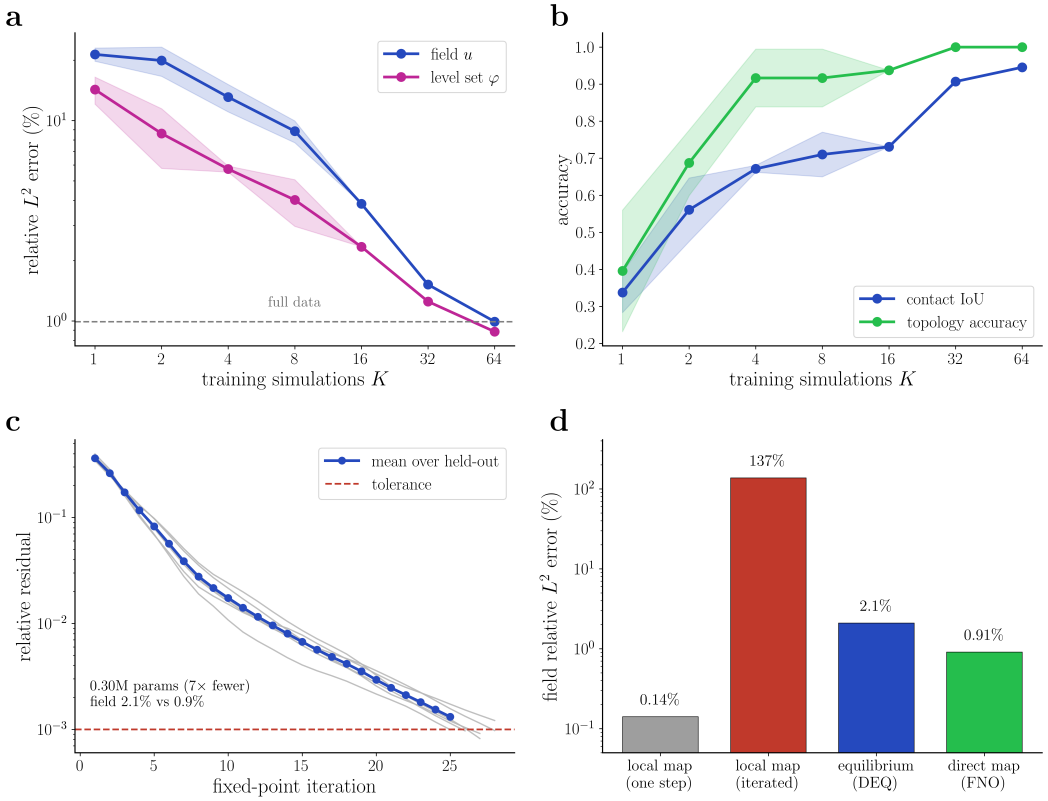

Figure 9: **Properties of the equilibrium operator, on the steady obstacle problem. (a)** Field and level-set error against the number of training simulations K. **(b)** Contact IoU and topology accuracy against K. **(c)** Geometric convergence of the fixed-point solve; the equilibrium form reaches comparable accuracy with far fewer parameters. **(d)** A one-step-accurate map carries a spurious fixed point; training through the fixed point removes it.