Supplementary Information for: Differentiable surrogate modeling for strain-engineered exciton localization in hybrid TMD-photonic nanostructures

Georgii Bulavin[1], Alexander Shorokhov[1], Andrey Fedyanin[1*]

[1*]Faculty of Physics, Lomonosov Moscow State University, Moscow, 119991, Russia.

*Corresponding author(s). E-mail(s): fedyanin@nanolab.phys.msu.ru;
Contributing authors: bulavin.ga21@physics.msu.ru;
shorokhovas@my.msu.ru;

Contents

Supplementary Methods

S1. Grid resolution selection and role of the 1024×1024 benchmark

All results reported in the main text — FDM training-dataset generation, surrogate training and evaluation (including the out-of-distribution test on the resonant waveguide geometry of Ref. [1]), the drift–diffusion exciton-transport calculation, and the topology optimization loop — were computed exclusively on the 256×256 grid (pixel size $\approx 46\,\mathrm{nm}$ over a 12×12 μm domain). The 1024×1024 grid quoted in Table 2 of the main text is used solely as a separate benchmarking point to characterize how the FDM-solver-to-surrogate speedup scales with grid size ($1868\times$ at 256×256 versus $5082\times$ at 1024×1024); it does not correspond to the resolution used to generate any training data, evaluate surrogate accuracy, or run the topology optimization reported elsewhere in the paper.

The choice of 256×256 for all substantive calculations was dictated primarily by computational cost rather than by numerical accuracy considerations. As reported in

Table 2 of the main text, the FDM solver runtime scales approximately as $\mathcal{O}(N^2)$ with grid size N: a single forward solve requires ≈ 5060 ms on a 256×256 grid versus $\approx 93\,000$ ms on a 1024×1024 grid (single CPU core, AMD EPYC 7742), an $\approx 18\times$ increase in per-sample cost for a $16\times$ increase in pixel count.

Generating the full training dataset of 6 300 samples at 256×256 resolution requires approximately 9 hours of single-core compute time. Repeating the same dataset-generation procedure at 1024×1024 resolution would require approximately one week of single-core compute time, which was judged impractical given that the dataset must be regenerated whenever the geometry sampling scheme or material parameters are revised during model development.

Because the strain fields produced by the linear Kirchhoff–Love plate solver are numerically well-resolved already at moderate grid densities (the biharmonic operator in Eq. (5) of the main text is solved directly via LU factorization, with no iterative-convergence error), refining the grid beyond 256×256 does not measurably reduce the numerical error of the computed strain field; it only increases spatial sampling density. The 256×256 resolution is therefore already sufficient to resolve the strain features relevant to exciton localization, and was adopted uniformly across the entire computational pipeline described in the main text.

S2. Boundary treatment for the 13-point biharmonic stencil

The simply supported condition ($w = 0$) at substrate contact points and at the domain edges is imposed by excluding all such nodes from the linear system entirely: only interior, non-contact grid points are treated as free unknowns of the biharmonic equation (Eq. (6) of the main text), and their fixed neighbors are dropped from the assembled 13-point stencil rather than being carried as explicit ghost values. This is exact rather than approximate, since the dropped neighbors are known to equal zero and therefore contribute nothing to the stencil sum regardless of whether they are included or omitted; no additional moment or curvature condition is imposed beyond this pointwise Dirichlet constraint. The same rule applies uniformly to arbitrarily shaped substrate contours, since contact/free classification is read directly from the rasterized support map $\mathcal{S}(x, y)$ rather than from an analytic boundary description.

Second derivatives used to compute stress and strain (Eqs. (7)–(8) of the main text) are likewise evaluated only at interior points via central differences, leaving the outermost node layer at zero by construction. The resulting loss of stencil accuracy is confined to the immediate vicinity of contact boundaries, consistent with the largest surrogate residuals (Fig. 3e) and the largest deviation from literature Raman data (Table 1) both occurring at structural edges in the main text.

S3. Calibration and cross-validation of the load parameter against literature strain values

The comparison values $\varepsilon_{\max}$(literature) in Table 1 of the main text are peak in-plane strain values taken directly from the cited studies [2–4], already reported there together with their own measurement uncertainty. The residual discrepancy between FDM and literature values reported below should therefore not be interpreted as evidence of

model failure, since part of it may simply reflect this pre-existing uncertainty in the reference values rather than an error in the solver.

The single scalar load parameter q in the Kirchhoff–Love model (main-text Eq. (2)) was calibrated using the array-of-nanopillars configuration of Ref. [2]: q was adjusted so that the FDM-predicted peak strain matched the corresponding literature value for that configuration, yielding the exact agreement reported in the first row of Table 1 ($\varepsilon_{\max} = 1.0\%$ for both literature and FDM). The same calibrated value of q, obtained without any further adjustment, was then applied to the two remaining configurations in Table 1 (the nanopillar array of Ref. [3] and the array of holes of Ref. [4]), and the resulting FDM peak strain was compared against the corresponding literature value for each case. This procedure tests whether a single fixed load parameter, calibrated once on one configuration, transfers to structurally different substrate configurations of the *same* film–substrate material pair without refitting. The resulting discrepancies (FDM $\varepsilon_{\max} = 0.78\%$ vs. literature 0.6% for Ref. [3]; FDM 0.2% vs. literature 0.3% for Ref. [4]) give the mean absolute error of 0.093% of peak strain quoted in the main text.

We additionally tested the transferability of this same calibrated load parameter to literature datasets involving different film and substrate materials than those used for calibration. In these cross-material tests, the fixed q value produced substantially larger deviations from the corresponding literature strain values than observed for the same-material configurations above. We conclude that the calibrated load parameter is specific to the particular film–substrate material pair on which it was fitted, and does not transfer across material systems; applying the model to a different film or substrate material therefore requires re-calibrating q against a reference measurement for that material pair, rather than reusing the value calibrated in this work.

Supplementary Notes

Supplementary Note 1. Choice of encoder–decoder depth for the U-Net surrogate

The five-stage encoder–decoder depth of the U-Net surrogate (Table S1) was selected empirically by comparing networks with 4, 5, and 6 downsampling stages under otherwise identical training settings. Architectures with 4 stages or fewer consistently achieved lower validation accuracy than the 5-stage network, indicating that the reduced receptive field and bottleneck capacity were insufficient to capture the strain-localization features relevant to the task. Increasing the depth to 6 stages substantially increased the number of trainable parameters and per-epoch training time without a corresponding improvement in accuracy that would justify the added computational cost. The 5-stage architecture was therefore adopted as the best trade-off between predictive accuracy and training efficiency for all results reported in the main text.

Supplementary Note 2. Choice of loss weighting in the surrogate training objective

The weights $\lambda_{\mathrm{MSE}} = 1000$ and $\lambda_{\mathrm{grad}} = 10$ in the composite loss function (main-text Eq. (4)) were selected empirically after testing several weighting combinations

during model development. Because the raw strain field values, and consequently the unweighted MSE and gradient-difference terms, are naturally on the order of 10^{-3}–10^{-4} (thousandths to ten-thousandths), leaving these terms unscaled would make the loss magnitude too small to produce meaningful gradient updates during training. The chosen weights rescale both loss components to a numerically convenient range closer to unity, which was found to give stable and effective training; no attempt was made to further tune the relative ratio $\lambda_{\mathrm{MSE}}/\lambda_{\mathrm{grad}}$ beyond this rescaling.

Supplementary Note 3. Sensitivity of topology optimization to filter radius and fill fraction

The optimized substrate configuration is sensitive to the density-filter radius r and the target fill fraction ρ_0 (main-text Eq. (8)). This sensitivity arises because the log-ratio and symmetry loss terms alone are not sufficient to suppress high-frequency noise in the density map $\mathcal{S}(x, y)$ during optimization: without adequate filtering, the optimizer tends to converge to a noisy, speckled distribution rather than a smooth, fabricable configuration. We tested filter radii of $r = 4$, $r = 8$, 16, 32, and 64 pixels; $r = 8$ pixels, the value adopted in the main text, was found to give the most visually clean and well-balanced configuration, whereas $r = 64$ pixels already over-smooths the design and noticeably degrades the strain-localization objective. The filter radius r and fill-factor penalty therefore act as the primary regularizers controlling geometric smoothness in the optimized support map. See Supplementary Fig. S3 for representative optimized configurations obtained at different (r, ρ_0) settings.

Supplementary Figures

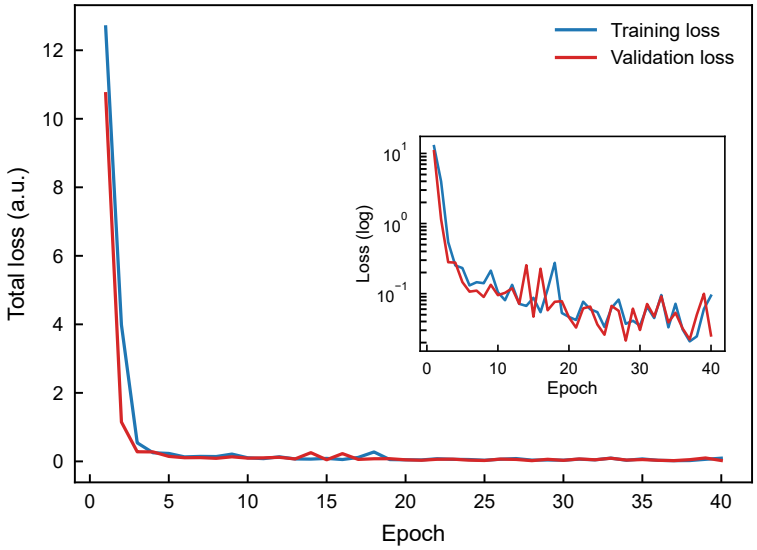

Fig. S1 Training dynamics of the U-Net surrogate model. Total training and validation loss (main-text Eq. (4)) versus epoch, shown on a linear scale (main panel) and on a logarithmic scale (inset) to resolve behavior at low loss values. Both curves drop sharply from their initial values (≈ 12.7 for training, ≈ 10.8 for validation) to ≈ 0.3 within the first 3 epochs, and continue to decrease gradually until reaching a noisy plateau by approximately epoch 20, after which the loss fluctuates within a narrow band (≈ 0.04–0.1, inset) for the remainder of training without further systematic decrease. Training and validation curves remain closely superimposed throughout, with no divergence indicative of overfitting. No early-stopping criterion was applied; training was run for the full 40 epochs as reported in the main text.

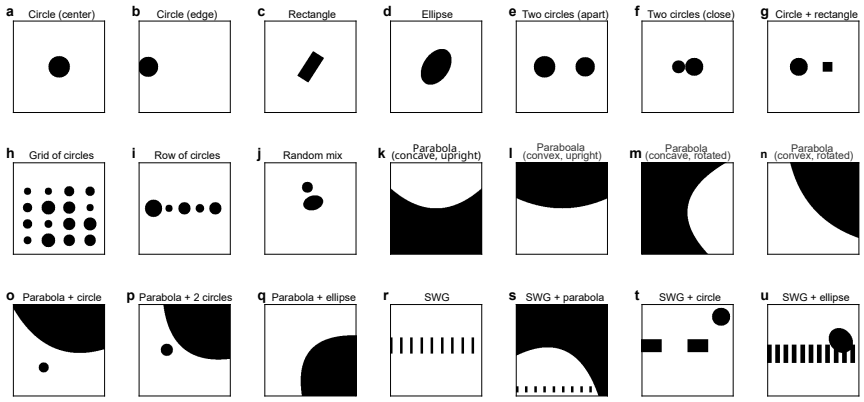

Fig. S2 Complete set of substrate configuration classes. [PLACEHOLDER: caption text – grid of representative binary support maps for each of the 21 classes listed in main-text Table 2 (single-feature, multi-feature, curved-boundary categories).]

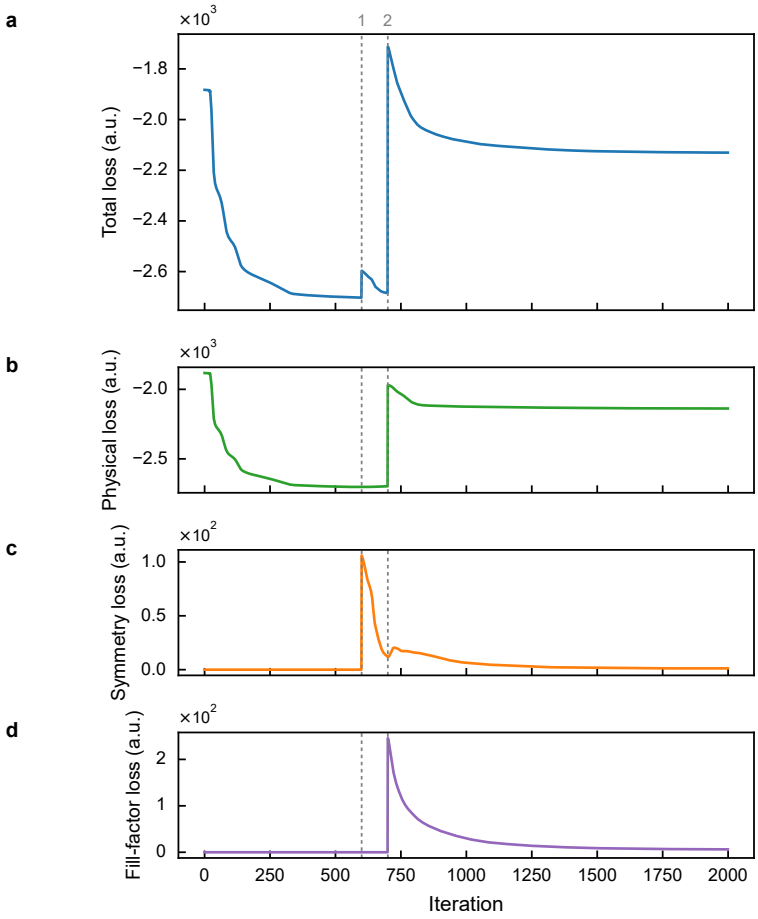

Fig. S3 Convergence of the three-phase topology optimization loss. a Total optimization loss versus iteration (main-text Eq. (8)); dashed vertical lines mark the transitions between the three optimization phases, labeled "1" (Phase I → Phase II, iteration 600) and "2" (Phase II → Phase III, iteration 700). **b–d** Corresponding weighted contributions of the physical loss term (**b**, log-ratio plus absolute-strain objective), the symmetry loss term (**c**), and the fill-factor loss term (**d**) to the total loss in **a**. The symmetry and fill-factor terms are zero-weighted during Phase I and are activated only after their respective phase boundaries, consistent with the curriculum schedule described in the main text.

Supplementary Tables

Supplementary Discussion

Table S1 Full hyperparameters for U-Net surrogate training.

Hyperparameter	Value
Encoder stages	5
Channels per stage	64, 128, 256, 512, 1024
Bottleneck channels	1024
Kernel size	3×3
Activation	ReLU
Pooling	2×2 max-pool
Upsampling	bilinear, $2\times$
Optimizer	Adam
Learning rate	10^{-3}
Optimizer betas	(0.9, 0.999) (PyTorch default)
Weight decay	0 (PyTorch default)
Batch size	32
Epochs	40
λ_{MSE}	1000
λ_{grad}	10
Total trainable parameters	28,999,393
GPU	NVIDIA A100-SXM4-80GB

References

[1] Gartman, A.D., Shorokhov, A.S., Fedyanin, A.A.: Efficient light coupling and purcell effect enhancement for interlayer exciton emitters in 2d heterostructures combined with sin nanoparticles. Nanomaterials **13**(12), 1821 (2023) https://doi.org/10.3390/nano13121821

[2] Palacios-Berraquero, C., *et al.*: Large-scale quantum-emitter arrays in atomically thin semiconductors. Nature Communications **8**, 15093 (2017) https://doi.org/10.1038/ncomms15093

[3] Branny, A., Kumar, S., Proux, R., Gerardot, B.D.: Deterministic strain-induced arrays of quantum emitters in a two-dimensional semiconductor. Nature Communications **8**, 15053 (2017) https://doi.org/10.1038/ncomms15053

[4] Kumar, S., Kaczmarczyk, A., Gerardot, B.D.: Strain-induced spatial and spectral isolation of quantum emitters in mono- and bilayer wse2. Nano Letters **15**, 7567–7573 (2015) https://doi.org/10.1021/acs.nanolett.5b03312