Neural Contrast Expansion for Explainable Structure-Property Prediction and Random Microstructure Design

Guangyu Nie[1] **Yang Jiao**[2] **Yi Ren**[1]
[1]Mechanical Engineering, Arizona State University
[2]Materials Science, Arizona State University
Tempe, AZ 85287, USA
{gnie1,yjiao13,yiren}@asu.edu

Abstract

Effective properties of composite materials are defined as the ensemble average of property-specific PDE solutions over the underlying microstructure distributions. Traditionally, predicting such properties can be done by solving PDEs derived from microstructure samples or building data-driven models that directly map microstructure samples to properties. The former has a higher running cost, but provides explainable sensitivity information that may guide material design; the latter could be more cost-effective if the data overhead is amortized, but its learned sensitivities are often less explainable. With a focus on properties governed by linear self-adjoint PDEs (e.g., Laplace, Helmholtz, and Maxwell curl-curl) defined on bi-phase microstructures, we propose a structure-property model that is both cost-effective and explainable. Our method is built on top of the strong contrast expansion (SCE) formalism, which analytically maps N-point correlations of an unbounded random field to its effective properties. Since real-world material samples have finite sizes and analytical PDE kernels are not always available, we propose Neural Contrast Expansion (NCE), an SCE-inspired architecture to learn surrogate PDE kernels from structure-property data. For static conduction and electromagnetic wave propagation cases, we show that NCE models reveal accurate and insightful sensitivity information useful for material design. Compared with other PDE kernel learning methods, our method does not require measurements about the PDE solution fields, but rather only requires macroscopic property measurements that are more accessible in material development contexts.

1 Introduction

This paper is concerned with *effective properties* of composite materials, which are macroscopic measurements affected by the materials' random microstructure [1]. We seek a model trained on structure–property data that *explains* what and how microstructural features affect the effective property, providing insights for microstructural design. Some structure–property relations are intuitive (e.g., the effective conductivity of a composite is dominated by the connectivity of its conductive filler [2, 3]); others have been derived from first principles by materials scientists, e.g., minimizing energy dissipation of light propagation within an optical sensor during wave propagation can be achieved by a hyperuniform arrangement of the random microstructure [4, 5]. However, for novel materials and target properties, such human-comprehensible knowledge may not be accessible from first principles due to the multiscale and nonlinear nature of the governing equations.

39th Conference on Neural Information Processing Systems (NeurIPS 2025) Workshop: Accelerated Materials Design.

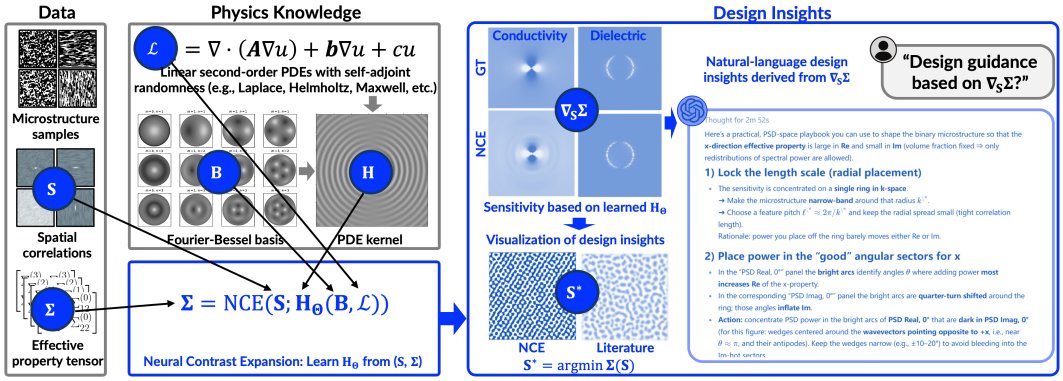


Figure 1: This paper improves the workflow for deriving composite material design insights from measurement data regarding microstructures ("S") and their effective properties ("$\boldsymbol{\Sigma}$"). We propose Neural Contrast Expansion (NCE), which learns PDE kernels ("$\mathbf{H}$") based on data and physics knowledge (PDE "$\mathscr{L}$" and basis "$\mathbf{B}$") to provide accurate sensitivity information ("$\nabla_S \boldsymbol{\Sigma}$"). The sensitivity derived from NCE is much more accurate than that of purely data-driven models (see Fig. 4), and in turn, provides explainable design insights through LLM and optimal microstructure design ("S^*"). For optical material design with minimal energy loss, design insights derived from NCE lead to hyperuniform microstructures, consistent with recent literature [16].

Conventionally, property prediction is achieved by either solving the governing PDEs or building data-driven surrogates: The former requires knowledge about the PDEs and often a significant amount of compute. Although when differentiable, the solution process yields explainable sensitivities; the latter is more cost-effective with amortized data overhead, but generally lacks explainable sensitivities because it is detached from the physics. Structure–property mappings that leverage advantages of both have been developed in the context of effective-medium theorems [6, 4]; among these, we focus on the strong contrast expansion (SCE), which is accurate and explicit for linear PDEs. SCE states that *linear* effective properties of *unbounded* random fields are analytical functions given by spatial convolutions between the PDE kernel (i.e., Green's functions or their Hessians) and the infinite series of N-point correlation functions (NPCFs) [1, 7, 8], which characterize random fields at increasing orders and in semantic ways (Fig. 2 and [9]). For example, connectivity of a contrast phase within a composite material is captured by its 2-point correlation function [10, 11] (Fig. 2). However, SCE has practical limitations: real samples are finite (introducing boundary effects not captured by SCE), and the relevant kernels may lack closed forms or may not exist.

Method. To address these limitations, we investigate Neural Contrast Expansion (NCE), a learnable model with an SCE-guided architecture. We encode physics by mapping structure to property as a convolution of NPCFs with a learnable kernel, and estimate that kernel from structure–property data. We use two-dimensional static conduction and electromagnetic wave propagation in bi-phase composites as case studies, whose governing equations (Laplace and Maxwell curl-curl, respectively) have analytic kernels under mild conditions. Numerical results show that NCE generalizes comparably to convolutional neural networks and neural operators, while producing sensitivities that are substantially more explainable and accurate than purely data-driven counterparts. Compared with supervised methods that learn Green's functions or closures [12, 13, 14], NCE does not require measurements of PDE solution fields; it only needs macroscopic property measurements that are more accessible in practice. Compared with physics-informed methods for learning Green's functions [15], NCE avoids numerical issues introduced by the singularity of Green's functions through a cavity technique inherited from SCE.

Scope and limitations. In this work we restrict attention to randomness entering the zeroth- or second-order terms of linear second-order PDEs. These two classes admit a unified SCE because their perturbations appear as self-adjoint modifications to a coercive baseline operator and generate even convolutional kernel structures that close under ensemble averaging. In contrast, first-order randomness (random drift/advection) produces skew-adjoint, odd-kernel perturbations that, in general, break this closure: the resulting effective operator depends on nonlocal flow statistics rather than solely on local correlations, and the convergence of SCE is generally not guaranteed (see App. A.2 for explanation). We note that there are niche yet valuable applications of NCE for the focused

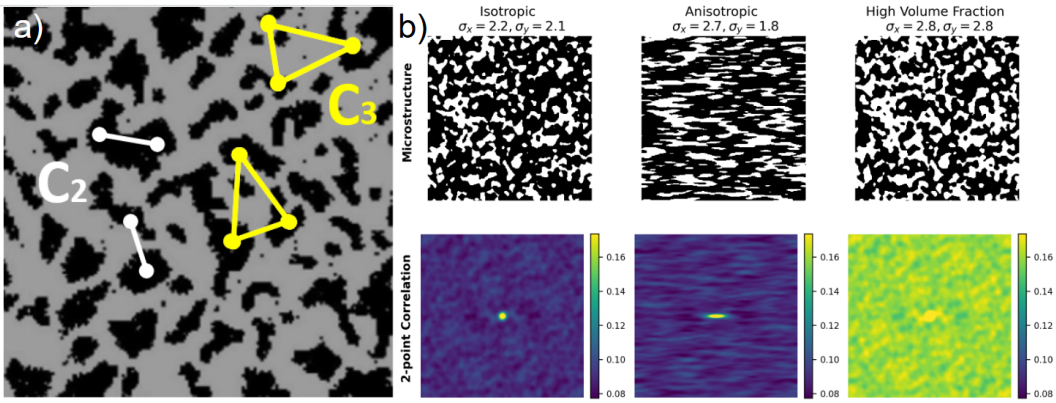


Figure 2: (a) Illustration of two- and three-point configurations ($\mathcal{C}_2$ and $\mathcal{C}_3$) in a binary microstructure. The black and gray regions represent two distinct phases of the medium, while the white line segments and yellow triangles mark representative 2-point and 3-point configurations. (b) Binary microstructures with distinct features: (left) isotropic, (middle) anisotropic with horizontal elongation, and (right) high-volume fraction. The parameters σ_x and σ_y indicate the effective directional conductivity along the x and y axes. The second row shows the corresponding 2PCFs.

randomness classes beyond materials science: For climate science, elliptic/parabolic PDEs for heat, moisture, or tracer transport (e.g., diffusion or advection–diffusion) include spatially heterogeneous parameter fields (e.g., diffusivities, reaction rates, subgrid closures). NCE could be used to learn kernels that predict how fast heat, moisture, or pollution spreads and mixes when tiny swirls of air or water cannot be resolved. For neuroscience, the quasi-static electromagnetic forward problem is modeled by Laplace/Poisson equations with spatially varying and anisotropic tissue conductivities. When solution fields (e.g., MRI) are unavailable, NCE could be used to estimate how electrical signals travel through different parts of the brain using sensor-level measurements. For cosmology, the gravitational potential obeys the Poisson equation and the matter density is a random field with rich N-point correlations. NCE can learn kernels that map these correlations to large-scale effective observables (e.g., summary statistics related to lensing or clustering) without requiring full field solutions. The advantage of NCE against other kernel-learning models with stronger cosmology priors (e.g., effective field theorem [17] and Halo models [18]) is yet to be understood.

2 Related Work

Learning Green's functions from field-level data. A substantial line of work learns Green's function of linear PDEs directly from source–solution data, using either physics-regularized setups or supervised excitation–response pairs. Representative approaches include multiscale neural networks tailored to the singular/multiscale structure of the Green's function [15], rational-network regression of Green's functions [12], and low-rank/operator–SVD interpolation of learned kernels [13]; extensions like DeepGreen address nonlinear BVPs via latent linearization [14]. These methods assume access to dense fields or high-fidelity simulations for training.

Nonlocal closures and operator learning. In parallel, closure and operator-learning work fits nonlocal convolution kernels (e.g., eddy-diffusivity [19, 20], nonlocal constitutive laws [21]) so a coarse model matches resolved simulations, enforcing physics constraints such as symmetry/reciprocity, decay, and positivity. Examples include systematic constructions of nonlocal eddy-diffusivity operators in fluids and oceanography [19, 20], and data-driven learning of nonlocal constitutive/transport kernels from high-fidelity simulations [22, 17].

Our setting and gap. By contrast, many experimental regimes (notably in materials) lack field-level supervision but do provide structure–property datasets and statistics of heterogeneous coefficients. Our NCE targets this regime: We (i) learn the PDE kernel at a chosen coarse resolution from ensemble/effective measurements (rather than field snapshots), and (ii) plug that kernel into SCE to obtain closed-form maps and sensitivities from coefficient correlations to effective properties.

3 Preliminaries

We shall start with the formal definitions of random fields, effective properties, NPCFs, and SCE.

Notations. Let the relative position between two vertices $(\mathbf{i}, \mathbf{j})$ be $\mathbf{r}(\mathbf{i}, \mathbf{j}) = \mathbf{i} - \mathbf{j}$, $r = \|\mathbf{r}\|_2$, and $\mathbf{n} = \mathbf{r}/r$. $\delta(\mathbf{r})$ is the Dirac delta function. $\mathbb{E}[\cdot]$ denotes the ensemble average and $\langle \cdot \rangle$ the spatial average. ∇, $\nabla\cdot$, and $\nabla\times$ are gradient, divergence, and curl operators, respectively. $\mathcal{A}/a$ is set $\mathcal{A}$ subtracts element a.

Random field. We consider stationary and ergodic random fields defined on a grid that describe binary-phase microstructures: Let a random field be $X = \{X_{\mathbf{i}} \in \{0, 1\} : \mathbf{i} \in \mathcal{G}\}$, where $X_{\mathbf{i}}$ is a binary random variable located at vertex $\mathbf{i}$ on a d-dimensional grid $\mathcal{G} \subset \mathbb{Z}^d$ where $G := |\mathcal{G}|$. The grid $\mathcal{G}$ is taken as a uniform discretization of a unit d-dimensional hypercube $[0, 1]^d$. Here $X_{\mathbf{i}} = 1$ indicates the inclusion (contrast) phase and $X_{\mathbf{i}} = 0$ the matrix (reference) phase. X is *stationary* if all of its moments are invariant to linear translations, and is *ergodic* if its spatial and ensemble averages are equal. Ergodicity allows us to focus the analysis on a single realization $x \sim X$ defined on a large enough grid. The volume fraction of the contrast phase is $\phi = \mathbb{E}[X_{\mathbf{i}}] = \langle X_{\mathbf{i}} \rangle$. For a binary medium $\phi_1 = \phi$, $\phi_0 = 1 - \phi$ denote the phase 1 and phase 0 volumetric proportions, respectively.

Effective property. A linear effective property of the composite is defined via a homogenization of the governing PDE solution. Here we use static conduction and wave propagation as two representative examples, where randomness of microstructures affect the second- and zeroth-order terms of the respective PDEs. **Static conduction:** In a static conduction problem with a reference conductivity σ_0 and inclusion conductivity σ_1, if an electric field $\mathbf{E}_0$ is applied at the boundary $\partial\mathcal{G}$, the local electric potential $\Phi(\mathbf{i}, x)$ in the sample x satisfies the elliptic equation

$$\nabla \cdot [\sigma(x_{\mathbf{i}}) \nabla \Phi] = 0 \quad \text{in } \mathcal{G},$$

with $-\nabla\Phi|_{\partial\mathcal{G}} = \mathbf{E}_0$. The induced current density is $\mathbf{J}(\mathbf{i}, x) = \sigma(x_{\mathbf{i}})\mathbf{E}(\mathbf{i}, x)$ for $\mathbf{i} \in \partial\mathcal{G}$, whose spatial average $\langle \mathbf{J} \rangle(x)$ equals the ensemble-averaged flux due to ergodicity. The effective conductivity tensor $\boldsymbol{\Sigma_e}$ is then defined by Ohm's law $\langle \mathbf{J} \rangle(x) = \boldsymbol{\Sigma}_e(x)\mathbf{E}_0$. In practice, one can compute $\boldsymbol{\Sigma}_e(x)$ by solving the PDE (or measuring $\langle \mathbf{J} \rangle(x)$ experimentally) under d linearly independent boundary conditions and averaging fluxes. **Wave propagation:** We now consider the time-harmonic Maxwell equations. In a nonmagnetic composite ($\mu = \mu_0$) with spatially varying permittivity $\epsilon(x_{\mathbf{i}})$ and $e^{-i\omega t}$ convention, the electric field satisfies the curl–curl equation

$$\nabla \times \nabla \times \mathbf{E}(\mathbf{i}, x) \; - \; \omega^2 \mu_0 \, \epsilon(x_{\mathbf{i}}; \mathbf{k}) \, \mathbf{E}(\mathbf{i}, x) = \mathbf{0} \quad \text{in } \mathcal{G},$$

with appropriate radiation/periodic boundary conditions and an incident plane wave of wave vector $\mathbf{k}$. The effective dynamic dielectric tensor $\boldsymbol{\epsilon}_e(x, \mathbf{k})$ is defined by the macroscopic constitutive relation $\langle \mathbf{D} \rangle(x) \; = \; \boldsymbol{\epsilon}_e(x, \mathbf{k}) \, \langle \mathbf{E} \rangle(x)$ where $\mathbf{k}$ is the wave vector of the incident radiation. Unless stated otherwise, we assume a homogeneous reference medium on $\mathbb{R}^d$ with translation-invariant boundary conditions; consequently the Green's function is homogeneous. We note that in these examples, randomness appears in the second-order term for conduction and in the zeroth-order term for wave propagation. As a result, the PDE kernel for the former will turn out to be the Hessian of Green's function, while that of the latter be the Green's function itself.

N-point correlation function. The NPCF of a random field describes the probability of simultaneously finding a specific N-tuple of points in the inclusion phase. Formally, for any configuration $c = \{\mathbf{i}_1, \ldots, \mathbf{i}_N\}$ of N points (with $\mathbf{i}_1 = 0$ as a reference origin), the N-point function is defined as

$$S_N(X; c) \; = \; \mathbb{E}\big[X_{\mathbf{i}_1} X_{\mathbf{i}_2} \cdots X_{\mathbf{i}_N}\big] \; = \; \frac{1}{G} \sum_{\{\mathbf{j}_2, \ldots, \mathbf{j}_N\}} x_{\mathbf{j}_1} x_{\mathbf{j}_2} \cdots x_{\mathbf{j}_N}, \tag{2}$$

where the second equality is enabled by ergodicity of X and the sum runs over all placements $\{\mathbf{j}_1, \ldots, \mathbf{j}_N\}$ congruent to c in the domain. Dependency on the sample x will be omitted when possible. For $N = 2$, $S_2(\mathbf{r})$ (with $\mathbf{r} = \mathbf{i}_2 - \mathbf{i}_1$) is the two-point correlation giving the probability that two points separated by vector $\mathbf{r}$ are both in the inclusion phase. The lower-order correlations are contained in higher-order ones, i.e., $\mathcal{C}_N \subset \mathcal{C}_{N+1}$ where $\mathcal{C}_N$ is the set of all N-point configurations. In particular, knowledge of all N-point functions up to $N = \infty$ fully characterizes the microstructure. In practice, low-order correlations (e.g. $N = 2, 3$) already capture important morphological features such as phase connectivity and clustering.

Strong Contrast Expansion. SCE treats the contrast medium as a perturbation about the reference medium and systematically accounts for multiple scattering or interactions via successive convolution integrals, resulting in an analytical expression of effective properties as series expansions in terms of

NPCFs convolved with a PDE kernel [23]. Here we briefly outline SCE in terms of static conduction and electromagnetic wave propagation. See detailed derivation in App. A.1. **Static conduction:** Briefly, SCE introduces a polarization field based on the spatial perturbation of the contrast phase:

$$\mathbf{P}(\mathbf{i}, x) = \big(\sigma(x_{\mathbf{i}}) - \sigma_0\big)\mathbf{E}(\mathbf{i}, x),$$

and derives a fixed-point iteration of the polarization field from the Laplace equation:

$$\mathbf{P}(\mathbf{i}, x) = a(x_{\mathbf{i}})\,\mathbf{E}_0 + a(x_{\mathbf{i}})\,\Delta V \sum_{\mathbf{j}\in\mathcal{G}/\mathbf{i}} \mathbf{H}(\mathbf{i}-\mathbf{j})\,\mathbf{P}(\mathbf{j}, x), \tag{1}$$

where ΔV is the d-dimensional unit volume occupied by each vertex of $\mathcal{G}$,

$$a(x_{\mathbf{i}}) = d\sigma_0\beta_\sigma\, x_{\mathbf{i}}, \quad \beta_\sigma = \frac{\sigma_1 - \sigma_0}{\sigma_1 + (d-1)\sigma_0} \in \Big[-\frac{1}{d-1}, 1\Big), \tag{2}$$

and

$$\mathbf{H}(\mathbf{r}) = \frac{d\,\mathbf{n}\mathbf{n}^\top - \mathbf{I}}{\Omega_d\,\sigma_0\,r^d}$$

is the kernel (Hessian of Green's) for 2D conductivity with Ω_d being the d-dimensional total solid angle. Iterative substitution of Eq. (1) into itself and a spatial averaging lead to the following series expansion of the effective conductivity:

$$\beta_\sigma^2\phi^2(\mathbf{\Sigma}_e - \sigma_0\mathbf{I})^{-1}\Big(\mathbf{\Sigma}_e + (d-1)\sigma_0\mathbf{I}\Big) = \beta_\sigma\phi\,\mathbf{I} - \sum_{n=2}^{G}\mathbf{A}_n\,\beta_\sigma^n, \tag{3}$$

where each $\mathbf{A}_n$ is an n-th order convolution integral of total correlation functions (constructed from S_n) with $(n-1)$ factors of the operator kernel $\mathbf{T} = \Omega_d\sigma_0\mathbf{H}$ [1]:

$$\mathbf{A}_n = \left(-\frac{1}{\rho}\right)^{n-2}\left(\frac{d}{\Omega_d}\right)^{n-1}\int d2\cdots dn\Delta(1,\cdots,n)\mathbf{T}(1,2)\mathbf{T}(2,3)\cdots\mathbf{T}(n-1,n), \tag{4}$$

where $\Delta(1,\cdots,n)$ is the n-order total correlation defined by the configuration $\{1,\cdots,n\}$:

$$\Delta(1,\cdots,n) = \begin{vmatrix} S_2(x,\{1,2\}) & S_1(x,2) & \cdots & 0 \\ S_3(x,\{1,2,3\}) & S_2(x,\{2,3\}) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ S_n(x,\{1,\ldots,n\}) & S_{n-1}(x,\{1,\ldots,n\}) & \cdots & S_2(x,\{n-1,n\}) \end{vmatrix}$$

At second order, the leading correction involves the two-point function S_2 via an integral of $S_2(\mathbf{r})$ against $\mathbf{H}(\mathbf{r})$, linking phase connectivity to effective conductivity. **Wave propagation:** For a macroscopically anisotropic bi-phase medium with matrix permittivity ϵ_0 and inclusion permittivity ϵ_1, one obtains an SCE formulation for the dielectric constant $\boldsymbol{\epsilon}_e$ dependent on the wave number $\mathbf{k}$. This formulation turns out to be identical to that of the static conduction problem [24]:

$$\beta_\epsilon^2\phi^2(\boldsymbol{\epsilon}_e(\mathbf{k}) - \epsilon_0\mathbf{I})^{-1}\Big(\boldsymbol{\epsilon}_e(\mathbf{k}) + (d-1)\sigma_0\mathbf{I}\Big) = \beta_\epsilon\phi\,\mathbf{I} - \sum_{n=2}^{G}\mathbf{A}(\mathbf{k})_n\,\beta_\epsilon^n, \tag{5}$$

where

$$\beta_\epsilon \;=\; \frac{\epsilon_1 - \epsilon_0}{\epsilon_1 + (d-1)\epsilon_0}. \tag{6}$$

Different from the conduction case, since microstructural randomness in permittivity only affects the zeroth-order term of the Maxwell equation, the kernel becomes the Green's function itself. In 2D and when local material properties (μ, ϵ_0, and ϵ_1) are isotropic, this Green's function $H^{(0)}(\mathbf{r})$ has been derived as:

$$H_{ij}^{(0)}(\mathbf{r}) \;=\; \frac{i}{4\,\epsilon_0}\left[\Big(k_0^2\mathcal{H}_0^{(1)}(k_0r) - \frac{k_0}{r}\mathcal{H}_1^{(1)}(k_0r)\Big)\delta_{ij} \;+\; k_0^2\mathcal{H}_2^{(1)}(k_0r)\,\mathbf{n}_i\mathbf{n}_j\right] \tag{7}$$

[1] For conciseness, we use "1" instead $\mathbf{i}_1$ when appropriate.

where $\mathbf{k}_0 = \omega\sqrt{\mu_0\varepsilon_0}$ and ω are the wave vector and frequency of the incident radiation respectively, $k_0 = \|\mathbf{k}_0\|$, and $\mathcal{H}_\nu^{(1)}$ is the Hankel function of the first kind of order ν. It should be noted that when local permittivity constants are anisotropic, this analytical Green's does not exist [25]. We also note that in both cases, the convolution integral $\mathbf{A}_n$ is defined on $\mathcal{G}/\mathcal{N}(0)$, where $\mathcal{N}(0)$ is a cavity around around $r = 0$, and is therefore proper. This is because with proper choice of $\mathcal{N}(0)$, e.g., sphere or cylinder, the improper convolution integral over $\mathcal{N}(0)$ can have an analytical form and moved out of the integral. In plain language, for the purpose of predicting effective properties, we do not need to compute an accurate convolution integral within $\mathcal{N}(0)$. This sets our method apart from existing works on learning Green's functions for predicting PDE solution fields [15], where singularity of the Green's function becomes a learning challenge.

4 Neural Contrast Expansion

Building upon the SCE framework, we propose Neural Contrast Expansion (NCE), a data-driven learnable architecture for approximating PDE kernels through structure-property data, circumventing the theoretical assumptions of SCE and allowing explainable sensitivity analysis for structure-property mappings where the kernel cannot be analytically derived (e.g., in the case of anisotropic permittivity in Maxwell's) or does not exist (e.g., when a PDE is mildly nonlinear). Our key insights are: (1) When microstructural randomness only affects one of the second- and zeroth-order terms of the governing PDE, SCE provides a universal functional relationship between the NPCFs and the effective property, evidenced by Eq. (3) and Eq. (5), and (2) the derivation of this functional relationship does not require an explicit kernel. In the following we explain the choice of a hypothesis space for learning kernels, and regularization considerations for promoting explainability and physics consistency.

Bessel–Fourier Kernel Parameterization. NCE approximates $\mathbf{H}(\mathbf{r})$ via a surrogate $\hat{\mathbf{H}}(\mathbf{r})$ with radial-angular decomposition. For 2D problems, we have

$$\hat{\mathbf{H}}(\mathbf{r}) = \begin{pmatrix} \hat{\mathbf{H}}_{11}(r,\theta) & \hat{\mathbf{H}}_{12}(r,\theta) \\ \hat{\mathbf{H}}_{21}(r,\theta) & \hat{\mathbf{H}}_{22}(r,\theta) \end{pmatrix}, \tag{8}$$

where $\mathbf{r} = (x, y)$ and $\theta = \arctan_2(y, x)$. We consider a Bessel-Fourier space for each entry:

$$\hat{\mathbf{H}}_{ij}(r,\theta) \;=\; r^{-\alpha_{\text{env}}} \sum_{n=0}^{N} \sum_{m=1}^{M} \left(C^{(R)}_{ij,n,m} \;+\; i\, C^{(I)}_{ij,n,m} \right) J_n(\alpha_{n,m}\, r)\, \psi_n(\theta), \tag{9}$$

where J_n is the Bessel function of order n, and $\psi_n(\theta)$ is an angular Fourier basis mode $e^{in\theta}$. This basis is chosen because solutions to isotropic PDEs in 2D naturally separate into radial Bessel and angular Fourier modes [26]. The coefficients $C^{(R)}_{ij,n,m}$ and $C^{(I)}_{ij,n,m}$ are real learnable parameters representing the cosine- and sine-phase weights of each basis mode. The coefficient $C^{(R)} + iC^{(I)}$ thus encode complex amplitudes for modes $J_n(\alpha_{n,m} r)e^{in\theta}$. A multiplicative envelope $r^{-\alpha_{\text{env}}}$ imposes the appropriate physical decay, with a learnable α_{env}. The radial wavenumbers $\alpha_{n,m}$ are treated as additional learnable parameters, initialized to span a range of physically relevant length scales.

Regularization for explainability. An ℓ_1-regularization penalty is applied to all coefficients $C^{(R)}$ and $C^{(I)}$, encouraging sparsity in the active Bessel–Fourier modes. This improves interpretability by forcing only the most salient radial and angular components to represent the kernel, while reducing overfitting.

Regularization for physics consistency. We impose a physics-informed regularization to ensure that $\hat{\mathbf{H}}(\mathbf{r})$ is consistent with the target PDE. Recall that the Green's function $\mathbf{G}(\mathbf{r})$ of a linear differential operator $\mathscr{L}$ satisfies

$$\mathscr{L}[\mathbf{G}(\mathbf{r})] = \delta(\mathbf{r}), \tag{10}$$

and the relationship between $\hat{\mathbf{H}}(\mathbf{r})$ and $\mathbf{G}(\mathbf{r})$ is known: When randomness affects the second-order PDE term, $\mathbf{H}(\mathbf{r})$ is the Hessian of $\mathbf{G}(\mathbf{r})$, and when randomness affects the zeroth-order term, $\mathbf{H}(\mathbf{r}) = \mathbf{G}(\mathbf{r})$. Therefore the residual of Eq. (10) is considered a regularization term in the learning of $\hat{\mathbf{H}}(\mathbf{r})$ to embed the PDE's defining property directly into the learning, ensuring that $\hat{\mathbf{H}}(\mathbf{r})$ remains within the physically admissible family determined by the target PDE.

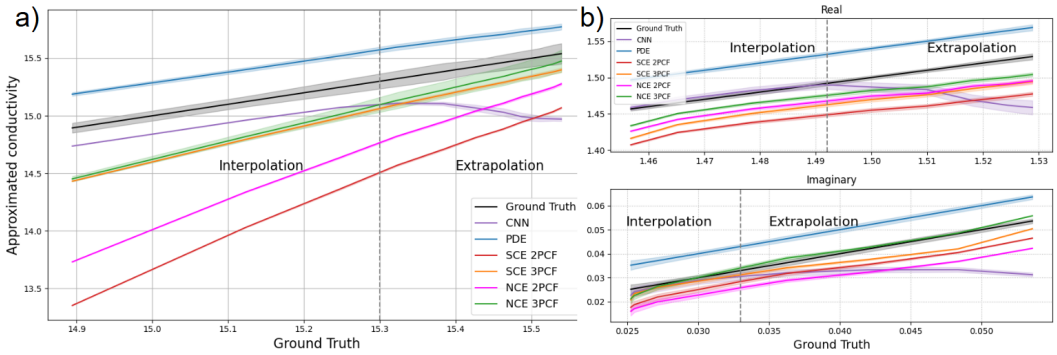


Figure 3: (a) Approximated conductivity and (b) dielectric constant versus ground truth for CNN, low-resolution PDE solver, SCE and NCE methods based on 2- and 3PCF. The vertical dashed line marks the transition from interpolation (left) to extrapolation (right). Shaded regions indicate one standard deviation.

To summarize, with dataset $\mathcal{D} = \{(S_n^{(m)}, \mathbf{\Sigma}_e^{(m)})\}_{m=1}^M$, the loss is defined as:

$$L(\Theta) = \frac{1}{M}\sum_{m=1}^{M}\left\|D(S_n^{(m)}) - \hat{D}(S_n^{(m)};\Theta)\right\|_2^2 + \lambda_1\|C\|_1 \;+\; \lambda_2\|\mathscr{L}[\hat{\mathbf{G}}(\mathbf{r})] - \delta(\mathbf{r})\|^2 \\ + \mathbf{1}_{\text{Hessian}}\,\lambda_3\,\frac{1}{|\mathcal{G}|}\sum_{(x,y)\in\mathcal{G}}\sum_{i,j,k}\left|\partial_x\hat{\mathbf{H}}_{ij}(x,y) - \partial_y\hat{\mathbf{H}}_{ik}(x,y)\right|^2, \tag{11}$$

where D and $\hat{D}$ are defined as:

$$D(S_n) = \beta^2\phi^2\left(\mathbf{\Sigma}_e(S_n) - \sigma_0\mathbf{I}\right)^{-1}\left(\mathbf{\Sigma}_e(S_n) + \sigma_0\mathbf{I}\right), \quad \hat{D}(S_n;\Theta) \;= \beta\phi\mathbf{I} - \sum_{n=2}^{N}\mathbf{A}_n(S_n;\Theta)\beta^n.$$

We choose to fit D rather than regressing $\mathbf{\Sigma}_e$ to avoid numerical instabilities, since the inversion in $(\mathbf{\Sigma}_e(S_n) - \sigma_0\mathbf{I})^{-1}$ can become singular or ill-conditioned during optimization. The indicator $\mathbf{1}_{\text{Hessian}}$ activates the mixed-partial regularization only when $\mathbf{H}(\mathbf{r})$ is the Hessian of $\mathbf{G}(\mathbf{r})$, emphasizing the curl-free conditions for hessian of a scalar potential. In a discretized setting, the Dirac delta function $\delta(\mathbf{r})$ is approximated as a Kronecker delta, which is non-zero only at the origin, making the residual computationally tractable. We also reiterate that the physics consistency loss $\|\mathscr{L}[\hat{\mathbf{G}}(\mathbf{r})] - \delta(\mathbf{r})\|^2$ only requires a numerical integral in $\mathcal{G}/\mathcal{N}(0)$ where the Green's function is smooth.

5 Experimental Results

Data. We use 10 distinct parameter settings of bi-phase 2D random fields with contrast-phase volume fraction $\phi = 0.5$, defined by correlation lengths in the x- and y-directions. The correlation length in the x-direction is fixed at 0.1% of the resolution, while in the y-direction, it varies from 5% to 60%. For each parameter settings, we generate 100 realizations of 1024 × 1024 microstructure of size and use the same set for both static conduction and wave propagation. For the static conduction problem, the effective conductivity of the reference phase and contrast phase is set to 5 and 20 respectively; for the wave propagation problem, the dielectric constants are 1 and 2. These property values are chosen arbitrarily to ensure a strong contrast between the two media. For the wave-propagation problem, we fix the reference-phase wave number $k = 10$. Volume fraction and correlation lengths are chosen so that the ergodicity assumption is satisfied: i.e., the effective conductivities and dielectric constant computed from all 100 microstructure samples using a PDE solver have a small within-group standard deviation (see the gray shade in Fig. 3). For each microstructure x, we collect $(S_2(x), S_3(x), \mathbf{\Sigma}_e^*(x))$, where $S_2(x)$ (resp. $S_3(x)$) contains all 2-point (resp. 3-point) correlations averaged over 256 patches (64 × 64 each) randomly sampled from the 1024 × 1024 microstructure. This mimics realistic settings where microstructure reconstruction is only affordable on small material samples yet effective properties can be experimentally measured from a relatively large sample.

Baselines and the training setting. We generate a high-resolution (1024 × 1024) ground truth benchmark using PDE solvers tailored to specific physics. For static conduction, we employ a

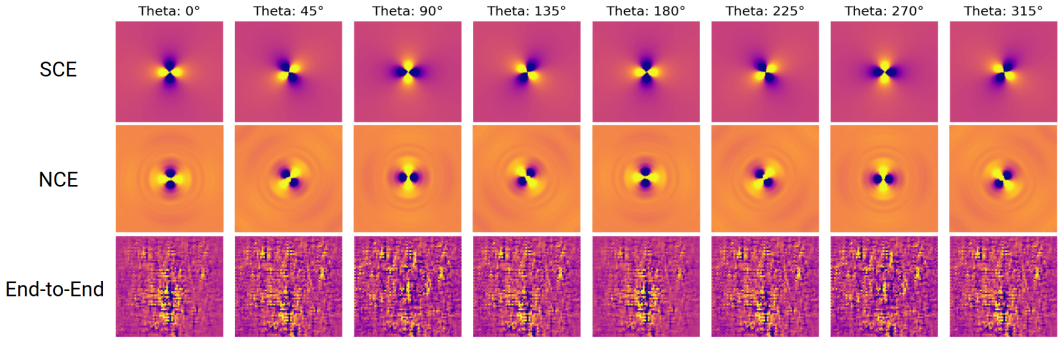


Figure 4: Sensitivity analysis of directional effective conductivity for SCE(top), NCE (middle), and end-to-end approach (bottom)—shown across different orientations θ. The color scale (yellow to dark blue) indicates regions of positive versus negative influence.

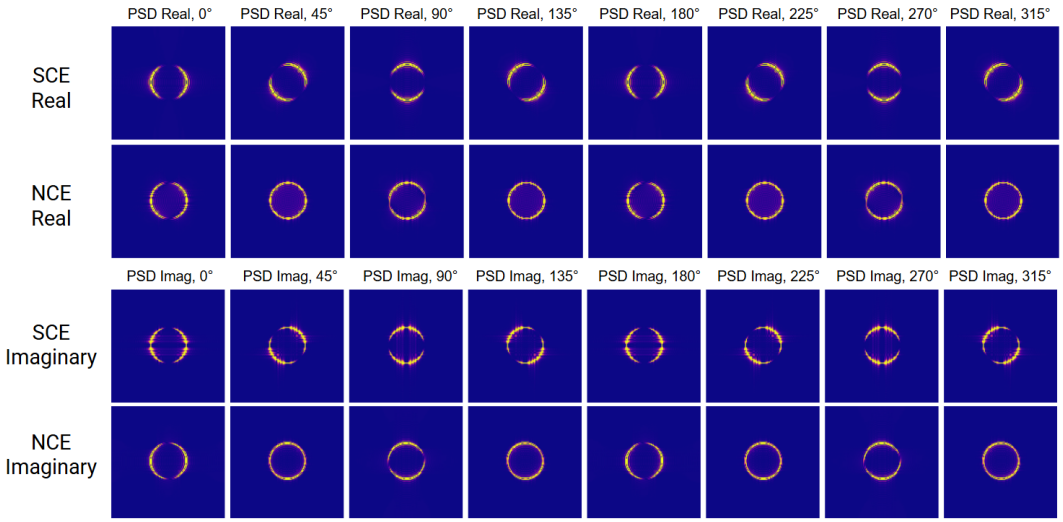


Figure 5: Sensitivity analysis of the real and imaginary parts of the directional effective dielectric constant with respect to power spectral density is shown for both the SCE and NCE approaches across different orientation angles. The color scale (yellow to dark blue) indicates regions of strong positive influence versus negligible influence. (See the time-domain counterpart in Fig. 6)

conventional steady-state solver, while for wave propagation, we use the Finite-Difference Time-Domain (FDTD) method [27]. We then assess three low-resolution (64×64) methods against this benchmark: a down-sampled solver serving as a baseline, and two machine learning models. Both machine learning models are trained to predict effective property from spatial correlations. The first is an end-to-end (E2E) convolutional neural network using 2PCF as input. The second is our Neural Contrast Expansion model, which is trained on either 2-point (2PCF) or both 2- and 3-point (3PCF) correlations and learns its kernel corrections via Eq. (11).

Interpolation and extrapolation accuracy. Fig. 3 compares the predicted conductivities from all models with the ground truth. Indicated by the vertical dashed line, the test uses random field parameters that are both within and out of the training distribution. The standard deviation in the prediction across all 100 realizations are characterized by the shades. We have the following observations: While both NCE and the E2E model performed well on random fields within the training distribution (interpolation), the E2E model's accuracy deteriorated significantly when presented with novel random fields (extrapolation). In contrast, NCE maintained high accuracy in both regimes for the static conduction and wave propagation problems, demonstrating a robust generalization capability that the purely data-driven approach lacks. The analytical SCE models showed inherent prediction errors, confirming that their truncation error and infinite-domain assumption is a significant limitation for finite-sized samples, a gap that NCE's learned kernel mitigates.

Explainability for static conduction. This superior generalization is a direct consequence of NCE's ability to learn physically meaningful relationships, which we validated through sensitivity analysis. Specifically, we compute the partial derivative of directional electrical conductivity with respect to corresponding 2PCF. The directional physical property is measured along an angle θ from the x-axis. For static conduction, Fig. 4 shows that NCE correctly learns the characteristic quadrupole sensitivity pattern provided by the analytical SCE model. This pattern physically signifies that directional conductivity is enhanced by phase connectivity parallel to the applied field and impeded by connectivity in the orthogonal direction. The localized nature of the influence indicates that short-range spatial correlations dominate charge transport mechanisms, providing a clear design implications on synthesizing composite materials. On the other hand, the sensitivity map of the E2E model (Fig. 4, bottom) is not physically interpretable since the model does not utilize physics-informed inductive biases. Thus while the E2E model can achieve reasonable in-distribution prediction accuracy (Fig. 3), it lacks physical explainability to provide material design guidelines.

Design insights from wave propagation. We analyze sensitivity in the Fourier domain by differentiating the directional component of the effective dielectric tensor with respect to the Fourier transform of the two-point correlation (equivalently the power spectral density). Both SCE and NCE produce the same physical picture: sensitivity concentrates on the k-shell $|\mathbf{k}| = k_0$, forming a ring in Fig. 5. Crucially, the imaginary part $\mathrm{Im}(\epsilon_e)$ controls scattering loss; the maps show that loss is driven by spectral power placed on this ring (and, directionally, on its angular sectors). Therefore, to reduce scattering loss at a target frequency k_0, design the microstructure to suppress spectral power on $|\mathbf{k}| = k_0$ (isotropically or in selected sectors). This directly serves low-loss applications such as transparent electromagnetic windows [28] and on-chip photonic waveguides [29], where minimizing $\mathrm{Im}(\epsilon_e)$ is essential.

This design insight connects directly to the concept of hyperuniformity: a novel class of exotic disordered patterns possessing hidden long-range order [30, 16, 31]. Hyperuniform microstructures strongly suppress long-wavelength scattering and thus possess an exclusion zone in their spectral density [16]. By shaping the exclusion to cover or intersect the k_0 ring, one suppresses the long-wavelength scattering channels that feed $\mathrm{Im}(\epsilon_e)$, achieving low loss while retaining control of $\mathrm{Re}(\epsilon_e)$ via modes away from the ring. Indeed, the resulting microstructures shown in Fig. 1(a) are verified to be hyperuniform and consistent with those constructed based on specific spectral densities with anisotropic and isotropic exclusion zones [32, 33]. Because NCE learns a kernel consistent with the governing physics, its sensitivity maps become actionable design levers: weighting a candidate power spectrum by the learned sensitivity identifies the loss-dominant bands to remove and the angular sectors in which to remove them. In this way, Fig. 5 functions not only as diagnosis but as a compact, quantitative recipe for engineering low-loss, anisotropic dielectrics at prescribed (ω, k_0).

6 Conclusion

In this work, we propose Neural Contrast Expansion (NCE), a data-driven model that preserves the analytical architecture of strong-contrast expansion while learning a surrogate PDE kernel directly from structure–property data. Across static conduction and electromagnetic wave propagation, NCE matched or exceeded the predictive accuracy of purely data-driven baselines while retaining physically interpretable sensitivities. In conduction, NCE recovered the quadrupolar sensitivity pattern connecting directional effective conductivity to phase connectivity; in waves, Fourier-domain sensitivities concentrated on the $|\mathbf{k}| = k_0$ shell, cleanly separating how spectral content controls $\mathrm{Re}(\epsilon_e)$ versus $\mathrm{Im}(\epsilon_e)$. This yields actionable design rules: suppress power on the k_0-ring to reduce scattering loss in $\mathrm{Im}(\epsilon_e)$ and adjust off-ring content to tailor anisotropy in $\mathrm{Re}(\epsilon_e)$—linking directly to (stealthy) hyperuniform microstructures that enforce spectral exclusion zones. Practically, NCE requires only macroscopic property measurements, not field-level supervision, and its physics-informed kernel parameterization mitigates SCE's finite-domain and truncation errors while enabling principled extrapolation to novel microstructures.

Limitations suggest clear next steps. Our analysis focuses on randomness entering zeroth- or second-order terms of linear, self-adjoint PDEs; extending to skew-adjoint and weakly nonlinear settings will require a careful re-examination of the SCE derivation and its solvability assumptions. A complementary direction is to jointly optimize over selected higher-order correlations which could boost sample and computational efficiency. Ultimately, NCE provides a practical, physics-grounded framework that translates microstructural statistics into actionable guidelines, accelerating more reliable material discovery and design.

7 Acknowledgment

This material is based upon work partially supported by the Engineering Research and Development Center - Information Technology Laboratory (ERDC-ITL) under Contract No. W912HZ24C0022 and NSF 2503018.

References

[1] S. Torquato. *Random Heterogeneous Materials: Microstructure and Macroscopic Properties*. Springer-Verlag, New York, 2002.

[2] M. Sahimi. *Heterogeneous Materials I: Linear transport and optical properties*, volume 1. Springer, New York, 2003.

[3] Wenxiang Xu, Yufeng Zhang, Jinyang Jiang, Zhiyong Liu, and Yang Jiao. Thermal conductivity and elastic modulus of 3d porous/fractured media considering percolation. *International Journal of Engineering Science*, 161:103456, 2021.

[4] Salvatore Torquato and Jaeuk Kim. Nonlocal effective electromagnetic wave characteristics of composite media: beyond the quasistatic regime. *Physical Review X*, 11(2):021002, 2021.

[5] O. Leseur, R. Pierrat, and R. Carminati. High-density hyperuniform materials can be transparent. *Optica*, 3(7):763–767, Jul 2016.

[6] S Torquato. Effective electrical conductivity of two-phase disordered composite media. *Journal of Applied Physics*, 58(10):3790–3797, 1985.

[7] Hessam Malmir, Muhammad Sahimi, and Yang Jiao. Higher-order correlation functions in disordered media: Computational algorithms and application to two-phase heterogeneous materials. *Physical Review E*, 98(6):063317, 2018.

[8] Sheng Cheng, Yang Jiao, and Yi Ren. Data-driven learning of 3-point correlation functions as microstructure representations. *Acta Materialia*, 229:117800, 2022.

[9] Sihao Cheng, Rudy Morel, Erwan Allys, Brice Ménard, and Stéphane Mallat. Scattering spectra models for physics. *PNAS Nexus*, 3(4):pgae103, 03 2024.

[10] Yang Jiao, FH Stillinger, and S Torquato. Modeling heterogeneous materials via two-point correlation functions: Basic principles. *Physical Review E*, 76(3):031110, 2007.

[11] Y Jiao, FH Stillinger, and S Torquato. Modeling heterogeneous materials via two-point correlation functions. ii. algorithmic details and applications. *Physical Review E*, 77(3):031135, 2008.

[12] Nicolas Boullé, Christopher J. Earls, and Alex Townsend. Data-driven discovery of green's functions with human-understandable deep learning. *Scientific Reports*, 12(1), March 2022.

[13] Harshwardhan Praveen, Nicolas Boullé, and Christopher Earls. Principled interpolation of green's functions learned from data. *Computer Methods in Applied Mechanics and Engineering*, 409:115971, 2023.

[14] Craig R. Gin, Daniel E. Shea, Steven L. Brunton, and J. Nathan Kutz. Deepgreen: Deep learning of green's functions for nonlinear boundary value problems, 2020.

[15] Wenrui Hao, Rui Peng Li, Yuanzhe Xi, Tianshi Xu, and Yahong Yang. Multiscale neural networks for approximating green's functions, 2024.

[16] S. Torquato. Hyperuniform states of matter. *Phys. Rep.*, 745:1–95, 2018.

[17] Giovanni Cabass, Mikhail M. Ivanov, Matthew Lewandowski, Mehrdad Mirbabayi, and Marko Simonović. Snowmass white paper: Effective field theories in cosmology, 2022.

[18] Marika Asgari, Alexander J. Mead, and Catherine Heymans. The halo model for cosmology: a pedagogical review. *The Open Journal of Astrophysics*, 6, November 2023.

[19] Jessie Liu, Hannah H. Williams, and Ali Mani. Systematic approach for modeling a nonlocal eddy diffusivity. *Phys. Rev. Fluids*, 8:124501, Dec 2023.

[20] Mingyue Liu, Ru Chen, Glenn R. Flierl, Wenting Guan, Hong Zhang, and Qianqian Geng. Scale-dependent eddy diffusivities at the kuroshio extension: A particle-based estimate and comparison to theory. *Journal of Physical Oceanography*, 53(8):1851 – 1869, 2023.

[21] Benjamin Sanderse, Panos Stinis, Romit Maulik, and Shady E. Ahmed. Scientific machine learning for closure models in multiscale problems: a review, 2024.

[22] Huaiqian You, Yue Yu, Stewart Silling, and Marta D'Elia. Data-driven learning of nonlocal models: from high-fidelity simulations to constitutive laws, 2020.

[23] Salvatore Torquato. Random heterogeneous materials: microstructure and macroscopic properties. *Appl. Mech. Rev.*, 55(4):B62–B63, 2002.

[24] Salvatore Torquato and Jaeuk Kim. Nonlocal effective electromagnetic wave characteristics of composite media: Beyond the quasistatic regime. *Phys. Rev. X*, 11:021002, Apr 2021.

[25] Said Mikki. Generalized current green's function formalism for electromagnetic radiation by composite systems. *Progress In Electromagnetics Research B*, 87:171–191, 2020.

[26] George B. Arfken, Hans J. Weber, and Frank E. Harris. Chapter 14 - bessel functions. In George B. Arfken, Hans J. Weber, and Frank E. Harris, editors, *Mathematical Methods for Physicists (Seventh Edition)*, pages 643–713. Academic Press, Boston, seventh edition edition, 2013.

[27] Allen Taflove, A. Oskooi, and S. Johnson. *Advances in FDTD Computational Electrodynamics: Photonics and Nanotechnology*. Artech, 01 2013.

[28] Ni Yang, Shaofan Xu, and Chengying Xu. Highly electromagnetic transparent ceramic composite made of boron nitride nanotubes and silicon oxynitride via perhydropolysilazane infiltration method. *Scientific Reports*, 12:14374, 2022.

[29] Minh A. Tran, Duanni Huang, Tin Komljenovic, Jonathan Peters, Aditya Malik, and John E. Bowers. Ultra-low-loss silicon waveguides for heterogeneously integrated silicon/iii-v photonics. *Applied Sciences*, 8(7), 2018.

[30] S. Torquato and F. H. Stillinger. Local density fluctuations, hyperuniformity, and order metrics. *Phys. Rev. E*, 68(4):041113, 2003.

[31] Duyu Chen, Houlong Zhuang, Mohan Chen, Pinshane Y Huang, Vojtech Vlcek, and Yang Jiao. Disordered hyperuniform solid state materials. *Applied Physics Reviews*, 10(2), 2023.

[32] Yaopengxiao Xu, Shaohua Chen, Pei-En Chen, Wenxiang Xu, and Yang Jiao. Microstructure and mechanical properties of hyperuniform heterogeneous materials. *Physical Review E*, 96(4):043301, 2017.

[33] Wenlong Shi, David Keeney, Duyu Chen, Yang Jiao, and Salvatore Torquato. Computational design of anisotropic stealthy hyperuniform composites with engineered directional scattering properties. *Physical Review E*, 108(4):045306, 2023.

A Appendix

A.1 Derivation of SCE for 2D conductivity and discussion about its computational complexity

We now provide the derivation of the effective linear conductivity under the assumption of an infinite boundary (i.e., a static electric field at infinity) [23]. For clarity, we focus on a continuous two-dimensional domain $\mathcal{G}$. Dependence on the random field realization x is omitted when possible. The governing equation is

$$\nabla \cdot \big(\sigma \nabla \phi\big) = 0, \tag{12}$$

with the Neumann boundary condition $\nabla\phi(x_{\mathbf{i}}) = -\mathbf{E}_0$ and with $\sigma(x_{\mathbf{i}}) = \sigma_0$ for all $\mathbf{i} \in \partial\mathcal{G}$. Writing the local conductivity as

$$\sigma(x_{\mathbf{i}}) = \sigma_0 + (\sigma_1 - \sigma_0)x_{\mathbf{i}} = \sigma_0 + \sigma'\, x_{\mathbf{i}},$$

the PDE becomes:

$$\nabla^2\phi = -\frac{1}{\sigma_0}\nabla \cdot \big(\sigma' \nabla\phi\big). \tag{13}$$

This can be solved by introducing an appropriate Green's function for the Laplace operator. When boundary is at infinity and in 2D, we have:

$$G(\mathbf{i}, \mathbf{i}') = -\frac{1}{2\pi\sigma_0} \ln \|\mathbf{i} - \mathbf{i}'\|_2, \tag{14}$$

which satisfies

$$\nabla \cdot \nabla G(\mathbf{i}, \mathbf{i}') = -\frac{1}{\sigma_0}\delta(\mathbf{i} - \mathbf{i}'),$$

where $\delta(\cdot)$ is the 2D Dirac delta function. The solution to Eq. (13) in terms of the Green's function is shown to be:

$$\begin{aligned}
\phi(\mathbf{i}) &= \phi_0(\mathbf{i}) - \int_{\mathcal{G}} G(\mathbf{i}, \mathbf{i}')\, \nabla' \cdot \mathbf{P}(\mathbf{i}')\, d\mathbf{i}' \\
&= \phi_0(\mathbf{i}) + \int_{\mathcal{G}} \nabla' G(\mathbf{i}, \mathbf{i}')^\top\, \mathbf{P}(\mathbf{i}')\, d\mathbf{i}', \\
&\text{(integral by part and } X = 0 \text{ at } \partial\mathcal{G}) \\
&= \phi_0(\mathbf{i}) - \int_{\mathcal{G}} \nabla G(\mathbf{i}, \mathbf{i}')^\top \mathbf{P}(\mathbf{i}') d\mathbf{i}', \\
&(\nabla' G(\mathbf{i}, \mathbf{i}') = -\nabla G(\mathbf{i}, \mathbf{i}')),
\end{aligned} \tag{15}$$

where $\phi_0(\mathbf{i}) = -\langle \mathbf{E}_0, \mathbf{i}\rangle$ is the potential for the homogeneous reference phase, $\mathbf{P} = -\sigma'\nabla\phi$ is the polarization field, and $\nabla' \cdot$ is the divergence with respect to $\mathbf{i}'$.

Since $\nabla\phi(\mathbf{i}) = -\mathbf{E}(\mathbf{i})$ and $\mathbf{P}$ is proportional to $\mathbf{E}$, we take the gradient of both sides of Eq. (15) to create a fixed-point iteration with respect to $\mathbf{P}$. However, it should be first noted that

$$\nabla G(\mathbf{i}, \mathbf{i}') = -\frac{1}{2\pi\sigma_0}\frac{\mathbf{r}}{r^2}, \quad \text{with} \quad \mathbf{r} = \mathbf{i} - \mathbf{i}',$$

is singular at $\mathbf{r} = \mathbf{0}$. To handle this, we decompose the integration domain by separating a small disk

$$\mathcal{B}_\epsilon = \{\mathbf{i}' \in \mathcal{G} \mid \|\mathbf{i} - \mathbf{i}'\|_2 \le \epsilon\}$$

from the rest of the domain $\mathcal{G}/\epsilon$. Applying the divergence theorem shows that the contribution from $\mathcal{B}_\epsilon$ yields

$$\nabla \int_{\partial\mathcal{B}_\epsilon} \langle G(\mathbf{i}, \mathbf{i}')\, \mathbf{P}(\mathbf{i}'), \mathbf{n}'\rangle ds' = -\frac{1}{2\sigma_0}\mathbf{P}(\mathbf{i}), \tag{16}$$

where $\mathbf{n}'$ is the outward normal on $\partial\mathcal{B}_\epsilon$. For $d > 2$, the gradient in Eq. (16) is in general $-\frac{1}{d\sigma_0}\mathbf{P}$. Therefore, differentiating Eq. (15) gives

$$\mathbf{E}(\mathbf{i}) = \mathbf{E}_0 - \frac{1}{d\sigma_0}\mathbf{P}(\mathbf{i}) + \int_{\mathcal{G}/\epsilon} \nabla^2 G(\mathbf{i}, \mathbf{i}')\, \mathbf{P}(\mathbf{i}')\, d\mathbf{i}', \tag{17}$$

where for 2D we set $d = 2$.

Defining

$$\mathbf{F} = \Big(1 + \frac{\sigma'}{d\sigma_0}\Big)\mathbf{E} \quad \text{and} \quad l = \frac{\sigma'}{1 + \frac{\sigma'}{d\sigma_0}},$$

so that $\mathbf{P} = l\,\mathbf{F}$, and letting $\mathbf{H} = \nabla^2 G$, Eq. (17) can be rewritten as

$$\mathbf{F} = \mathbf{E}_0 + \mathbf{H}\,\mathbf{P}. \tag{18}$$

Iteratively applying this relation yields

$$\begin{aligned} \mathbf{P} &= l\,\mathbf{E}_0 + l\,\mathbf{H}\,l\,\mathbf{E}_0 + l\,\mathbf{H}\,l\,\mathbf{H}\,l\,\mathbf{E}_0 + \cdots \\ &= \Big(l + l\,\mathbf{H}\,l + l\,\mathbf{H}\,l\,\mathbf{H}\,l + \cdots\Big)\mathbf{E}_0 \\ &= \mathbf{S}\,\mathbf{E}_0. \end{aligned} \tag{19}$$

Ensemble averaging Eq. (19) along with Eq. (18) leads to

$$\langle \mathbf{F} \rangle = \Big(\langle \mathbf{S}^{-1} \rangle + \mathbf{H}\Big)\langle \mathbf{P} \rangle.$$

A further expansion (involving ensemble averages of the products of l and $\mathbf{H}$) yields an expression for $\langle \mathbf{F} \rangle$ in terms of the microstructural correlation functions. Introducing the expansion parameter

$$\beta = \frac{\sigma_1 - \sigma_0}{\sigma_1 + (d-1)\sigma_0},$$

one finds that

$$\langle l(\mathbf{i}) \rangle = d\sigma_0 \beta\,\phi, \quad \langle l(\mathbf{i})\,l(\mathbf{i}') \rangle = (d\sigma_0\beta)^2 S_2(\mathbf{i}, \mathbf{i}'),$$

for all points in $\mathcal{G}$ (with ϕ denoting the phase density). With further algebra, the operator

$$\mathbf{Q} := \langle \mathbf{S}^{-1} \rangle + \mathbf{H}$$

can be recast as

$$\mathbf{Q} = \frac{1}{d\sigma_0\beta^2\phi^2}\Big(\beta\phi\,\mathbf{I} - \sum_{n=2}^{\infty} \mathbf{A}_n\,\beta^n\Big),$$

where $\mathbf{A}_n$ is defined in Eq. (4).

Relating the polarization field to the effective conductivity via

$$\langle \mathbf{P} \rangle = \mathbf{L}_e \langle \mathbf{F} \rangle$$

with

$$\mathbf{L}_e = d\sigma_0\big(\mathbf{\Sigma}_e - \sigma_0\mathbf{I}\big)\Big[(d-1)\sigma_0\mathbf{I} + \mathbf{\Sigma}_e\Big]^{-1},$$

one finally arrives at the SCE series for the effective conductivity in Eq. (3). Note that Eq. (19) is an expansion of the fixed-point iteration for solving the linear system of equation defined in Eq. (17) with respect to $\mathbf{E}$. By focusing on the polarization field $\mathbf{P}$, we introduce an expansion parameter β that has value between 0 and 1 when $\sigma_1 > \sigma_0$. This derivation underpins the SCE method, naturally leading to the infinite series representation for the effective conductivity, as given in Eq. (3).

Critically, Eq. (3) shows that effective linear properties such as conductivity can be analytically explained by two decoupled factors: the Green's function that is only related to the property-dependent physics, and the NPCFs that are only related to the random field. Lem. 1 characterizes the computational complexity of SCE for approximating effective linear properties:

Lemma 1. *Let $\mathbf{\Sigma}_e(X)$ be a ground-truth effective linear property for X and $\hat{\mathbf{\Sigma}}_e(X)$ be its approximation up to an order N. Let the approximation error be defined by the Frobenius norm: $\epsilon(X) := \|\mathbf{\Sigma}_e(X) - \hat{\mathbf{\Sigma}}_e(X)\|_2$. Given $\epsilon > 0$, $\epsilon(X) \leq \epsilon$ can be achieved with a computational cost of $\mathcal{O}(\epsilon^{-|\beta|\ln(G)})$.*

Proof. By reorganizing (3), we have:

$$\mathbf{\Sigma}_e = (\mathbf{A}+\mathbf{D})^{-1}(\mathbf{B}+\mathbf{D}), \tag{20}$$

where $\mathbf{A} := (\beta\phi - \beta^2\phi^2)\mathbf{I} - \Delta_N$, $\mathbf{B} := \beta^2\phi^2(d-1) + \beta\phi\mathbf{I} - \Delta_N$, $\mathbf{D} := \sum_{n=N+1}^{G} \mathbf{A}_n\beta^n$, and $\Delta_N := \sum_{n=2}^{N} \mathbf{A}_n\beta^n$. The finite-order SCE approximation is

$$\hat{\mathbf{\Sigma}}_e = \mathbf{A}^{-1}\mathbf{B}, \tag{21}$$

and the approximation error is

$$\epsilon^{prop} := \|\hat{\mathbf{\Sigma}}_e - \mathbf{\Sigma}_e\|_2. \tag{22}$$

We now derive an upper bound on ϵ^{prop}. First

$$\begin{aligned}
\epsilon^{prop} &= \|(\mathbf{A}+\mathbf{D})^{-1}(\mathbf{B}+\mathbf{D}) - \mathbf{A}^{-1}\mathbf{B}\|_2 \\
&= \|((\mathbf{A}+\mathbf{D})^{-1} - \mathbf{A}^{-1})\mathbf{B} + (\mathbf{A}+\mathbf{D})^{-1}\mathbf{D}\|_2 \\
&\leq \|\mathbf{A}^{-1}(\mathbf{A}+\mathbf{D}-\mathbf{A})(\mathbf{A}+\mathbf{D})^{-1}\mathbf{B}\|_2 \\
&+ \|(\mathbf{A}+\mathbf{D})^{-1}\mathbf{D}\|_2 \\
&\leq \|(\mathbf{A}+\mathbf{D})^{-1}\|_2\|\mathbf{D}\|_2(\|\mathbf{A}^{-1}\|_2\|\mathbf{B}\|_2 + 1).
\end{aligned} \tag{23}$$

Since $\mathbf{D}$ is a bounded sum of high-order residuals, $\|\mathbf{D}\|_2 \leq \|\mathbf{A}\|_2$ for large enough N. Under this condition, we have

$$\begin{aligned}
(\mathbf{A}+\mathbf{D})^{-1} &= \left[\mathbf{A}\left(\mathbf{I}+\mathbf{A}^{-1}\mathbf{D}\right)\right]^{-1} \\
&= \left(\mathbf{I}+\mathbf{A}^{-1}\mathbf{D}\right)^{-1}\mathbf{A}^{-1} \\
&= \sum_{k=0}^{\infty}\left(-\mathbf{A}^{-1}\mathbf{D}\right)^{k}\mathbf{A}^{-1}.
\end{aligned} \tag{24}$$

Therefore

$$\epsilon^{prop} \leq \frac{\|\mathbf{A}^{-1}\|_2}{1 - \|\mathbf{A}^{-1}\|_2\|\mathbf{D}\|_2}\|\mathbf{D}\|_2(\|\mathbf{A}^{-1}\|_2\|\mathbf{B}\|_2 + 1). \tag{25}$$

ϵ^{prop} monotonically increases with $\|\mathbf{D}\|_2$, i.e., to achieve $\epsilon^{prop} \leq \epsilon$, we need $\|\mathbf{D}\|_2 \leq C_1\epsilon$ with some $C_1 > 0$.

Inspecting the structure of $\mathbf{D}$ to have

$$\mathbf{D} = \sum_{n=N+1}^{G} \beta^n\delta_n, \tag{26}$$

where the matrix δ_n absorbs the convolution between the field kernel and the NPCF at order n. Let C_2 be the largest element of δ_n for any $n \in \{N+1, ..., G\}$ and notice that $|\beta| < 1$, we can bound $\|\mathbf{D}\|_2$ by

$$\|\mathbf{D}\|_2 \leq C_2 \sum_{n=N+1}^{G} |\beta|^n = C_3|\beta|^N. \tag{27}$$

Therefore, to achieve $\|\mathbf{D}\|_2 \leq C_1\epsilon$, we need $N \geq \frac{\ln(C_1\epsilon/C_3)}{\ln(|\beta|)}$. Since the computational complexity of the SCE approximation is determined by that of the highest order NPCF, this complexity is $\mathcal{O}\left(G^{\frac{\ln(C_1\epsilon/C_3)}{\ln(|\beta|)}-1}\right) = \mathcal{O}\left(\epsilon^{-|\beta|\ln(G)}\right)$. □

Remarks. For regular grid $\mathcal{G}$, we have $\ln(G) = d\ln(s^{-1})$. Therefore the worst-case complexity of the SCE approximation is $\mathcal{O}(\epsilon^{|\beta|d\ln(s)})$. While the exact computational complexities of the PDE and SCE approaches are not directly comparable due to their involvement of problem specific paramters, the following observations are useful: (1) Both approaches are affected by the phase-wise property gap $\sigma_1 - \sigma_0$. For PDE, the gap affects the condition number of the resultant system of equations, which in turn affects the convergence rate of the solver. For SCE, the gap directly controls β. In the trivial case of $\sigma_1 = \sigma_0$, SCE directly provides the analytical solution, making PDE unnecessary.

(2) Real-world random fields often have limited correlation lengths, i.e., there exists $r_0 > 0$ such that for any $\mathbf{i}_1, \mathbf{i}_2 \in \mathcal{G}$ and $\|\mathbf{i}_1 - \mathbf{i}_2\|_2 \geq r_0$, $\mathbb{E}[X_{\mathbf{i}_1} X_{\mathbf{i}_2}] = \mathbb{E}[X_{\mathbf{i}_1}]\mathbb{E}[X_{\mathbf{i}_2}]$. With this property, Lem. 2 shows that the proportion of non-zero Δ_N among $\mathcal{C}_N$ reduces exponentially along N. Therefore, we hypothesize that it is possible to replace $\{\mathcal{C}_n\}_{n=1}^N$ with a significantly smaller subset $\mathcal{C}^*$ to achieve both good approximation of $\mathbf{\Sigma}_e$ and a much lower computational complexity.

Lemma 2. *Given correlation order N and the grid $\mathcal{G}$, let $\gamma(N)$ be the fraction of N-point configurations $c \in \mathcal{C}_N$ for which the total correlation $\Delta_N(c) > 0$, and let $\Gamma(\cdot)$ be the the Gamma function. If $r_0 < \left(\Gamma(d/2+1)/\pi^{d/2}\right)^{1/d}$, $\gamma(N)$ decreases exponentially with N.*

Proof. The total number of possible N-point configurations (considered as ordered tuples) is G^{N-1} considering that the first point is fixed to $\mathbf{0}$. An N-point configuration c is said to be *connected* if every pair of points in c is connected via a path of points within c such that consecutive points in the path are within distance r_0 of each other. The number of connected configurations of size N can be approximated as μ^{N-1}, where μ is the average number of ways to add a new site to a connected cluster, also known as the *branching factor*. For a regular grid, μ can be approximated as $V(r_0)/l^d$, where $V(r_0)$ is the volume of a sphere with radius r_0 in $\mathbb{R}^d$, and l^d is the volume occupied by each grid point.

We now show that if c is not connected, i.e., there exists clusters $c_1, ...c_K$ so that $c = \bigcup_{k\in[K]} c_k$ and $c_k \bigcap c_{k'} =$ for all $k \neq k'$, then $\Delta(c) = 0$. To show this, first notice that $\Delta(c)$ is a joint cumulant and is invariant to the permutation of points in c. We can now consider $c = c_0 \bigcup c_1$ and $c_0 \bigcap c_1 =$, so that for any $X_0 \in c_0$ and $X_1 \in c_1$, $\mathbb{E}[X_0 X_1] = \mathbb{E}[X_0]\mathbb{E}[X_1]$. We can reorder the elements of $\Delta(c)$ so that for some $k \in [N]$, points involved in the rows above the kth row all belong to c_0, and the new elements introduced on and after the kth row belong to c_1. Thus the elements on the $k-1$ and kth rows are linearly dependent and therefore the determinant is 0.

Let the side length of the grid be 1. The fraction $\gamma(N)$ is therefore

$$\begin{aligned}\gamma(N) &= \frac{\text{Number of connected configurations}}{\text{Total configurations}} \\ &\approx \frac{\mu^{N-1}}{G^{N-1}} = \left(\frac{\pi^{d/2} r_0^d}{\Gamma(d/2+1)}\right)^{N-1}.\end{aligned} \tag{28}$$

Therefore, when $r_0 < \left(\Gamma(d/2+1)/\pi^{d/2}\right)^{1/d}$, $\gamma(N)$ decreases exponentially with N. The threshold for r_0 is 0.61 for $d = 2$ and 0.64 for $d = 3$. □

A.2 SCE breaks down for first–order randomness

Here we investigate the case where the randomness appears *only in the first–order term* (e.g., a drift/advection field) of a linear second-order PDE.[2] We derive the corresponding Lippmann–Schwinger representation, identify the associated kernels and their singular structure, and then show precisely why SCE cannot be carried out: the perturbation is non–self–adjoint, there is no local "polarizability" linking the polarization to the driving field, and interface distributions enter in a way that breaks the usual correlation–function framework. A narrow exception occurs when the drift is a potential field, which can be removed by a similarity transform that converts the problem into one with zeroth–order randomness, where SCE applies.

Fix $d \geq 2$. Consider the linear PDE

$$\mathcal{L}u := -\nabla\cdot(\mathbf{A}_0\nabla u) \;+\; \mathbf{b}(\mathbf{x})\cdot\nabla u \;+\; c_0\, u \;=\; s, \tag{29}$$

where $\mathbf{A}_0 \in \mathbb{R}^{d\times d}$ is constant, symmetric positive definite (SPD), $c_0 \geq 0$ is constant, and the *randomness* is only in the drift $\mathbf{b}(\mathbf{x})$. We take the baseline

$$\mathcal{L}_0 u := -\nabla\cdot(\mathbf{A}_0\nabla u) + c_0 u, \qquad \mathcal{L}_0 G_0(\mathbf{r}) = \delta(\mathbf{r}). \tag{30}$$

We will impose a uniform macroscopic gradient $\boldsymbol{E}_0$ in the baseline medium so that the baseline solution is $u_0(\mathbf{x}) = -\boldsymbol{E}_0\cdot\mathbf{x}$ (or a periodic surrogate), and denote the total gradient ("field") by

$$\boldsymbol{E}(\mathbf{x}) := \nabla u(\mathbf{x}), \qquad \langle \boldsymbol{E}\rangle = \boldsymbol{E}_0.$$

[2]The derivation in this subsection is assisted by GPT5 Pro and verified by the authors.

Self–adjointness. The formal L^2 adjoint (integration by parts) of the drift term is

$$(\mathbf{b}\cdot\nabla)^* = -\mathbf{b}\cdot\nabla - (\nabla\cdot\mathbf{b})\,,$$

hence the adjoint of $\mathcal{L}$ is

$$\mathcal{L}^* = -\nabla\cdot(\mathbf{A}_0\nabla\cdot) \;-\; \mathbf{b}\cdot\nabla \;-\; (\nabla\cdot\mathbf{b}) \;+\; c_0. \tag{31}$$

Therefore $\mathcal{L}$ is self–adjoint in L^2 *iff* $\mathbf{b}\equiv\mathbf{0}$. Even if $\nabla\cdot\mathbf{b}=0$ (divergence–free drift), the operator remains non–self–adjoint because of the $-\mathbf{b}\cdot\nabla$ in (31).

Lippmann–Schwinger representation for first–order randomness. Subtract $\mathcal{L}_0 u_0 = s$ from $\mathcal{L}u = s$ to get

$$\mathcal{L}_0(u-u_0) = -\mathbf{b}\cdot\nabla u.$$

Convolving with G_0 and using $\mathcal{L}_0 G_0 = \delta$ gives the *resolvent identity*

$$u = u_0 - G_0 * (\mathbf{b}\cdot\nabla u). \tag{32}$$

Use the product rule to expose only *multiplications* (no derivatives on u) inside the convolution:

$$\mathbf{b}\cdot\nabla u = \nabla\cdot(\mathbf{b}u) - (\nabla\cdot\mathbf{b})\,u.$$

Integrate by parts inside the convolution:

$$\int_{\mathbb{R}^d} G_0(\mathbf{x}-\mathbf{y})\,\nabla\cdot_{\mathbf{y}}\,(\mathbf{b}(\mathbf{y})u(\mathbf{y}))\,d\mathbf{y} = -\int_{\mathbb{R}^d}\nabla G_0(\mathbf{x}-\mathbf{y})\cdot(\mathbf{b}(\mathbf{y})u(\mathbf{y}))\,d\mathbf{y}.$$

Thus

$$\boxed{u(\mathbf{x}) = u_0(\mathbf{x}) + \int_{\mathbb{R}^d}\nabla G_0(\mathbf{x}-\mathbf{y})\cdot\underbrace{\big[\mathbf{b}(\mathbf{y})\,u(\mathbf{y})\big]}_{:=\boldsymbol{P}(\mathbf{y})}\,d\mathbf{y} - \int_{\mathbb{R}^d} G_0(\mathbf{x}-\mathbf{y})\,\underbrace{\big[(\nabla\cdot\mathbf{b})(\mathbf{y})\,u(\mathbf{y})\big]}_{:=q(\mathbf{y})}\,d\mathbf{y}.} \tag{33}$$

Differentiating (33) yields the gradient equation

$$\boxed{\boldsymbol{E}(\mathbf{x}) = \boldsymbol{E}_0 + \int_{\mathbb{R}^d}\underbrace{\nabla\nabla G_0(\mathbf{x}-\mathbf{y})}_{:=\mathbf{K}_0(\mathbf{x}-\mathbf{y})} : \boldsymbol{P}(\mathbf{y})\,d\mathbf{y} - \int_{\mathbb{R}^d}\nabla G_0(\mathbf{x}-\mathbf{y})\,q(\mathbf{y})\,d\mathbf{y}.} \tag{34}$$

Singular structure of the kernels. The Hessian kernel $\mathbf{K}_0(\mathbf{r}) = \nabla\nabla G_0(\mathbf{r})$ has a short–range δ–singular part:

$$\boxed{\mathbf{K}_0(\mathbf{r}) = -\boldsymbol{\Lambda}_2\,\delta(\mathbf{r}) \;+\; \mathbf{T}^{\,L}(\mathbf{r}), \qquad \boldsymbol{\Lambda}_2 := \tfrac{1}{d}\,\mathbf{A}_0^{-1}.} \tag{35}$$

The remainder $\mathbf{T}^{\,L}$ is integrable and satisfies $\int_{\mathbb{R}^d}\mathbf{T}^{\,L}(\mathbf{r})\,d\mathbf{r} = \boldsymbol{\Lambda}_2$. In contrast, the vector kernel $\nabla G_0(\mathbf{r})$ is *odd* and has

$$\int_{\mathbb{R}^d}\nabla G_0(\mathbf{r})\,d\mathbf{r} = \mathbf{0} \quad\text{and}\quad \text{no } \delta \text{ part.} \tag{36}$$

(An easy Fourier–domain check: $\widehat{\nabla G_0}(\mathbf{k}) = i\mathbf{k}/(\mathbf{k}\cdot\mathbf{A}_0\mathbf{k}+c_0)$ is odd in $\mathbf{k}$, so it cannot produce a constant term, i.e., a δ in real space.)

Plugging (35) into (34) gives

$$\boxed{\boldsymbol{E} = \boldsymbol{E}_0 \;-\; \boldsymbol{\Lambda}_2\,\boldsymbol{P} \;+\; \mathbf{T}^{\,L} * \boldsymbol{P} \;-\; (\nabla G_0) * q, \qquad \boldsymbol{P} := \mathbf{b}\,u, \quad q := (\nabla\cdot\mathbf{b})\,u.} \tag{37}$$

How SCE works in the self–adjoint cases. For comparison, when the randomness is in the *second–order* term, such as in the case of conduction, write $\mathbf{A}(\mathbf{x}) = \mathbf{A}_0 + X(\mathbf{x})$ and define

$$\boldsymbol{P} = X\,\boldsymbol{E}.$$

Equation (37) then becomes $\boldsymbol{E} = \boldsymbol{E}_0 - \boldsymbol{\Lambda}_2 X\boldsymbol{E} + \mathbf{T}^{\,L} * (X\boldsymbol{E})$, so one can *move the local term* to the left and *invert it pointwise*:

$$(\mathbf{I} + \boldsymbol{\Lambda}_2 X)\,\boldsymbol{E} = \boldsymbol{E}_0 + \mathbf{T}^{\,L} * (X\boldsymbol{E}) \quad\Longrightarrow\quad \boldsymbol{P} = M\,\boldsymbol{E}_0 + M\,\mathbf{T}^{\,L} * \boldsymbol{P},$$

with the *local polarizability*

$$M := X\,(\mathbf{I} + \boldsymbol{\Lambda}_2 X)^{-1}.$$

This produces the SCE series for the effective tensor in terms of N-point correlation functions of the microstructure. A similar (scalar) story holds when the randomness is purely in the zeroth–order term.

Why SCE breaks for first–order randomness. Return to (37), where

$$\boldsymbol{P} = \mathbf{b}\, u \quad \text{and} \quad q = (\nabla\cdot \mathbf{b})\, u.$$

The barriers to SCE are

- No local relation $\boldsymbol{P} = M\,\boldsymbol{E}$: In the self–adjoint cases, $\boldsymbol{P}$ depends on $\boldsymbol{E}$ by *multiplication*. Here, $\boldsymbol{P}$ depends on the *potential* u, not on its gradient $\boldsymbol{E}$. There is no local matrix $M(\mathbf{x})$ so that $\mathbf{b}(\mathbf{x})\, u(\mathbf{x}) = M(\mathbf{x})\, \boldsymbol{E}(\mathbf{x})$ for all admissible u.

  Lemma 3 (No local polarizability)**.** *Suppose there exists a (measurable) matrix field $M(\mathbf{x})$ with $\boldsymbol{P}(\mathbf{x}) = M(\mathbf{x})\, \boldsymbol{E}(\mathbf{x})$ for all solutions u. Then $\mathbf{b} \equiv \mathbf{0}$.*

  Proof. Fix $\mathbf{x}_0$ and choose a test solution u that is locally constant near $\mathbf{x}_0$ (e.g., a smooth bump function flattened near $\mathbf{x}_0$). Then $\boldsymbol{E}(\mathbf{x}_0) = \mathbf{0}$ but $u(\mathbf{x}_0)$ can be nonzero, hence $\boldsymbol{P}(\mathbf{x}_0) = \mathbf{b}(\mathbf{x}_0)\, u(\mathbf{x}_0)$. The assumed relation gives $\boldsymbol{P}(\mathbf{x}_0) = M(\mathbf{x}_0)\boldsymbol{E}(\mathbf{x}_0) = \mathbf{0}$, forcing $\mathbf{b}(\mathbf{x}_0) = \mathbf{0}$. Since $\mathbf{x}_0$ is arbitrary, $\mathbf{b} \equiv \mathbf{0}$. □

- The q–term has no δ to renormalize: The last term in (37) involves the kernel ∇G_0, which has no δ part (36). Thus, unlike the $\mathbf{K}_0$ term, there is no local singularity to "absorb" into a pointwise transform. Any attempt to eliminate q would be nonlocal (via $\mathcal{L}_0^{-1}$), defeating the SCE philosophy.
- Interface distributions pollute the microstructure coefficients: In many microstructure models, properties are piecewise constant by phase. If $\mathbf{b}(\mathbf{x}) = \sum_{p=1}^{M} \mathcal{I}_p(\mathbf{x})\, \boldsymbol{b}_p$ with phase indicators $\mathcal{I}_p \in \{0,1\}$ and constant vectors $\boldsymbol{b}_p$, then

  $$\nabla\cdot \mathbf{b} = \sum_{p=1}^{M} \boldsymbol{b}_p\cdot \nabla \mathcal{I}_p.$$

  Distributionally, $\nabla\mathcal{I}_p$ is a *surface measure* concentrated on the phase boundaries:

  $$\nabla\mathcal{I}_p = -\,\boldsymbol{n}\,\delta_{\partial\Omega_p},$$

  where $\boldsymbol{n}$ is the unit normal and $\delta_{\partial\Omega_p}$ is a Dirac mass on the interface. Consequently $q = (\nabla\cdot \mathbf{b})\, u$ is *supported on surfaces*, and ensemble averages of the form $\langle (\nabla G_0) * q\rangle$ depend on *surface–surface* statistics (geometry of interfaces), not on the usual *volume* N-point correlation functions $\{S_n\}$ that power SCE. This is a structural mismatch.
- No symmetric energy (variational) structure: Self–adjoint randomness leads to a positive energy form $\int (\nabla u)\cdot \mathbf{A}(\mathbf{x})\, \nabla u + \int c(\mathbf{x})\, u^2$, from which bounds and renormalized series follow. The drift contribution satisfies

  $$\int \phi\, (\mathbf{b}\cdot\nabla u) = -\int u\, \big(\mathbf{b}\cdot\nabla\phi + (\nabla\cdot \mathbf{b})\, \phi\big),$$

  which is not sign–definite and provides no coercive variational principle. This removes the key positivity used in SCE resummations.

A narrow exception: potential drift can be removed. There is one structurally important exception:

Theorem 1 (Self–adjointization by similarity transform)**.** *If $\mathbf{b}(\mathbf{x}) = \mathbf{A}_0\, \nabla\psi(\mathbf{x})$ for some scalar ψ, and $\mathbf{A}_0$ is constant SPD, then the change of variables $u = e^{\psi/2} v$ converts $\mathcal{L}$ into a self–adjoint operator with a* random zeroth–order potential*:*

$$\mathcal{L}u = e^{\psi/2}\left[-\nabla\cdot (\mathbf{A}_0\nabla v) + \Big(c_0 \;+\; \tfrac{1}{4}\,\nabla\psi\cdot\mathbf{A}_0\nabla\psi \;-\; \tfrac{1}{2}\,\nabla\cdot (\mathbf{A}_0\nabla\psi)\Big)\, v\right]. \tag{38}$$

Proof sketch. Compute $\nabla u = e^{\psi/2}\big(\nabla v + \frac{1}{2} v\, \nabla\psi\big)$ and use the product rule with $\mathbf{A}_0$ constant:

$$\begin{aligned}\nabla\cdot (\mathbf{A}_0\nabla u) &= \nabla\cdot\Big(e^{\psi/2}\mathbf{A}_0\big(\nabla v + \tfrac{1}{2} v\, \nabla\psi\big)\Big)\\ &= e^{\psi/2}\Big[\nabla\cdot (\mathbf{A}_0\nabla v) + (\nabla v)\cdot\mathbf{A}_0\nabla\psi + \tfrac{1}{2} v\, \nabla\cdot (\mathbf{A}_0\nabla\psi) + \tfrac{1}{4} v\, \nabla\psi\cdot\mathbf{A}_0\nabla\psi\Big].\end{aligned}$$

Also $\mathbf{b}\cdot\nabla u = e^{\psi/2}\big((\mathbf{A}_0\nabla\psi)\cdot\nabla v + \frac{1}{2} v\, \nabla\psi\cdot\mathbf{A}_0\nabla\psi\big)$. Substituting in $\mathcal{L}u = -\nabla\cdot(\mathbf{A}_0\nabla u) + \mathbf{b}\cdot\nabla u + c_0 u$ cancels the $(\nabla v)\cdot\mathbf{A}_0\nabla\psi$ terms and yields (38). □

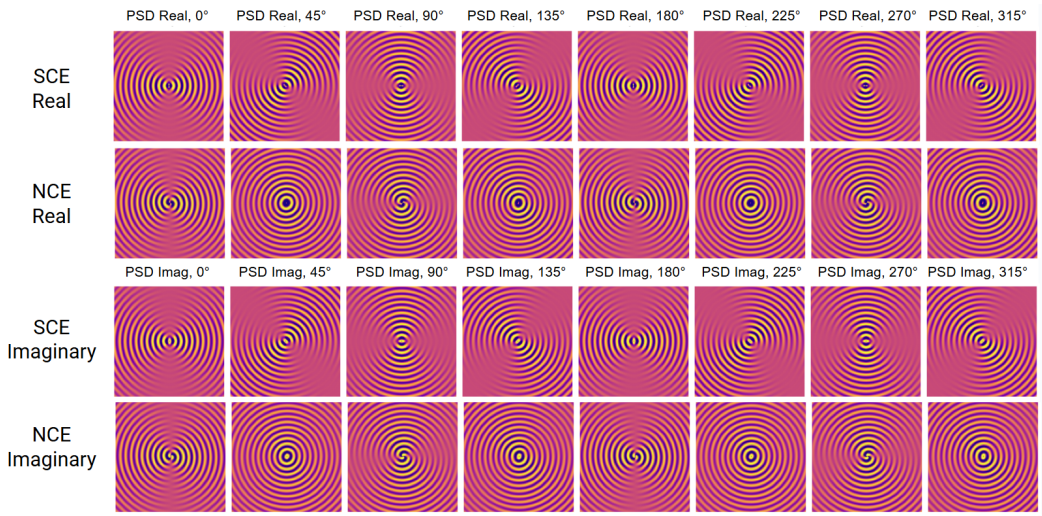
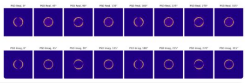


Figure 6: Sensitivity analysis of the real and imaginary parts of the directional effective dielectric constant with respect to 2PCF is shown for both the SCE and NCE approaches across different orientation angles. The color scale (yellow to dark blue) indicates regions of strong positive influence versus negligible influence.

Remark 2. In one spatial dimension ($d = 1$), every scalar drift is a gradient; therefore the transform always exists and reduces the problem to one with zeroth–order randomness, where SCE applies. In $d \geq 2$, a generic stationary random drift is *not* a gradient ($\nabla \times \mathbf{b} \neq \mathbf{0}$), so the transform is unavailable.

A.3 NCE–LLM Integration: Extracting Physical Insights

In this section, we present a blinded study with the GPT-5 Thinking model: memory and prior-conversation referencing were disabled, the model was shown only the sensitivity maps (without disclosing the underlying physical property), and it was tasked with proposing designs that meet a specified engineering objective under a fixed volume fraction. Screenshots of the conversation are provided below.



Context: The accompanying figure presents the sensitivity analysis results for a binary microstructure. It specifically details how the real part and the imaginary part of the directional effective physical property respond to variations in the microstructure's power spectral density (PSD). The color scale (yellow to dark blue) indicates regions of strong positive influence versus negligible influence.

Inquiry: Based on a thorough analysis of this figure, please provide a set of actionable design principles for tailoring the binary microstructure. The primary engineering goal is to optimize the effective property in the x-direction by simultaneously maximizing its real component while minimizing its imaginary component, subject to a constant volume fraction.

Thought for 2m 52s

Here's a practical, PSD-space playbook you can use to shape the binary microstructure so that the **x-direction effective property** is large in **Re** and small in **Im** (volume fraction fixed ⇒ only redistributions of spectral power are allowed).

1) Lock the length scale (radial placement)

- The sensitivity is concentrated on a **single ring in k-space**.
  ➜ Make the microstructure **narrow-band** around that radius $k^{\backslash *}$.
  ➜ Choose a feature pitch $\ell^{\backslash *} \approx 2\pi/k^{\backslash *}$ and keep the radial spread small (tight correlation length).
  Rationale: power you place off the ring barely moves either Re or Im.

2) Place power in the "good" angular sectors for x

- In the "PSD Real, 0°" panel the **bright arcs** identify angles θ where adding power **most increases Re** of the x-property.
- In the corresponding "PSD Imag, 0°" panel the bright arcs are **quarter-turn shifted** around the ring; those angles **inflate Im**.
- **Action:** concentrate PSD power in the bright arcs of **PSD Real, 0°** that are **dark in PSD Imag, 0°** (for this figure: wedges centered around the **wavevectors pointing opposite to +x**, i.e., near $\theta \approx \pi$, and their antipodes). Keep the wedges narrow (e.g., ±10–20°) to avoid bleeding into the Im-hot sectors.

3) Use symmetry to suppress Im automatically

- Impose **mirror symmetry about the x-axis** in the angular spectrum: for every wedge at θ include the same weight at $-\theta$.
- Also keep the inherent π-pairing (wedge at θ and $\theta + \pi$).
- This **even-in-θ** distribution cancels the sine-odd contributions that feed the imaginary part, while the cosine-even contributions (driving Re) add constructively.

4) Take power away from the "bad" sectors

- Because volume fraction is fixed, any increase in the good wedges must be offset by decreases elsewhere.
- **Deplete** the wedges that are bright in **PSD Imag, 0°** (the ones roughly **90° away** from your chosen Re-maximizing wedges).
- In real space, this means **reducing** features whose orientation produces those "bad" wavevectors (e.g., avoid edges/grains whose normals point along +x when those sectors are Im-hot).

5) Favor orthotropy over diagonals

- Across the panels, the diagonals (≈45°, 135°, …) neither help Re strongly nor suppress Im.
- **Reduce diagonal orientation content** in the PSD; bias the orientation distribution toward the x–y axes (with the x-aligned choice prioritized per #2).

6) Keep the spectrum clean

- Avoid multimodal length scales or broad angular scatter. Extra sidebands tend to **re-introduce Im** (because some of them will land in the Im-sensitive quadrants) while offering little Re gain.

7) Translate to real-space shapes (if you prefer geometry rules)

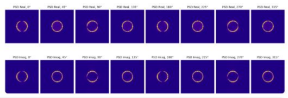
- Build an **anisotropic but mirror-symmetric** texture with a **single dominant pitch** $\ell^{\backslash *}$.
- Two robust families that realize the PSD rules above:
  - **Lamellar/striped** patterns with normals concentrated near $\pm x$ (interfaces roughly orthogonal to x) but arranged with **up–down symmetry** about the x-axis.
  - **Elliptical/prolate inclusions** whose **major axes are parallel to x**, distributed with the same x-axis mirror symmetry.
- In either case, keep the orientation spread tight and avoid diagonal feature orientations.

8) A minimal optimization recipe you can implement

1. Parameterize the PSD's **angular weight** $w(\theta)$ on the ring (e.g., as a sum of a few Gaussian wedges centered at chosen θ_i with widths σ_i).
2. Constrain $\int w(\theta)\, d\theta = \text{const}$ (volume fraction).
3. Maximize $J = \int w(\theta)\, R_x(\theta)\, d\theta - \lambda \int w(\theta)\, I_x(\theta)\, d\theta$
   where R_x, I_x are the measured sensitivities from the **PSD Real/Imag, 0°** maps.
4. Enforce $w(\theta) = w(-\theta) = w(\theta + \pi)$ (mirror + antipodal symmetry).
   This will push all the power into the **Re-bright/Im-dark** arcs and keep Im suppressed by construction.

In short: make the spectrum **narrow-band** at the ring radius, **pack power into the Real-bright wedges for the x-direction** while **removing it from the Imag-bright wedges**, and **enforce mirror symmetry** of the angular distribution. That combination maximizes the real part of the x-effective property and simultaneously keeps the imaginary part minimal, without changing the volume fraction.

Figure 7: The primary engineering goal is to optimize the effective property in the x-direction by simultaneously maximizing its real component while minimizing its imaginary component, subject to a constant volume fraction.



Context: The accompanying figure presents the sensitivity analysis results for a binary microstructure. It specifically details how the real part and the imaginary part of the directional effective physical property respond to variations in the microstructure's power spectral density (PSD). The color scale (yellow to dark blue) indicates regions of strong positive influence versus negligible influence. Inquiry: Based on a thorough analysis of this figure, please provide a set of actionable design principles for tailoring the binary microstructure. The primary engineering goal is to optimize the effective property in the x-direction by simultaneously maximizing its real component while minimizing its imaginary component, subject to a constant volume fraction.

Inquiry: Based on a thorough analysis of this figure, please provide a set of actionable design principles for tailoring the binary microstructure. The primary engineering goal is to optimize the effective property in all directions by simultaneously maximizing its real component while minimizing its imaginary component, subject to a constant volume fraction.

Thought for 52s

Great—now the target is **isotropic** performance: big **Re** and tiny **Im** of the effective property **in every direction**, with volume fraction fixed. Reading the figure (bright sensitivity sitting on a ring in k-space, with Real/Imag hot sectors rotating with direction) leads to these concrete, PSD-space rules:

Design principles (all-directions)

1. **Concentrate power on the ring at $k^{\backslash *}$.**
   Keep a narrow radial bandwidth (e.g., $\Delta k/k^{\backslash *} \lesssim 0.1$). Power off the ring moves neither Re nor Im much—reallocate it onto the ring.
2. **Make the angular spectrum uniform.**
   Set the PSD weight $w(\theta)$ around the ring as flat as possible (rotational invariance). This maximizes the **directional average of Real** while preventing any single direction from lagging.
3. **Kill odd angular harmonics.**
   Enforce $w(\theta) = w(-\theta) = w(\theta + \pi)$. In Fourier terms, keep $m = 0, 2, 4, \ldots$ and suppress $m = 1, 3, 5, \ldots$. This symmetry annihilates the first-order drivers of the **Imag** part in every direction.
4. **Avoid angular notches or wedges.**
5. **Single length scale beats multi-scale.**
   Multimodal or broad-band spectra create sidebands that inevitably land in Imag-sensitive zones for some directions.
6. **Use high rotational symmetry if perfect isotropy is impractical.**
   With unit cells, enforce ≥8-fold rotational symmetry (equal wedges every 45°); keep all wedge weights equal.
7. **Respect the volume-fraction constraint by redistribution, not addition.**
   Pull power from the PSD interior/outer halo and any residual anisotropic lobes; push it onto the uniform ring.
8. **Keep the microstructure centrosymmetric and non-chiral.**
   Any parity breaking reintroduces odd harmonics ⇒ extra Im.
9. **Connectivity tuning without anisotropy.**
   If the property benefits from percolation, improve it via **isotropic** topology changes (e.g., raise interface density uniformly at $k^{\backslash *}$), not by adding oriented paths.

Practical recipe (PSD space)

- Parameterize the ring PSD as
  $S(k, \theta) = A\, \exp[-(k - k^{\backslash *})^2/(2\sigma_k^2)] \left[a_0 + \sum_{m=2,4,\ldots}^{M} a_m \cos(m\theta)\right]$, with $a_m \geq 0$.
- Enforce $\int S\, dk\, d\theta$ = constant (volume fraction).
- Optimize the **directional average** objective
  $J = \langle \mathrm{Re} \rangle_\theta - \lambda \langle \mathrm{Im} \rangle_\theta$.
- In practice the optimum drives a_0 high and $a_{m>0} \to 0$ ⇒ **uniform ring**.

Real-space constructions that realize this PSD

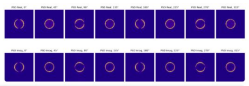
- **Poisson-disk (blue-noise) placement of nearly circular inclusions** with diameter $\ell^{\backslash *} \approx 2\pi/k^{\backslash *}$.
- **Isotropic band-pass textures**: generate Gaussian noise, band-pass filter at $k^{\backslash *}$, then threshold to the desired volume fraction.
- **Isotropic Voronoi/foam-like networks** scaled to $\ell^{\backslash *}$, with no preferred edge orientation.

Quick acceptance tests

- Circular variance of angular PSD on the ring < 0.05.
- Odd-harmonic energy fraction $E_{\mathrm{odd}}/E_{\mathrm{total}} < 1\%$.
- Radial bandwidth $\Delta k/k^{\backslash *} \leq 0.1$.
- Directional sweep (0–315°): Re nearly flat, Im near noise floor.

Follow these and you'll get a **uniformly high real response with minimal imaginary leakage** in every direction—while keeping the volume fraction unchanged.

Figure 8: The primary engineering goal is to optimize the effective property in all directions by simultaneously maximizing its real component while minimizing its imaginary component, subject to a constant volume fraction.



Context: The accompanying figure presents the sensitivity analysis results for a binary microstructure. It specifically details how the real part and the imaginary part of the directional effective physical property respond to variations in the microstructure's power spectral density (PSD). The color scale (yellow to dark blue) indicates regions of strong positive influence versus negligible influence. Inquiry: Based on a thorough analysis of this figure, please provide a set of actionable design principles for tailoring the binary microstructure. The primary engineering goal is to optimize the effective property in the x-direction by simultaneously maximizing its real component while minimizing its imaginary component, subject to a constant volume fraction.

Inquiry: Based on a thorough analysis of this figure, please provide a set of actionable design principles for tailoring the binary microstructure. The primary engineering goal is to optimize the effective property in all directions by simultaneously minimizing its real component while maximizing its imaginary component, subject to a constant volume fraction.

Got it — now we want the **opposite/isotropic** target: across **all directions**, make the **imaginary part large** while keeping the **real part small**, with the volume fraction fixed. Reading the sensitivity maps (bright ring in k-space; Imag "hot" sectors rotate with direction and are roughly in quadrature with the Real ones) suggests the following PSD-space rules.

Design principles (maximize Im, minimize Re, all directions)

1. **Stay on the ring at k^***
   Keep a narrow radial bandwidth (concentrate power on the sensitivity ring). Power off the ring barely moves either part.
2. **Break mirror symmetry, keep 180° pairing**
   A real microstructure forces $S(\theta) = S(\theta + \pi)$ (antipodal pairing), but mirror symmetry $S(\theta) = S(-\theta)$ is optional. **Intentionally violate mirror symmetry** around the ring; this is what drives a net **Imag** response. (Mirror symmetry suppresses Im.)
3. **Excite the odd–under–reflection angular modes only**
   Expand the angular PSD on the ring as
   $$S(\theta) = a_0 + \sum_{m=2,4,\ldots} \big[c_m \cos(m\theta) + s_m \sin(m\theta)\big],$$
   where π-periodicity removes odd m.
   - **Set** $c_m \approx 0$ (these feed the Real part).
   - **Put your budget into** s_m, starting with the **dominant Imag mode** (typically $m = 2$, i.e., a quadrupolar sine).
     This keeps the Real projection (even/cosine) small and maximizes overlap with Imag (odd under reflection/sine).
4. **Use a single handedness globally**
   Choose one sign of the s_m coefficients (all positive or all negative) so the induced Imag contribution doesn't cancel when directions rotate. This creates an **isotropic–chiral** spectrum: rotationally uniform power level with a consistent sense of twist.
5. **Avoid a uniform ring component**
   Minimize or remove a_0 (the "flat" ring). The a_0 term couples strongly to Real and raises it in every direction.
6. **Don't notch the ring; modulate it**
   Large hard angular notches introduce even (cosine) content and leak Real. Prefer **smooth sinusoidal modulation** on the ring: concentrate where the Imag map is bright for a reference direction and taper elsewhere smoothly.
7. **Keep it single-scale**
   Multimodal k content adds stray sectors that inevitably overlap Real-hot regions for some directions. A **single** k^* with clean angular modulation maximizes "Im without Re".

Real-space constructions that realize this PSD

- **Isotropic-chiral labyrinths / vortex tilings**: band-limited labyrinthine patterns where interfaces locally **swirl with one handedness** (no mirror twins).
- **Spiral-inclusion ensembles**: nearly circular inclusions replaced by **single-handed spirals** of pitch $\ell^* \approx 2\pi/k^*$.
- **Pinwheel/rotating-chevron textures**: elements arranged so local edges prefer a rotational sense (CW or CCW) uniformly over space.

These look rotationally uniform at large scale (no preferred axis) but **lack mirror symmetry**, which is exactly what you want.

A minimal recipe (you can implement directly)

1. **Band-limit**: start from Gaussian noise, apply a narrow radial band-pass at k^*.
2. **Angular mask** on the ring:
   $$M(\theta) = \max\big[0,\ 1 + \beta \sin(m^*(\theta - \phi_0))\big],$$
   with $m^* = 2$ (or whatever mode the Imag map shows strongest), $0.6 \lesssim \beta \leq 1$.
   Multiply the Fourier field by $M(\theta)$. (This injects sine—mirror-odd—content while keeping $S \geq 0$.)
3. **Inverse FFT → threshold** to hit the target volume fraction (re-threshold to maintain it if you iterate).
4. **Tune**: increase β to boost Im; reduce any residual a_0 by subtracting a small uniform ring component before rectification; keep $c_m \approx 0$.

Quick acceptance checks

- **Odd-under-reflection energy fraction** on the ring,
  $\frac{\sum s_m^2}{\sum(c_m^2 + s_m^2)} \geq 0.8.$
- **Uniform component** a_0 near zero (e.g., ≤5% of total ring power).
- **Directional sweep (0–315°)**: Im almost flat and high; Re remains close to the noise floor.

Do / Don't (fast summary)

- **Do:** narrow band at k^*; smooth **sine-type** angular modulation; one global handedness.
- **Don't:** mirror-symmetric spectra, uniform ring power, multiple length scales, or sharp angular notches.

Follow these and you'll get a microstructure that is **directionally uniform**, **Imag-dominant**, and **Real-suppressed**—while keeping the volume fraction fixed.

Figure 9: The primary engineering goal is to optimize the effective property in all directions by simultaneously minimizing its real component while maximizing its imaginary component, subject to a constant volume fraction.