Uncertainty Propagation Through Green's Kernels and Gaussian Process Inference Dynamics

Chi-Jen Roger Lo and Joan Lasenby

Department of Engineering, University of Cambridge, United Kingdom

Abstract

We study uncertainty propagation under Gaussian process (GP) priors whose covariance is induced by Green or resolvent operators associated with latent generators. Rather than specifying non-stationary dependence through input warping or other coordinate deformations, we construct covariance from discounted propagation under an underlying semigroup. This yields a separation between covariance geometry and transport: the self-adjoint dissipative part of the generator provides the operator from which prior covariance is constructed, while the skew-adjoint part contributes transport within the full propagation without being encoded directly as covariance. We derive observation, posterior, and propagation identities in operator form, and compare this framework with pullback, stationary, and Matérn baselines. Experiments are organised as a four-stage ladder ascending from three-dimensional electrostatics through a frequency-domain Helmholtz proxy and time-domain semigroup propagation to an operator-mismatch ablation. At each stage, the Green prior is realised as a Gaussian Markov random field (GMRF) whose precision is the discretised PDE operator; after assembly, the posterior mean is obtained by a single sparse linear solve. Across the reported experiments, Green's kernel priors often outperform tuned Matérn-3/2 and RBF baselines on boundary-aware and PDE-residual metrics, with larger gains under heterogeneous material coefficients and semigroup propagation. In the late-time propagation setting the normalised RMSE gap exceeds $4\times$. The central claim is that, when an informative governing operator is available, covariance induced by that operator yields boundary-respecting and dynamically interpretable uncertainty—especially on operator-sensitive diagnostics—under the reported synthetic conditions.

1. Introduction

Green's functions first emerged as a powerful analytical device in classical field theory, especially electromagnetism, where they encode the response of a field to a localised source under a governing differential operator; closely related propagator constructions later became central in quantum theory, statistical physics, and stochastic dynamics, where they describe how disturbances, amplitudes, or probability mass evolve under an underlying generator (Green, 1828; Jackson, 1999; Arfken et al., 2013; Feynman and Hibbs, 2010; Evans, 2010). The common thread is that structure should be derived from the propagation law itself rather than imposed externally. The same principle motivates our inferential viewpoint: when a physically or statistically informative propagation law is available, covariance construction should, where possible, reflect that law rather than an extrinsic coordinate deformation. In this paper, 'inference dynamics' refers to posterior adaptation and forecast propagation from a fixed observation set, not to recursive sequential data assimilation. Under that scope, prior uncertainty is acted upon by observations, transported through latent evolution,

and reshaped into posterior uncertainty; this motivates an operator-theoretic viewpoint in which uncertainty adaptation is studied through Green operators and semigroup propagation.

A persistent weakness of many non-stationary Gaussian process constructions is that the covariance is modified by *coordinate design* rather than derived from an underlying dynamical law. Input warping, deep feature maps, and ad hoc non-stationary kernels can be effective empirically, but they often obscure the mechanism by which uncertainty propagates through the system (Snelson et al., 2004; Snoek et al., 2014; Higdon, 1998). In particular, a pullback kernel $k_{\mathrm{warp}}(x, x') = k_0(z(x), z(x'))$ encodes similarity after reparameterisation, but does not in general reflect the propagation law of a latent Markov or semigroup dynamics.

In this paper we take a different route. Rather than imposing geometry through warping, we derive covariance from the generator of a stochastic or physical evolution. Let L denote a linear operator governing latent propagation on a state space X. The central object is the Green or resolvent operator $R_\beta = (\beta I - L)^{-1}$ for $\beta > 0$, or, when only the dissipative part of the dynamics is to define covariance, the corresponding symmetric Green operator $C_\beta = (\beta I - S)^{-1}$, where $L = S + A$ with S self-adjoint dissipative and A skew-adjoint. This construction yields a Gaussian prior whose covariance is induced by discounted propagation under the latent dynamics rather than by an externally chosen deformation of coordinates.

The resulting viewpoint has three advantages. First, it is structurally aligned with Markov semigroups, SPDE formulations, covariance-operator viewpoints, and Green-function theory (Pazy, 1983; Engel and Nagel, 2000; Berlinet and Thomas-Agnan, 2004). Secondly, it separates covariance geometry from transport within the full non-equilibrium dynamics: the self-adjoint part of the generator provides the operator from which covariance is constructed, while the skew-adjoint part contributes transport or circulation without being encoded directly as covariance. Thirdly, it makes uncertainty propagation analytically transparent. Observation, filtering, and forward evolution all act on the prior through operator compositions, so posterior and predictive uncertainty can be written directly in terms of Green operators and semigroup actions.

This perspective is closely related to the SPDE view of Gaussian fields, in which Matérn-type covariances arise as inverse elliptic operators (Whittle, 1963; Lindgren et al., 2011; Rue and Held, 2005), to latent-force constructions that derive kernels from differential equations (Álvarez et al., 2009, 2013), and to the state-space literature in which certain Gaussian process priors admit finite-dimensional Markov representations (Hartikainen and Särkkä, 2010; Särkkä and Hartikainen, 2012; Särkkä and Solin, 2019; Li et al., 2024). However, our emphasis is not merely on computational acceleration. The primary goal is to formulate uncertainty propagation itself in operator-theoretic terms: how prior uncertainty induced by a Green operator is transmitted through observation maps, propagated under semigroup evolution, and transformed into posterior uncertainty over latent and predictive quantities.

Main message.

When the law governing latent propagation is known, covariance should respect that law.

This yields a mathematically cleaner framework for non-stationary modelling, inverse problems, and dynamical inference when an informative latent operator is available.

Contributions.

1. We define Gaussian process priors from Green/resolvent operators associated with latent generators, emphasising the distinction between self-adjoint covariance structure and skew-adjoint transport.

2. We derive explicit formulas for uncertainty propagation through linear observation operators, semigroup evolution, and downstream quantities of interest.

3. We compare Green's kernel propagation with input-warped kernels and show that even when a warp variable is itself Markovian, a pure pullback kernel remains a quotient or coordinate kernel rather than the Green's kernel of the full latent dynamics; a non-injective pure pullback cannot distinguish points on the same fibre, although coordinate-aware augmentations can remove this invariance.

4. We identify conditions under which the framework connects to classical SPDE, state-space, and filtering constructions.

5. We present a four-stage experimental ladder on 3D electromagnetic problems—electrostatics, frequency-domain Helmholtz, time-domain semigroup propagation, and operator-mismatch ablation—showing in our reported experiments that Green's kernel priors often deliver lower RMSE and better boundary and PDE-residual diagnostics compared with tuned stationary baselines, with advantages that are largest under heterogeneous material coefficients and that persist, in degraded form, even under operator misspecification.

2. Related Work and Positioning

This paper sits between several literatures that are often discussed separately. First, in Gaussian process modelling, non-stationarity has commonly been introduced through coordinate deformation, output warping, process convolutions, or other kernel-engineering devices rather than through an explicit latent propagator (Snelson et al., 2004; Snoek et al., 2014; Higdon, 1998; Paciorek and Schervish, 2004; Booth et al., 2024). Our critique is not that these constructions are useless, but that their covariance semantics are usually geometric or phenomenological rather than dynamical.

Secondly, the SPDE and GMRF literature provides the clearest precedent for deriving covariance from an operator rather than choosing it directly. In this line of work, Matérn-type Gaussian fields arise as solutions of stochastic elliptic equations, and the associated covariance is the Green operator of the governing differential operator (Whittle, 1963; Lindgren et al., 2011; Rue and Held, 2005; Krainski et al., 2018). Our framework is broader in one respect and narrower in another: broader because we emphasise abstract generator-induced covariance and uncertainty propagation beyond the Matérn family; narrower because we focus on the operator semantics of posterior adaptation rather than on mesh-based approximation alone.

Thirdly, state-space and filtering methods show that certain temporal and spatio-temporal Gaussian process priors admit finite- or infinite-dimensional Markov representations, enabling Kalman filtering and smoothing (Hartikainen and Särkkä, 2010; Särkkä and Hartikainen, 2012; Särkkä and Solin, 2019; Solin and Särkkä, 2014; Zhu et al., 2023; Li et al., 2024; Chen et al., 2025). That literature is directly relevant to scalable implementations of our framework, but the present paper is not primarily a complexity paper. The main emphasis here is the structural question of how uncertainty should propagate when covariance is induced by a generator, rather than on state-space representation alone.

Fourthly, covariance-operator viewpoints connect the work to Bayesian inverse problems, Gaussian measures on function spaces, and operator-theoretic uncertainty quantification (Bogachev, 1998; Stuart, 2010; Sullivan, 2015; Jorgensen and Tian, 2024). There is also a structurally parallel connection to Koopman and transfer-operator methods: both use semigroups and resolvents, but Koopman theory approximates eigenfunctions and transport operators, whereas the present work uses related resolvent objects as covariance operators for Bayesian uncertainty propagation (Mezić, 2005; Klus et al., 2020; Hamzi et al., 2026; Hertel et al., 2025). Recent GP-based operator-learning work has revived interest in uncertainty-aware mappings between function spaces (Cowperthwaite and Moss, 2025); this is complementary to, rather than a replacement for, identifying the posterior covariance itself with the Green or resolvent operator of the latent dynamics.

3. Green's Kernels

We work with a latent generator L on a state space X, its associated strongly continuous semigroup $(P_t)_{t\geq 0}$, and a symmetric dissipative part S that induces covariance through the resolvent $(\beta I - S)^{-1}$. Throughout, the central distinction is between *coordinate-based* covariance construction and *propagator-based* covariance construction. In the former, correlation is specified by choosing a similarity function, possibly after a reparameterisation of the inputs. In the latter, covariance is induced by the same operator that governs latent propagation. This paper adopts the second viewpoint. When the latent dynamics admit a generator–semigroup description, uncertainty may be transported, conditioned, and summarised by explicit operator composition.

3.1. Gaussian Process Priors from Green's Kernels

Let $(\mathsf{X}, \mathcal{X}, \mu)$ be a measurable state space, and let L be the generator of a strongly continuous semigroup $(P_t)_{t\geq 0}$ on $L^2(\mu)$ (Pazy, 1983; Engel and Nagel, 2000; Ethier and Kurtz, 1986). We write $L = S + A$, where S is self-adjoint and negative semidefinite, and A is skew-adjoint, whenever such a decomposition is available on the domain of interest. The operator S represents dissipative geometry, while A represents conservative transport or circulation.

For $\beta > 0$, define the symmetric Green operator $C_\beta := (\beta I - S)^{-1}$. Whenever C_β is positive, self-adjoint, and sufficiently regular to admit an integral kernel, we write $(C_\beta f)(x) = \int_{\mathsf{X}} k_\beta(x, x') f(x')\, \mu(\mathrm{d}x')$. We then define the latent field $f \sim \mathcal{GP}(m, k_\beta)$ with mean m and covariance kernel k_β, in the same Gaussian process sense used throughout the standard GP literature (Rasmussen and Williams, 2006) and compatible with operator-valued covariance viewpoints in probability and RKHS theory (Berlinet and Thomas-Agnan, 2004; Da Prato and Zabczyk, 1992).

Proposition 1 (Validity of the Green covariance operator). *Suppose S is self-adjoint and negative semidefinite on $L^2(\mu)$. Then for every $\beta > 0$, the resolvent*

$$C_\beta = (\beta I - S)^{-1}$$

is bounded, self-adjoint, and positive. Hence C_β defines a valid covariance operator on $L^2(\mu)$.

Proof. Since S is self-adjoint and $\sigma(S) \subset (-\infty, 0]$, the spectrum of $\beta I - S$ lies in $[\beta, \infty)$. Therefore $(\beta I - S)^{-1}$ is bounded and self-adjoint by the spectral theorem. Positivity follows from $\langle f, (\beta I - S)^{-1} f\rangle \geq 0$ for all $f \in L^2(\mu)$. □

The resolvent C_β coincides, up to norm equivalence, with the canonical RKHS kernel associated with S for uniformly elliptic generators on bounded domains (Hamzi et al., 2026); this also explains why Dirichlet conditions encoded by S are inherited by the GP prior. A concrete one-dimensional kernel and the corresponding sparse precision calculation are given in Appendix B.3.

3.2. Uncertainty under Observation

Let $y = Hf + \varepsilon$, where H is a bounded linear operator from the latent function space to an observation space Y, and $\varepsilon \sim \mathcal{N}(0, \Sigma_\varepsilon)$ is independent Gaussian noise. Since f is Gaussian under the Green prior, the observation y is also Gaussian.

Proposition 2 (Marginal uncertainty propagation). *Let $f \sim \mathcal{N}(m, C_\beta)$, $y = Hf + \varepsilon$, $\varepsilon \sim \mathcal{N}(0, \Sigma_\varepsilon)$. Then*

$$y \sim \mathcal{N}\big(Hm,\ HC_\beta H^* + \Sigma_\varepsilon\big).$$

The observation covariance $\Sigma_y = HC_\beta H^* + \Sigma_\varepsilon$ is the Green covariance propagated through the observation geometry.

Proposition 3 (Posterior uncertainty). *Under the assumptions of Proposition 2, the posterior covariance is*

$$C_{\beta|y} = C_\beta - C_\beta H^*(HC_\beta H^* + \Sigma_\varepsilon)^{-1} HC_\beta.$$

Uncertainty reduction is controlled by the extent to which observed directions intersect the propagated prior geometry.

3.3. Uncertainty over Downstream Quantities of Interest

Many inverse and forecasting problems require not the full latent field but a downstream quantity of interest $q = Bf$, where B is a bounded linear functional or operator (local field values, boundary fluxes, spatial averages).

Proposition 4 (Uncertainty propagation to quantities of interest). *If $f \mid y \sim \mathcal{N}(m_{\beta|y}, C_{\beta|y})$, then*

$$q \mid y \sim \mathcal{N}\big(Bm_{\beta|y},\ BC_{\beta|y}B^*\big).$$

Observation, posterior conditioning, and downstream decision layers thus all operate by explicit operator composition on the Green covariance.

3.4. Semigroup Propagation of Uncertainty

Let $(T_t)_{t\geq 0}$ be a strongly continuous semigroup acting on the latent state space. In the simplest case, $T_t = e^{tL}$ or $T_t = e^{tS}$, depending on whether one evolves under the full dynamics or only the dissipative part. Define the propagated latent field $f_t := T_t f$.

Proposition 5 (Propagation of Gaussian uncertainty). *If $f \sim \mathcal{N}(m, C_\beta)$, then for each $t \geq 0$,*

$$f_t \sim \mathcal{N}(T_t m,\ T_t C_\beta T_t^*), \qquad C_\beta(t) := T_t C_\beta T_t^*.$$

Prior uncertainty is not merely shifted in input space; it is transported by the semigroup itself. When $T_t{=}e^{tS}$, one obtains dissipative contraction of uncertainty along stable directions. When $T_t{=}e^{tL}$ with $L{=}S{+}A$, one obtains a decomposition into dissipative shaping and conservative transport; the dynamics under L are generally not diagonal in the S-eigenbasis once A is active, so the separation is cleanest when viewed in terms of energy dissipation rates rather than individual modal trajectories.

Remark 6 (What is specific to Green's kernels). *Proposition 5 is a Gaussian–semigroup identity that holds for any covariance operator, not only Green's kernels. The empirical question is not whether $C_t = T_t C T_t^*$ is algebraically valid, but whether the covariance one starts from encodes the right geometry before propagation begins. The experiments below test calibration and structural robustness under a fixed propagation law, rather than the transport identity itself.*

3.5. Comparison with Pure Input-Warped and Factor-Space Kernels

For comparison, consider an input-warped kernel $k_{\text{warp}}(x, x') = k_0\big(z(x), z(x')\big)$ (Snoek et al., 2014; Snelson et al., 2004; Higdon, 1998). Even if the latent warp variable z is itself Markov-generated, k_{warp} is a pullback of a chosen base kernel rather than the Green kernel of the full latent dynamics on x. The two coincide only when z is a bijective semigroup conjugacy with k_0 the Green's kernel of the transformed generator; if z is a non-invertible Markov factor, then k_{warp} represents the Green's kernel of a quotient dynamics and discards information along fibres of the factor map. This distinction matters for uncertainty propagation: a Green's kernel answers *how does latent uncertainty propagate under the actual dynamics?* whereas a warped kernel answers *how similar do two points appear after reparameterisation?* The comparison below is deliberately narrow: it applies to pure pullback kernels, not to all non-stationary or coordinate-aware GP constructions.

Setting	Method	RMSE	Cov95 (%)	Bnd Cov95 (%)	Gauss Res.	BC Viol.
Homo	Green	**0.0072** (14)	**99.7** (0.3)	**100.0** (0.0)	7.2	**0.0000**
	Mat-3/2	0.0127 (07)	92.6 (1.5)	86.1 (4.2)	18.2	0.0063
	RBF	0.0096 (06)	92.4 (1.4)	83.5 (4.8)	**6.2**	0.0052
Hetero	Green	**0.0059** (09)	**99.9** (0.1)	**100.0** (0.0)	**6.1**	**0.0000**
	Mat-3/2	0.0128 (07)	92.8 (1.4)	86.5 (3.9)	19.4	0.0064
	RBF	0.0094 (06)	92.7 (1.3)	84.0 (4.5)	7.0	0.0053

Table 1. Stage I 3D electrostatic inverse problem (10 seeds, mean±std in last two digits). Across the reported seeds, the Green prior achieves the lowest mean RMSE, essentially perfect boundary coverage, zero BC violation for the discretised posterior mean, and consistently low Gauss-law residuals, with the smallest mean residual in the heterogeneous setting. The advantage is larger under heterogeneous ε.

Proposition 7 (Fibre invariance of pure pullback kernels). *Let $z : \mathsf{X} \to \mathsf{Z}$ and $k_z(x, x') := k_0(z(x), z(x'))$. If $z(x) = z(x')$, then $k_z(x, u) = k_z(x', u)$ for every $u \in \mathsf{X}$. Consequently, for any data collected through point evaluations, the posterior mean and variance at x and x' coincide. A non-injective pure pullback kernel cannot distinguish states on the same fibre of z.*

Proof. If $z(x) = z(x')$, then $k_z(x, u) = k_0(z(x), z(u)) = k_0(z(x'), z(u)) = k_z(x', u)$ for every u. Posterior means and variances under point evaluations are built from kernel sections $k_z(\cdot, u_i)$ and the diagonal $k_z(\cdot, \cdot)$, so equal kernel sections imply equal posteriors at x and x'. □

Coordinate-aware augmentations and spatially varying lengthscales can remove this invariance; Appendix A.1 discusses this limitation and its implications for baselines.

4. Experiments: 3D Electrostatics-to-Maxwell Ladder

We evaluate the Green's kernel framework on a four-stage experimental ladder that ascends from three-dimensional electrostatics to a frequency-domain Helmholtz proxy and time-domain semigroup propagation. Each stage adds physical complexity while the underlying GMRF inference machinery remains unchanged: the Green prior is always the resolvent $C_\beta = \sigma^2(\beta I + K)^{-1}$ of the relevant PDE operator, realised as a sparse precision matrix. The posterior mean is then obtained by a single sparse LU solve after assembly; posterior variances are estimated separately via the Hutchinson stochastic trace estimator (Appendix B.3). Baselines (Matérn-3/2 and RBF) are tuned by the same observed-data log marginal likelihood criterion at every stage. They are diagnostic rather than exhaustive baselines; physics-aware and non-stationary alternatives are discussed in Appendix A.1. All experiments use 10 random seeds, 200 random observation sites on $[0, 1]^3$ discretised on a 14^3 grid (M=2,744 interior DOF), observation noise σ_n=0.02 (Stages I–II) or σ_n=0.005 (Stage III), and homogeneous Dirichlet boundary conditions. The ground-truth source is a random superposition of sine modes with spectral decay $(l^2+m^2+n^2)^{-\alpha}$, α=1.

4.1. Stage I: 3D Electrostatic Inverse Problem

The governing operator is $K = -\mathrm{div}(\varepsilon\nabla)$, discretised by a second-order finite-difference stencil with harmonic-mean averaging of ε at cell faces. Two material settings are tested: (i) *homogeneous*, $\varepsilon \equiv 1$; and (ii) *heterogeneous*, with a radius-0.25 dielectric inclusion of permittivity ε=20. The Green hyperparameters are selected by grid search; the stationary baselines are tuned over length scale and amplitude.

Table 1 summarises the results. Three observations are consistent across seeds. First, the Green prior achieves the lowest mean RMSE in both settings, with the advantage more pronounced under

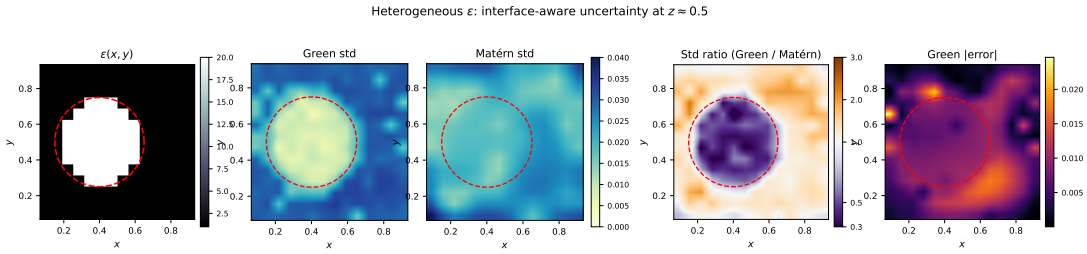

Figure 1. Stage I heterogeneous interface (z≈0.5). Left to right: $\varepsilon(x, y)$ map, Green std, Matérn std, Green / Matérn std ratio, Green |error|. The Green std drops sharply inside the high-ε inclusion (red dashed circle), reflecting the operator's material structure, while the Matérn std is spatially uniform. The std-ratio panel (orange >1, blue <1) shows that the Green prior redistributes uncertainty according to the material interface—the central claim of this paper.

Setting	Method	RMSE	Cov95 (%)	Bnd Cov95 (%)	BC Viol.
Homo	Green	**0.0072**	**99.7**	**100.0**	**0.0000**
	Mat-3/2	0.0125	92.5	85.8	0.0061
	RBF	0.0094	92.3	83.2	0.0050
Hetero	Green	**0.0075**	**99.8**	**100.0**	**0.0000**
	Mat-3/2	0.0125	92.6	85.9	0.0062
	RBF	0.0093	92.5	83.8	0.0051

Table 2. Stage II Helmholtz inverse problem (representative seed). In the seed shown here, the qualitative pattern resembles Stage I: the Green prior leads on RMSE and boundary metrics. The mass term $\omega^2\varepsilon$ does not disrupt the GMRF formulation.

heterogeneous ε where the operator encodes the material interface. Second, the discretised Green posterior mean has zero boundary-condition violation (up to solver tolerance), since the GMRF precision inherits the Dirichlet structure of K. Third, the Gauss-law residual $\|\mathrm{div}(\varepsilon\nabla\hat{\phi})\|$ is smallest for the Green posterior. A representative homogeneous-ε slice is shown in Figure 4 (Appendix C.3). The most striking visual result from Stage I is the heterogeneous setting: Figure 1 shows that the Green posterior standard deviation drops sharply inside the high-ε inclusion, reflecting the operator's material structure, whereas the Matérn standard deviation is spatially uniform. The std-ratio panel confirms that the Green prior allocates relatively lower uncertainty inside the inclusion where the operator concentrates correlation.

4.2. Stage II: Frequency-Domain Helmholtz (Scalar Proxy)

We replace the electrostatic operator with the scalar Helmholtz operator $H_\omega = -\mathrm{div}(\mu^{-1}\nabla) + \omega^2\varepsilon$, which can arise as a simplified reduction of the time-harmonic Maxwell system under appropriate polarisation and symmetry assumptions, and serves here as a scalar proxy for frequency-domain electromagnetic problems. The frequency ω=2 is chosen in a regime where the discretised operator remained positive definite for both material settings tested. Two settings are again tested: homogeneous (ε=1, μ=1) and heterogeneous (ε=10, μ=5 in a spherical inclusion).

Table 2 shows that the qualitative pattern from Stage I carries over. The mass term $\omega^2\varepsilon$ adds a positive diagonal to the stiffness matrix, keeping the discretised operator positive definite in this setup and the GMRF formulation unchanged. Figure 6 (Appendix C.3) shows the posterior slice with PDE-residual panels; in the displayed slice, the Matérn mean exhibits large-scale sign bias, while the Green mean appears to capture the oscillatory pattern more faithfully.

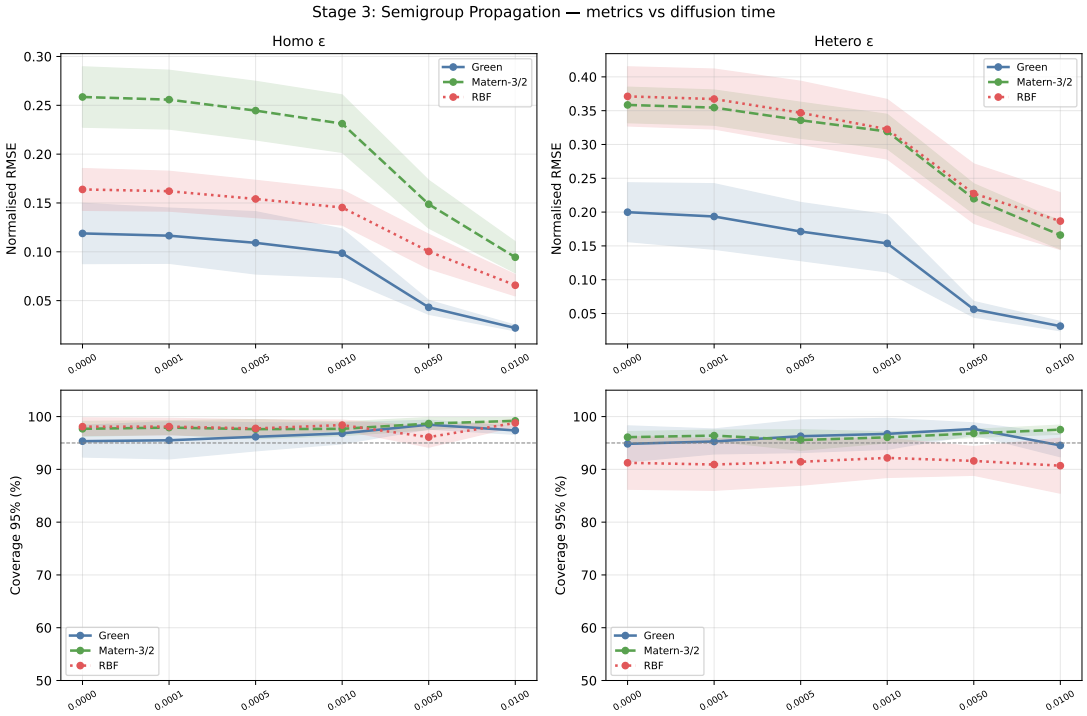

Figure 2. Stage III semigroup propagation. Normalised RMSE vs. time for Green and Matérn-3/2 under homogeneous and heterogeneous operators. The Green prior tracks the decaying field more closely in the reported run; by t=0.01 the normalised RMSE gap exceeds 4× (homo) and 5× (hetero).

4.3. Stage III: Time-Domain Semigroup Propagation

This stage probes the paper's main empirical implication: that covariance induced by a governing operator can track forward propagation more faithfully than covariance chosen independently of the dynamics. We observe the electrostatic potential at t=0 and propagate the posterior under the semigroup $T_t = e^{-tK}$, comparing against the reference evolved field $\phi(t) = T_t\phi_0$ at times $t \in \{0, 10^{-4}, 5{\times}10^{-4}, 10^{-3}, 5{\times}10^{-3}, 10^{-2}\}$. Here $K = -S$ is the positive-definite stiffness matrix, with eigenvalues $\lambda_k > 0$ (note the sign convention: these are eigenvalues of K, not S, so larger λ_k corresponds to faster decay). In the truncated spectral basis (200 eigenvalues), the Green prior covariance propagates as

$$C_\beta(t)_{kk} = \frac{\sigma^2 e^{-2\lambda_k t}}{\beta + \lambda_k},$$

which matches the modal decay factor $e^{-2\lambda_k t}$ in the truncated spectral basis used here. Stationary baselines are re-fitted at each time t using the propagated observations. This comparison is structurally asymmetric—the Green prior exploits the semigroup structure while the stationary baselines do not—which is intentional: the question is whether encoding the propagation law yields measurably better uncertainty under an operator-aligned synthetic mechanism, not whether both methods use the same information.

Figure 2 shows that the Green prior's normalised RMSE remains low as the field decays, while the Matérn-3/2 baseline deteriorates sharply at late times. At t=0.01, the homogeneous Green NRMSE is 0.022 vs. 0.094 for Matérn (4.3× advantage), widening to 0.031 vs. 0.166 under the heterogeneous operator (5.4× advantage). In the operator-aligned synthetic setting tested here, propagation-aligned covariance produces measurably different and materially better predictive behaviour under forward evolution.

Method	RMSE	Cov95 (%)	Gauss Res.	BC Viol.
Green (correct)	**0.0059**	**99.9**	7.2	**0.0000**
Green (misspec.)	0.0065	99.4	28.3	**0.0000**
Matérn-3/2	0.0128	92.8	18.2	0.0064
RBF	0.0094	92.7	**6.2**	0.0053

Table 3. Stage IV operator mismatch ablation (10 seeds, mean values). Even with the wrong ε, the misspecified Green prior has lower RMSE than both stationary baselines in this experiment and maintains zero BC violation for the discretised posterior mean. The Gauss-law residual usefully diagnoses the mismatch in this setup.

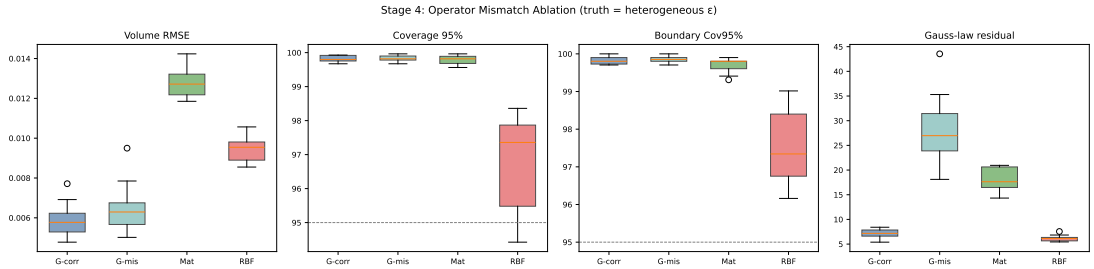

Figure 3. Stage IV operator mismatch ablation. Boxplots over 10 seeds comparing RMSE, Gauss-law residual, and BC violation for the four methods. In this experiment, the misspecified Green prior has lower RMSE than both stationary baselines.

4.4. Stage IV: Operator Mismatch Ablation

The previous stages assumed that the true PDE operator is known. In practice, the material coefficients may be uncertain. We test robustness by comparing four methods on the heterogeneous electrostatic problem: (i) Green with the *correct* operator K_{hetero}, (ii) Green with a *misspecified* operator K_{homo} (assuming $\varepsilon{\equiv}1$ when the truth has a $\varepsilon{=}20$ inclusion), (iii) Matérn-3/2, and (iv) RBF.

Table 3 reveals two findings. First, even the misspecified Green prior (RMSE 0.0065) outperforms both Matérn-3/2 (0.0128) and RBF (0.0094) on reconstruction error, suggesting that, in this setting, a "wrong but structured" operator prior can still be preferable to no operator at all. Second, the Gauss-law residual usefully diagnoses the mismatch: 28.3 for the misspecified Green vs. 7.2 for the correct one, providing a built-in model-checking signal. Both Green variants maintain zero boundary-condition violation for the discretised posterior mean, since the Dirichlet structure is inherited from K regardless of whether ε is correctly specified.

5. Discussion

In the reported experiments, the Green's kernel viewpoint shows three advantages that are most pronounced on operator-sensitive diagnostics.

Physics-encoded boundaries. The GMRF precision $Q = \sigma^{-2}(\beta I + K)$ inherits the Dirichlet structure of the stiffness matrix K, so the discretised posterior mean has zero boundary-condition violation up to solver tolerance (Tables 1–3). Translation-invariant kernels do not explicitly encode this; their boundary uncertainty is governed primarily by nearby observations, which explains the systematic coverage gap at the domain boundary in our experiments.

Material-aware covariance. Under heterogeneous ε, the Green prior induces uncertainty patterns aligned with the material interface, producing lower RMSE and a smaller Gauss-law residual

in this setting than stationary baselines that do not explicitly encode the material discontinuity but can respond to it indirectly through observations. The Stage IV ablation (Table 3) shows that, in this setting, even a misspecified operator prior has lower RMSE than no-operator baselines, while the Gauss-law residual provides a built-in diagnostic for operator mismatch.

Propagation fidelity. Under semigroup evolution (Stage III), the Green covariance tracks physical modal decay mode by mode, whereas stationary baselines do not encode that damping through the governing semigroup. In our late-time propagation setting, the resulting 4–5× NRMSE advantage (Figure 2) is the clearest empirical signature of propagation-aligned covariance; we emphasise that this comparison is structurally asymmetric by design (see Section 4.3).

The resolvent also admits a stochastic extension via the Feynman–Kac formula: when L generates a diffusion (X_t), the Green's kernel is $G_\beta(x, x') = \mathbb{E}_x\big[\int_0^\infty e^{-\beta t}\,\delta(X_t - x')\,\mathrm{d}t\big]$, making Monte Carlo kernel approximation straightforward and opening a route to generators too complex for spectral decomposition (Hamzi et al., 2026; Hertel et al., 2025).

Several extensions remain: scalable SPDE discretisation and iterative solvers (Lindgren et al., 2011); state-space Kalman formulations for sequential assimilation; a fuller error decomposition separating modelling, discretisation, and semigroup-mismatch contributions; non-symmetric generators with advective transport; and extension to the full vector Maxwell system. The broader motivation is that, when an informative latent operator is available, covariance construction can be aligned with the physical and geometric structure of the problem rather than with an extrinsic coordinate design, making covariance, propagation, and posterior adaptation part of a single dynamical calculus (Stuart, 2010; Sullivan, 2015).

6. Conclusion and Prospects

When an informative governing operator is available, its physical and geometric structure can be reflected directly in the prior. Green's kernel priors realise this principle by inducing covariance from a governing PDE operator, yielding a compact operator calculus for observation, posterior adaptation, semigroup evolution, and downstream uncertainty quantification. The four-stage 3D electromagnetic ladder supports this principle empirically: in the reported synthetic experiments, Green priors often outperform tuned stationary baselines on RMSE and are strongest on boundary-aware and PDE-residual diagnostics, with advantages that are largest under material heterogeneity and forward evolution. Even under operator misspecification, the structured prior can outperform unstructured alternatives in the reported synthetic setting, while the Gauss-law residual provides a built-in diagnostic for model error. These results support a broader principle: when an informative latent propagation law is available, covariance should, where possible, be induced by it. Encoding that structure into the prior yields boundary-respecting, dynamically interpretable uncertainty, while flexible baselines remain useful for effects not captured by the chosen operator.

References

Mauricio Álvarez, David Luengo, and Neil D. Lawrence. Latent force models. In *Proceedings of the Twelfth International Conference on Artificial Intelligence and Statistics*, volume 5 of *Proceedings of Machine Learning Research*, pages 9–16. PMLR, 2009. URL https://proceedings.mlr.press/v5/alvarez09a.html.

Mauricio A. Álvarez, David Luengo, and Neil D. Lawrence. Linear latent force models using Gaussian processes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 35(11):2693–2705, 2013. doi: 10.1109/TPAMI.2013.86.

George B. Arfken, Hans J. Weber, and Frank E. Harris. *Mathematical Methods for Physicists*. Academic Press, 7 edition, 2013.

Alain Berlinet and Christine Thomas-Agnan. *Reproducing Kernel Hilbert Spaces in Probability and Statistics*. Kluwer Academic, 2004.

Vladimir I. Bogachev. *Gaussian Measures*. American Mathematical Society, 1998.

Annie S. Booth, Andrew Cooper, and Robert B. Gramacy. Nonstationary Gaussian process surrogates. *arXiv preprint arXiv:2305.19242*, 2024. doi: 10.48550/arXiv.2305.19242. URL https://arxiv.org/abs/2305.19242.

Wenlong Chen, Naoki Kiyohara, Harrison Bo Hua Zhu, Jacob Curran-Sebastian, Samir Bhatt, and Yingzhen Li. Recurrent memory for online interdomain Gaussian processes. *arXiv preprint arXiv:2502.08736*, 2025. doi: 10.48550/arXiv.2502.08736. URL https://arxiv.org/abs/2502.08736. Published at the 39th Conference on Neural Information Processing Systems (NeurIPS 2025).

Thomas Cowperthwaite and Henry Moss. GPGreen: Learning linear operators with Gaussian processes. OpenReview, 2025. URL https://openreview.net/forum?id=eNFMYj2ZG6. Differentiable Systems and Scientific Machine Learning Workshop at NeurIPS 2025.

Giuseppe Da Prato and Jerzy Zabczyk. *Stochastic Equations in Infinite Dimensions*. Cambridge University Press, 1992.

Klaus-Jochen Engel and Rainer Nagel. *One-Parameter Semigroups for Linear Evolution Equations*. Springer, 2000. doi: 10.1007/b97696.

Stewart N. Ethier and Thomas G. Kurtz. *Markov Processes: Characterization and Convergence*. Wiley, 1986.

Lawrence C. Evans. *Partial Differential Equations*. American Mathematical Society, 2 edition, 2010.

Richard P. Feynman and Albert R. Hibbs. *Quantum Mechanics and Path Integrals*. Dover Publications, emended edition edition, 2010.

George Green. *An Essay on the Application of Mathematical Analysis to the Theories of Electricity and Magnetism*. T. Wheelhouse, 1828.

Boumediene Hamzi, Houman Owhadi, and Umesh Vaidya. Kernel methods for stochastic dynamical systems with application to Koopman eigenfunctions: Feynman–Kac representations and RKHS approximation. *arXiv preprint arXiv:2603.01077*, 2026. doi: 10.48550/arXiv.2603.01077. URL https://arxiv.org/abs/2603.01077.

Jouni Hartikainen and Simo Särkkä. Kalman filtering and smoothing solutions to temporal Gaussian process regression models. In *2010 IEEE International Workshop on Machine Learning for Signal Processing*, pages 379–384. IEEE, 2010. doi: 10.1109/MLSP.2010.5589113. URL https://ieeexplore.ieee.org/document/5589113.

Maximiliano Hertel, Friedrich M. Philipp, Manuel Schaller, and Karl Worthmann. Koopman for stochastic dynamics: Error bounds for kernel extended dynamic mode decomposition. *arXiv preprint arXiv:2512.20247*, 2025. doi: 10.48550/arXiv.2512.20247. URL https://arxiv.org/abs/2512.20247.

Dave Higdon. A process-convolution approach to modelling temperatures in the North Atlantic Ocean. *Environmental and Ecological Statistics*, 5(2):173–190, 1998. doi: 10.1023/A:1009665427990.

John David Jackson. *Classical Electrodynamics*. Wiley, 3 edition, 1999.

Palle E. T. Jorgensen and James Tian. Operator-valued Gaussian processes and their covariance kernels. *Infinite Dimensional Analysis, Quantum Probability and Related Topics*, 27(5):2350020, 2024. doi: 10.1142/S0219025723500200.

Stefan Klus, Feliks Nüske, Sebastian Peitz, Jan H. Niemann, Cecilia Clementi, and Christof Schütte. Data-driven approximation of the Koopman generator: Model reduction, system identification, and control. *Physica D: Nonlinear Phenomena*, 406:132416, 2020. doi: 10.1016/j.physd.2020.132416.

Elias T. Krainski, Virgilio Gómez-Rubio, Haakon Bakka, Amanda Lenzi, Daniela Castro-Camilo, Daniel Simpson, Finn Lindgren, and Håvard Rue. *Advanced Spatial Modeling with Stochastic Partial Differential Equations Using R and INLA*. Chapman and Hall/CRC, 2018. doi: 10.1201/9780429031892. URL https://doi.org/10.1201/9780429031892.

Weihan Li, Yule Wang, Chengrui Li, and Anqi Wu. Markovian Gaussian process: A universal State-Space representation for stationary temporal Gaussian process. *arXiv preprint arXiv:2407.00397*, 2024. doi: 10.48550/arXiv.2407.00397. URL https://arxiv.org/abs/2407.00397.

Zongyi Li, Nikola Kovachki, Kamyar Azizzadenesheli, Burigede Liu, Kaushik Bhattacharya, Andrew Stuart, and Anima Anandkumar. Fourier Neural Operator for parametric Partial Differential Equations. In *International Conference on Learning Representations*, 2021. URL https://openreview.net/forum?id=c8P9NQVtmnO.

Finn Lindgren, Håvard Rue, and Johan Lindström. An explicit link between Gaussian fields and Gaussian Markov Random Fields: The Stochastic Partial Differential Equation approach. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 73(4):423–498, 2011. doi: 10.1111/j.1467-9868.2011.00777.x.

Lu Lu, Pengzhan Jin, Guofei Pang, Zhongqiang Zhang, and George Em Karniadakis. Learning nonlinear operators via DeepONet based on the universal approximation theorem of operators. *Nature Machine Intelligence*, 3(3):218–229, 2021. doi: 10.1038/s42256-021-00302-5.

Igor Mezić. Spectral properties of dynamical systems, model reduction and decompositions. *Nonlinear Dynamics*, 41(1–3):309–325, 2005. doi: 10.1007/s11071-005-2824-x.

Christopher J. Paciorek and Mark J. Schervish. Nonstationary covariance functions for Gaussian process regression. In *Advances in Neural Information Processing Systems 16*, pages 273–280. MIT Press, 2004. URL https://proceedings.neurips.cc/paper/2003/hash/326a8c055c0d04f5b06544665d8bb3ea-Abstract.html.

Amnon Pazy. *Semigroups of Linear Operators and Applications to Partial Differential Equations*. Springer, 1983. doi: 10.1007/978-1-4612-5561-1.

Maziar Raissi, Paris Perdikaris, and George Em Karniadakis. Physics-Informed Neural Networks: A deep learning framework for solving forward and inverse problems involving nonlinear Partial Differential Equations. *Journal of Computational Physics*, 378:686–707, 2019. doi: 10.1016/j.jcp.2018.10.045.

Carl Edward Rasmussen and Christopher K. I. Williams. *Gaussian Processes for Machine Learning*. MIT Press, 2006.

Håvard Rue and Leonhard Held. *Gaussian Markov Random Fields: Theory and Applications*. Chapman and Hall/CRC, 2005.

Simo Särkkä and Jouni Hartikainen. Infinite-dimensional Kalman filtering approach to spatio-temporal Gaussian process regression. In *Proceedings of the 15th International Conference on Artificial Intelligence and Statistics*, volume 22 of *Proceedings of Machine Learning Research*, pages 993–1001. PMLR, 2012. URL https://proceedings.mlr.press/v22/sarkka12.html.

Edward Snelson, Zoubin Ghahramani, and Carl E. Rasmussen. Warped Gaussian processes. In *Advances in Neural Information Processing Systems 16*, pages 337–344. MIT Press, 2004. URL https://proceedings.neurips.cc/paper_files/paper/2003/hash/6b5754d737784b51ec5075c0dc437bf0-Abstract.html.

Jasper Snoek, Kevin Swersky, Richard S. Zemel, and Ryan P. Adams. Input warping for Bayesian optimization of non-stationary functions. In *Proceedings of the 31st International Conference on Machine Learning*, volume 32 of *Proceedings of Machine Learning Research*, pages 1674–1682. PMLR, 2014. URL https://proceedings.mlr.press/v32/snoek14.html.

Arno Solin and Simo Särkkä. Explicit link between periodic covariance functions and state space models. In *Proceedings of the Seventeenth International Conference on Artificial Intelligence and Statistics*, volume 33 of *Proceedings of Machine Learning Research*, pages 904–912. PMLR, 2014. URL https://proceedings.mlr.press/v33/solin14.html.

Andrew M. Stuart. Inverse problems: A Bayesian perspective. *Acta Numerica*, 19:451–559, 2010. doi: 10.1017/S0962492910000061.

Timothy J. Sullivan. *Introduction to Uncertainty Quantification*. Springer, 2015. doi: 10.1007/978-3-319-23395-6.

Simo Särkkä and Arno Solin. *Applied Stochastic Differential Equations*. Cambridge University Press, 2019. doi: 10.1017/9781108186735.

Peter Whittle. Stochastic processes in several dimensions. *Bulletin of the International Statistical Institute*, 40:974–994, 1963.

Andrew Gordon Wilson, Zhiting Hu, Ruslan Salakhutdinov, and Eric P. Xing. Deep Kernel Learning. In *Proceedings of the 19th International Conference on Artificial Intelligence and Statistics*, volume 51 of *Proceedings of Machine Learning Research*, pages 370–378. PMLR, 2016. URL https://proceedings.mlr.press/v51/wilson16.html.

Harrison Zhu, Carles Balsells-Rodas, and Yingzhen Li. Markovian Gaussian process variational autoencoders. In *Proceedings of the 40th International Conference on Machine Learning*, volume 202 of *Proceedings of Machine Learning Research*, pages 42938–42961. PMLR, 2023. URL https://proceedings.mlr.press/v202/zhu23b.html.

A. Scope, Baselines, and Misspecification

A.1. Baseline Scope and Interpretation

The main experiments compare against tuned Matérn-3/2 and RBF kernels because they are canonical GP baselines and make the operator effect easy to isolate. This is not an exhaustive benchmark against all possible non-stationary or physics-aware methods. Several stronger alternatives are relevant: non-stationary covariance functions with spatially varying lengthscales (Paciorek and Schervish, 2004; Booth et al., 2024), SPDE/GMRF Matérn models fit on meshes (Lindgren et al., 2011; Krainski et al., 2018), latent force models (Álvarez et al., 2009, 2013), deep kernel learning (Wilson et al., 2016), physics-informed neural networks (Raissi et al., 2019), and neural operators such as DeepONet and Fourier neural operators (Lu et al., 2021; Li et al., 2021).

These methods answer different questions. An SPDE Matérn model can be viewed as a physics-aware or geometry-aware baseline when the operator family is learned or chosen flexibly; it is closest in spirit to the present approach. Latent force models derive output covariance from differential-equation responses to latent driving processes; they are also closely related, although the emphasis is usually multi-output modelling rather than posterior covariance as a Green operator of the state itself. Deep kernel learning and coordinate-aware pullback kernels can preserve spatial distinguishability by augmenting the warped representation with the original coordinates, for example through $k_0(z(x), z(x')) + \tau^2 k_x(x, x')$. This directly avoids the fibre invariance in Proposition 7; the proposition is therefore best read as a statement about pure factor-space kernels, not as a limitation of all non-stationary GP practice. PINNs and neural operators target solution or operator learning and can exploit PDE residual losses or paired input-output functions, but they are not direct drop-in replacements for a calibrated GP posterior without additional uncertainty quantification machinery.

The synthetic experiments should therefore be interpreted as a mechanism test: when the data-generating operator is known and the inferential target is an operator-sensitive field, does covariance induced by that operator improve boundary, residual, and propagation diagnostics? The results support this narrower claim. A broader benchmark should include at least one spatially varying SPDE or coordinate-aware deep-kernel baseline and, for solution-operator tasks, a neural operator with uncertainty calibration.

A.2. Unknown Operators and Misspecification

The main theory assumes that an informative operator is available. In applications this operator may be imperfect or only partially known. There are several natural extensions. First, one can parameterise the operator, for example $K_\theta = -\mathrm{div}(\varepsilon_\theta \nabla) + V_\theta$, and infer θ by marginal likelihood, cross-validation, or a hierarchical Bayesian prior. Secondly, when only coarse physics are known, the Green prior can be used as a structured base prior while residual diagnostics such as Gauss-law violation or PDE residuals are used for model checking. This is the role of Stage IV: the misspecified operator still improves RMSE in the reported setting, but its residual is much worse than the correctly specified operator, making the mismatch visible. Thirdly, the operator itself could be learned from data using system identification, Koopman-generator estimation, or operator-learning methods, then inserted into the same resolvent construction.

These extensions change the statistical problem. If the operator is learned from the same data used for posterior inference, the uncertainty in K_θ should propagate into the field posterior rather than being treated as fixed. The present paper does not solve that hierarchical problem. Its contribution is the conditional statement: given an informative operator, the Green/resolvent covariance gives a transparent way to align prior geometry, posterior conditioning, and forward uncertainty propagation.

A.3. Practical Scope and Takeaways

The experiments are synthetic and operator-aligned: the proposed prior receives the governing operator, while the main baselines are intentionally simple stationary GP kernels. The results should therefore be read as evidence for a conditional claim rather than a universal benchmark: when a trusted operator is available, inducing covariance from that operator improves boundary, residual, and propagation diagnostics in these controlled settings. They do not rule out stronger non-stationary or physics-aware alternatives, such as spatially varying SPDE models, latent force models, coordinate-aware deep kernels, PINNs, or neural operators with calibrated uncertainty.

For an applied problem, the proposed construction should be tried when three conditions hold: the field of interest lives on a domain with meaningful boundary conditions; a sparse discretisation of an informative operator is available; and downstream decisions care about operator-sensitive diagnostics such as residuals, fluxes, boundary behaviour, or propagated uncertainty. If these conditions fail, flexible non-stationary baselines or operator-learning extensions are better suited to discovery rather than operator-aligned uncertainty propagation.

A decisive follow-up benchmark should vary both mesh resolution and observation count, include at least one spatially varying SPDE or coordinate-aware deep kernel baseline, and test an application where the operator comes from measured or independently estimated physical parameters rather than from the simulator used to generate the truth.

B. Operator, Spectral, and Computational Details

B.1. Resolvent and RKHS Remarks

Remark 8 (Resolvent form). *When the full semigroup admits a transition density $p_t(x, x')$, the Green's kernel associated with the full generator L is formally $G_\beta(x, x') = \int_0^\infty e^{-\beta t} p_t(x, x')\,\mathrm{d}t$; in the reversible/SPDE setting this is the operator-theoretic bridge exploited in the Matérn–SPDE correspondence (Whittle, 1963; Lindgren et al., 2011; Krainski et al., 2018). If L is reversible, or if one restricts to the symmetric part S, then G_β or its symmetric counterpart yields a valid covariance kernel, subject to the conditions of Proposition 1. In non-reversible settings, it is often preferable to let S define covariance and reserve A for propagation.*

Remark 9 (Kernel equivalence). *For uniformly elliptic generators on bounded domains, Hamzi et al. (2026) show that three constructions yield equivalent reproducing-kernel Hilbert spaces (as sets, with equivalent norms): (i) the Lions variational (energy) inner product, (ii) the Green's function convolution kernel, and (iii) the resolvent operator $(\beta I - S)^{-1}$. C_β is therefore, up to norm equivalence, the canonical RKHS kernel associated with S. This equivalence also explains why the Dirichlet boundary conditions encoded by S are inherited by the GP prior.*

B.2. Spectral Structure of Posterior Uncertainty

The Green's kernel viewpoint becomes particularly transparent in the spectral domain. Suppose the symmetric operator S admits an orthonormal eigensystem $\{(\lambda_j, e_j)\}_{j\geq 1}$ on $L^2(\mu)$ with $Se_j = \lambda_j e_j$, $\lambda_j \leq 0$. Then the Green covariance operator acts diagonally:

$$C_\beta e_j = \frac{1}{\beta - \lambda_j} e_j, \qquad \mathrm{Var}[\langle f, e_j\rangle] = \frac{1}{\beta - \lambda_j}.$$

Modes with λ_j near zero therefore receive substantially larger prior variance, while rapidly damped high-frequency modes with $\lambda_j \ll 0$ are more strongly regularised. The Green prior is thus a multi-scale prior induced directly by the dissipative operator rather than by an externally chosen similarity metric.

After conditioning on observations, posterior uncertainty contracts only in the directions interrogated by the observation operator H. Modes orthogonal to the range of H retain their prior variance, while modes that are visible through H are selectively shrunk. Projecting the posterior covariance onto the eigenbasis of S therefore makes explicit which spatial frequencies have actually been learned from the data.

Under semigroup propagation, $C_t = T_t C_{\beta|y} T_t^*$, and if $T_t = e^{tS}$ and the posterior covariance is viewed in the same modal basis, with posterior modal variances $\sigma^2_{j|y}$, then each mode damps independently:

$$C_t e_j = e^{2t\lambda_j}\, \sigma^2_{j|y}\, e_j.$$

Dissipative propagation therefore removes uncertainty at rates dictated by the spectrum of S: low-frequency components persist, while high-frequency uncertainty vanishes rapidly. For the full non-self-adjoint generator $L = S + A$, the skew-adjoint component may mix or transport modes, so that the dynamics are no longer diagonal in the S-eigenbasis; however, the energy dissipation envelope is still controlled by the spectrum of S in a quadratic sense. For the prior itself, $\sigma^2_{j|y}$ is replaced by $(\beta - \lambda_j)^{-1}$.

This modal viewpoint links directly to the three-dimensional experiments in Section 4. There the electrostatic operator $-\mathrm{div}(\varepsilon\nabla)$ on $[0,1]^3$ is discretised on a 14^3 grid (2,744 DOF), and the electrostatic potential is observed at 200 random locations with noise σ_n=0.02. The random sine-mode source $f = \sum a_{lmn} \sin(l\pi x)\sin(m\pi y)\sin(n\pi z)$ with spectral decay $(l^2{+}m^2{+}n^2)^{-1}$ places appreciable energy in mid-to-high modes. That is exactly the regime in which spectral allocation matters: the Green prior, whose modal variance is $\sigma^2/(\beta{+}\lambda_j)$, retains substantially more prior mass in the modes that carry signal than the stationary baselines, whose isotropic spectra decay too aggressively. This spectral advantage is amplified under semigroup propagation (Stage III, Section 4.3), where each mode damps at rate $e^{-2\lambda_j t}$ and the Green prior tracks this physical decay while stationary baselines cannot.

The spectral perspective therefore does more than decorate the theory: it explains why propagated calibration succeeds or fails. Prior allocation, posterior shrinkage, and temporal damping all become explicit in eigenmode coordinates. Once an eigen-decomposition of the discretised operator S_h has been computed, propagation and modal diagnostics reduce to diagonal operations in that basis, so the cost per additional time step is only $O(N)$ after the upfront decomposition.

B.3. Computational Realisation and One-Dimensional Example

A one-dimensional Dirichlet problem shows the same construction in closed form. Let $\mathsf{X} = (0,1)$, take $S = \mathrm{d}^2/\mathrm{d}x^2$ with $\mathcal{D}(S) = H^2(0,1) \cap H^1_0(0,1)$, and set $C_\beta = (\beta I - S)^{-1}$. Then k_β solves

$$\left(\beta - \frac{\mathrm{d}^2}{\mathrm{d}x^2}\right) k_\beta(x,x') = \delta(x-x'), \qquad k_\beta(0,x') = k_\beta(1,x') = 0.$$

Writing $\gamma = \sqrt{\beta}$,

$$k_\beta(x,x') = \frac{\sinh(\gamma \min\{x,x'\}) \sinh(\gamma(1-\max\{x,x'\}))}{\gamma \sinh\gamma}.$$

Equivalently, with $e_j(x) = \sqrt{2}\sin(j\pi x)$ and $-Se_j = (j\pi)^2 e_j$,

$$k_\beta(x,x') = \sum_{j=1}^{\infty} \frac{e_j(x) e_j(x')}{\beta + (j\pi)^2}.$$

This makes visible the two properties used in the experiments: the covariance inherits the boundary condition, and modal prior variance is allocated according to the spectrum of the differential operator.

On a grid with M interior nodes and spacing $h = 1/(M+1)$, the corresponding positive stiffness matrix for $K = -S$ is

$$K = \frac{1}{h^2}\begin{bmatrix} 2 & -1 & & \\ -1 & 2 & -1 & \\ & \ddots & \ddots & \ddots \\ & & -1 & 2 \end{bmatrix},$$

which is the one-dimensional analogue of the sparse precision matrices used below.

The present experiments use second-order finite-difference discretisation of the PDE operator on $[0,1]^3$ with homogeneous Dirichlet boundary conditions. Rather than forming the dense resolvent C_β, we exploit the GMRF precision formulation: the prior precision is $Q_{\text{prior}} = \sigma^{-2}(\beta I + K)$ where K is the (sparse) stiffness matrix of $-\text{div}(\varepsilon\nabla)$ or the Helmholtz operator. The posterior precision is $Q_{\text{post}} = Q_{\text{prior}} + \sigma_n^{-2} D_{\text{obs}}$ where D_{obs} is a diagonal indicator at observation sites, and the posterior mean requires a single sparse LU solve after assembly. Posterior variances are estimated via the Hutchinson stochastic trace estimator with 120 Rademacher probes. For semigroup propagation (Stage III), we use a truncated spectral decomposition (200 eigenvalues via ARPACK) so that the propagated covariance is diagonal in the eigenbasis. This approach is adequate for the mechanism-revealing 14^3 grid used here; for larger problems, iterative solvers, SPDE-on-mesh formulations, or Monte Carlo Feynman–Kac approximations provide scalable alternatives (Hamzi et al., 2026; Hertel et al., 2025; Lindgren et al., 2011).

B.4. Computational Scaling and Implementation Details

The experiments are intentionally small enough to make repeated ablations and seed averages inexpensive, but the computational bottleneck differs from a dense GP. Let M denote the number of grid or mesh degrees of freedom and N_{obs} the number of observations. A dense GP baseline forms an $N_{\text{obs}} \times N_{\text{obs}}$ covariance matrix and is dominated by the usual $O(N_{\text{obs}}^3)$ Cholesky factorisation. The Green GMRF formulation instead works with an $M \times M$ sparse precision. Its cost is controlled by the sparse solve or factorisation of Q_{post}, together with the cost of estimating marginal variances.

Method	Dominant object	Scaling interpretation
Dense RBF/Matérn GP	$N_{\text{obs}} \times N_{\text{obs}}$ dense covariance	Cholesky cost $O(N_{\text{obs}}^3)$; independent of mesh size unless predictions are requested at many grid sites.
Green GMRF, direct solve	$M \times M$ sparse precision Q_{post}	Sparse factorisation depends on graph dimension and ordering; for regular 3D grids the fill-in can grow substantially, so direct LU is a moderate-scale baseline rather than the final scalable implementation.
Green GMRF, iterative solve	Sparse matrix-vector products with Q_{post}	Each iteration costs $O(\text{nnz}(Q_{\text{post}}))$; preconditioning and multigrid become the relevant scalability tools.
Spectral propagation	Truncated eigenbasis of K	After r modes are computed, each additional time point is diagonal and cheap; the upfront eigensolve is the bottleneck.

Table 4. Computational comparison for the mechanisms used in this paper. The Green method shifts the main cost from the number of observations to the discretised operator size and sparsity structure.

In the reported experiments, $M = 14^3 = 2{,}744$ and $N_{\text{obs}} = 200$. This is enough to test the covariance mechanism, boundary behaviour, and operator-mismatch diagnostics, but it is not intended as a large-scale computational benchmark. Increasing N_{obs} mainly affects the diagonal observation term

$\sigma_n^{-2} D_{\text{obs}}$ in the Green precision when observations lie on grid nodes; increasing the grid resolution increases M and is therefore the dominant cost for the operator method. For off-grid observations, $H^* \Sigma_\varepsilon^{-1} H$ is no longer purely diagonal but remains sparse or low-rank if local interpolation is used.

All experiments were implemented in Python using NumPy, SciPy sparse linear algebra, and Matplotlib; the run used for this camera-ready version was checked with Python 3.12.12, NumPy 2.3.5, SciPy 1.16.3, and Matplotlib 3.10.8. The scripts are the four experiments_maxwell_stage*.py files. The reported posterior means use sparse direct solves through SciPy's sparse linear algebra interface, while Stage III uses eigsh for the 200-mode truncated spectral propagation.

C. Experimental Details and Supplementary Results

C.1. Experimental Protocol

The experiment scripts are experiments_maxwell_stage{1,2,3,4}.py. All methods are tuned by observed-data log marginal likelihood on the same grid. For the Green prior, $\beta \in \{0.001, 0.01, 0.05, 0.1, 0.5, 1, 5, 10\}$ and $\sigma^2 \in \{1, 2, 5, 10, 20, 50, 100, 200, 500, 1000\}$. For the stationary baselines, $\ell \in \{0.05, 0.08, 0.1, 0.15, 0.2, 0.3\}$ and amplitude $\in \{0.1, 0.5, 1, 2, 5, 10\}$. Posterior variances are estimated via the Hutchinson trace estimator with 120 Rademacher random vectors applied to Q_{post}^{-1}. Stage III uses a truncated spectral decomposition (200 eigenvalues via scipy.sparse.linalg.eigsh) for semigroup propagation. Each stage is repeated over 10 independent random seeds controlling both the source realisation and the observation locations.

C.2. Grid Resolution, Observation Count, and Real Data

The grid size and observation count were chosen to make the four-stage ladder repeatable over seeds and hyperparameter grids. The resulting 14^3 grid is not meant to prove large-scale performance. It is sufficient for the specific mechanism tested here: whether a sparse operator prior inherits boundary conditions, reflects material coefficients, and propagates uncertainty according to the semigroup spectrum. Those are structural effects that should be visible before moving to larger meshes.

The number of observations plays a different role for Green and dense GP baselines. For the dense baselines, increasing N_{obs} increases the Cholesky cost directly. For the grid-aligned Green GMRF, additional observations mainly add diagonal precision at observed nodes, so the solve is still governed by the sparse operator size M. This difference is favourable when many observations lie on a fixed mesh, but it also means that refinement studies should vary M and N_{obs} separately. A useful future benchmark would report error and wall-clock time across a grid of pairs (M, N_{obs}), not just a single resolution.

Real data introduce additional complications absent from the synthetic ladder. Observation locations may be off-grid, boundary conditions may be uncertain, material coefficients may be only partially observed, and the forward operator may be an approximation to the experimental system. In that setting, the residual diagnostics used here become model-checking tools rather than proof of correctness. For example, a Green prior with the wrong ε can still improve boundary behaviour, but a large PDE residual should be interpreted as evidence that the operator, coefficient field, or observation model needs revision.

The most direct real-data applications are therefore those in which a trusted operator is already part of the scientific workflow: elliptic inverse problems with known boundary conditions, diffusion or heat-transfer systems with measured material maps, and frequency-domain field interpolation where the operating frequency and medium model are known. Applications with unknown constitutive laws require the hierarchical or learned operator extensions discussed in Appendix A.2. We leave such validation to future work rather than presenting an under-specified real-data example that would confound operator learning with uncertainty propagation.

C.3. Supplementary Experimental Results

This subsection collects supplementary figures and tables that complement the main-text results. Figures 4–7 provide posterior-slice visualisations and boxplots for Stages I–III. Tables 5–6 give the full Stage III time-series metrics underlying Figure 2 (10 seeds, mean±std), and Table 7 gives per-seed metrics for the Stage IV operator-mismatch ablation.

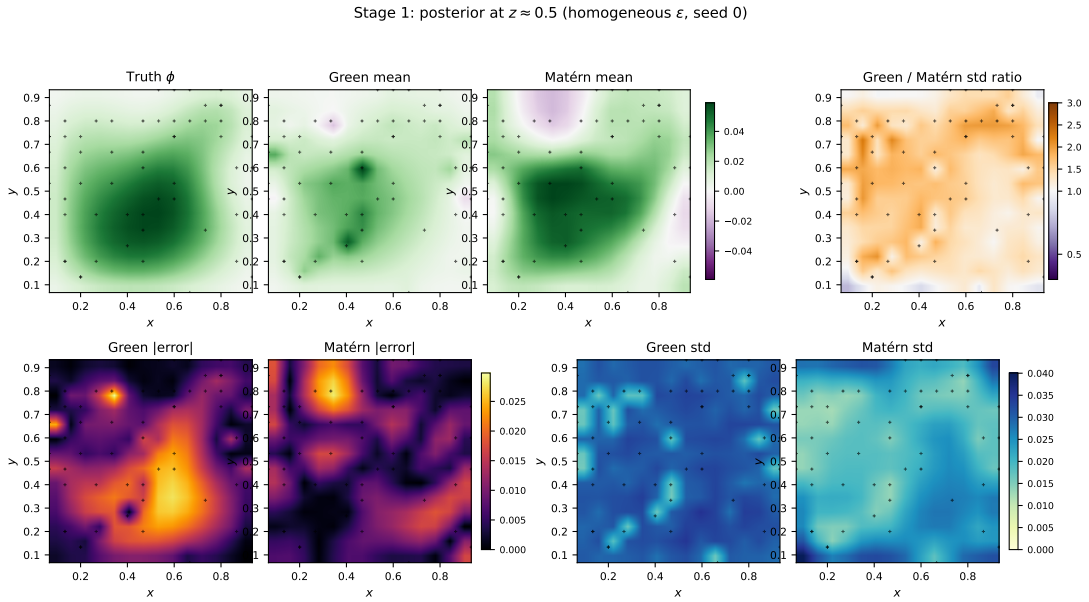

Figure 4. Stage I posterior slice ($z \approx 0.5$, homogeneous ε, seed 0). Top: truth, Green mean, Matérn mean, Green / Matérn std ratio. Bottom: Green |error|, Matérn |error|, Green std, Matérn std. The Matérn mean appears smoother, but the shared error colorbar shows its error is more uniformly spread, yielding higher aggregate RMSE. The std-ratio panel (orange >1) highlights where the Green prior assigns greater uncertainty, especially near boundaries.

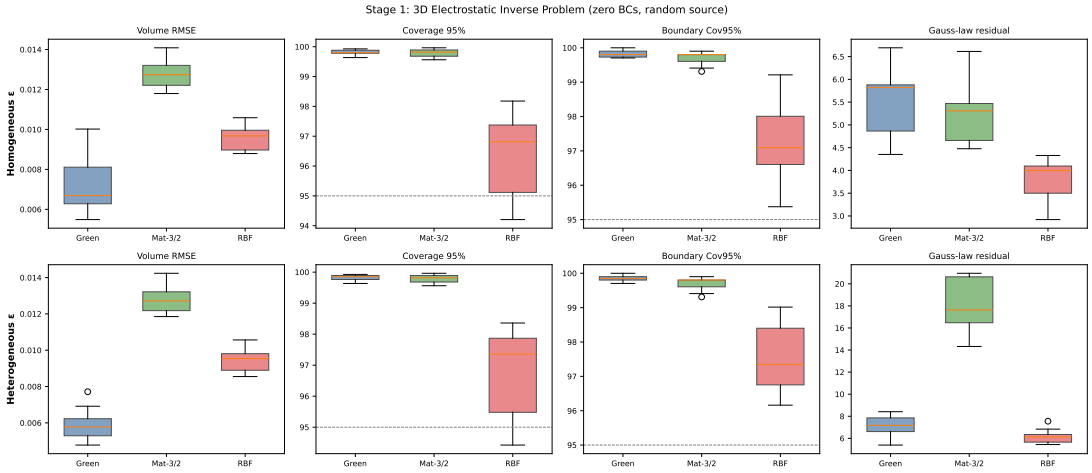

Figure 5. Stage I boxplots over 10 seeds: RMSE, coverage, Gauss-law residual, and BC violation for homogeneous and heterogeneous settings.

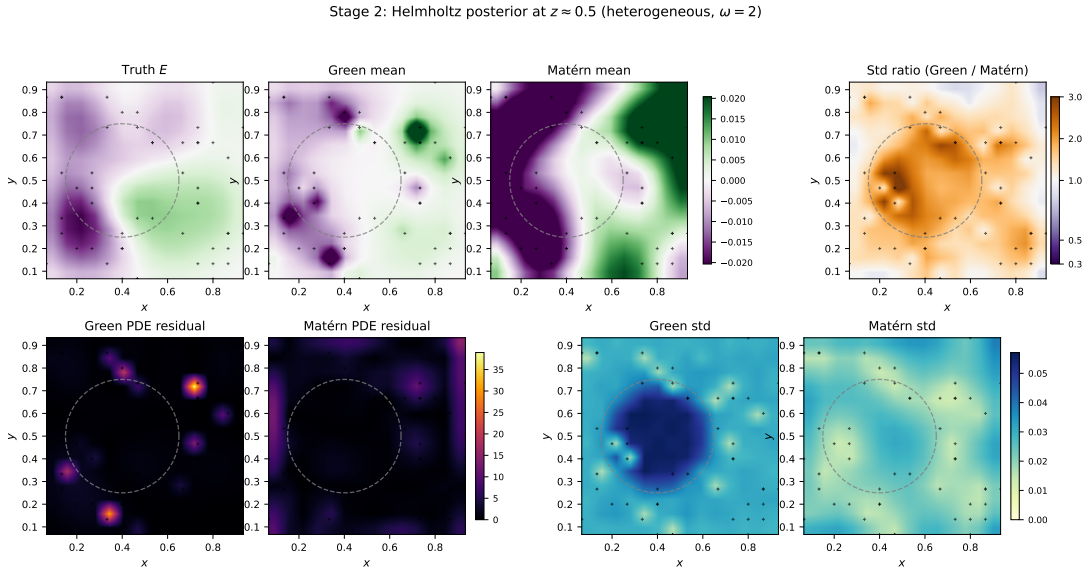

Figure 6. Stage II Helmholtz posterior slice ($z \approx 0.5$, representative seed). Top: truth E, Green mean, Matérn mean, Green / Matérn std ratio. Bottom: Green PDE residual, Matérn PDE residual, Green std, Matérn std. The Matérn mean shows large-scale bias (sign structure does not match the truth), while the Green mean captures the oscillatory pattern more faithfully despite appearing patchier.

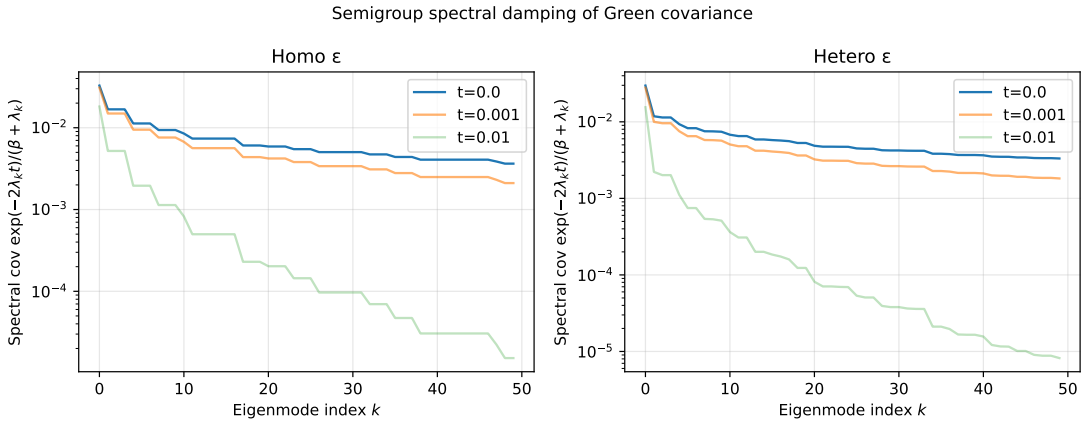

Figure 7. Stage III spectral decay. Posterior modal variance as a function of eigenvalue index at several propagation times, confirming the $e^{-2\lambda_k t}/(\beta + \lambda_k)$ law.

t	Method	NRMSE	Cov95 (%)	Bnd Cov95 (%)
0	Green	**0.119** (31)	95.3 (3.0)	95.3 (3.1)
	Mat-3/2	0.259 (31)	97.7 (1.4)	99.4 (0.5)
	RBF	0.164 (21)	98.1 (1.7)	99.2 (0.6)
10^{-3}	Green	**0.099** (25)	96.8 (2.0)	96.5 (2.4)
	Mat-3/2	0.231 (29)	97.7 (1.3)	99.5 (0.3)
	RBF	0.145 (18)	98.4 (0.9)	99.4 (0.3)
5×10^{-3}	Green	**0.043** (7)	98.5 (0.8)	98.7 (0.8)
	Mat-3/2	0.149 (24)	98.7 (1.1)	99.5 (0.4)
	RBF	0.100 (18)	96.1 (1.9)	96.5 (1.8)
10^{-2}	Green	**0.022** (3)	97.4 (0.7)	97.4 (1.4)
	Mat-3/2	0.094 (16)	99.2 (0.6)	99.3 (0.4)
	RBF	0.066 (11)	98.8 (0.7)	98.7 (1.0)

Table 5. Stage III semigroup propagation: homogeneous ε (10 seeds). Selected time points from Figure 2. By t=0.01 the Green NRMSE is 4.3× lower than Matérn-3/2.

t	Method	NRMSE	Cov95 (%)	Bnd Cov95 (%)
0	Green	**0.200** (43)	94.8 (3.4)	95.1 (3.3)
	Mat-3/2	0.359 (26)	96.1 (1.0)	99.4 (0.3)
	RBF	0.371 (44)	91.2 (5.0)	97.9 (2.1)
10^{-3}	Green	**0.154** (42)	96.7 (2.9)	97.0 (3.1)
	Mat-3/2	0.319 (25)	96.1 (1.0)	99.4 (0.4)
	RBF	0.323 (44)	92.2 (3.7)	98.5 (0.8)
5×10^{-3}	Green	**0.056** (12)	97.7 (1.1)	98.7 (0.7)
	Mat-3/2	0.220 (22)	96.8 (0.7)	99.2 (0.4)
	RBF	0.227 (44)	91.6 (2.7)	98.4 (0.9)
10^{-2}	Green	**0.031** (7)	94.5 (2.1)	96.3 (1.8)
	Mat-3/2	0.166 (20)	97.5 (0.9)	99.1 (0.3)
	RBF	0.187 (42)	90.7 (5.2)	97.2 (2.8)

Table 6. Stage III semigroup propagation: heterogeneous ε (10 seeds). By t=0.01 the Green NRMSE is 5.4× lower than Matérn-3/2.

Method	RMSE	Cov95 (%)	Gauss Res.	BC Viol.	Width
Green (correct)	**0.0059** (09)	**99.8** (0.1)	7.2 (1.0)	**0.003** (001)	0.105
Green (misspec.)	0.0065 (13)	**99.8** (0.1)	28.3 (7.4)	**0.003** (001)	0.116
Matérn-3/2	0.0128 (07)	99.8 (0.1)	18.2 (2.3)	0.013 (002)	0.090
RBF	0.0094 (05)	97.0 (1.5)	**6.2** (0.6)	0.011 (002)	0.038

Table 7. Stage IV per-seed summary (10 seeds, mean±std in last digits). The Gauss-law residual for the misspecified Green prior (28.3 ± 7.4) is clearly separated from the correct Green prior (7.2 ± 1.0), providing a reliable mismatch diagnostic.