CAUSAL FIELD THEORY: CAUSAL SEMANTICS FOR PDE-BASED SPATIO-TEMPORAL SYSTEMS

Arash Mehrjou
Max Planck Institute for Intelligent Systems & GSK.ai
arash@distantvantagepoint.com

Bernhard Schölkopf
Max Planck Institute for Intelligent Systems
bs@tuebingen.mpg.de

ABSTRACT

Partial differential equations model complex spatio-temporal phenomena from fluid dynamics to biological systems, yet conventional PDE solvers do not typically offer causal semantics for interventions. We introduce Causal Field Theory (CFT), a framework that endows PDE-based systems with explicit causal structure through regional mechanism interventions, which modify evolution laws within spatial regions, and causal response kernels that quantify how influence propagates. CFT bridges mechanistic simulation and causal reasoning, enabling quantitative intervention comparison via causal flux and causal cones. We prove that macro-scale interventions aligned with coherent propagation modes generate greater causal flux than fragmented micro-scale perturbations; biological systems such as tissue-scale drug delivery and morphogen dynamics provide suitable examples of this principle. Experiments on reaction-diffusion PDEs illustrate the framework and its potential relevance for neural operators and physics-informed learning. CFT may enable interpretable causal analysis of PDE dynamics and principled intervention design for scientific AI.

1 INTRODUCTION

Partial differential equations lie at the heart of modeling spatio-temporal phenomena across physics, biology, and engineering. Recent advances in neural operators (Li et al., 2020; Lu et al., 2019), physics-informed neural networks (Mehrjou et al., 2021; 2023), and data-driven PDE discovery (Brunton et al., 2016; Rudy et al., 2017) have transformed how we simulate such systems. Yet an important gap remains: PDE simulators describe dynamics faithfully but do not typically offer causal semantics for interventions. When we ask how a localized perturbation propagates through the domain, or what would happen under an alternative intervention, standard frameworks provide no principled formalism to answer such questions.

This gap is particularly apparent when interventions are spatially localized while their effects propagate through continuous domains. Biological systems offer a suitable illustration: drug delivery, genetic perturbations, and mechanical modulation act locally, yet their consequences diffuse through tissue via signaling, transport, and mechanical coupling. Classical causal frameworks assume no interference between units (SUTVA) or treat variables as discretely indexed (Rubin, 1974; Pearl, 2009), and thus do not naturally capture geometric propagation. Mechanistic simulators, by contrast, reproduce dynamics with high fidelity but typically support neither counterfactual reasoning nor quantitative comparison of intervention strategies. These limitations motivate a framework that combines the geometric structure of spatial propagation with explicit causal semantics.

We propose **Causal Field Theory (CFT)**, which treats causality as a property of continuous fields evolving under PDE dynamics (see Figure 3 in the Appendix for a conceptual overview). The framework introduces regional mechanism interventions that modify evolution laws within spatial regions, causal response kernels that quantify how influence propagates, and derived quantities, namely causal flux and causal cones, that enable quantitative intervention comparison. It connects to Green's functions and adjoint methods (Evans, 2010) while extending them with explicit intervention semantics. We present a causal formalism for PDE-based spatio-temporal systems that generalizes classical causal inference (Pearl, 2009; Peters et al., 2017) to field-based dynamics, a theoretical result showing that under certain conditions macro-interventions aligned with propaga-

tion modes yield greater causal flux than fragmented micro-interventions, and computational validation on reaction-diffusion systems. CFT may be relevant for neural operator learning (Li et al., 2020; Lu et al., 2019), since once $\mathcal{F}$ is learned, causal quantities can be computed directly, and for interpretable scientific AI more broadly.

2 Framework

2.1 System Model

We consider a spatial domain $\mathcal{D} \subset \mathbb{R}^d$ whose state is represented by a field $X(x,t)$, which may be scalar-valued (e.g., concentration) or vector-valued (e.g., multiple species or mechanical displacements). The field evolves according to

$$\partial_t X(x,t) = \mathcal{F}[X](x,t) + \eta(x,t), \tag{1}$$

where $\mathcal{F}$ encodes the mechanism (combining local dynamics with spatial coupling) and η represents fluctuations. For reaction-diffusion, $\mathcal{F}[X] = D\nabla^2 X + f(X)$ combines diffusion with production and degradation. CFT applies whenever $\mathcal{F}$ is Fréchet-differentiable and exhibits spatial coupling. These assumptions remain valid in many physical and practical systems. As one example, although biological systems are discrete at the cellular level, the continuum approximation holds when cell density is high and signaling diffuses over length scales much larger than intercellular spacing; in such regimes, concentration fields and mechanical properties vary smoothly in space, enabling the framework to describe tissue-scale dynamics.

2.2 Regional Mechanism Interventions

Classical causal inference relies on interventions that set variables to specific values, $\mathbf{do}(X = x)$. In many application domains, however, interventions modify *mechanisms* rather than variable values. As a few examples, in biological systems a drug alters reaction rates locally, a gene knockout removes terms from the dynamics in a specific region, and mechanical modulation changes boundary conditions. We formalize a *regional mechanism intervention* as

$$\mathbf{do}_{\Omega,\tau}[\mathcal{F}]: \quad \mathcal{F}(x,\cdot) \mapsto \mathcal{F}'(x,\cdot) \quad \text{for } x \in \Omega,\ t \in [\tau, \tau+\Delta], \tag{2}$$

which modifies the operator within region Ω during the intervention window while leaving $\mathcal{F}$ unchanged elsewhere. This generalizes the do-operator from variables to spatially localized mechanisms, embedding interventions directly into the system's dynamics.

2.3 Causal Response Kernel

The *causal response kernel* $G(x,t;\Omega,\tau)$ quantifies how an intervention affects the field at each location and time. Formally, it is the functional derivative

$$G(x,t;\Omega,\tau) := \frac{\delta X(x,t)}{\delta\, \mathbf{do}_{\Omega,\tau}[\mathcal{F}]}, \tag{3}$$

which answers the question of how much the intervention on the mechanism in Ω affected the field at (x,t). The response kernel is itself a field over space-time, measuring the sensitivity of the state to the intervention.

Theorem 1 (Well-definedness)**.** *Under mild regularity, G exists, is unique, and satisfies*

$$\partial_t G(x,t;\Omega,\tau) = D\mathcal{F}[X] \cdot G(x,t;\Omega,\tau) + S_\Omega(x,t), \tag{4}$$

where $D\mathcal{F}[X]$ is the Fréchet derivative and S_Ω is the localized source from the intervention.

A proof sketch is in Appendix A.2. This establishes G as a well-defined mathematical object, connecting CFT to Green's functions (Evans, 2010) and adjoint sensitivity analysis.

2.4 Causal Flux and Causal Cone

From the response kernel we derive quantities that make causal influence measurable and comparable. The *causal flux* from Ω to target region Γ is the total integrated influence:

$$\Phi_{\Omega\to\Gamma} = \int_{\Gamma}\int_{\tau}^{\infty} G(x,t;\Omega,\tau)\,dt\,dx. \tag{5}$$

One can compare $\Phi_{\Omega_1\to\Gamma}$ and $\Phi_{\Omega_2\to\Gamma}$ to determine which intervention region exerts stronger causal influence on the target. The *causal cone* is the support of G, that is, the region of space-time where the intervention has non-negligible effect.

Theorem 2 (Finite-Speed Propagation)**.** *If $\mathcal{F}$ satisfies locality and finite-speed propagation, then $G(x,t;\Omega,\tau) = 0$ for $|x - x_\Omega| > c(t-\tau)$ with propagation speed c. For parabolic systems, influence decays rapidly with distance.*

Proof in Appendix A.3. This theorem justifies the causal cone concept and distinguishes CFT from DAG-based causality (which has no geometry) and potential outcomes (which assumes no propagation). The causal cone reveals which regions can be causally affected by an intervention and by when, information that is useful for experimental and intervention design. A glossary of key objects in CFT is provided in Appendix A.7.

3 Scale Alignment Theorem

In spatially extended systems, interventions can be defined at multiple scales, from fine-grained local perturbations to coarse-grained regional modifications. A natural question arises as to when macro-interventions yield stronger causal influence than the aggregation of equivalent micro-interventions. CFT addresses this through *causal coherence*, that is, regimes where interventions at a coarser scale induce more structured or spatially extensive responses than fragmented fine-scale interventions. The key insight is that causal effectiveness depends on alignment between the intervention scale and the system's dominant propagation modes. Micro-level mechanisms generate the dynamics, but macro-level effectiveness arises from geometric and dynamical alignment with coherent propagation structures.

Proposition 1 (Scale Alignment)**.** *Let Ω be partitioned into subregions $\{\Omega_i\}$. Consider a macro-intervention over Ω versus micro-interventions on each Ω_i. Under: (1) linearized regime; (2) macro perturbation equals sum of micro perturbations; (3) spatial coherence with correlation length ℓ_c; (4) $\mathrm{diam}(\Omega) \gtrsim \ell_c$; then*

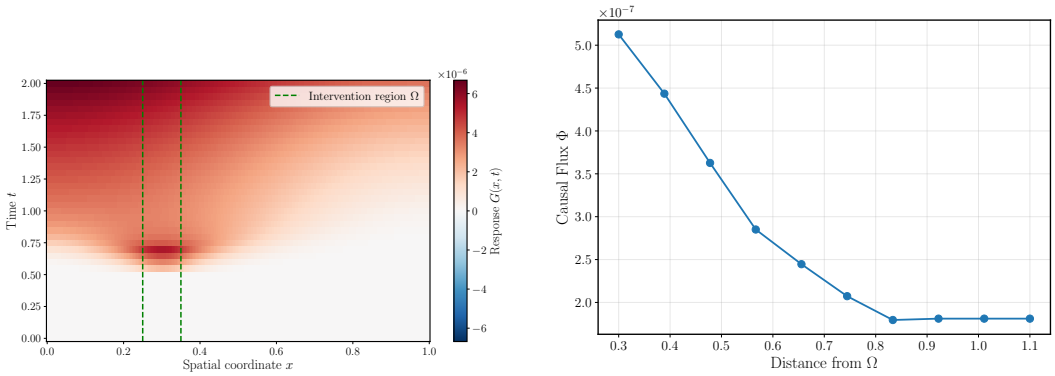
$$\Phi^{\mathrm{macro}}_{\Omega\to\Gamma} \geq \sum_{i=1}^{n} \Phi^{\mathrm{micro},i}_{\Omega_i\to\Gamma}. \tag{6}$$

The proof leverages constructive interference. When micro-interventions are spatially correlated over scales comparable to the correlation length, their effects combine coherently, whereas isolated micro-interventions may suffer destructive interference or decay before reaching the target. As one example, in biological systems regional drug delivery in tissue, when aligned with diffusion length scales, may yield greater causal flux than fragmented local applications. Full proof in Appendix A.4; extended discussion in Appendix A.5.

4 Experiments

We validate the framework on reaction-diffusion PDEs of the form $\partial_t X = D\nabla^2 X + rX(1 - X/K) - \lambda X$, which arise across a wide range of systems, including morphogen gradients and cytokine dynamics in biological systems, chemical reaction networks, and pattern formation in physics.

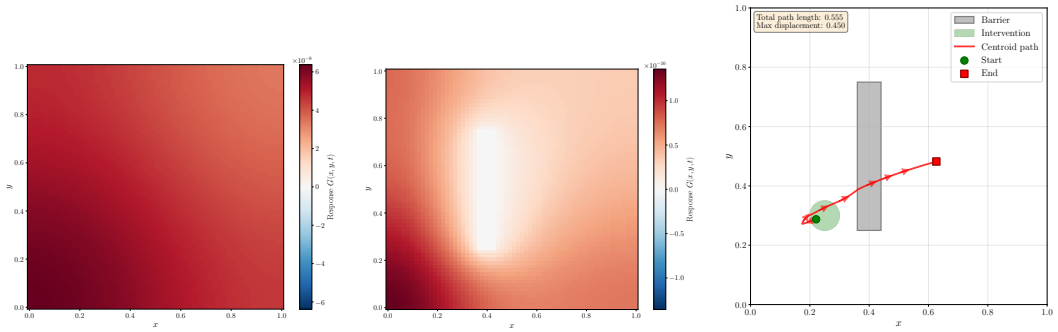
In the first experiment, we apply a regional mechanism intervention by locally increasing the production rate within a region Ω in a 1D system, then track how the causal effect propagates through space and time. Figure 1 shows the response kernel as a heatmap, revealing the causal cone, that is, the region of space-time where the intervention has non-negligible effect. The cone expands over time as influence diffuses outward from the intervention region. The right panel shows how causal flux



(a) Response kernel: causal cone

(b) Causal flux vs. distance

Figure 1: Experiment 1: Causal propagation in 1D reaction-diffusion.



(a) Homogeneous

(b) With barrier

(c) Centroid trajectory

Figure 2: Experiment 2: Geometry alters causal propagation.

decreases with distance, quantifying how effectively causal influence reaches distant target regions and enabling comparison of intervention strategies.

The second experiment examines how geometry affects causal propagation. We compare a homogeneous 2D domain with one containing an impermeable barrier, which may model obstacles in various settings. As one example, in biological systems a scar or impermeable region in tissue can act as such a barrier. Figure 2 shows that causal influence deforms around the obstacle and cannot penetrate it, illustrating that propagation is geometrically mediated. Note that the centroid of the causal flux can pass through the obstacle because it is the center of mass of the distributed influence (which propagates around the barrier), not a physical particle trajectory. Barriers can shield some regions from causal influence while channeling it to others, a feature not naturally captured by graph-based causal frameworks, which lack explicit spatial structure.

5 Discussion and Outlook

CFT integrates naturally with modern approaches to learning PDE dynamics from data. When the mechanism operator $\mathcal{F}$ is learned from spatio-temporal observations (Li et al., 2020; Lu et al., 2019), the response kernel can be computed by solving the linearized evolution equation, yielding interpretability by revealing *how* interventions propagate and supporting intervention design via optimization over Ω and τ to maximize causal flux to target regions. Algorithm 1 in the Appendix outlines a practical workflow for this data-driven causal inference. Similarly, when $\mathcal{F}$ is discovered in interpretable form (Brunton et al., 2016; Rudy et al., 2017), CFT adds causal analysis. In relation to existing causal frameworks, CFT reduces to standard causal effects in the well-mixed limit and to dynamic structural causal models under spatial discretization; the response kernel coincides with Green's functions in the linearized regime (Evans, 2010).

Current experiments use finite-difference PDE solving; extension to learned operators is straightforward. The scale alignment result requires coherence assumptions, and empirical validation in diverse systems remains ongoing. CFT may offer a useful framework for endowing learned or discovered PDE models with causal semantics and enabling interpretable intervention analysis and design.

REFERENCES

Steven L Brunton, Joshua L Proctor, and J Nathan Kutz. Discovering governing equations from data by sparse identification of nonlinear dynamical systems. *Proceedings of the national academy of sciences*, 113(15):3932–3937, 2016.

Chi-Tsong Chen. *Linear system theory and design*. Saunders college publishing, 1984.

Lawrence C Evans. *Partial differential equations*, volume 19. American Mathematical Society, 2010.

Verônica A Grieneisen, Ben Scheres, Paulien Hogeweg, and Athanasius F M Marée. Morphogengineering roots: comparing mechanisms of morphogen gradient formation. *BMC systems biology*, 6(1):37, 2012.

Zongyi Li, Nikola Kovachki, Kamyar Azizzadenesheli, Burigede Liu, Kaushik Bhattacharya, Andrew Stuart, and Anima Anandkumar. Fourier neural operator for parametric partial differential equations. *arXiv preprint arXiv:2010.08895*, 2020.

Lu Lu, Pengzhan Jin, and George Em Karniadakis. Deeponet: Learning nonlinear operators for identifying differential equations based on the universal approximation theorem of operators. *arXiv preprint arXiv:1910.03193*, 2019.

Arash Mehrjou, Mohammad Ghavamzadeh, and Bernhard Schölkopf. Neural lyapunov redesign. In *Proceedings of the 3rd Conference on Learning for Dynamics and Control*, volume 144 of *Proceedings of Machine Learning Research*, pp. 459–470. PMLR, 2021.

Arash Mehrjou, Andrea Iannelli, and Bernhard Schölkopf. Learning dynamical systems using local stability priors. *Journal of Computational Dynamics*, 10(1):175–198, 2023.

Alan V Oppenheim, Alan S Willsky, and Syed Hamid Nawab. *Signals & systems*. Pearson Educación, 1997.

A Wayne Orr, Brian P Helmke, Brett R Blackman, and Martin A Schwartz. Mechanisms of mechanotransduction. *Developmental cell*, 10(1):11–20, 2006.

Judea Pearl. *Causality*. Cambridge university press, 2009.

Jonas Peters, Dominik Janzing, and Bernhard Schölkopf. *Elements of causal inference: foundations and learning algorithms*. The MIT press, 2017.

Donald B Rubin. Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of educational Psychology*, 66(5):688, 1974.

Samuel H Rudy, Steven L Brunton, Joshua L Proctor, and J Nathan Kutz. Data-driven discovery of partial differential equations. *Science advances*, 3(4):e1602614, 2017.

A APPENDIX

A.1 CONCEPTUAL OVERVIEW

Figure 3 illustrates how CFT bridges classical causal frameworks (DAGs, PDE simulation) with spatial intervention analysis. Biological systems provide suitable examples where localized interventions propagate through continuous domains.

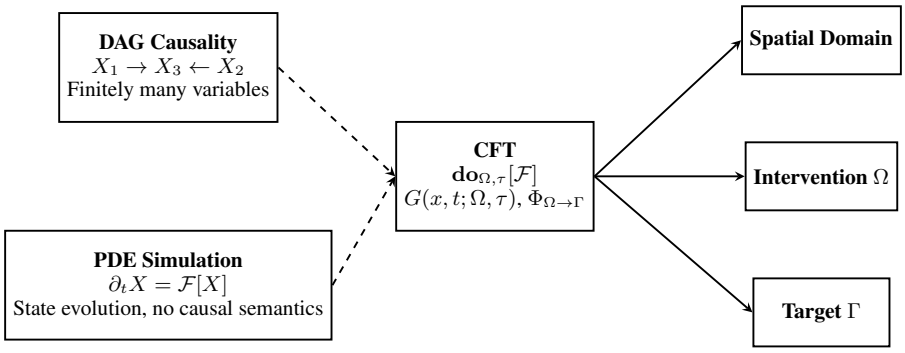


Figure 3: Conceptual overview: CFT extends DAG causality and PDE simulation with regional mechanism interventions and causal response fields.

A.2 Proof of Theorem 1: Well-definedness of the Causal Response Kernel

Let $X_\epsilon(x,t)$ denote the solution under the perturbed mechanism with intervention magnitude ϵ. The causal response kernel is $G(x,t;\Omega,\tau) = \lim_{\epsilon\to 0}(X_\epsilon - X_0)/\epsilon$. Under Fréchet-differentiability of $\mathcal{F}$, the perturbation satisfies a linearized variational equation. Taking $\epsilon \to 0$, we obtain $\partial_t G = D\mathcal{F}[X_0]\cdot G + S_\Omega$, where S_Ω is a localized source in $\Omega \times [\tau, \tau+\Delta]$. By standard results on linear PDEs (Evans, 2010), G exists and is unique under mild regularity. This connects CFT to Green's functions and linear response theory (Oppenheim et al., 1997; Chen, 1984).

A.3 Proof of Theorem 2: Finite-Speed Propagation

The response kernel G satisfies $\partial_t G = D\mathcal{F}[X_0]\cdot G + S_\Omega$ with S_Ω localized in $\Omega \times [\tau, \tau+\Delta]$. For hyperbolic systems, standard propagation bounds (Evans, 2010) give $G = 0$ for $|x - x_\Omega| > c(t-\tau)$. For parabolic systems (e.g., reaction-diffusion), influence spreads with Gaussian decay $G \sim \exp(-|x-x_\Omega|^2/(4D(t-\tau)))$; a soft causal boundary is determined by the decay rate. This geometric structure distinguishes CFT from graph-based frameworks.

A.4 Proof of Proposition 1: Scale Alignment

In the linearized regime, the macro-intervention response kernel is $G_{\text{macro}}(x,t) = \int_\Omega K(x,x',t-\tau)\delta\mathcal{F}_{\text{macro}}[X_0](x',\tau)dx'$. By assumption (2), $\delta\mathcal{F}_{\text{macro}} = \sum_i \chi_{\Omega_i}\delta\mathcal{F}_{\text{micro},i}$, hence $G_{\text{macro}} = \sum_i G_{\text{micro},i}$ at the kernel level. The inequality $\Phi^{\text{macro}}_{\Omega\to\Gamma} \geq \sum_i \Phi^{\text{micro},i}_{\Omega_i\to\Gamma}$ arises when computing flux: assumption (4) ensures constructive interference for simultaneous (macro) interventions, while isolated micro-interventions lose coherence.

A.5 Scale Alignment: Extended Discussion

The result provides sufficient conditions under which macro-interventions yield greater causal flux; it does not claim macro-interventions are always superior. As examples, morphogen gradients, tissue-scale mechanical interventions, and cytokine signaling in biological systems exhibit coherent propagation modes with characteristic length scales. Interventions aligned with these scales may exploit coherent propagation; see (Orr et al., 2006; Grieneisen et al., 2012) for biological context. The framework suggests that intervention scale relative to system correlation lengths may inform experimental and therapeutic design.

A.6 Data-Driven Causal Inference Algorithm

A.7 Glossary of Key Objects

- **State field** $X(x,t)$: System state at (x,t).
- **Mechanism operator** $\mathcal{F}$: $\partial_t X = \mathcal{F}[X]$.

Algorithm 1 Data-Driven Causal Inference with CFT

Require: Data $\mathcal{D} = \{(X_i(x,t), \Omega_i, \tau_i)\}_{i=1}^n$
Ensure: Response kernels $G(x,t;\Omega,\tau)$, causal flux $\Phi_{\Omega\to\Gamma}$
1: Learn $\mathcal{F}$ via neural operators or sparse discovery
2: Solve $\partial_t X_0 = \mathcal{F}[X_0]$ for baseline trajectory
3: Evaluate $D\mathcal{F}[X_0]$ along baseline
4: **for** each intervention (Ω, τ) **do**
5: Solve $\partial_t G = D\mathcal{F}[X_0] \cdot G + S_\Omega$
6: **end for**
7: **for** each target Γ **do**
8: Compute $\Phi_{\Omega\to\Gamma} = \int_\Gamma \int_{t>\tau} |G| \, dt \, dx$
9: **end for**

- **Regional intervention** $\mathbf{do}_{\Omega,\tau}[\mathcal{F}]$: Modifies $\mathcal{F}$ in Ω during $[\tau, \tau + \Delta]$.
- **Causal response kernel** $G(x,t;\Omega,\tau)$: $\delta X / \delta \, \mathbf{do}_{\Omega,\tau}[\mathcal{F}]$.
- **Causal flux** $\Phi_{\Omega\to\Gamma}$: $\int_\Gamma \int_\tau^\infty G \, dt \, dx$.
- **Causal cone**: Support of G.