A General Framework for Knowledge Integration in Machine Learning for Electromagnetic Scattering Using Quasinormal Modes

Downloaded from: https://research.chalmers.se, 2026-10-11 07:51 UTC

Citation for the original published paper (version of record):
Aadland Lilja, V., Jonasson Svärdsby, A., Gahlmann, T. et al (2026). A General Framework for Knowledge Integration in Machine Learning for Electromagnetic Scattering Using Quasinormal Modes. Laser and Photonics Reviews, 20(11). http://dx.doi.org/10.1002/lpor.202502769

N.B. When citing this work, cite the original published paper.

research.chalmers.se offers the possibility of retrieving research publications produced at Chalmers University of Technology. It covers all kind of research output: articles, dissertations, conference papers, reports etc. since 2004. research.chalmers.se is administrated and maintained by Chalmers Library

(article starts on next page)

Check for updates

Laser & Photonics Reviews

LASER & PHOTONICS REVIEWS

www.lpr-journal.org

RESEARCH ARTICLE OPEN ACCESS

A General Framework for Knowledge Integration in Machine Learning for Electromagnetic Scattering Using Quasinormal Modes

Viktor A. Lilja | Albin J. Svärdsby | Timo Gahlmann | Philippe Tassin

Department of Physics, Chalmers University of Technology, GöteborgSE-412 96, Sweden

Correspondence: Philippe Tassin (philippe.tassin@chalmers.se)

Received: 14 October 2025 | **Revised:** 24 February 2026 | **Accepted:** 25 February 2026

Keywords: inverse design | machine learning | nanophotonics | neural network | physics-informed | quasinormal modes

ABSTRACT

Neural networks have been demonstrated to be able to accelerate the modeling and inverse design of optical and electromagnetic devices by serving as fast surrogates for electromagnetic solvers. Nevertheless, such neural networks can be unreliable and normally require extreme amounts of data to train. Here it is shown that these limitations can be alleviated by constraining neural-network models using prior knowledge about the governing physics. We propose a universal physics-informed neural network framework for electromagnetic scattering based on the quasinormal mode expansion of the scattering matrix. The neural networks learn the resonant structure underlying the scattering spectrum, are guaranteed to obey energy conservation and causality, and are shown to have significantly improved data efficiency for photonic-crystal slabs and all-dielectric free-form metasurfaces. Furthermore, the framework allows additional problem-specific constraints, such as losslessness, symmetries, and number of modes, to be imposed manually when they are available. The method can be applied to a wide range of optical and electromagnetic devices owing to the generality of the quasinormal mode formalism.

1 | Introduction

Machine learning [1, 2] is a powerful tool for inverse design in electromagnetism, mechanics, materials science, quantum many-body systems, and other paradigms of physics [3–14]. The goal of inverse design is to find a device design that exhibits some predefined properties [15–19]. For linear electromagnetic devices, the properties are often quantities that describe the relation between incoming and outgoing amplitudes of electromagnetic modes, such as transmittance, reflectance, or absorption [6–8, 14, 18–41]. These quantities can be derived from the frequency-dependent scattering matrix $S(\omega)$, which encodes the complete scattering response of a general linear N-port device by relating the amplitudes $\mathbf{s}^- = (s_1^-, \ldots, s_N^-)$ of the incoming modes to the amplitudes $\mathbf{s}^+ = (s_1^+, \ldots, s_N^+)$ of the outgoing modes through $\mathbf{s}^- = S(\omega)\mathbf{s}^+$. The forward problem of finding $S(\omega)$ for a given design can be easily solved with full-wave electrodynamics solvers. The inverse problem of finding a design with a desired $S(\omega)$ is, on the other hand, significantly more challenging and fundamentally ill-posed, since a given $S(\omega)$ can have zero or several corresponding designs. Traditional inverse-design methods use optimization algorithms to search through the design space, often requiring hundreds to thousands of sequential full-wave simulations to converge. Another approach is to first replace the simulator with a neural network (NN) trained to predict $S(\omega)$ given a design parametrization, summarized schematically as

$$\text{Design} \xrightarrow{\text{NN}} S(\omega).$$

Once trained, this forward model can typically predict $S(\omega)$ orders of magnitude faster than an electrodynamics solver and its gradient can easily be calculated using backpropagation. This

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2026 The Author(s). *Laser & Photonics Reviews* published by Wiley-VCH GmbH

allows for efficient exploration of the design space through iterative optimization [19, 42] or the training of inverse models [22, 30, 32–34, 38, 43]. The downside of this approach is that the performance of the final design becomes dependent on the accuracy of the NN. Although it may be possible to improve accuracy by increasing the number of trainable parameters and the number of training samples, computational resources are often a limiting factor [13]. Furthermore, the black box nature of NNs makes them unreliable and very limited insights can be gained from their predictions [44].

These issues can be alleviated by utilizing prior knowledge about the governing physics [44, 45], a strategy often referred to as physics-informed machine learning. One physics-informed method, which has recently gathered attention in electromagnetic scattering problems [14, 20, 25, 29], is to include a physics layer as a final step in the forward model. The augmented forward model can be schematically summarized as

$$\text{Design} \xrightarrow{\text{NN}} \begin{matrix}\text{Physics}\\ \text{parameters}\end{matrix} \xrightarrow{\text{Physics model}} S(\omega), \tag{1}$$

where the physics model is a known computable expression for $S(\omega)$ with the physics parameters as arguments. Now, any constraints imposed on the physics model will also be satisfied by the full forward model. Furthermore, the task of the NN has become to predict the physics parameters rather than $S(\omega)$ directly, which may be easier to learn. Indeed, previous works have demonstrated that using a physics layer reduces prediction error for a given NN size [14, 25, 29], improves generalization [20, 25, 29], and reduces the required amount of training data [20, 29]. Another benefit is that the physics parameters become accessible as a learned latent description [20, 29] that can be used as a design target in the inverse problem [14]. The forward model also becomes less of a black box in the sense that the predicted spectra can be explained in terms of the physics parameters. Implementing the physics model in a numerical framework that supports automatic differentiation enables the full forward model to be trained directly on simulated $S(\omega)$. This is in contrast to the methods proposed in References [28, 42, 46–48], where the physics parameters are extracted from $S(\omega)$ before training.

Various physics models that exploit the resonant structure of the scattering matrix have been proposed in the literature [14, 25, 29], but are either tailored to specific device geometries [14, 29], making their range of applicability limited, or rely on spectral decompositions without a formal theoretical basis [25], making the physics parameters hard to interpret. However, the general and exact relation between resonances and scattering parameters is, in fact, known and is formalized by the theory of quasinormal modes (QNMs) [49, 50]. We propose a physics-informed NN architecture based on the QNM expansion of the scattering matrix. We will refer to this architecture as the QNM-Net. Owing to the generality of the QNM formalism, the QNM-Net is universal in the sense that it can be applied to linear electromagnetic devices with an arbitrary number of resonances and scattering ports. Furthermore, by implementing the QNM-Net as a modular framework rather than a fixed architecture, we are able to use the rich physics of the QNM formalism to manually tailor the model implementation to the unique physics of each problem. This makes our method simultaneously more general than previous methods while also allowing for stronger problem-specific constraints to be imposed.

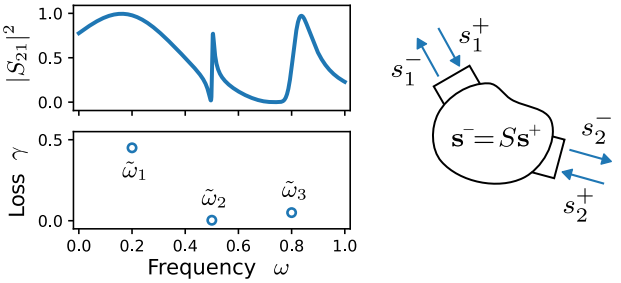

FIGURE 1 | Relation between the S matrix and QNMs of a fictitious 2-port scattering device. Right: The scattering matrix S relates the amplitudes $s_n^{\pm}$ of incoming (+) and outgoing (−) waves. Top left: Resonances are observed when plotting elements of S as a function of frequency. Bottom left: Each resonance corresponds to a complex pole of $S(\omega)$. The poles are eigenfrequencies $\tilde{\omega}_m$ of QNMs.

2 | Quasinormal-Mode-Based Neural Network

QNMs are a generalization of normal modes to lossy systems. They are defined as eigenmodes of the source-free Maxwell's equations

$$\begin{aligned}\nabla \times \tilde{\mathbf{E}}_m &= i\tilde{\omega}_m \mu \tilde{\mathbf{H}}_m,\\ \nabla \times \tilde{\mathbf{H}}_m &= -i\tilde{\omega}_m \varepsilon \tilde{\mathbf{E}}_m,\end{aligned} \tag{2}$$

with outgoing radiation boundary conditions, where $\tilde{\omega}_m$ is the QNM eigenfrequency, m enumerates the QNMs, and ε and μ material parameters describing a linear and local material response. Due to the open boundary and material absorption, the eigenproblem is non-Hermitian and $\tilde{\omega}_m = \omega_m + i\gamma_m$ complex-valued. The imaginary part γ_m corresponds to a decay in time and can be decomposed as $\gamma_m = \gamma_{\mathrm{r},m} + \gamma_{\mathrm{nr},m}$, where $\gamma_{\mathrm{r},m}$ is due to radiative losses and $\gamma_{\mathrm{nr},m}$ is due to nonradiative losses [51, 52].

A connection can be made between $S(\omega)$ and the QNMs by expanding the scattered electromagnetic field as a sum of the QNM fields $\tilde{\mathbf{E}}_m$ and $\tilde{\mathbf{H}}_m$ [49, 50, 53]. This so-called quasinormal mode expansion of the scattering matrix makes it clear that the QNM frequencies $\tilde{\omega}_m$ correspond to resonances in the scattering spectrum, as illustrated in Figure 1. Many different versions of the expansion can be found in the literature [49, 50, 53, 54]. In this work, we use the form

$$S(\omega) = e^{i\omega\tau}[C(\omega) + D(i\omega - i\tilde{\Omega})^{-1}M^{-1}D^{\dagger}C(\omega)]e^{i\omega\tau}, \tag{3}$$

which is approximate but specifically formulated to respect energy conservation even when a finite number of QNMs are included in the expansion [54], making it suitable for our application. In this expression, for an N-port device with P modes, $C(\omega)$ is an $N \times N$ matrix representing a slowly varying background, D is an $N \times P$ matrix with columns $\mathbf{d}_m$, $\tilde{\Omega}$ is a diagonal matrix of QNM frequencies $\tilde{\omega}_m$, and M is a $P \times P$ matrix with elements

$$M_{mp} = \frac{\mathbf{d}_m^{\dagger}\mathbf{d}_p}{i\tilde{\omega}_{\mathrm{r},p} - i\tilde{\omega}_{\mathrm{r},m}^{*}},$$

18638899, 2026, 11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/lpor.202502769 by Chalmers University Of Technology, Wiley Online Library on [16/07/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

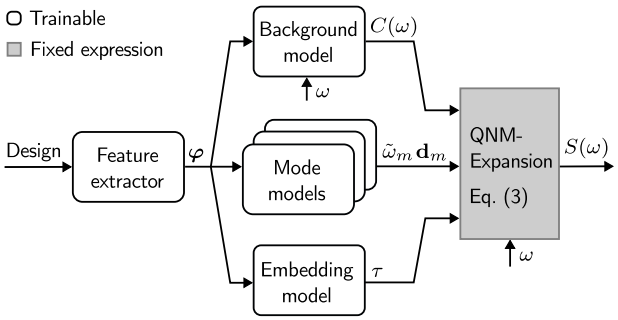

FIGURE 2 | Overview of the QNM-Net architecture. White boxes are submodels with trainable parameters, the gray box is an analytical physics model with no trainable parameters. Arrows indicate the relations between inputs and outputs of each submodel.

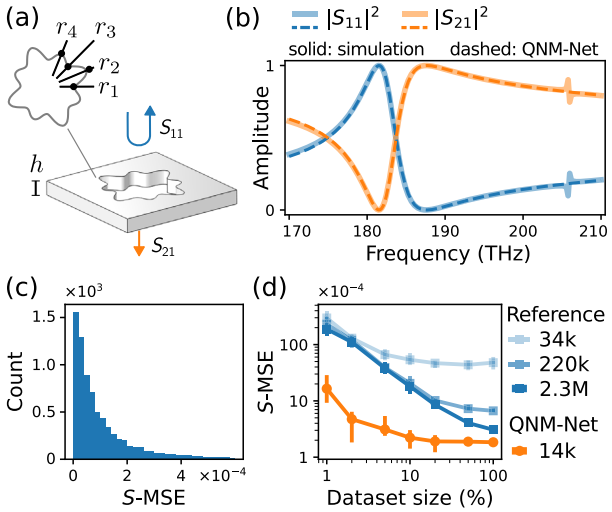

FIGURE 3 | Applying the QNM-Net to PhC slabs. (a) Illustration of the PhC parametrization and geometry. (b) Simulated and QNM-Net-predicted S parameters for a sample from the validation set. (c) Histogram of test losses for the QNM-Net when using 20 % of the dataset. (d) Validation loss as a function of dataset size for the QNM-Net and standard NNs. Markers show cross-validation average, vertical bars show lowest and highest values. Numbers in the legend refer to the number of trainable parameters.

where $\tilde{\omega}_{r,m} = \omega_m + i\gamma_{r,m}$. The N complex-valued elements of $\mathbf{d}_m$ should be interpreted as the amplitudes of the m:th QNM on the ports. Finally, τ is a diagonal $N \times N$ matrix of real numbers τ_n that determine the phase accumulated between the scatterer and the ports. We refer to Reference [54] for a more in-depth description, derivation, and error analysis.

Our proposed physics-informed NN architecture incorporates the QNM expansion as a physics model according to the principle shown in the schematic in Equation (1). The physics parameters are in our case $C(\omega)$ and τ, as well as $\tilde{\omega}_m$ and $\mathbf{d}_m$ for each mode m. We use a modular architecture composed of multiple submodels as illustrated in Figure 2 to facilitate adaptation to a wide range of systems. The first submodel maps the design description to an abstract feature vector $\boldsymbol{\varphi}$. The subsequent submodels use $\boldsymbol{\varphi}$ to predict the physics parameters evaluated at frequency ω. Finally, $S(\omega)$ is computed from the physics parameters using the QNM expansion (3). The internal structure of the submodels is left unspecified and can be tailored to the specific physics of each system.

3 | Results and Discussion

3.1 | Photonic Crystal Slab

We demonstrate the QNM-Net on two example systems. The first example is a photonic crystal (PhC) slab, consisting of a sheet of lossless dielectric with a repeating pattern of holes. We study transmission and reflection at normal incidence. The design is described by five parameters (h, r_1, r_2, r_3, r_4) as illustrated in Figure 3a. Fourfold rotational symmetry is enforced to ensure polarization independence, resulting in a S matrix of size 2×2. The scattering spectrum in the frequency range of interest is dominated by a single Fano resonance on an otherwise smooth background, as exemplified in Figure 3b.

Having defined the system, we now choose the implementation of the QNM-Net submodels based on our prior knowledge about its physics. Consulting the diagram in Figure 2, the background model should take $\boldsymbol{\varphi}$ and ω as inputs and return the corresponding background matrix $C(\omega)$. Because the PhC is reciprocal, lossless, and mirror symmetric, C satisfies $C^T = C$, $C^\dagger C = I$, and $C_{11} = C_{22}$ [54], so we set

$$C = \begin{pmatrix} i\sin\alpha & \cos\alpha \\ \cos\alpha & i\sin\alpha \end{pmatrix} e^{i\alpha}. \tag{4}$$

To allow for a design-dependent slow variation with frequency, we let $\alpha = \alpha(\omega)$ be a linear function of ω with values at 170 THz and 210 THz predicted from $\boldsymbol{\varphi}$ using a single-layer NN. Next, the mode models should predict the QNM frequencies $\tilde{\omega}_m$ and port amplitudes $\mathbf{d}_m$ from $\boldsymbol{\varphi}$, with each mode model corresponding to one mode in the QNM expansion. Since the scattering spectrum of the PhC is well approximated by a single resonance, we only use one mode model in this case. Furthermore, because the PhC is reflection symmetric, the QNMs are either even or odd, corresponding to $\mathbf{d}_m = (1, 1)$ or $\mathbf{d}_m = (1, -1)$, respectively [49, 54]. The observed Fano mode is even, so we set $\mathbf{d}_1 = (1, 1)$. The material of the PhC being lossless implies $\gamma_{nr,1} = 0$. We let the remaining unknown and design-dependent mode parameters ω_1 and $\gamma_{r,1}$ be predicted from $\boldsymbol{\varphi}$ using a single-layer NN, with causality enforced by applying a strictly positive activation function to $\gamma_{r,1}$ [54]. The embedding model should predict τ_n from $\boldsymbol{\varphi}$. We use a single-layer NN to predict τ_1 from $\boldsymbol{\varphi}$ and set $\tau_2 = \tau_1$ to respect reflection symmetry. Finally, we use a fully-connected feed-forward NN in the feature extractor to provide the bulk of the trainable weights of the network. A linear skip connection was added to improve training robustness and convergence speed [14]. A detailed schematic of all the submodels is provided in the Supporting Information.

To train the QNM-Net, we generated a dataset of 10000 random PhC designs and their corresponding scattering spectra $S(\omega)$ at 200 equally spaced frequency points using frequency-domain finite-element method simulations. The QNM-Net was trained to minimize the S-parameter mean-squared-error (S-MSE), defined

18638899, 2026, 11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/lpor.202502769 by Chalmers University Of Technology, Wiley Online Library on [16/07/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

as

$$S-\mathrm{MSE}=\frac{1}{N_S N_\omega}\sum_{i,j,k}|S_{ij}^{\mathrm{predicted}}(\omega_k)-S_{ij}^{\mathrm{true}}(\omega_k)|^2,$$

where N_S is the number of elements S_{ij} in the scattering matrix and N_ω is the number of sampled frequencies ω_k. Gradients were automatically backpropagated through the entire model, including the QNM-expansion, using automatic differentiation. Models were trained on 80 % of the dataset, while the remaining 20 % was used for validation. The validation error was averaged over five dataset splits. More details about the neural-network implementation and training procedures are provided in the Supporting Information.

The trained QNM-Net is able to reproduce the simulated scattering accurately, as exemplified by the sample shown in Figure 3b. The sample has an S-MSE of 2.1×10^{-4}, which is close to the mean of the loss histogram shown in Figure 3c. The histogram was evaluated on a test set separate from both the validation and training sets to avoid bias. For reference, we compare the QNM-Net to three standard fully-connected feed-forward NNs with varying numbers of trainable parameters, similar to architectures used in forward models in the literature [19, 22, 23, 27, 38]. Figure 3d shows the validation loss for all models at different dataset sizes. The parameter count of all models are listed in the legend of the figure. The QNM-Net requires only 2 % of the dataset, corresponding to 160 training samples, to achieve an S-MSE less than 10^{-3}. The reference models require approximately one order of magnitude more data and trainable parameters to achieve a similar loss. This shows that the QNM-Net can reduce the need for training data significantly by learning the physics of the system in terms of its QNMs.

Another benefit of physics-informed models operating according to the principle shown in the schematic in Equation (1) is that their predictions are explainable in terms of the network-predicted physics parameters [29], which for the QNM-Net are $C(\omega)$, $\tilde{\omega}_m$, $\mathbf{d}_m$, and τ_n. Based on the theoretical foundations of the QNM expansion, we know that $\tilde{\omega}_m$ correspond exactly to eigenfrequencies of Maxwell's equations (2). Thus, the accuracy of the learned physics can be verified by comparing the network-predicted $\tilde{\omega}_m$ to eigenfrequencies calculated using full-wave eigenmode simulations. Figure 4 shows the network-predicted $\tilde{\omega}_1$ and the closest simulated eigenfrequency for 100 new randomly generated designs for the PhC slab example. The two methods agree to high precision, indicating that the network training has converged to a representation of the spectra that matches the correct underlying physics for this system. It should be noted that, although finding the correct QNM parameters should lead to vanishing S-MSE and thus correspond to minima in the NN loss landscape, there is no guarantee that these minima are unique or that the training procedure will always find them. Thus, validating the trained model with a reference eigenmode solver as shown here is necessary if the accuracy of the learned physics is of importance.

Since the QNM-Net is able to accurately predict the QNM parameters of the PhC, these can now be used as design targets for inverse design. To demonstrate this, we inverse design five

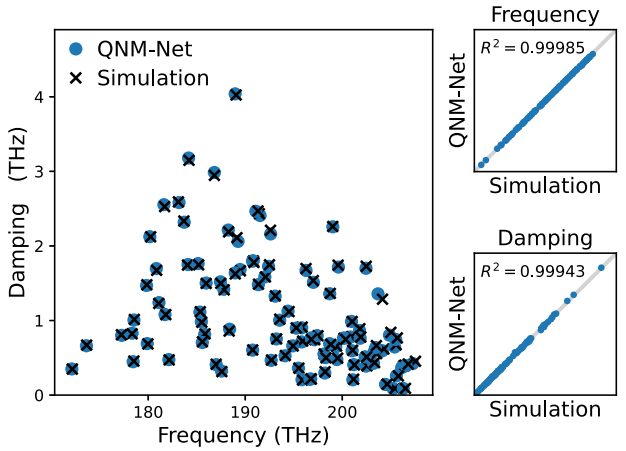

FIGURE 4 | Comparison between $\tilde{\omega}_1$ learned by the QNM-Net and calculated by full-wave eigenmode simulations for the PhC geometry. Left: QNM frequency and damping for 100 randomly generated designs. Each dot/cross corresponds to a separate design. Right: Parity plots for frequency and damping with corresponding coefficient of determination.

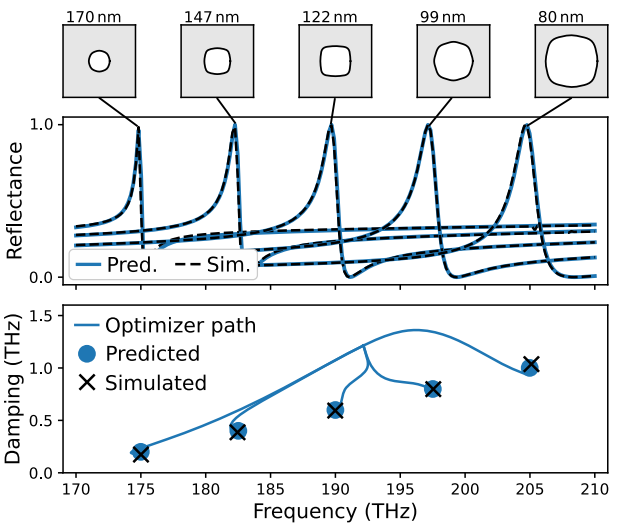

FIGURE 5 | Inverse design of PhC slabs. The PhC geometries in the top row were optimized for linearly spaced complex eigenfrequencies. The thickness of the PhC slab is displayed above each design. QNM-Net predictions (blue) and full-wave simulations (black) of the PhC reflectivity spectra and eigenmodes are shown in the middle and lower panels.

PhC slabs with linearly increasing eigenfrequency and loss rate. Starting from an initial design with all design parameters set to zero, the difference between model-predicted and desired value of $\tilde{\omega}_1$ was minimized using the Adam optimizer [55] with gradients backpropagated through the network using automatic differentiation. The optimization converged to the desired eigenfrequency after a few hundred steps in less than one second. The final designs were simulated using the full-wave solver as verification. The predicted and simulated results shown in Figure 5 match closely, demonstrating that a QNM-Net trained solely on spectral data can be used for fast and accurate inverse design of eigenfrequencies.

18638899, 2026, 11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/lpor.202502769 by Chalmers University Of Technology, Wiley Online Library on [16/07/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

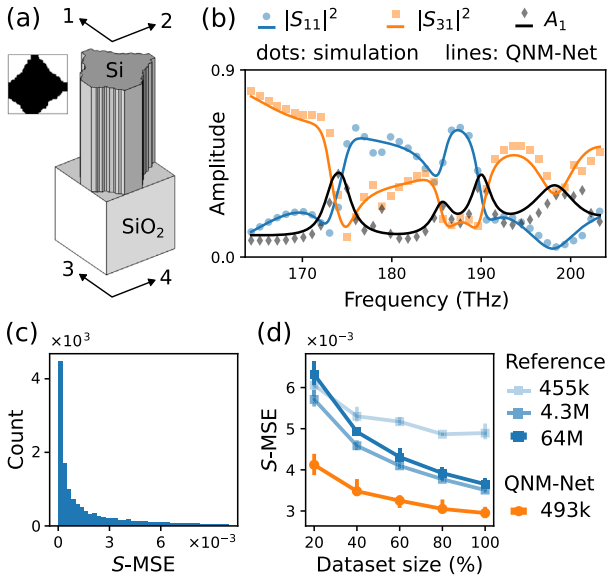

FIGURE 6 | QNM-Net applied to free-form metasurface. (a) Illustration of a metasurface unit cell and the corresponding bitmap. Numbered arrows indicate the locations and polarizations of the ports. (b) Simulated and QNM-Net-predicted spectrum. $A_1 = 1 - \Sigma_n |S_{n1}|^2$ is the absorbed power fraction for incidence on port 1. (c) Histogram of QNM-Net test losses for dataset size 80 %. (d) Validation loss as a function of dataset size for the QNM-Net and reference models. Numbers in the legend refer to the number of trainable parameters.

3.2 | Free-Form Dielectric Metasurface

To investigate if the QNM-Net approach is also beneficial for more complex geometries when less is known about the physics, we apply the QNM-Net to free-form all-dielectric metasurfaces. We use a dataset introduced in Reference [56] containing roughly 80000 meta-atoms and their corresponding scattering spectra (more details are available in the Supporting Information). The geometry is described by a 100×100 binary bitmap representing the etching mask for the metasurface unit cell, as illustrated in Figure 6a. As far as we know, no previous works have applied NNs with resonance-based physics models to free-form designs. This system has a vastly larger design space, lower frequency resolution, a polarization-dependent scattering response, multiple overlapping resonances, nonradiative loss, and lacks mirror symmetry due to the presence of a substrate. Furthermore, only two columns of the S matrix, corresponding to light incident from the top, are available in the dataset.

We again adapt the QNM-Net submodel implementation to our prior knowledge about the system. Because the design parameterization now has an inherent spatial structure, we use a version of the convolutional DenseNet architecture [57] which we previously found to perform well on this dataset [34] as the feature extractor. In this case, there is no clearly distinguishable background. We therefore set $C = -I$, which is commonly used in QNM expansions [49, 54] and let the QNM-Net fit the background using many broad resonances. The spectra have a varying number of modes within the sampled frequency range. We chose to include 20 mode models in the QNM-Net, which we deemed to be more than sufficient based on manual inspection of a few simulated spectra. If a smaller number of modes are necessary to explain a given spectrum, the QNM-Net can learn to place the superfluous modes outside the sampled range, so that they do not contribute to the predicted spectrum. For each mode, the QNM frequency and damping rates, as well as the real and imaginary parts of $\mathbf{d}_m$, are predicted from $\boldsymbol{\varphi}$ by small NNs. Each mode model has separate trainable weights. Fixed port delays τ_n are used to account for the distance between the metasurface boundary and the ports.

Subsequently, we compare the QNM-Net to reference models consisting of DenseNets followed by branched feed-forward NNs, similar to the architecture used in our previous work [34]. The same DenseNet architecture was used for all models, but the width of the fully-connected layers was varied to change the number of trainable parameters. Figure 6b shows a simulated spectrum and the corresponding QNM-Net prediction for a sample from the validation set. The S-MSE of the prediction is 2.4×10^{-3}, which is typical compared to the histogram shown in Figure 6c. We observe that the QNM-Net reproduces most resonances, but fails to predict a few weaker resonances. We believe this is due to the design space being so large that even tens of thousands of training samples are not enough to accurately learn all features of the spectra. Nevertheless, Figure 6d shows that the QNM-Net requires only around one third of the data to achieve a loss as low as that of the best reference models, demonstrating that the method is useful for spectrum prediction also for this system even if all predicted resonances are not accurate.

To further investigate the accuracy of the learned QNMs, we conduct an eigenmode study for three metasurface designs from the validation set with varying number of resonances in the sampled frequency range. The QNM-Net-predicted eigenfrequencies together with all eigenfrequencies returned by an eigenmode solver with outgoing radiation boundary conditions are shown in Figure 7. Most modes predicted by the QNM-Net inside the sampled frequency range correspond to true eigenmodes found by the simulator. Modes outside this range and high-damping modes close to the edges are generally not accurate, which is expected since there is very little information available in the data to learn these modes. Interestingly, designs that connect the edges of the unit cell (the leftmost and rightmost designs in Figure 7 exhibit several low-damping modes with no clear corresponding features in the scattering spectra and no corresponding modes predicted by the QNM-Net. We interpret these as modes that are very weakly coupled to the incoming radiation and therefore do not contribute significantly to the scattering spectrum. This illustrates that the QNM-Net automatically identifies the modes that are the most important for the scattering response, whereas eigenmode solvers return very large numbers of modes that would otherwise be cumbersome to filter manually.

4 | Conclusion and Further Research

In conclusion, we have demonstrated a modular physics-informed NN architecture for electromagnetic scattering based on the QNM expansion of the scattering matrix. We have shown that this architecture requires significantly less trainable parameters and is substantially more data efficient than conventional NNs for prediction of scattering spectra. Compared to previous physics-informed NNs for electromagnetic scattering [14, 25, 29], we

18638899, 2026, 11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/lpor.202502769 by Chalmers University Of Technology, Wiley Online Library on [16/07/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

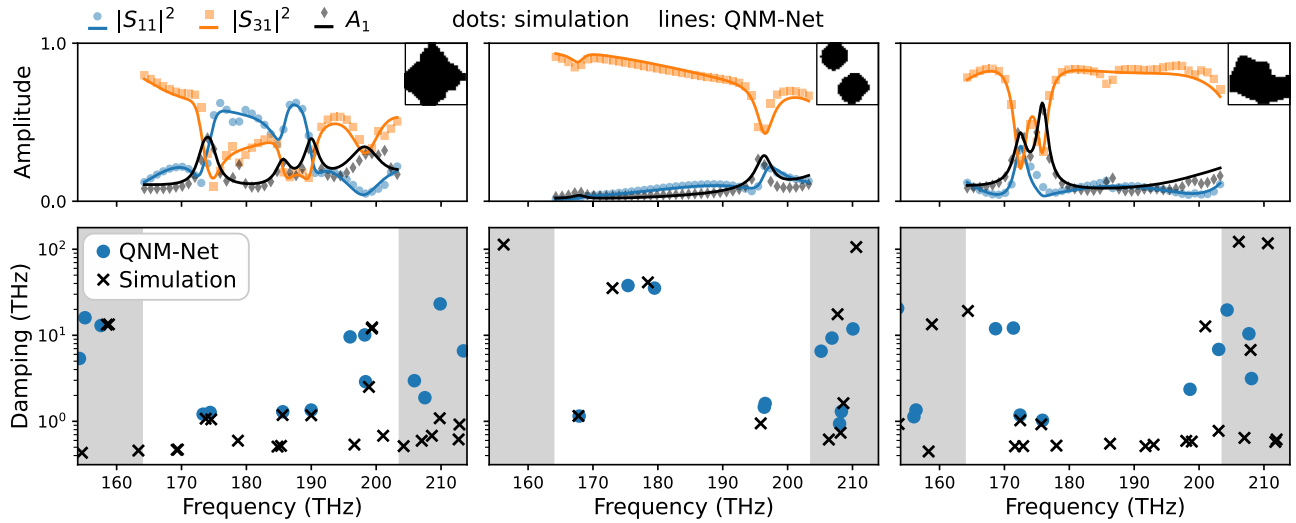

FIGURE 7 | Eigenmode study for the metasurface example. Each column shows the QNM-Net-predicted and simulated spectrum (top row) and eigenfrequencies (bottom row) of a different design (top-right inset) from the validation set. Shaded regions in the bottom row correspond to frequencies outside the training range.

expect the QNM-Net to perform similarly in terms of accuracy and data efficiency while having the benefit of being applicable to a wider range of systems and a more well-founded theoretical basis in the formalism of quasinormal modes. In particular, the QNM frequencies learned by the model correspond directly to the eigenmodes of the scatterer, which was verified by comparing the learned physics to eigenmode simulations. We believe that the method will be useful for modeling a wide range of electromagnetic systems where the scattering spectrum is accurately described by a set of resonances.

The method presented here opens the door to further developments in machine learning for electromagnetic scattering, which previously was limited to overly simplified systems or required an amount of training data that was difficult to obtain in practice. The significantly improved data efficiency can make it practically feasible to train models on experimental data for simple designs. In the Supporting Information, we add simulated noise to the training data to demonstrate that the QNM-Net is more robust to noise than standard NNs, which indicates that the QNM-Net may have additional advantages in this scenario. The QNM-Net can also be combined with methods such as those presented in References [58, 59] to account for fabrication-induced performance degradation in inverse design. Furthermore, investigating the learned parameters can provide insights into the underlying physics for automated knowledge discovery. For example, if the reflection symmetry of the PhC mode had not been included as a prior constraint, it could have been discovered by looking at the model predictions for $\mathbf{d}_1$.

In future studies, it would be interesting to investigate the effect of individual constraints to determine which inductive biases are most important for data efficiency. Another possible development is to use transfer learning of the QNM-Net submodels for further improvements to data efficiency. For example, the feature extractor could be pretrained as the encoder network of an autoencoder. The rigorous connection between the QNM expansion and electromagnetic eigenmodes also makes it possible to train the QNM-Net on eigenmode simulations. This approach could be particularly useful for sharp resonances that are not easily resolved in the frequency domain.

Acknowledgements

We acknowledge partial financial support from Chalmers' area of advance Nano (Excellence Grant), from the Swedish Research Council under Grant No. 2020-05284, and from the Knut och Alice Wallenberg Stiftelse under Grant No. 2022.0090. Training data generation and NN training were performed on resources provided by the Swedish National Infrastructure for Computing (NAISS), at the Chalmers/C3SE and KTH/PDC sites, partially funded by the Swedish Research Council under Grant No. 2022-06725. The work was performed in part within the framework of the Excellence Center META-PIX (Grant No. 2022-06575).

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

References

1. I. N. Da Silva, D. H. Spatti, R. A. Flauzino, L. H. B. Liboni, and S. F. dos Reis Alves, *Artificial Neural Networks: A Practical Course* (Springer, 2017).

2. B. Mehlig, *Machine Learning with Neural Networks: An Introduction for Scientists and Engineers* (Cambridge University Press, 2021).

3. P. R. Wiecha, A. Lecestre, N. Mallet, and G. Larrieu, "Pushing the Limits of Optical Information Storage Using Deep Learning," *Nature Nanotechnology* 14 (2019): 237–244.

4. R. Iten, T. Metger, H. Wilming, L. Del Rio, and R. Renner, "Discovering Physical Concepts with Neural Networks," *Physical Review Letters* 124 (2020): 010508.

18638899, 2026, 11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/lpor.202502769 by Chalmers University Of Technology, Wiley Online Library on [16/07/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

5. V. Peano, F. Sapper, and F. Marquardt, "Rapid Exploration of Topological Band Structures Using Deep Learning," *Physical Review X* 11 (2021): 021052.

6. P. R. Wiecha, A. Arbouet, C. Girard, and O. L. Muskens, "Deep Learning in Nano-Photonics: Inverse Design and Beyond," *Photonics Research* 9 (2021): B182.

7. O. Khatib, S. Ren, J. Malof, and W. J. Padilla, "Deep Learning the Electromagnetic Properties of Metamaterials–A Comprehensive Review," *Advanced Functional Materials* 31 (2021): 2101748.

8. J. Jiang, M. Chen, and J. A. Fan, "Deep Neural Networks for the Evaluation and Design of Photonic Devices," *Nature Reviews Materials* 6 (2021): 679–700.

9. M. Zandehshahvar, Y. Kiarashi, M. Chen, R. Barton, and A. Adibi, "Inverse Design of Photonic Nanostructures Using Dimensionality Reduction: Reducing the Computational Complexity," *Optics Letters* 46 (2021): 2634–2637.

10. M. Zandehshahvar, Y. Kiarashinejad, M. Zhu, H. Maleki, T. Brown, and A. Adibi, "Manifold Learning for Knowledge Discovery and Intelligent Inverse Design of Photonic Nanostructures: Breaking the Geometric Complexity," *ACS Photonics* 9 (2022): 714–721.

11. N. Mohseni, T. Fösel, L. Guo, C. Navarrete-Benlloch, and F. Marquardt, "Deep Learning of Quantum Many-Body Dynamics via Random Driving," *Quantum* 6 (2022): 714.

12. J. Lee, D. Park, M. Lee, et al., "Machine Learning-Based Inverse Design Methods Considering Data Characteristics and Design Space Size in Materials Design and Manufacturing: A Review," *Materials Horizons* 10 (2023): 5436–5456.

13. M. Sanchez, C. Everly, and P. A. Postigo, "Advances in Machine Learning Optimization for Classical and Quantum Photonics," *Journal of the Optical Society of America B* 41 (2024): A177–A190.

14. Y. Xu, J. Q. Yang, K. Fan, et al., "Physics-Informed Inverse Design of Programmable Metasurfaces," *Advanced Science* 11 (2024): 2406 878.

15. L. Su, D. Vercruysse, J. Skarda, N. V. Sapra, J. A. Petykiewicz, and J. Vučković, "Nanophotonic Inverse Design with SPINS: Software Architecture and Practical Considerations," *Applied Physics Reviews* 7 (2020): 011407.

16. O. Tsilipakos, G. Perrakis, M. Farsari, and M. Kafesaki, "Polymeric Optical Metasurfaces by Two-Photon Lithography: Practical Designs for Beam Steering," in *Metamaterials 2024* (2024): 1–3.

17. R. S. Hegde, "Photonics Inverse Design: Pairing Deep Neural Networks with Evolutionary Algorithms," *IEEE Journal of Selected Topics in Quantum Electronics* 26 (2019): 1–8.

18. S. So, T. Badloe, J. Noh, J. Bravo-Abad, and J. Rho, "Deep Learning Enabled Inverse Design in Nanophotonics," *Nanophotonics* 9 (2020): 1041–1057.

19. Y. Deng, S. Ren, K. Fan, J. M. Malof, and W. J. Padilla, "Neural-Adjoint Method for the Inverse Design of All-Dielectric Metasurfaces," *Optics Express* 29 (2021): 7526–7534.

20. Y. Yan, F. Li, J. Shen, et al., "Highly Intelligent Forward Design of Metamaterials Empowered by Circuit-Physics-Driven Deep Learning," *Laser Photonics Reviews* 19 (2024): 2400724.

21. C. C. Nadell, B. Huang, J. M. Malof, and W. J. Padilla, "Deep Learning for Accelerated All-Dielectric Metasurface Design," *Optics Express* 27 (2019): 27 523–27 535.

22. S. Sarkar, A. Ji, Z. Jermain, et al., "Physics-Informed Machine Learning for Inverse Design of Optical Metamaterials," *Advanced Photonics Research* 4 (2023): 2300158.

23. L. Xu, M. Rahmani, Y. Ma, et al., "Enhanced Light–Matter Interactions in Dielectric Nanostructures via Machine-Learning Approach," *Advanced Photonics* 2 (2020): 026003.

24. Y. Jing, H. Chu, B. Huang, J. Luo, W. Wang, and Y. Lai, "A Deep Neural Network for General Scattering Matrix," *Nanophotonics* 12 (2023): 2583–2591.

25. A. P. Blanchard-Dionne and O. J. F. Martin, "Teaching Optics to a Machine Learning Network," *Optics Letters* 45 (2020): 2922–2925.

26. T. Zhang, C. Y. Kee, Y. S. Ang, E. Li, and L. K. Ang, "Symmetry Enhanced Network Architecture Search for Complex Metasurface Design," *IEEE Access* 10 (2022): 73 533–73 547.

27. S. So, J. Mun, and J. Rho, "Simultaneous Inverse Design of Materials and Structures via Deep Learning: Demonstration of Dipole Resonance Engineering Using Core–Shell Nanoparticles," *ACS Applied Materials & Interfaces* 11 (2019): 24264–24268.

28. X. Ma, Y. Ma, P. Cunha, et al., "Strategical Deep Learning for Photonic Bound States in the Continuum," *Laser Photonics Reviews* 16 (2022): 2100658.

29. O. Khatib, S. Ren, J. Malof, and W. J. Padilla, "Learning the Physics of All-Dielectric Metamaterials with Deep Lorentz Neural Networks," *Advanced Optical Materials* 10 (2022): 2200097.

30. W. Ma, F. Cheng, Y. Xu, Q. Wen, and Y. Liu, "Probabilistic Representation and Inverse Design of Metamaterials Based on a Deep Generative Model with Semi-Supervised Learning Strategy," *Advanced Materials* 31 (2019): 1901111.

31. J. Guimbao, L. Sanchis, L. Weituschat, et al., "Numerical Optimization of a Nanophotonic Cavity by Machine Learning for Near-Unity Photon Indistinguishability at Room Temperature," *ACS Photonics* 9 (2022): 1926–1935.

32. Z. Liu, D. Zhu, S. P. Rodrigues, K. T. Lee, and W. Cai, "Generative Model for the Inverse Design of Metasurfaces," *Nano Letters* 18 (2018): 6570–6576.

33. J. Xu, P. Xu, Z. Yang, et al., "Freeform Metasurface Design with a Conditional Generative Adversarial Network," *Applied Physics A* 130 (2024): 530.

34. T. Gahlmann and P. Tassin, "Deep Neural Networks for the Prediction of the Optical Properties and the Free-Form Inverse Design of Metamaterials," *Physical Review B* 106 (2022): 085408.

35. C. Yeung, B. Pham, R. Tsai, K. T. Fountaine, and A. P. Raman, "DeepAdjoint: An All-in-One Photonic Inverse Design Framework Integrating Data-Driven Machine Learning with Optimization Algorithms," *ACS Photonics* 10 (2023): 884–891.

36. M. H. Tahersima, K. Kojima, T. Koike-Akino, et al., "Deep Neural Network Inverse Design of Integrated Photonic Power Splitters," *Scientific Reports* 9 (2019): 1368.

37. Y. Qian, B. Ni, Z. Feng, et al., "Deep Learning for the Design of Random Coding Metasurfaces," *Plasmonics* 18 (2023): 1941–1948.

38. D. Liu, Y. Tan, E. Khoram, and Z. Yu, "Training Deep Neural Networks for the Inverse Design of Nanophotonic Structures," *ACS Photonics* 5 (2018): 1365–1369.

39. A. V. Kildishev, O. Yesilyurt, V. Mkhitaryan, J. S. Huang, and I. Shadrivov, "Inverse-Designed Superchiral Dielectric Metasurfaces," in *Metamaterials, Metadevices, and Metasystems 2024* (2024): PC1310916.

40. G. You, C. Qian, S. Tan, E. Li, and H. Chen, "Driving Deep-Learning-Based Metasurface Design with Kramers-Kronig Relations," *Physical Review Applied* 22 (2024): L041002.

41. Y. Gao, W. Chen, F. Li, et al., "Meta-Attention Deep Learning for Smart Development of Metasurface Sensors," *Advanced Science* 11 (2024): 2405750.

42. J. L. Su, J. W. You, L. Chen, et al., "MetaPhyNet: Intelligent Design of Large-Scale Metasurfaces Based on Physics-Driven Neural Network," *Journal of Physics: Photonics* 6 (2024): 035010.

43. Z. A. Kudyshev, A. V. Kildishev, V. M. Shalaev, and A. Boltasseva, "Machine Learning–Assisted Global Optimization of Photonic Devices," *Nanophotonics* 10 (2021): 371–383.

18638899, 2026, 11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/lpor.202502769 by Chalmers University Of Technology, Wiley Online Library on [16/07/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

44. S. W. Kim, I. Kim, J. Lee, and S. Lee, "Knowledge Integration into Deep Learning in Dynamical Systems: An Overview and Taxonomy," *Journal of Mechanical Science and Technology* 35 (2021): 1331–1342.

45. G. E. Karniadakis, I. G. Kevrekidis, L. Lu, P. Perdikaris, S. Wang, and L. Yang, "Physics-Informed Machine Learning," *Nature Reviews Physics* 3 (2021): 422–440.

46. J. Zhang, J. W. You, F. Feng, et al., "Physics-Driven Machine-Learning Approach Incorporating Temporal Coupled Mode Theory for Intelligent Design of Metasurfaces," *IEEE Transactions on Microwave Theory and Techniques* 71 (2023): 2875–2887.

47. X. Ding, V. Devabhaktuni, B. Chattaraj, et al., "Neural-Network Approaches to Electromagnetic-Based Modeling of Passive Components and Their Applications to High-Frequency and High-Speed Nonlinear Circuit Optimization," *IEEE Transactions on Microwave Theory and Techniques* 52 (2004): 436–449.

48. S. Liu, F. Feng, X. Huang, et al., "Novel Neuro-Coupling Matrix Technique for Parametric Modeling of Microwave Filters," *IEEE Microwave and Wireless Technology Letters* 34 (2024): 871–874.

49. F. Alpeggiani, N. Parappurath, E. Verhagen, and L. Kuipers, "Quasinormal-Mode Expansion of the Scattering Matrix," *Physical Review X* 7 (2017): 021035.

50. H. Zhang and O. D. Miller, "Quasinormal Coupled Mode Theory," (2020): 2010.08650.

51. P. Lalanne, W. Yan, K. Vynck, C. Sauvan, and J. P. Hugonin, "Light Interaction with Photonic and Plasmonic Resonances," *Laser Photonics Reviews* 12 (2018): 1700113.

52. P. T. Kristensen, K. Herrmann, F. Intravaia, and K. Busch, "Modeling Electromagnetic Resonators Using Quasinormal Modes," *Advances in Optics and Photonics* 12 (2020): 612–708.

53. T. Weiss and E. A. Muljarov, "How to Calculate the Pole Expansion of the Optical Scattering Matrix from the Resonant States," *Physical Review B* 98 (2018): 085433.

54. M. Benzaouia, J. D. Joannopoulos, S. G. Johnson, and A. Karalis, "Quasi-Normal Mode Theory of the Scattering Matrix, Enforcing Fundamental Constraints for Truncated Expansions," *Physical Review Research* 3 (2021): 033228.

55. D. P. Kingma and J. Ba, "Adam: A Method for Stochastic Optimization," *arXiv* (2017): 1412.6980.

56. T. Gahlmann and P. Tassin, "Evaluation of Machine Learning Techniques for Conditional Generative Adversarial Networks in Inverse Design," *arXiv* (2025): 2502.11934.

57. G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely Connected Convolutional Networks," in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (2017): 2261–2269.

58. S. Raza, M. Hammood, N. A. F. Jaeger, and L. Chrostowski, "Fabrication-Aware Inverse Design with Shape Optimization for Photonic Integrated Circuits," *Optics Letters* 50, no. 1 (2025): 117–120.

59. R. P. Jenkins, S. D. Campbell, and D. H. Werner, "Establishing Exhaustive Metasurface Robustness Against Fabrication Uncertainties Through Deep Learning," *Nanophotonics* 10, no. 18 (2021): 4497–4509.

Supporting Information

Additional supporting information can be found online in the Supporting Information section.

Supporting File: lpor71072-sup-0001-SuppMat.pdf.

18638899, 2026, 11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/lpor.202502769 by Chalmers University Of Technology, Wiley Online Library on [16/07/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License