



HAL
open science

From PDEs on standard domains to self-similar particle systems on fractals

Georgi Medvedev, Emmanuel Trélat

► **To cite this version:**

Georgi Medvedev, Emmanuel Trélat. From PDEs on standard domains to self-similar particle systems on fractals. 2026. <hal-05601143>

HAL Id: hal-05601143

<https://hal.science/hal-05601143v1>

Preprint submitted on 23 Apr 2026

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Copyright - All rights reserved

From PDEs on standard domains to self-similar particle systems on fractals

Georgi S. Medvedev*

Emmanuel Trélat †

Abstract

We construct transported PDEs on self-similar fractal domains from reference equations posed on the unit interval, and derive explicit self-similar interacting particle systems that approximate the resulting dynamics. The construction combines a measure-preserving isometry between L^2 -spaces on $[0, 1]$ and on the fractal [20], a nonlocal-to-local approximation of differential operators [24], and a Galerkin discretization on the canonical self-similar partitions. This yields a two-parameter approximation scheme whose error separates a nonlocal consistency term from a Galerkin network term. We work out the transport, Burgers, and heat equations, discuss the relation with intrinsic operators on fractals, and outline extensions to local charts and to pullbacks of nonlocal equations on fractal domains.

Moreover, the reverse mapping transforms a nonlocal evolution equation on a fractal domains into the evolution equation on the unit interval, where the methods of classical numerical analysis can be applied. This suggests a promising direction for the development of numerical methods for nonlocal models on fractals, including fractional heat equation, fractal scattering, and related models.

1 Introduction

Partial differential equations on fractal domains arise naturally in models of transport, diffusion, and wave propagation in geometrically irregular media. Classical examples include diffusion in porous or ramified materials, signal propagation on hierarchical structures, and scattering by fractal interfaces. The difficulty is that a fractal set typically has no smooth atlas in the usual sense, so differential operators cannot be introduced by direct geometric differentiation.

A large part of the existing theory therefore proceeds intrinsically, either through Dirichlet forms and diffusion processes, or through renormalized graph approximations of the Laplacian and related operators. For post-critically finite self-similar sets, this viewpoint leads to Kigami's Laplacian and to a rich analytic theory of elliptic, parabolic, and variational problems on fractals; see, for instance, [9, 18, 25, 15, 11, 23, 5]. Our objective here is different. We do not attempt to recover an intrinsic differential calculus on the fractal. Instead, we construct *transported* PDEs on a self-similar set K from a reference PDE posed on a simpler domain Q , and we derive from this construction explicit finite-dimensional interacting particle systems (IPS) adapted to the self-similar structure of K .

The basic mechanism is simple. Recent work of [20] provides an explicit measure-preserving isomorphism between a self-similar set K endowed with a stationary self-similar measure ν and an isomorphic iterated function system on the unit interval $Q = [0, 1]$, endowed with the corresponding stationary measure λ . In the uniform case relevant for the concrete PDE examples below, λ is simply the Lebesgue measure on $[0, 1]$. This isomorphism induces a linear isometry

$$T : L^2(K, \nu) \longrightarrow L^2(Q, \lambda),$$

*Department of Mathematics, Drexel University, Philadelphia, PA 19104, USA, medvedev@drexel.edu

†Sorbonne Université, Université Paris Cité, CNRS, Inria, Laboratoire Jacques-Louis Lions, LJLL, F-75005 Paris, France, emmanuel.trelat@sorbonne-universite.fr

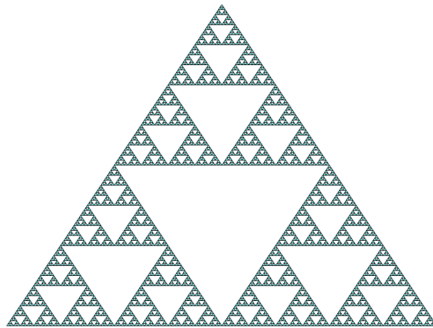


Figure 1: The Sierpinski gasket, which serves as a prototype throughout the paper.

together with its inverse T^{-1} . Therefore, if \mathbf{A} is the generator of a well-posed evolution equation on Q , then the conjugated operator $T^{-1}\mathbf{A}T$ defines a corresponding evolution on K .

This operator-theoretic lifting, however, is only the first part of the story. In order to connect the transported PDE to explicit particle systems, we insert an intermediate nonlocal regularization step, following the nonlocal-to-local approximation framework of [24]. Thus, starting from a PDE on Q , we first approximate its generator by a family of integral operators \mathbf{A}_ε , then transport these operators to K , and finally discretize the lifted nonlocal equation on the canonical level- m partition of the self-similar set. The resulting ODE system is a self-similar interacting particle system whose nodes are the cells K_w of generation m .

The paper has three main objectives. First, we formulate an abstract two-scale framework that combines the nonlocal-to-local approximation on Q with the Galerkin/IPS approximation on K . Second, we turn this framework into explicit constructions for the transport, Burgers, and heat equations, obtaining concrete self-similar particle systems and quantitative error estimates. Third, we clarify the scope of the construction: the operators transported to K are not, in general, intrinsic operators in the sense of fractal analysis, but they are natural continuum limits of hierarchical networks built directly from the underlying iterated function system (IFS).

This distinction between transported and intrinsic operators is conceptually important. If one aims at the canonical Laplacian attached to the geometry of the fractal, then one must retain the adjacency structure of prefractal graphs and the renormalized Dirichlet form. The global coding map used here preserves measure and symbolic cylinders, but not the adjacency relations between cells in Euclidean space. On the other hand, if the aim is to model dynamics on self-similar networks, or to import a well-understood PDE from a simple reference domain to a hierarchical one, then the present construction gives an explicit and flexible alternative. We return to this issue in the final section, where we also discuss two possible continuations of the present work: local charts on the Sierpinski gasket, and the pullback of nonlocal fractal equations (for instance fractional powers of an intrinsic Laplacian) to the reference interval.

Let us also emphasize the relation with the intrinsic first-order calculus on fractals. The Dirichlet-form approach of [6] and the subsequent vector-analysis framework of [15, 13] construct gradients, divergences and quasilinear PDEs from energy measures and resistance forms. Recent works on Burgers-type equations, transport/continuity equations, graph approximations and hydrodynamic limits on fractals further show that first-order and particle-system questions can be developed intrinsically [11, 14, 12, 5]. The present paper is complementary to that line of work: our operators are transported through a symbolic measure isomorphism and are designed to yield explicit self-similar Galerkin/IPS approximations, whereas the intrinsic theory is governed by energy, harmonic coordinates, and the adjacency structure of prefractal graphs.

The paper is organized as follows. Section 2 recalls the self-similar and measure-theoretic background. Section 3 reviews the canonical measure-preserving isomorphism between the fractal and the unit interval and

the induced isometries on function spaces. Section 4 summarizes the abstract nonlocal evolution problem on a self-similar set and its deterministic and random Galerkin approximations. Section 5 contains the general lifting framework from PDEs on Q to nonlocal equations and self-similar IPS on K . Section 6 is devoted to the transport, Burgers, and heat equations. Finally, Section 7 discusses extensions, limitations, and possible bridges with intrinsic operators and native nonlocal equations on fractals.

2 Self-similar sets, symbolic coding, and self-similar measures

In this section, we collect the measure-theoretic background on self-similar sets and stationary self-similar measures that will be used throughout the paper. Standard references are [1, 3, 16].

Iterated function systems and attractors. Let $\mathcal{F} = \{f_i : \mathbb{R}^d \rightarrow \mathbb{R}^d \mid i \in [k]\}$ be a finite collection of *similitudes*

$$|f_i(x) - f_i(y)| = r_i|x - y| \quad \text{for all } x, y \in \mathbb{R}^d, i \in [k],$$

with $0 < r_i < 1$. A unique compact set K satisfying $K = \bigcup_{i \in [k]} f_i(K)$ is called the *attractor* of the IFS $\mathcal{F} = \{f_i\}_{i \in [k]}$ (cf. [3, Theorem 2.1.1]). K is also called a *self-similar* set.

Symbolic space and Bernoulli measures. There is a natural family of self-similar measures associated with K . These measures are most conveniently described through Bernoulli measures on the symbolic space $\Sigma \doteq [k]^{\mathbb{N}}$ which we review first.

For each $n \geq 1$, let Σ_n be the set of words of length n , and let $\Sigma^* = \bigcup_{n=0}^{\infty} \Sigma_n$ denote the collection of all finite words over $[k]$.

We equip Σ with the metric $\rho(i, j) = k^{-|i \wedge j|}$, for $i = (i_1, i_2, \dots)$ and $j = (j_1, j_2, \dots)$, where $i \wedge j$ stands for the common prefix of the two sequences i and j , and $|i \wedge j|$ denotes its length. Thus, $|i \wedge j| = \sup\{l \in \mathbb{N} : i_l = j_l\}$ if $i \wedge j \neq \emptyset$ and 0 otherwise. This metric defines a topology on Σ that is consistent with the product topology. (Σ, ρ) is a compact metric space (cf. [1]).

Let $i_1, \dots, i_n \in [k]$ and define the cylinders

$$[i_1 i_2 \dots i_n] \doteq \{j = (j_1 j_2 \dots) \in \Sigma : j_l = i_l, 1 \leq l \leq n\}.$$

The collection of all cylinders generates the Borel σ -algebra in Σ , $\mathcal{B}(\Sigma)$.

We say that $p = (p_1, p_2, \dots, p_k)$ is a *probability vector*, if $p_i > 0$ for all $i \in [k]$ and $\sum_{i=1}^k p_i = 1$. The Bernoulli measure corresponding to p is first defined on cylinders $\mu_p([i_1 i_2 \dots i_n]) = p_{i_1} p_{i_2} \dots p_{i_n}$, and is then extended uniquely to a Borel probability measure on $(\Sigma, \mathcal{B}(\Sigma))$.

Natural projection and self-similar measures. The symbolic representation of K is provided by the natural projection $\pi_K : \Sigma \rightarrow K$.

For $w = (w_1 \dots w_n) \in \Sigma_n$, we set $f_w \doteq f_{w_1} \circ \dots \circ f_{w_n}$ and $K_w \doteq f_w(K)$. The natural projection is defined by

$$\pi_K(w) = \bigcap_{n \geq 1} K_{w_1 \dots w_n}, \quad w = (w_1 w_2 \dots) \in \Sigma.$$

Since $\{K_{w_1 \dots w_n}\}_{n \geq 1}$ is a nested sequence of compact sets with diameters tending to zero, their intersection consists of a single point. The map π_K is continuous and surjective (cf. [18, Theorem 1.2.3]). However, it may not be injective, i.e.,

$$\mathcal{N}_{\pi_K} = \{w \in \Sigma : \text{card} \{\pi_K^{-1}(\pi_K(w))\} > 1\}.$$

may not be empty. Indeed, this happens for many interesting examples of self-similar sets such as SG.

Throughout this paper, we make the following standing assumption.

Assumption 2.1. *There holds $\mu_p(\mathcal{N}_{\pi_K}) = 0$.*

We are now ready to introduce the stationary self-similar measure $\nu_p = (\pi_K)_*\mu_p$:

$$\nu_p(A) = \mu_p(\pi_K^{-1}(A)), \quad A \in \mathcal{B}(K).$$

The pushforward measure $\nu_p = (\pi_K)_*\mu_p$ is *the stationary measure* of the probabilistic IFS (K, \mathcal{F}, ν_p) , which means that

$$\nu_p(A) = \sum_{j=1}^k p_j \nu_p(f_j^{-1}(A)), \quad A \in \mathcal{B}(K).$$

The stationary measure ν_p is unique [3, Theorem 2.1.1]. It is called *a self-similar measure*.

Under Assumption 2.1, we have the following explicit formula for the pushforward measure ν_p on the cylinders:

$$\nu_p(K_w) = p_{w_1} p_{w_2} \dots p_{w_n}, \quad \forall w \in \Sigma_n, n \in \mathbb{N}. \quad (2.1)$$

3 Measure-preserving isomorphism with the unit interval

The key tool for transporting PDEs from Q to K is the measure-preserving isomorphism between the function spaces $L^2(Q, \lambda_p)$ and $L^2(K, \nu_p)$. In this section we recall the construction from [20] and the induced transport of functions and kernels.

3.1 The model IFS on the unit interval

Given a probabilistic IFS $(K, \{f_i\}_{i \in [k]}, \nu_p)$, we construct an isomorphic IFS on the unit interval $Q = [0, 1]$, denoted $(Q, \{g_i\}_{i \in [k]}, \lambda_p)$.

To this end, note that Q is the attractor of the following IFS

$$Q = \bigcup_{i=1}^k g_i(Q),$$

where $g_i(x) = \frac{x}{k} + x_{i-1}$ for $i \in [k]$, and $x_i = \frac{i}{k}$. Denote $\mathcal{G} = \{g_i\}_{i \in [k]}$.

We denote the natural projection $\pi_Q : \Sigma \rightarrow Q$:

$$\Sigma \ni w = (w_1 w_2 \dots) \mapsto \bigcap_{j=1}^{\infty} g_{w_1 w_2 \dots w_j}(Q).$$

Let $\lambda_p = (\pi_Q)_*\mu_p$ be the pushforward of the Bernoulli measure. On cylinders $Q_{w_1 w_2 \dots w_n} = g_{w_1 w_2 \dots w_n}(Q)$, we have

$$\lambda_p(Q_{w_1 w_2 \dots w_n}) = \mu_p(\pi_Q^{-1}(Q_{w_1 w_2 \dots w_n})) = p_{w_1} p_{w_2} \dots p_{w_n}.$$

In particular, for the uniform $p = (k^{-1}, k^{-1}, \dots, k^{-1})$, λ_p coincides with the Lebesgue measure.

For $(Q, \mathcal{G}, \lambda_p)$, one can easily verify

$$\mu_p(\mathcal{N}_{\pi_Q}) = 0. \quad (3.1)$$

Theorem 3.1. *Under Assumption 2.1, there exist full-measure subsets $Q_0 \subset Q$ and $K_0 \subset K$ and a measure-preserving bijection $\Phi : Q_0 \rightarrow K_0$ such that $\Phi \circ g_i = f_i \circ \Phi$ on $Q_0 \cap g_i^{-1}(Q_0)$, for $i \in [k]$. Equivalently, the probabilistic IFS (K, \mathcal{F}, ν_p) and $(Q, \mathcal{G}, \lambda_p)$ are isomorphic.*

This is a reformulation of [20, Theorem 5.7].

Remark 3.2. *The correspondence $f_i \leftrightarrow g_i$ used in Theorem 3.1 is not unique: any permutation $\sigma \in \mathfrak{S}_k$ yields an equally valid isomorphism Φ_σ satisfying $\Phi_\sigma \circ g_i = f_{\sigma(i)} \circ \Phi_\sigma$. Since each choice preserves the measure structure and the self-similar partition, the transported operators constructed in the sequel depend on this choice only through the labeling of the cells K_w . For the error estimates in Section 5, all such choices lead to the same convergence rates, so the particular convention adopted here (the identity permutation) entails no loss of generality.*

3.2 Transport of function spaces

Having constructed an isomorphic IFS $(Q, \mathcal{G}, \lambda_p)$, we now transport the measure-space isomorphism to the function spaces $L^1(K, \nu_p)$ and $L^1(Q, \lambda_p)$.

By Assumption 2.1 and (3.1), the image measures $\nu = (\pi_K)_*\mu$ and $\lambda = (\pi_Q)_*\mu$ are measure-preserving¹. Define the pullback operators $U_K : L^1(K, \nu) \rightarrow L^1(\Sigma, \mu)$ and $U_Q : L^1(Q, \lambda) \rightarrow L^1(\Sigma, \mu)$ by

$$U_K f = f \circ \pi_K, \quad U_Q g = g \circ \pi_Q.$$

Since π_K and π_Q are measure-preserving, both operators are linear isometries onto their respective ranges. Here, the term *isometry* may be understood in the L^1 - or L^2 -sense, namely as a linear map preserving the corresponding norm. In the L^2 -setting, the isometries U_K and U_Q preserve the inner product as well.

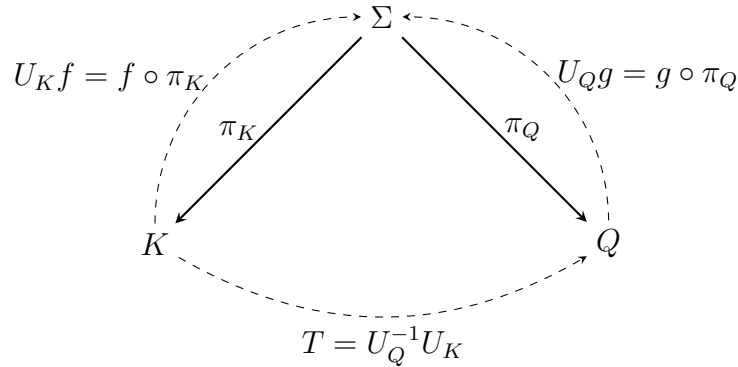


Figure 2: The isomorphism between measure spaces (K, ν) and (Q, λ) induces a corresponding isomorphism between the function spaces $L^1(K, \nu)$ and $L^1(Q, \lambda)$.

Moreover, since π_Q is bijective μ -a.e., the inverse of U_Q is well-defined (up to null sets) on $L^1(\Sigma, \mu)$ by $U_Q^{-1}F = F \circ \pi_Q^{-1}$. This allows us to define a canonical operator

$$T : L^1(K, \nu) \longrightarrow L^1(Q, \lambda), \quad Tf := U_Q^{-1}(U_K f), \quad (3.2)$$

Equivalently, Tf is the unique (up to λ -null sets) function satisfying

$$f \circ \pi_K = (Tf) \circ \pi_Q \quad \mu\text{-a.e.} \quad (3.3)$$

By construction, T is a surjective linear isometry preserving positivity and integrals, i.e., $\int_K f d\nu = \int_Q Tf d\lambda$ for every $f \in L^1(K, \nu)$.

3.3 Martingale representation and transport of kernels

The following martingale representation of Tf is useful both conceptually and computationally.

Given $f \in L^1(K, \nu)$, we generate a martingale sequence $(f_n, n \in \mathbb{N})$ as follows:

$$f_n \doteq \mathbb{E}(f | \mathcal{K}_n)(x) = \sum_{w \in \Sigma_n} \nu_w(f) \mathbf{1}_{K_w}(x), \quad \nu_w(f) = \int_{K_w} f(y) d\nu(y),$$

where $\mathbb{E}(f | \mathcal{K}_n)$ stands for the conditional expectation with respect to the σ -algebra of subsets of K generated by the partition, $\mathcal{K}_n = \sigma(\{K_w, w \in \Sigma_n\})$. By construction, $\mathbb{E}(f_{n+1} | \mathcal{K}_n) = f_n$. Thus, (f_n) form a martingale sequence converging to f ν -a.e. and in $L^1(K, \nu)$ by [26, Theorem 7.1.3].

¹Since p is fixed, we henceforth omit it from the notation of the measures μ , ν , and λ .

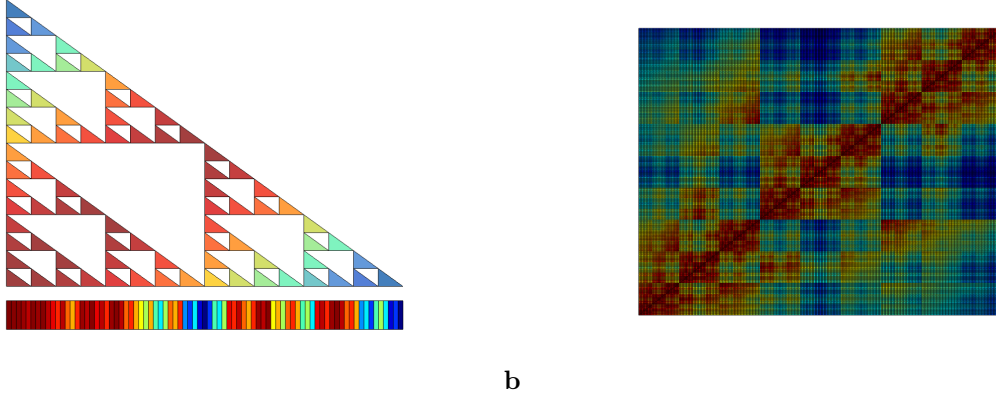


Figure 3: **a.** The heat map of the function $f(x) = e^{-|x_1 - x_2|}$ on SG $G \subset \mathbb{R}^2$ (above) and its representation as a function on $[0, 1]$ (below). **b.** Graphon representation of $W(x, y) = e^{-\|x-y\|}$, $(x, y) \in G \times G \subset \mathbb{R}^2 \times \mathbb{R}^2$.

Next, define $\tilde{f}_n : Q \rightarrow \mathbb{R}$ by

$$\tilde{f}_n(x) = \sum_{w \in \Sigma_n} \nu_w(f) \mathbf{1}_{Q_w}(x).$$

It is shown in [20], that the sequence (\tilde{f}_n) converges to $\tilde{f} \in L^1(Q, \lambda)$ λ -a.e. and in $L^1(Q, \lambda)$. Moreover, $\tilde{f} = Tf$.

Under the isomorphism constructed above, f and \tilde{f} generate the same projections, i.e., $\nu_w(f) = \lambda_w(\tilde{f})$ for every $w \in \Sigma^*$.

The same construction applies to kernels. Given $W \in L^1(K \times K, \nu \times \nu)$, define

$$W_{wv} := \int_{K_w \times K_v} W d(\nu \times \nu),$$

and set

$$\tilde{W}_n(x, y) \doteq \sum_{w, v \in \Sigma_n} W_{wv} \mathbf{1}_{Q_w \times Q_v}(x, y).$$

Then (\tilde{W}_n) converges in $L^1(Q \times Q, \lambda \times \lambda)$ (and a.e. along a subsequence) to a function \tilde{W} , which is precisely the image of W under the induced isomorphism between $L^1(K \times K, \nu \times \nu)$ and $L^1(Q \times Q, \lambda \times \lambda)$.

4 Nonlocal dynamics on self-similar sets and Galerkin/IPS approximations

4.1 A nonlocal evolution problem on K

Let $(K, \{f_i\}_{i=1}^k, \nu)$ be a probabilistic IFS and consider the following nonlocal evolution equation on K

$$\partial_t u(t, x) = F(t, u(t, x)) + \int_K W(x, y) D(u(t, x), u(t, y)) d\nu(y), \quad (4.1)$$

$$u(0, x) = g(x), \quad x \in K. \quad (4.2)$$

We assume $W \in L^2(K \times K, \nu \times \nu)$ and $g \in L^2(K, \nu)$. In addition,

$$\bar{W} \doteq \max \left\{ \text{ess sup}_{x \in K} \int_K |W(x, y)| d\nu(y), \text{ess sup}_{y \in K} \int_K |W(x, y)| d\nu(x) \right\} < \infty.$$

Functions $F(t, u)$ and $D(u, v)$ are jointly continuous and

$$|F(t, u) - F(t, u')| \leq L_F |u - u'|, \quad \forall t \in \mathbb{R}, u, u' \in \mathbb{R}, \quad (4.3)$$

$$|D(u, v) - D(u', v')| \leq L_D (|u - u'| + |v - v'|), \quad \forall u, v, u', v' \in \mathbb{R}, \quad (4.4)$$

where L_F and L_D are positive Lipschitz constants. In addition,

$$\sup_{K \times K} |D(u, v)| \leq 1. \quad (4.5)$$

Existence and uniqueness of solutions to (4.1)–(4.2) on finite time intervals, with $u \in C([-T, T]; L^2(K, \nu))$, follow from a standard contraction mapping argument; see [20, Theorem 4.2].

In many examples one also has an a priori bound of the form

$$\sup_{t \in [0, T]} \|u(t, \cdot)\|_{L^\infty(K, \nu)} \leq M.$$

In that case, the global assumptions (4.4) and (4.5) may be replaced by their local counterparts. More precisely, it is enough to assume that D is locally Lipschitz on \mathbb{R}^2 and that for every $M > 0$ there exists $C_D(M)$ such that $|D(u, v)| \leq C_D(M)$ whenever $|u|, |v| \leq M$. We shall implicitly use this observation in some of the PDE examples below.

4.2 Galerkin discretization and deterministic self-similar IPS

We begin by discretizing the self-similar domain K . To this end, fix $m \in \mathbb{N}$ and consider the self-similar partition $(K_w)_{w \in \Sigma_m}$ of K , defined by

$$\mathcal{K}^m \doteq \{K_w : K_w = f_w(K), |w| = m\}, \quad K = \bigcup_{|w|=m} K_w, \quad (4.6)$$

where we recall that f_w stands for $f_{w_1} \circ f_{w_2} \circ \dots \circ f_{w_m}$.

The partition (4.6) defines a finite-dimensional subspace of $\mathcal{H} \doteq L^2(K, \nu)$:

$$\mathcal{H}^m := \text{span}\{\mathbf{1}_A : A \in \mathcal{K}^m\}.$$

Throughout, $\mathbf{1}_A$ denotes the indicator function of the set A . Although the sets K_w and K_v corresponding to distinct $w, v \in \Sigma_m$ may intersect, their intersections have ν -measure zero by (2.1). Consequently, the resulting overlaps in the supports of the basis functions do not affect the convergence properties of the L^2 -projections onto \mathcal{H}^m .

We now construct the Galerkin approximation. The solution of the initial value problem (4.1)–(4.2) is approximated by a piecewise constant function of the form

$$u^m(t, x) = \sum_{|w|=m} u_w(t) \mathbf{1}_{K_w}(x). \quad (4.7)$$

Inserting (4.7) into (4.1) and projecting the resulting equation onto \mathcal{H}^m yields

$$\partial_t u^m(t, x) = F(t, u^m(t, x)) + \int_K W^m(x, y) D(u^m(t, x), u^m(t, y)) d\nu(y), \quad (4.8)$$

where the discretized kernel W^m is given by

$$W^m := \sum_{|w|, |v|=m} W_{wv} \mathbf{1}_{K_w \times K_v}, \quad W_{wv} = \int_{K_w \times K_v} W d(\nu \times \nu). \quad (4.9)$$

The initial condition (4.2) is approximated as follows

$$g^m := \sum_{|w|=m} g_w \mathbf{1}_{K_w}, \quad g_w := \int_{K_w} g \, d\nu. \quad (4.10)$$

It is instructive to rewrite the discrete problem (4.8), (4.10) as an initial value problem for a system of ODEs:

$$\dot{u}_w = F(t, u_w) + \sum_{|v|=m} W_{wv} D(u_w, u_v) \nu(K_v), \quad |w| = m, \quad (4.11)$$

$$u_w(0) = g_w, \quad |w| = m, \quad (4.12)$$

where

$$W_{wv} = \int_{K_w \times K_v} W(x, y) \, d(\nu \times \nu)(x, y), \quad g_w = \int_{K_w} g \, d\nu.$$

In this formulation, the Galerkin approximation of (4.1)–(4.2) admits a natural interpretation as an IPS. $D(u_w, u_v)$ defines the pairwise interaction between particles labeled by the words $w, v \in \Sigma_m$, while W_{wv} specifies the corresponding interaction weight.

As a system obtained by discretizing a PDE on a self-similar domain, (4.11) inherits the self-similar structure of the underlying continuum model.

4.3 Random self-similar networks and continuum limit

Along with the deterministic model (4.11), we consider an IPS on a W -random graph:

$$\dot{u}_w = F(t, u_w) + \sum_{|v|=n} \xi_{wv} D(u_w, u_v) \nu(K_v), \quad w \in \Sigma_n, \quad (4.13)$$

where ξ_{wv} are independent Bernoulli random variables such that

$$\mathbb{P}(\xi_{wv} = 1) = \int_{K_w \times K_v} W \, d(\nu \times \nu), \quad \mathbb{P}(\xi_{wv} = 0) = 1 - \mathbb{P}(\xi_{wv} = 1). \quad (4.14)$$

For the random model, in addition to integrability of W we also assume $W \geq 0$ and $\int_{K \times K} W \, d(\nu \times \nu) = 1$.

Both the deterministic model (4.11) and its random counterpart (4.13) converge to the continuum limit (4.1). The use of the random graph in (4.13) may be viewed as a Monte–Carlo approximation of the nonlocal term. This viewpoint is particularly useful numerically, and it suggests sparse random approximations of the lifted kernels as a natural extension; see [17].

In sparse variants one should expect a three-way balance between the regularity of W (or of the lifted kernel), the sampling density of the random graph, and the Galerkin scale m . This is the self-similar analogue of the regularity–sparsity tradeoff familiar in graphon-based nonlocal dynamics, and it would be natural to treat it as a separate refinement of the deterministic theory below.

The convergence of the deterministic and random Galerkin schemes is summarized by the following theorem, which is a reformulation of the main continuum-limit estimate proved in [20].

Theorem 4.1. *Let $u(t, x)$ be the solution of (4.1)–(4.2) with $g \in L^2(K, \nu)$. Likewise, suppose*

$$u^n(t, x) = \sum_{|j|=n} u_j(t) \mathbf{1}_{K_j}(x), \quad \mathbf{u}^n(t, x) = \sum_{|j|=n} \mathbf{u}_j(t) \mathbf{1}_{K_j}(x),$$

solve the IVPs for (4.11) and (4.13), respectively, and satisfy $u^n(0, \cdot) = \mathbf{u}^n(0, \cdot) = g^n$. Then, for $v \in \{u^n, \mathbf{u}^n\}$,

$$\|u - v\|_{C([0, T]; L^2(K, \nu))} \leq C_T \left(\|g - g^n\|_{L^2(K, \nu)} + \|W - W^n\|_{L^2(K \times K, \nu \times \nu)} \right). \quad (4.15)$$

In the random case $v = \mathbf{u}^n$, estimate (4.15) holds almost surely, i.e., for almost every realization of the W -random graph (4.14).

The rate of convergence of the right-hand side in (4.15) can be quantified by taking into account the regularity of the kernel W and the initial data g . For Euclidean domains, convergence rates were established using L^p -Lipschitz regularity in [17]. For self-similar domains, analogous estimates were obtained in [20]. In the self-similar setting, the definition of L^p -Lipschitz spaces must be suitably adapted to account for the underlying self-similar structure.

The randomized model (4.13) can be further extended to incorporate sparse graphs. From a computational perspective, sparsity can be used to reduce cost without sacrificing accuracy; see [17] for the Euclidean/graphon counterpart.

5 Abstract lifting framework and two-scale approximation

In this section we formulate the abstract transport-and-discretization procedure that underlies the rest of the paper. Starting from a reference PDE on Q , we first introduce a nonlocal regularization at scale ε , then transport it to the self-similar set K , and finally discretize the lifted equation on the level- m partition of K . The outcome is a two-parameter approximation scheme and an explicit self-similar interacting particle system.

5.1 Transporting a reference PDE from Q to K

We fix once and for all a self-similar set K equipped with a stationary self-similar probability measure ν , and, in the framework of Section 3, we assume Assumption 2.1 so that the canonical isomorphism of Section 3 applies. In particular, we use the linear isometry

$$T : L^2(K, \nu; \mathbb{R}^p) \longrightarrow L^2(Q, \lambda; \mathbb{R}^p)$$

introduced in (3.2)–(3.3) (and still denoted by T in the L^2 -setting), together with its inverse T^{-1} . Recall that $Q = [0, 1]$. Here λ denotes the measure $\lambda_p = (\pi_Q)_* \mu_p$ introduced in Section 3.1; in the general probabilistic IFS setting it is the stationary measure on $Q = [0, 1]$ associated with the same probability vector p as on K . In particular, if $p = (k^{-1}, \dots, k^{-1})$, then λ coincides with the Lebesgue measure on $[0, 1]$. Since the concrete PDE examples below use standard Sobolev spaces and convolution kernels on Q , we henceforth work in this uniform case, so that $\lambda = dx$.

By Theorem 3.1, there exist subsets $Q_0 \subset Q$ and $K_0 \subset K$ of full measure and an invertible measure-preserving map $\Phi : Q_0 \rightarrow K_0$ such that for every $f \in L^2(K, \nu; \mathbb{R}^p)$ and every $g \in L^2(Q, \lambda; \mathbb{R}^p)$,

$$(Tf)(x) = f(\Phi(x)) \quad \text{for } \lambda\text{-a.e. } x \in Q, \quad (T^{-1}g)(x) = g(\Phi^{-1}(x)) \quad \text{for } \nu\text{-a.e. } x \in K.$$

We will use Φ (or Φ^{-1}) only inside functions or kernels, always modulo null sets.

Reference PDE on Q . We start from an evolution equation on $Q = [0, 1]$ of the form

$$\partial_t u = \mathbf{A}u, \quad u(0, \cdot) = u^0(\cdot), \tag{5.1}$$

where $u(t, \cdot) \in L^2(Q, \lambda; \mathbb{R}^p)$ and \mathbf{A} is an (unbounded) linear operator with dense domain $\mathcal{D}(\mathbf{A}) \subset L^2(Q, \lambda; \mathbb{R}^p)$, as in [24]. We endow $\mathcal{D}(\mathbf{A})$ with the graph norm $\|\varphi\|_{\mathcal{D}(\mathbf{A})} := \|\varphi\|_{L^2(Q)} + \|\mathbf{A}\varphi\|_{L^2(Q)}$.

For the purpose of constructing a corresponding evolution on K , it is convenient to assume that \mathbf{A} is the generator of a C_0 -semigroup $(S(t))_{t \geq 0}$ on $L^2(Q, \lambda; \mathbb{R}^p)$, so that (5.1) is well-posed in the mild sense, and the solution of (5.1) is $u(t) = S(t)u^0$.

One can treat semilinear/quasilinear perturbations $\partial_t u = \mathbf{A}u + \mathcal{N}(u)$ with \mathcal{N} locally Lipschitz by the usual variation-of-constants approach; we focus on the linear part for the moment to keep the exposition light.

Transport of the dynamics from Q to K . Define the transported state \tilde{u} on K by

$$\tilde{u}(t, \cdot) := T^{-1}u(t, \cdot) \in L^2(K, \nu; \mathbb{R}^p). \quad (5.2)$$

Then the operator

$$\tilde{\mathbf{A}} := T^{-1} \mathbf{A} T, \quad \mathcal{D}(\tilde{\mathbf{A}}) := T^{-1}(\mathcal{D}(\mathbf{A}))$$

is the generator of a C_0 -semigroup on $L^2(K, \nu; \mathbb{R}^p)$ given by $\tilde{S}(t) := T^{-1}S(t)T$, and \tilde{u} is the (mild) solution of

$$\partial_t \tilde{u} = \tilde{\mathbf{A}} \tilde{u}, \quad \tilde{u}(0, \cdot) = \tilde{u}^0 := T^{-1}u^0. \quad (5.3)$$

This provides a purely operator-theoretic way of transporting the dynamics from Q to K .

Remark 5.1. *The evolution (5.3) should be interpreted as the transport of the Q -dynamics through the symbolic/IFS identification. In general, $\tilde{\mathbf{A}}$ is not the “intrinsic” (local) operator one may define on K via analysis on fractals (e.g., Kigami’s Laplacian, see [9, 18, 25], which is the self-adjoint operator associated with the canonical renormalized energy/Dirichlet form on a post critically finite self-similar set—equivalently, the renormalized limit of graph Laplacians on prefractal graphs). It is instead the natural operator associated with the hierarchical partition induced by the IFS and the measure isomorphism T .*

5.2 Nonlocal regularization on Q

To connect (5.1) to interacting particle systems, we insert an intermediate nonlocal integral approximation.

Integral approximation. Following [24], we approximate \mathbf{A} by a family of integral operators

$$\mathbf{A}_\varepsilon u(x) := \int_Q \sigma_\varepsilon(x, y) u(y) dy, \quad x \in Q,$$

where σ_ε is a kernel depending on $\varepsilon > 0$. For simplicity, and because this is the setting directly covered by Theorem 4.1 after lifting to K , we often keep the sufficient hypothesis $\sigma_\varepsilon \in L^2(Q \times Q)$. Conceptually, however, the approximation is better understood at the operator level: it is enough to require that $\mathbf{A}_\varepsilon \in \mathcal{L}(L^2)$ and that the associated semigroups satisfy the uniform stability estimate stated below. In boundary-adapted constructions one often obtains this boundedness by Schur’s test even when the kernel is not in L^2 ; see [24].

In many PDEs of interest (transport, diffusion, conservation laws), it is often more convenient to use the “zero-mass” form

$$\mathbf{A}_\varepsilon u(x) := \int_Q \gamma_\varepsilon(x, y) (u(y) - u(x)) dy$$

(integral terms plus bounded local terms), as will be illustrated in the examples treated further.

We denote by u_ε the solution of

$$\partial_t u_\varepsilon = \mathbf{A}_\varepsilon u_\varepsilon, \quad u_\varepsilon(0) = u^0. \quad (5.4)$$

Stability and consistency requirements. In PDE examples, the operator norms $\|\mathbf{A}_\varepsilon\|_{\mathcal{L}(L^2)}$ blow up as $\varepsilon \rightarrow 0$ (typically like ε^{-1} for first order terms and ε^{-2} for diffusion), because \mathbf{A}_ε is a bounded approximation of the unbounded operator \mathbf{A} . What is needed for well-posedness and error estimates is a *uniform stability* of the associated semigroups. Moreover, the consistency estimate usually involves one extra derivative, so we allow a stronger norm.

Let X be a Banach space such that $X \hookrightarrow L^2(Q, \lambda; \mathbb{R}^p)$ densely and continuously, and $X \subset \mathcal{D}(\mathbf{A})$. We denote its norm by $\|\cdot\|_X$. The space X is the regularity space in which the consistency error is measured. Concretely, it encodes the number of derivatives needed to control the Taylor remainder in the nonlocal-to-local approximation. It also ensures that the reference solution of (5.1) remains regular enough on $[0, T]$

for the Duhamel estimate leading to (5.5). In the examples treated further, $X = H^2(\mathbb{T})$ is sufficient for first-order operators because one remainder derivative is needed, whereas $X = H^4(\mathbb{T})$ is convenient for the heat equation because the second-order approximation leaves a remainder involving fourth derivatives.

In view of obtaining a clean two-step limit, we make the following standing assumption, which will have to be checked on each application example:

Assumption 5.2. *There exist a dense subspace $\mathcal{D}_0 \subset X$ which is a core for \mathbf{A} and a function $r(\varepsilon) \rightarrow 0$ as $\varepsilon \rightarrow 0$ such that:*

1. **Uniform semigroup stability:** *for any $\varepsilon > 0$, \mathbf{A}_ε generates a C_0 -semigroup $(S_\varepsilon(t))_{t \geq 0}$ on $L^2(Q, \lambda; \mathbb{R}^p)$ and there exist constants $M \geq 1$ and $\omega \in \mathbb{R}$, independent of ε , such that*

$$\|S_\varepsilon(t)\|_{\mathcal{L}(L^2(Q))} \leq M e^{\omega t} \quad \forall t \geq 0.$$

2. **Consistency on the core:**

$$\|\mathbf{A}_\varepsilon \varphi - \mathbf{A} \varphi\|_{L^2(Q)} \leq r(\varepsilon) \|\varphi\|_X \quad \forall \varphi \in \mathcal{D}_0.$$

Under Assumption 5.2, standard Duhamel and Gronwall arguments yield convergence $u_\varepsilon \rightarrow u$ for $u^0 \in X$ on any finite time interval $[0, T]$ on which the reference solution remains in X (see [24]), with an estimate of the form

$$\|u_\varepsilon - u\|_{C([0, T]; L^2(Q))} \leq C_T r(\varepsilon) \|u^0\|_X. \quad (5.5)$$

We will not reproduce the functional analytic proof here; our focus is on constructing explicit kernels σ_ε for representative PDEs, and on their discretization on K .

5.3 Lifting the nonlocal equation to K and deriving a self-similar IPS

Lift of the nonlocal equation. Define

$$\tilde{u}_\varepsilon := T^{-1} u_\varepsilon, \quad \tilde{\mathbf{A}}_\varepsilon := T^{-1} \mathbf{A}_\varepsilon T.$$

Then \tilde{u}_ε is the unique mild solution of

$$\partial_t \tilde{u}_\varepsilon = \tilde{\mathbf{A}}_\varepsilon \tilde{u}_\varepsilon, \quad \tilde{u}_\varepsilon(0) = \tilde{u}^0. \quad (5.6)$$

Moreover, by construction, $\tilde{\mathbf{A}}_\varepsilon$ admits an integral kernel representation on K :

$$\tilde{\mathbf{A}}_\varepsilon \tilde{u}(x) = \int_K \tilde{\sigma}_\varepsilon(x, y) \tilde{u}(y) d\nu(y),$$

where one convenient choice is

$$\tilde{\sigma}_\varepsilon(x, y) := \begin{cases} \sigma_\varepsilon(\Phi^{-1}(x), \Phi^{-1}(y)), & (x, y) \in K_0 \times K_0, \\ 0, & \text{otherwise.} \end{cases} \quad (5.7)$$

With this choice, $\tilde{\sigma}_\varepsilon \in L^2(K \times K, \nu \times \nu)$ and $\|\tilde{\sigma}_\varepsilon\|_{L^2(K \times K)} = \|\sigma_\varepsilon\|_{L^2(Q \times Q)}$.

In the more general Schur-kernel setting mentioned above (where only $\mathbf{A}_\varepsilon \in \mathcal{L}(L^2)$), one should interpret the lifting primarily at the operator level, by defining $\tilde{\mathbf{A}}_\varepsilon := T^{-1} \mathbf{A}_\varepsilon T$ first and only then averaging it on the cells. The explicit L^2 kernel formula (5.7) is retained here because it is exactly the setting covered by Theorem 4.1.

Galerkin/IPS discretization on the self-similar partition. Let $\mathcal{K}_m = \{K_w : |w| = m\}$ be the level- m partition of K and $\mathcal{H}_m := \text{span}\{\mathbf{1}_{K_w} : |w| = m\} \subset L^2(K, \nu)$ as in Section 4.2. Denote by P_m the L^2 -orthogonal projector onto \mathcal{H}_m . Define the step-kernel approximation

$$\tilde{\sigma}_{\varepsilon,m}(x, y) := \sum_{|w|, |v|=m} \tilde{\sigma}_{wv}^{(\varepsilon)} \mathbf{1}_{K_w \times K_v}(x, y), \quad \tilde{\sigma}_{wv}^{(\varepsilon)} := \int_{K_w \times K_v} \tilde{\sigma}_\varepsilon d(\nu \times \nu).$$

Then the Galerkin approximation of (5.6) reads

$$\partial_t \tilde{u}_{\varepsilon,m} = P_m \tilde{\mathbf{A}}_\varepsilon \tilde{u}_{\varepsilon,m} = \int_K \tilde{\sigma}_{\varepsilon,m}(\cdot, y) \tilde{u}_{\varepsilon,m}(y) d\nu(y), \quad \tilde{u}_{\varepsilon,m}(0) = P_m \tilde{u}^0.$$

Writing $\tilde{u}_{\varepsilon,m}(t, x) = \sum_{|w|=m} u_w^\varepsilon(t) \mathbf{1}_{K_w}(x)$, we obtain the finite-dimensional ODE system

$$\dot{u}_w^\varepsilon(t) = \sum_{|v|=m} \tilde{\sigma}_{wv}^{(\varepsilon)} u_v^\varepsilon(t) \nu(K_v), \quad |w| = m. \quad (5.8)$$

This is an IPS on the level- m self-similar network.

It is worth emphasizing what (5.8) represents. The unknowns are not attached to Euclidean mesh points, but to the cells K_w of the level- m self-similar partition. Thus each degree of freedom u_w^ε is the average state carried by one fractal cell, and the coefficients $\tilde{\sigma}_{wv}^{(\varepsilon)} \nu(K_v)$ are the effective interaction rates between two such cells. In this sense, (5.8) is simultaneously a Galerkin semi-discretization of the lifted nonlocal equation and an IPS on a hierarchical network with $N_m = k^m$ particles. This is one of the main interests of the construction: the network is not chosen ad hoc, but is generated canonically by the IFS itself. Increasing m does not merely refine a uniform mesh; it reveals finer generations of the self-similar geometry. From the PDE viewpoint, this gives an explicit finite-dimensional approximation adapted to the fractal structure. From the network viewpoint, it provides a concrete family of multiscale dynamical systems whose continuum limit is a PDE-type evolution on K . This constructive bridge between the two levels seems to be a genuinely new feature of the present framework.

Generalization to nonlinear/nonlocal operators. More generally, if \mathbf{A}_ε is nonlinear but has a nonlocal form

$$(\mathbf{A}_\varepsilon u)(x) = \int_Q W_\varepsilon(x, y) D(u(x), u(y)) dy, \quad (5.9)$$

then its lift to K yields a kernel \tilde{W}_ε and the Galerkin scheme produces precisely the self-similar IPS

$$\dot{u}_w^\varepsilon(t) = \sum_{|v|=m} W_{wv}^{(\varepsilon)} D(u_w^\varepsilon(t), u_v^\varepsilon(t)) \nu(K_v), \quad (5.10)$$

which is of the same type as (4.11). Here,

$$W_{wv}^{(\varepsilon)} := \int_{K_w \times K_v} \tilde{W}_\varepsilon d(\nu \times \nu) = \int_{Q_w \times Q_v} W_\varepsilon d(\lambda \times \lambda).$$

Remark 5.3. *The assumptions in Section 4 (in particular, boundedness of D in (4.5)) are tailored to globally Lipschitz dynamics on a compact state space. For the PDE examples below, one can either work in a regime where solutions are a priori bounded (e.g., $u^0 \in L^\infty$ and maximum principles), or insert a bounded saturation or truncation as in [24] to fit the hypotheses of Theorem 4.1. A complementary (and useful) improvement would be to extend the estimate (4.15) to interaction functions D of at most linear growth.*

The two-scale limit $m \rightarrow \infty, \varepsilon \rightarrow 0$. For fixed $\varepsilon > 0$, the convergence of the Galerkin/IPS approximation $\tilde{u}_{\varepsilon,m} \rightarrow \tilde{u}_\varepsilon$ as $m \rightarrow \infty$ follows from Theorem 4.1 applied to the kernel \tilde{W}_ε in the nonlinear case. In the linear case, $\tilde{\mathbf{A}}_\varepsilon$ is a bounded operator on $L^2(K, \nu)$ (since $\tilde{\sigma}_\varepsilon \in L^2(K \times K)$), and the Galerkin convergence is a standard consequence of the approximation of a bounded operator on the nested subspaces H_m ; alternatively, it follows from Theorem 4.1 after truncating the interaction function on bounded sets, as noted in Remark 5.3. In both cases:

$$\|\tilde{u}_{\varepsilon,m} - \tilde{u}_\varepsilon\|_{C([0,T];L^2(K,\nu))} \leq C_T \left(\|\tilde{u}^0 - P_m \tilde{u}^0\|_{L^2(K,\nu)} + \|\tilde{\sigma}_\varepsilon - \tilde{\sigma}_{\varepsilon,m}\|_{L^2(K \times K, \nu \times \nu)} \right). \quad (5.11)$$

Combining (5.5), (5.2) and (5.11), we obtain the following result.

Theorem 5.4. *Let u solve (5.1), let $\tilde{u} = T^{-1}u$, and for each $\varepsilon > 0$ let $\tilde{u}_{\varepsilon,m}$ be the Galerkin/IPS approximation defined above. Under Assumption 5.2, for every $T > 0$,*

$$\|\tilde{u}_{\varepsilon,m} - \tilde{u}\|_{C([0,T];L^2(K,\nu))} \leq C_T \left(r(\varepsilon) \|u^0\|_X + \|\tilde{u}^0 - P_m \tilde{u}^0\|_{L^2(K,\nu)} + \|\tilde{\sigma}_\varepsilon - \tilde{\sigma}_{\varepsilon,m}\|_{L^2(K \times K, \nu \times \nu)} \right). \quad (5.12)$$

Thus the derivation of the PDE on K may be viewed as a genuine two-step limit: first $m \rightarrow \infty$ at fixed ε (Galerkin/IPS to the lifted nonlocal equation), then $\varepsilon \rightarrow 0$ (nonlocal equation to the transported PDE), or any joint limit for which the three terms on the right-hand side of (5.12) vanish.

The last two terms in (5.12) are controlled by the approximation theory on K developed in [20] (in particular, in terms of generalized Lipschitz regularity relative to the self-similar partition), while the first term is the ‘‘nonlocal-to-local’’ error on Q coming from [24].

Balancing the two scales. Suppose, in addition, that the Galerkin terms are of order $q^{\alpha m}$, where $q \in (0, 1)$ is the common contraction ratio of the IFS and $\alpha > 0$ is the generalized Lipschitz exponent from the approximation theory on K developed in [20]. If the nonlocal consistency is $r(\varepsilon) = \mathcal{O}(\varepsilon^\beta)$, then (5.12) becomes

$$\|\tilde{u}_{\varepsilon,m} - \tilde{u}\|_{C([0,T];L^2(K,\nu))} \leq C_T (\varepsilon^\beta + q^{\alpha m}).$$

Since the number of cells at level m is $N_m = k^m$ and $q^m = N_m^{-1/s}$ with $s = \log k / \log(1/q)$, the network error is equivalently $N_m^{-\alpha/s}$. Hence the similarity dimension s plays the role of an effective dimension in the cost/error balance. The natural balanced choice is $q^{\alpha m} \asymp \varepsilon^\beta$, i.e.

$$N_m \asymp \varepsilon^{-\beta s / \alpha}, \quad \text{or equivalently} \quad \varepsilon \asymp N_m^{-\alpha / (\beta s)}.$$

We thus obtain the following corollary.

Corollary 5.5. *Assume, in addition to Theorem 5.4, that there exist $\alpha > 0$ and $C > 0$, independent of ε and m , such that*

$$\|\tilde{u}^0 - P_m \tilde{u}^0\|_{L^2(K,\nu)} + \|\tilde{\sigma}_\varepsilon - \tilde{\sigma}_{\varepsilon,m}\|_{L^2(K \times K, \nu \times \nu)} \leq C q^{\alpha m},$$

and that $r(\varepsilon) \leq C \varepsilon^\beta$ for some $\beta > 0$. Then, for every $T > 0$,

$$\|\tilde{u}_{\varepsilon,m} - \tilde{u}\|_{C([0,T];L^2(K,\nu))} \leq C_T (\varepsilon^\beta + q^{\alpha m}).$$

If $m = m(\varepsilon)$ is chosen so that $q^{\alpha m(\varepsilon)} \leq \varepsilon^\beta < q^{\alpha(m(\varepsilon)-1)}$, then the two contributions are balanced and

$$\|\tilde{u}_{\varepsilon,m(\varepsilon)} - \tilde{u}\|_{C([0,T];L^2(K,\nu))} \leq C_T \varepsilon^\beta.$$

Equivalently, since $N_m = k^m$ and $q^m = N_m^{-1/s}$ with $s = \log k / \log(1/q)$, the balanced choice $\varepsilon \asymp N_m^{-\alpha / (\beta s)}$ yields

$$\|\tilde{u}_{\varepsilon,m} - \tilde{u}\|_{C([0,T];L^2(K,\nu))} \leq C_T N_m^{-\alpha/s}.$$

Corollary 5.5 is the quantitative statement one can use in practice. It confirms the interpretation above: once the consistency order $r(\varepsilon)$ is known, the estimate prescribes how the nonlocalization scale ε and the network depth m should be coupled. For a target accuracy δ , one chooses $\varepsilon \asymp \delta^{1/\beta}$ and $N_m \asymp \delta^{-s/\alpha}$; for a fixed particle budget N_m , the balanced choice is $\varepsilon \asymp N_m^{-\alpha/(\beta s)}$. Taking ε much smaller than this is wasteful, because the network error $q^{\alpha m}$ then dominates. Taking m much larger is equally wasteful, because the nonlocalization error ε^β then dominates. Thus the rate is not only an asymptotic statement, but also a practical design rule for the approximation.

At a conceptual level, Theorem 5.4 is the key quantitative output of the construction: it combines, in a single estimate, the nonlocal-to-local approximation on the reference space Q and the Galerkin/self-similar network approximation on the fractal side K . In the present framework, this appears to give a new constructive and quantitative derivation of PDE-type dynamics on fractals from explicit interacting particle systems. It also makes visible the role of the similarity dimension s : it enters the cost/error balance as an effective dimension.

For the three model problems below one finds $\beta = 1$ for transport and Burgers, and $\beta = 2$ for heat, so diffusion allows a comparatively larger nonlocal scale ε for the same network size and target accuracy.

6 Representative examples

In all three examples below we work with periodic boundary conditions on $Q = [0, 1]$, i.e., we identify Q with the one-dimensional torus \mathbb{T} and periodize the kernels. The periodic setting isolates the basic mechanism and avoids boundary-layer terms.

The examples should be understood as transported models. In particular, the Burgers and transport equations below are not meant to coincide with the intrinsic first-order models built from Dirichlet-form vector fields; compare, for instance, the intrinsic viscous Burgers and transport/continuity theories in [11, 14]. Their purpose is to show how standard one-dimensional PDEs give rise, after nonlocalization, lifting and Galerkin projection, to explicit self-similar particle systems on K .

6.1 Transport equation

Let $b \in W^{1,\infty}(\mathbb{T})$. Consider the transport equation on \mathbb{T}

$$\partial_t u + b(x) \partial_x u = 0, \quad u(0) = u^0. \quad (6.13)$$

Nonlocal approximation on Q . Fix an arbitrary odd function $\eta \in L^1(\mathbb{R})$ such that $\int_{\mathbb{R}} \eta(z) dz = 0$ and $\int_{\mathbb{R}} z \eta(z) dz = -1$ and $C_\eta := \frac{1}{2} \int_{\mathbb{R}} |z|^2 |\eta(z)| dz < \infty$. Define the (periodized) kernel

$$\eta_\varepsilon(z) := \varepsilon^{-2} \eta(z/\varepsilon) \quad (6.14)$$

and the nonlocal approximate derivative

$$(D_\varepsilon u)(x) := \int_{\mathbb{T}} \eta_\varepsilon(x-y) u(y) dy = \int_{\mathbb{T}} \eta_\varepsilon(x-y) (u(y) - u(x)) dy. \quad (6.15)$$

Then, for smooth u , $D_\varepsilon u \rightarrow \partial_x u$ in $L^2(\mathbb{T})$ (more precisely, see (6.17) in the proof of Lemma 6.1 below).

One obtains an integral approximation of (6.13) by

$$\partial_t u_\varepsilon + b(x) D_\varepsilon u_\varepsilon = 0. \quad (6.16)$$

Equivalently, (6.16) has the form (5.4) with $\mathbf{A}_\varepsilon u = -b D_\varepsilon u$, i.e., $\sigma_\varepsilon(x, y) = -b(x) \eta_\varepsilon(x - y)$.

Lemma 6.1. *Assumption 5.2 is satisfied with $M = 1$, $\omega = \frac{1}{2} \|b'\|_{L^\infty} \int_{\mathbb{R}} |z| |\eta(z)| dz$, and $r(\varepsilon) = \mathcal{O}(\varepsilon)$ for the choice $X = H^2(\mathbb{T})$.*

Proof. Here $\mathbf{A}u = -b\partial_x u$ on $L^2(\mathbb{T})$, with $\mathcal{D}(\mathbf{A}) = H^1(\mathbb{T})$ and a core $\mathcal{D}_0 = C^\infty(\mathbb{T})$. We take $X = H^2(\mathbb{T})$.

Uniform semigroup stability. Although $\|D_\varepsilon\|_{\mathcal{L}(L^2)} = \|\eta_\varepsilon\|_{L^1} = \varepsilon^{-1}\|\eta\|_{L^1}$ blows up as $\varepsilon \rightarrow 0$, the L^2 -flow of (6.16) is uniformly stable. Since η_ε is odd, D_ε is skew-adjoint on $L^2(\mathbb{T})$. Hence, setting $[b, D_\varepsilon] := bD_\varepsilon - D_\varepsilon(b\cdot)$, for a smooth solution u_ε we have

$$\langle [b, D_\varepsilon]u_\varepsilon, u_\varepsilon \rangle = \langle bD_\varepsilon u_\varepsilon, u_\varepsilon \rangle - \langle D_\varepsilon(bu_\varepsilon), u_\varepsilon \rangle = \langle bD_\varepsilon u_\varepsilon, u_\varepsilon \rangle + \langle bu_\varepsilon, D_\varepsilon u_\varepsilon \rangle = 2\langle bD_\varepsilon u_\varepsilon, u_\varepsilon \rangle,$$

and thus

$$\frac{1}{2} \frac{d}{dt} \|u_\varepsilon(t)\|_{L^2(\mathbb{T})}^2 = -\langle bD_\varepsilon u_\varepsilon, u_\varepsilon \rangle = -\frac{1}{2} \langle [b, D_\varepsilon]u_\varepsilon, u_\varepsilon \rangle,$$

where $[b, D_\varepsilon]$ has the integral kernel representation

$$([b, D_\varepsilon]u)(x) = \int_{\mathbb{T}} K_\varepsilon(x, y) u(y) dy, \quad K_\varepsilon(x, y) := (b(x) - b(y)) \eta_\varepsilon(x - y).$$

Since $b \in W^{1,\infty}(\mathbb{T})$, one has $|b(x) - b(y)| \leq \|b'\|_{L^\infty(\mathbb{T})}|x - y|_{\mathbb{T}}$, where $|x - y|_{\mathbb{T}}$ denotes the periodic distance. Therefore

$$\sup_x \int_{\mathbb{T}} |K_\varepsilon(x, y)| dy \leq \|b'\|_{L^\infty(\mathbb{T})} \int_{\mathbb{T}} |x - y|_{\mathbb{T}} |\eta_\varepsilon(x - y)| dy \leq \|b'\|_{L^\infty(\mathbb{T})} \int_{\mathbb{R}} |z| |\eta(z)| dz,$$

after the change of variables $x - y = \varepsilon z$. The same estimate holds for $\sup_y \int_{\mathbb{T}} |K_\varepsilon(x, y)| dx$. Schur's test then gives

$$\begin{aligned} \|[b, D_\varepsilon]\|_{\mathcal{L}(L^2(\mathbb{T}))} &\leq \left(\sup_x \int |K_\varepsilon(x, y)| dy \right)^{1/2} \left(\sup_y \int |K_\varepsilon(x, y)| dx \right)^{1/2} \\ &\leq \|b'\|_{L^\infty(\mathbb{T})} \int_{\mathbb{R}} |z| |\eta(z)| dz. \end{aligned}$$

Consequently,

$$\|u_\varepsilon(t)\|_{L^2(\mathbb{T})} \leq \exp\left(\frac{t}{2} \|b'\|_{L^\infty} \int_{\mathbb{R}} |z| |\eta(z)| dz\right) \|u^0\|_{L^2(\mathbb{T})},$$

which is Assumption 5.2(1) with $M = 1$ and $\omega = \frac{1}{2} \|b'\|_{L^\infty} \int_{\mathbb{R}} |z| |\eta(z)| dz$. Note that $\int_{\mathbb{R}} |z| |\eta(z)| dz < +\infty$ by the assumptions done on η . Indeed, split the integral in two parts, $|z| \leq 1$ and $|z| > 1$; the first is finite because $\eta \in L^1(\mathbb{R})$, and the second is less than C_η because $|z| \leq |z|^2$.

Consistency. After periodization and the change of variables $x - y = \varepsilon z$, one has $(D_\varepsilon u)(x) = \frac{1}{\varepsilon} \int_{\mathbb{R}} \eta(z) u(x - \varepsilon z) dz$. Hence

$$D_\varepsilon u(x) - \partial_x u(x) = \frac{1}{\varepsilon} \int_{\mathbb{R}} \eta(z) \left(u(x - \varepsilon z) - u(x) + \varepsilon z \partial_x u(x) \right) dz.$$

Taylor's formula gives

$$u(x - \varepsilon z) - u(x) + \varepsilon z \partial_x u(x) = \varepsilon^2 z^2 \int_0^1 (1 - \theta) \partial_{xx} u(x - \theta \varepsilon z) d\theta.$$

Therefore, for $u \in H^2(\mathbb{T})$,

$$\|D_\varepsilon u - \partial_x u\|_{L^2(\mathbb{T})} \leq C_\eta \varepsilon \|\partial_{xx} u\|_{L^2(\mathbb{T})}, \quad (6.17)$$

by Minkowski's inequality and translation invariance of the $L^2(\mathbb{T})$ norm.

It follows from (6.17) that $\|\mathbf{A}_\varepsilon u - \mathbf{A}u\|_{L^2(\mathbb{T})} \leq C_\eta \|b\|_{L^\infty} \|u\|_{H^2(\mathbb{T})} \varepsilon$. In particular, Assumption 5.2(2) holds with $r(\varepsilon) = \mathcal{O}(\varepsilon)$ for the choice $X = H^2(\mathbb{T})$. \square

Lift to K and IPS. Define $\tilde{b} := T^{-1}b \in L^\infty(K, \nu)$ and the lifted kernel $\tilde{\sigma}_\varepsilon$ by (5.7). The lifted equation is

$$\partial_t \tilde{u}_\varepsilon(t, x) + \tilde{b}(x) \int_K \eta_\varepsilon(\Phi^{-1}(x) - \Phi^{-1}(y)) \tilde{u}_\varepsilon(t, y) d\nu(y) = 0.$$

The Galerkin scheme yields the particle system

$$\dot{u}_w^\varepsilon(t) = - \sum_{|v|=m} b_w \eta_{wv}^{(\varepsilon)} u_v^\varepsilon(t) \nu(K_v), \quad b_w := \int_{K_w} \tilde{b} d\nu = \int_{Q_w} b d\lambda \quad (6.18)$$

with weights

$$\eta_{wv}^{(\varepsilon)} := \int_{K_w \times K_v} \eta_\varepsilon(\Phi^{-1}(x) - \Phi^{-1}(y)) d(\nu \times \nu) = \int_{Q_w \times Q_v} \eta_\varepsilon(x - y) d(\lambda \times \lambda).$$

Let us note a few structural features of (6.18). First, the unknowns u_w^ε are attached to self-similar cells, so the transport dynamics is realized on a hierarchical network rather than on a uniform grid. Second, because $\int_{\mathbb{T}} \eta_\varepsilon = 0$, one has $\mathbf{A}_\varepsilon \mathbf{1} = 0$; correspondingly, the averaged coefficients satisfy

$$\sum_{|v|=m} \eta_{wv}^{(\varepsilon)} \nu(K_v) = 0,$$

so constant states are preserved, exactly as for the transport equation. Third, the coefficients may have both signs, which is the discrete signature of transport rather than diffusion. Finally, the scale ε determines how strongly two cells interact after they are compared through the coordinate Φ^{-1} on Q : for localized kernels, only cells that are close in the Q -parameter interact significantly. Thus (6.18) provides a concrete transport dynamics on a multiscale self-similar network, and this seems to be one of the most interesting new features of the present approach.

As $\varepsilon \rightarrow 0$ with m fixed, the averaged weights $\eta_{wv}^{(\varepsilon)}$ concentrate: since η_ε approximates a distributional derivative, the IPS (6.18) formally reduces to the Galerkin projection of the local transport equation $\partial_t u + b \partial_x u = 0$ onto the level- m partition of Q , transported to K . In the opposite regime $m \rightarrow \infty$ with ε fixed, the nonlocal structure is preserved but the spatial resolution on K refines. The joint limit $\varepsilon \rightarrow 0$, $m \rightarrow \infty$ is governed by Theorem 5.4 and Corollary 5.5, which prescribe the balanced coupling $\varepsilon \sim k^{-m}$.

6.2 Burgers equation

Consider the (inviscid) Burgers equation on \mathbb{T}

$$\partial_t u + \partial_x \left(\frac{1}{2} u^2 \right) = 0.$$

Since solutions may develop shocks, the approximation theory depends on whether one works with smooth solutions on short times, or with entropy solutions. Here, a robust strategy is to treat the viscous Burgers regularization and then discuss the inviscid limit.

A nonlocal conservative flux approximation. Let $F(s) = \frac{1}{2}s^2$ and use the same odd kernel η_ε as in (6.15). As in the transport case, we use the derivative scaling (6.14) so that D_ε approximates ∂_x (the overall sign can be adjusted by replacing η with $-\eta$ if needed). Consider the integral approximation

$$\partial_t u_\varepsilon = \mathbf{A}_\varepsilon(u_\varepsilon)$$

where (recall that $\int_{\mathbb{T}} \eta_\varepsilon = 0$)

$$(\mathbf{A}_\varepsilon(u))(x) := -D_\varepsilon(F(u)) = \int_{\mathbb{T}} \eta_\varepsilon(x - y) (F(u(x)) - F(u(y))) dy,$$

which fits (5.9) with $W_\varepsilon(x, y) = \eta_\varepsilon(x - y)$ and $D(a, b) = F(a) - F(b)$.

Lemma 6.2. Fix $R > 0$. There exists $C_R > 0$, independent of ε , such that for every $u \in H^2(\mathbb{T}) \cap L^\infty(\mathbb{T})$ with $\|u\|_{H^2(\mathbb{T})} + \|u\|_{L^\infty(\mathbb{T})} \leq R$,

$$\|\mathbf{A}_\varepsilon(u) + \partial_x F(u)\|_{L^2(\mathbb{T})} \leq C_R \varepsilon. \quad (6.19)$$

Moreover, if u_ε and v_ε solve $\partial_t u = \mathbf{A}_\varepsilon(u)$ on $[0, T]$ and remain bounded by R in $H^2(\mathbb{T})$ on this interval, then

$$\|u_\varepsilon(t) - v_\varepsilon(t)\|_{L^2(\mathbb{T})} \leq e^{C_R t} \|u_\varepsilon(0) - v_\varepsilon(0)\|_{L^2(\mathbb{T})}, \quad t \in [0, T]. \quad (6.20)$$

Consequently, in any smooth regime where the local Burgers solution and the nonlocal approximants remain uniformly bounded in $H^2(\mathbb{T}) \cap L^\infty(\mathbb{T})$ on $[0, T]$, the nonlinear analogue of Assumption 5.2 is satisfied with rate $r(\varepsilon) = \mathcal{O}(\varepsilon)$.

Proof. For the consistency estimate, apply (6.17) to $F(u)$:

$$\|\mathbf{A}_\varepsilon(u) + \partial_x F(u)\|_{L^2(\mathbb{T})} = \|(D_\varepsilon - \partial_x)(F(u))\|_{L^2(\mathbb{T})} \leq C_\eta \varepsilon \|\partial_{xx} F(u)\|_{L^2(\mathbb{T})}.$$

Since $F(s) = \frac{1}{2}s^2$, one has $\partial_{xx} F(u) = u \partial_{xx} u + (\partial_x u)^2$, and therefore

$$\|\partial_{xx} F(u)\|_{L^2} \leq \|u\|_{L^\infty} \|u\|_{H^2} + \|\partial_x u\|_{L^4}^2 \leq C(1 + \|u\|_{L^\infty}) \|u\|_{H^2}^2 \leq C_R,$$

using the 1D embeddings $H^1(\mathbb{T}) \hookrightarrow L^\infty(\mathbb{T}) \cap L^4(\mathbb{T})$. This proves (6.19).

For stability, let $w = u_\varepsilon - v_\varepsilon$ and $b = \frac{1}{2}(u_\varepsilon + v_\varepsilon)$. Since $F(u_\varepsilon) - F(v_\varepsilon) = bw$, the difference equation reads $\partial_t w + D_\varepsilon(bw) = 0$. Using again the skew-adjointness of D_ε , we obtain as in the transport case

$$\frac{1}{2} \frac{d}{dt} \|w(t)\|_{L^2(\mathbb{T})}^2 = \langle D_\varepsilon(bw), w \rangle = -\frac{1}{2} \langle [b, D_\varepsilon]w, w \rangle.$$

Since u_ε and v_ε are bounded in $H^2(\mathbb{T})$, the Sobolev embedding $H^2(\mathbb{T}) \hookrightarrow W^{1,\infty}(\mathbb{T})$ yields $\|b'\|_{L^\infty(\mathbb{T})} \leq C_R$. The same Schur estimate as in the proof of Lemma 6.1 therefore gives $\|[b, D_\varepsilon]\|_{\mathcal{L}(L^2(\mathbb{T}))} \leq C_R$, with C_R independent of ε . Hence

$$\frac{d}{dt} \|w(t)\|_{L^2(\mathbb{T})}^2 \leq C_R \|w(t)\|_{L^2(\mathbb{T})}^2,$$

and Gronwall's lemma yields (6.20). \square

Lift to K and Galerkin/IPS discretization. The corresponding lifted equation on K is

$$\partial_t \tilde{u}_\varepsilon(t, x) + \int_K \eta_\varepsilon(\Phi^{-1}(x) - \Phi^{-1}(y)) \left(F(\tilde{u}_\varepsilon(t, x)) - F(\tilde{u}_\varepsilon(t, y)) \right) d\nu(y) = 0. \quad (6.21)$$

The particle approximation of (6.21) is

$$\dot{u}_w^\varepsilon(t) = - \sum_{|v|=m} \eta_{wv}^{(\varepsilon)} \left(F(u_w^\varepsilon(t)) - F(u_v^\varepsilon(t)) \right) \nu(K_v),$$

where $\eta_{wv}^{(\varepsilon)}$ are the same averaged weights as in (6.18).

There is also a natural structural interpretation of this Burgers IPS. Because the averaged weights inherit the antisymmetry of the odd kernel, namely $\eta_{wv}^{(\varepsilon)} = -\eta_{vw}^{(\varepsilon)}$, the scheme is conservative: summing the equations over w shows that the discrete mass $\sum_{|w|=m} u_w^\varepsilon(t) \nu(K_w)$ is preserved. The nonlinearity acts only through pairwise flux differences $F(u_w^\varepsilon) - F(u_v^\varepsilon)$, so the model is the hierarchical analogue of a conservative finite-volume discretization. In particular, (6.21) and its IPS discretization show that even a nonlinear conservation law can be transported to a fractal domain and approximated there by an explicit self-similar particle system.

The same limiting regimes as for the transport equation apply: as $\varepsilon \rightarrow 0$ with m fixed, the nonlocal flux converges formally to the local one, while $m \rightarrow \infty$ refines the network. The balanced coupling is again $\varepsilon \sim k^{-m}$, since the smooth-regime consistency rate is $r(\varepsilon) = \mathcal{O}(\varepsilon)$.

Entropy regime: two practical routes. For inviscid Burgers, the central difficulty is not the particle approximation itself, but the selection of the entropy solution after shocks form. A first route is therefore to insert a small viscosity and to work with

$$\partial_t u + \partial_x F(u) = \kappa \partial_{xx} u, \quad \kappa > 0.$$

For each fixed κ , the equation is parabolic, so the smooth-regime estimates above, together with the heat approximation of Section 6.3, apply on finite time intervals. One may then justify the IPS approximation for the viscous problem first, and only afterwards let $\kappa \rightarrow 0$. By the classical vanishing-viscosity principle, this limit is the entropy solution of Burgers (see, for instance, [8]). The advantage of this route is that entropy selection is delegated to a well-understood parabolic regularization.

A second route is to build entropy selection directly into the nonlocal scheme. One replaces the centered odd kernel by a one-sided nonnegative kernel and a monotone two-point flux. Let $\rho \geq 0$ be supported in $[0, \infty)$, with $\int_0^\infty \rho(z) dz = 1$ and $\int_0^\infty z\rho(z) dz = 1$, and set $\rho_\varepsilon(z) := \varepsilon^{-2}\rho(z/\varepsilon)$. For Burgers, define

$$g(a, b) := \frac{1}{2}(a_+)^2 + \frac{1}{2}(b_-)^2, \quad a_+ := \max\{a, 0\}, \quad b_- := \min\{b, 0\}.$$

Then one may consider the one-sided nonlocal operator

$$(\mathcal{N}_\varepsilon^{\text{up}}u)(x) := \int_0^\infty \rho_\varepsilon(z) \left(g(u(x), u(x+z)) - g(u(x-z), u(x)) \right) dz,$$

with shifts understood modulo 1 on \mathbb{T} . For smooth u , a first-order Taylor expansion shows that $\mathcal{N}_\varepsilon^{\text{up}}(u) \rightarrow \partial_x F(u)$. After transport to K and Galerkin projection, this yields an IPS with oriented nonnegative interaction weights, that is, a nonlocal analogue of an upwind finite-volume scheme. This is precisely the structure in which one expects discrete maximum principles, L^1 -contraction, and convergence to the entropy solution; compare with the classical role of monotone schemes in scalar conservation laws [7].

6.3 Heat equation

Let $\kappa > 0$. Consider the heat equation on \mathbb{T}

$$\partial_t u = \kappa \partial_{xx} u.$$

Nonlocal diffusion kernel. Let $\rho \geq 0$ be an *even* function in $L^1(\mathbb{R})$ with $\int_{\mathbb{R}} \rho(z) dz = 1$ and finite fourth moment $m_4 := \int_{\mathbb{R}} z^4 \rho(z) dz < \infty$. Then $m_2 := \int_{\mathbb{R}} z^2 \rho(z) dz < \infty$. Set $\rho_\varepsilon(z) := \varepsilon^{-1}\rho(z/\varepsilon)$ (periodized on \mathbb{T}) and define

$$\partial_t u_\varepsilon(t, x) = \frac{2\kappa}{m_2 \varepsilon^2} \int_{\mathbb{T}} \rho_\varepsilon(x-y) (u_\varepsilon(t, y) - u_\varepsilon(t, x)) dy. \quad (6.22)$$

Equation (6.22) is of the form (5.9) with $D(a, b) = b - a$ and $W_\varepsilon(x, y) = \frac{2\kappa}{m_2 \varepsilon^2} \rho_\varepsilon(x - y)$, and

$$(\mathbf{A}_\varepsilon u)(x) = \frac{2\kappa}{m_2 \varepsilon^2} \int_{\mathbb{T}} \rho_\varepsilon(x-y) (u(y) - u(x)) dy, \quad (6.23)$$

For smooth u , the right-hand side of (6.22) converges to $\kappa \partial_{xx} u$ as $\varepsilon \rightarrow 0$ (see the last part of the lemma below).

Lemma 6.3. *Assumption 5.2 is satisfied with $M = 1$, $\omega = 0$ and $r(\varepsilon) = \mathcal{O}(\varepsilon^2)$ for the choice $X = H^4(\mathbb{T})$.*

Proof. Here $\mathbf{A} = \kappa \partial_{xx}$ on $L^2(\mathbb{T})$ with $\mathcal{D}(\mathbf{A}) = H^2(\mathbb{T})$ and core $\mathcal{D}_0 = C^\infty(\mathbb{T})$. For the consistency estimate we choose $X = H^4(\mathbb{T})$:

Uniform semigroup stability. Although $\|\mathbf{A}_\varepsilon\|_{\mathcal{L}(L^2)} \sim \varepsilon^{-2}$ because of the prefactor in (6.23), the operator \mathbf{A}_ε is symmetric and dissipative on $L^2(\mathbb{T})$:

$$\langle \mathbf{A}_\varepsilon u, u \rangle_{L^2(\mathbb{T})} = -\frac{\kappa}{m_2 \varepsilon^2} \iint_{\mathbb{T} \times \mathbb{T}} \rho_\varepsilon(x-y) |u(y) - u(x)|^2 dx dy \leq 0.$$

Hence \mathbf{A}_ε generates a contraction semigroup on $L^2(\mathbb{T})$, uniformly in ε , i.e., Assumption 5.2(1) holds with $M = 1$ and $\omega = 0$.

Consistency. After periodization and the change of variables $x - y = \varepsilon z$, one has

$$(\mathbf{A}_\varepsilon u)(x) = \frac{2\kappa}{m_2 \varepsilon^2} \int_{\mathbb{R}} \rho(z) (u(x - \varepsilon z) - u(x)) dz.$$

By Taylor's formula with integral remainder, one has

$$\begin{aligned} u(x - \varepsilon z) - u(x) &= -\varepsilon z \partial_x u(x) + \frac{\varepsilon^2 z^2}{2} \partial_{xx} u(x) - \frac{\varepsilon^3 z^3}{6} \partial_{xxx} u(x) \\ &\quad + \frac{\varepsilon^4 z^4}{6} \int_0^1 (1 - \theta)^3 \partial_{xxxx} u(x - \theta \varepsilon z) d\theta. \end{aligned}$$

Since ρ is even, the odd moments vanish, while the second moment equals m_2 . Therefore

$$\mathbf{A}_\varepsilon u(x) - \kappa \partial_{xx} u(x) = \frac{\kappa \varepsilon^2}{3m_2} \int_{\mathbb{R}} z^4 \rho(z) \int_0^1 (1 - \theta)^3 \partial_{xxxx} u(x - \theta \varepsilon z) d\theta dz.$$

Now use Minkowski's integral inequality together with the translation invariance of the $L^2(\mathbb{T})$ -norm, we get

$$\begin{aligned} \|\mathbf{A}_\varepsilon u - \kappa \partial_{xx} u\|_{L^2(\mathbb{T})} &\leq \frac{\kappa \varepsilon^2}{3m_2} \int_{\mathbb{R}} z^4 \rho(z) \int_0^1 (1 - \theta)^3 \|\partial_{xxxx} u(\cdot - \theta \varepsilon z)\|_{L^2(\mathbb{T})} d\theta dz \\ &= \frac{\kappa \varepsilon^2}{3m_2} \left(\int_{\mathbb{R}} z^4 \rho(z) dz \right) \left(\int_0^1 (1 - \theta)^3 d\theta \right) \|\partial_{xxxx} u\|_{L^2(\mathbb{T})} \\ &= \frac{\kappa m_4}{12m_2} \varepsilon^2 \|\partial_{xxxx} u\|_{L^2(\mathbb{T})}. \end{aligned}$$

Hence Assumption 5.2(2) holds with $r(\varepsilon) = \mathcal{O}(\varepsilon^2)$ and $X = H^4(\mathbb{T})$. \square

Lift to K and IPS. The lifted equation on K is

$$\partial_t \tilde{u}_\varepsilon(t, x) = \frac{2\kappa}{m_2 \varepsilon^2} \int_K \rho_\varepsilon(\Phi^{-1}(x) - \Phi^{-1}(y)) (\tilde{u}_\varepsilon(t, y) - \tilde{u}_\varepsilon(t, x)) d\nu(y),$$

and its Galerkin/IPS approximation reads

$$\dot{u}_w^\varepsilon(t) = \sum_{|v|=m} \rho_{wv}^{(\varepsilon)} (u_v^\varepsilon(t) - u_w^\varepsilon(t)) \nu(K_v), \quad (6.24)$$

where $\rho_{wv}^{(\varepsilon)} := \frac{2\kappa}{m_2 \varepsilon^2} \int_{K_w \times K_v} \rho_\varepsilon(\Phi^{-1}(x) - \Phi^{-1}(y)) d(\nu \times \nu) = \frac{2\kappa}{m_2 \varepsilon^2} \int_{Q_w \times Q_v} \rho_\varepsilon(x - y) d(\lambda \times \lambda)$.

The structure of (6.24) is especially transparent. Since ρ is even and nonnegative, the coefficients satisfy $\rho_{wv}^{(\varepsilon)} = \rho_{vw}^{(\varepsilon)} \geq 0$. Hence (6.24) is a *weighted graph Laplacian* on the level- m self-similar network. In particular, it preserves the discrete mass $\sum_{|w|=m} u_w^\varepsilon(t) \nu(K_w)$ and dissipates the discrete L^2 -energy:

$$\frac{1}{2} \frac{d}{dt} \sum_{|w|=m} |u_w^\varepsilon(t)|^2 \nu(K_w) = -\frac{1}{2} \sum_{|w|, |v|=m} \rho_{wv}^{(\varepsilon)} |u_v^\varepsilon(t) - u_w^\varepsilon(t)|^2 \nu(K_w) \nu(K_v) \leq 0.$$

Thus the IPS has the qualitative structure one expects from diffusion: symmetric couplings, nonnegative rates, and relaxation toward local equilibrium. In the present setting the important point is that these coefficients are not imposed by hand; they are obtained by transporting a classical nonlocal diffusion kernel through the IFS coding and then averaging it on the self-similar cells.

As $\varepsilon \rightarrow 0$ with m fixed, the weights $\rho_{wv}^{(\varepsilon)}$ diverge (reflecting the ε^{-2} prefactor) but concentrate on pairs of adjacent cells in Q ; the IPS (6.24) then formally approaches the Galerkin discretization of $\kappa \partial_{xx} u$ on the level- m partition. In the joint limit, Theorem 5.4 with $r(\varepsilon) = \mathcal{O}(\varepsilon^2)$ prescribes the balanced scaling $\varepsilon \sim k^{-m/2}$.

7 Intrinsic operators, local charts, and further directions

7.1 Scope of the construction

Sections 5–6 provide a four-step procedure: start from a PDE on a standard reference domain Q , replace it by a nonlocal approximation at scale ε , transport that approximation to the fractal domain K through the IFS isomorphism, and finally discretize the lifted problem on the level- m partition of K . The output is an explicit self-similar IPS, namely (5.8) in the linear case and (5.10) in the nonlinear/nonlocal case. Theorem 5.4 and Corollary 5.5 quantify the full approximation error.

The two-scale estimate is important because the two parameters have different meanings. The parameter ε measures the consistency error of the nonlocal surrogate on Q , whereas the level m (or equivalently the number of particles $N_m = k^m$) measures the resolution of the self-similar network on K . Once the order of consistency $r(\varepsilon)$ is known, the estimate prescribes how the nonlocalization scale and the network depth should be coupled. For the model problems treated here, one finds $\beta = 1$ for transport and Burgers in the smooth regime, and $\beta = 2$ for heat. In particular, diffusion allows a larger nonlocalization scale ε than first-order dynamics for the same network size and target accuracy.

The operator obtained on K is transported through the symbolic coding, not intrinsic in the sense of fractal analysis. This is both a limitation and a feature. It is a limitation if the objective is to recover an operator canonically attached to the geometry of K , such as a Kigami-type Laplacian. It is a feature if the objective is to model dynamics on hierarchical networks whose natural continuum description is inherited from a reference PDE posed on a simpler space Q . The present construction should therefore be viewed as complementary to the intrinsic theory of PDEs on fractals, not as a replacement for it.

There is nevertheless a useful interface with the Dirichlet-form framework. If a lifted kernel is symmetric and nonnegative, then it defines a quadratic nonlocal energy of the form

$$\mathcal{E}_\varepsilon(u, u) = \frac{1}{2} \iint_{K \times K} (u(x) - u(y))^2 J_\varepsilon(x, y) d\nu(x) d\nu(y),$$

and, under the standard closability and regularity assumptions, a symmetric Dirichlet form. In that case the calculus of [6, 15] applies at fixed ε . The intrinsic question is then a limiting one: after a suitable renormalization, do the forms \mathcal{E}_ε converge, in Mosco or Γ sense, to a local resistance form such as the Kigami energy? Related nonlocal-to-local limits on Sierpiński-type fractals were studied in [10, 28]. This question is orthogonal to the two-scale Galerkin estimate proved here, but it provides a natural bridge between transported nonlocal models and intrinsic fractal diffusion.

7.2 Nonlinear problems and boundary conditions

Several ingredients still deserve a systematic treatment before one can claim a fully general theory. The first is a genuinely nonlinear version of Theorem 5.4. For semilinear or quasilinear problems, the linear semigroup estimate should be replaced by a uniform well-posedness and Lipschitz-stability statement for the nonlocal flows on bounded subsets of a regularity space X , together with a consistency estimate for the nonlinear

operator on a core. The Burgers example strongly suggests that such an extension is feasible in the smooth regime, but it should be stated as an abstract theorem rather than left at the level of representative examples.

A second point is the treatment of boundary conditions on $[0, 1]$. The periodic case is the cleanest one because the kernels are translation invariant and no boundary layer appears. On a bounded interval, the nonlocal surrogate must encode the boundary condition through its interaction with the exterior of the domain. A natural guiding principle is to fix an extension operator \mathcal{E}_B associated with the boundary condition B , apply the whole-space nonlocal operator to $\mathcal{E}_B u$, and then restrict the result back to $[0, 1]$.

For homogeneous Dirichlet diffusion, one may take the zero extension $\mathcal{E}_D u$ and define

$$(\mathbf{A}_\varepsilon u)(x) := \frac{2\kappa}{m_2 \varepsilon^2} \int_{\mathbb{R}} \rho_\varepsilon(x-y) (\mathcal{E}_D u(y) - u(x)) dy,$$

that is,

$$(\mathbf{A}_\varepsilon u)(x) = \frac{2\kappa}{m_2 \varepsilon^2} \int_0^1 \rho_\varepsilon(x-y) (u(y) - u(x)) dy - \frac{2\kappa}{m_2 \varepsilon^2} u(x) \int_{\mathbb{R} \setminus [0,1]} \rho_\varepsilon(x-y) dy.$$

This preserves dissipativity and is the natural nonlocal counterpart of the Dirichlet heat operator. For homogeneous Neumann diffusion, one replaces zero extension by an even reflection $\mathcal{E}_N u$ at the endpoints and sets

$$(\mathbf{A}_\varepsilon u)(x) := \frac{2\kappa}{m_2 \varepsilon^2} \int_{\mathbb{R}} \rho_\varepsilon(x-y) (\mathcal{E}_N u(y) - u(x)) dy,$$

which preserves constants and is consistent with $\partial_x u = 0$ at the boundary. For first-order problems, the appropriate mechanism is typically one-sided: for instance, if $b > 0$ and one prescribes an inflow datum at $x = 0$, one should use a kernel supported in $[0, \infty)$ together with an exterior extension carrying the inflow value on $(-\infty, 0)$. This yields a boundary-adapted nonlocal derivative that approximates $b\partial_x u$ while enforcing the inflow condition. What remains to be proved in all these cases is the analogue of Assumption 5.2: one must establish uniform semigroup stability (or nonlinear stability) and control the boundary layer created by the loss of translation invariance near $x = 0, 1$.

A third point concerns genuinely hyperbolic regimes after shock formation. For inviscid Burgers, the centered nonlocal approximation is adequate only before shocks or after viscous regularization. To recover entropy solutions directly, one needs either a vanishing-viscosity limit or a monotone one-sided nonlocal flux. Thus entropy admissibility is not an automatic consequence of the present framework, but an additional structural requirement on the approximation.

7.3 Local charts on the Sierpinski gasket

The main geometric gap between the transported framework and intrinsic operators on fractals comes from adjacency. The global coding of the cylinders Q_w by subintervals of $[0, 1]$ imposes a total order, whereas the cells K_w in a fractal such as the Sierpinski gasket have a different adjacency graph in Euclidean space. A possible refinement is therefore to use the interval only *locally*, as a chart on each self-similar cell, rather than as a single global coordinate.

For every word $w \in \Sigma_m$, the conjugacy relation $\Phi \circ g_w = f_w \circ \Phi$ holds almost everywhere on Q_w . Hence one may regard

$$\Phi_w := f_w \circ \Phi \circ g_w^{-1} : Q_w \longrightarrow K_w \tag{7.1}$$

as a local chart from the interval cell Q_w to the fractal cell K_w . In the case of the Sierpinski gasket, the three first-generation cells K_1, K_2, K_3 already form a natural atlas, since their overlaps reduce to boundary points and therefore have ν -measure zero. For L^2 -based constructions, this means that one can work chartwise on each cell and then impose coupling conditions only on the lower-dimensional interfaces.

More concretely, let T_w denote the local isometry between $L^2(K_w, \nu|_{K_w})$ and $L^2(Q_w, \lambda|_{Q_w})$ induced by (7.1). Given a nonlocal approximation \mathbf{A}_ε on Q , one may define on each cell a local transported operator

$$\tilde{\mathbf{A}}_{\varepsilon, w} := T_w^{-1} \mathbf{A}_\varepsilon T_w.$$

A patchwise diffusion operator can then be assembled from the family $(\tilde{\mathbf{A}}_{\varepsilon,w})_{|w|=m}$ together with compatibility conditions across the vertices shared by neighboring cells. At least heuristically, this chartwise construction is closer in spirit to manifold atlases or finite-element assembly than the global coding used in the main body of the paper, and it may provide a bridge between the present transported viewpoint and intrinsic operators such as the Kigami Laplacian [18, 25]. Making this idea precise would require a careful analysis of the rescaling, of the matching conditions at common vertices, and of the limit $m \rightarrow \infty$.

A more intrinsic version of this idea would replace symbolic charts by harmonic coordinates, which are adapted to energy rather than to measure. On the Sierpiński gasket, harmonic coordinates and the associated Kusuoka-type energy measures provide a coordinate system closer to the resistance geometry [19, 27]. This is also consistent with recent works [22, 21] on harmonic maps from p.c.f. fractals to the circle and on the Kuramoto model on graphs approximating the Sierpiński gasket, where homotopy classes, harmonic extension, Dirichlet energy and Γ -convergence play a central role. These works suggest that an intrinsic variant of the present paper should probably be formulated in terms of energy-adapted local charts and prefractal adjacency, rather than through a single global symbolic coordinate.

7.4 Pullback of nonlocal equations on the fractal

The same isomorphism can be used in the reverse direction. Suppose that one starts from a nonlocal evolution equation on the fractal:

$$\partial_t u(t, x) = \int_K J(x, y)(u(t, y) - u(t, x)) d\nu(y), \quad (7.2)$$

where $J \in L^2(K \times K, \nu \times \nu)$ (or, more generally, satisfies suitable Schur-type bounds). Define the pullback kernel

$$J^\Phi(\xi, \eta) := \begin{cases} J(\Phi(\xi), \Phi(\eta)), & (\xi, \eta) \in Q_0 \times Q_0, \\ 0, & \text{otherwise.} \end{cases} \quad (7.3)$$

Then the dynamics on K is conjugate to a nonlocal equation on Q . The following statement then follows immediately from the measure-preserving property of Φ and the definitions of T and J^Φ .

Proposition 7.1. *Let \mathcal{L}_K be the operator defined by the right-hand side of (7.2), and let \mathcal{L}_Q be the operator on $L^2(Q, \lambda)$ defined by*

$$(\mathcal{L}_Q v)(\xi) = \int_Q J^\Phi(\xi, \eta)(v(\eta) - v(\xi)) d\lambda(\eta).$$

Then $T \mathcal{L}_K T^{-1} = \mathcal{L}_Q$. Consequently, u solves (7.2) on K if and only if $v := Tu$ solves $\partial_t v = \mathcal{L}_Q v$ on Q .

This simple observation suggests a complementary numerical program. Instead of transporting a classical PDE from Q to K , one may start from a native nonlocal equation on the fractal and pull it back to the interval, where one can apply one-dimensional quadrature, sparse approximation, or Monte-Carlo methods. A particularly attractive case is the fractional heat equation $\partial_t u + (-\Delta_K)^s u = 0$, for $0 < s < 1$, on the Sierpinski gasket, where Δ_K denotes an intrinsic Laplacian on the fractal and $(-\Delta_K)^s$ is defined through spectral calculus. Whenever such an operator admits an integral representation (either directly as a singular kernel or through a fractional Riesz kernel) its pullback through (7.3) produces a singular one-dimensional nonlocal operator on Q . Recent work on fractional Gaussian fields and fractional Riesz kernels on the Sierpinski gasket indicates that this direction is analytically meaningful [2]. Developing numerical schemes for these pulled-back operators would naturally complement the transport framework of the present paper.

This reverse use of the isomorphism is close in spirit to recent integral-equation methods for acoustic scattering by fractals, where the physical model is first written as a nonlocal integral equation on the fractal scatterer and then discretized by Galerkin and quadrature methods [4]. Our contribution in such a setting would be different: the symbolic isomorphism gives a systematic way of pulling native fractal integral equations back to a one-dimensional reference space, where classical quadrature, sparse approximation, randomized sampling, or the self-similar cell averages developed in this paper can be applied.

7.5 Additional directions

The previous subsections point to several further questions. One may compare the transported operators obtained here with intrinsic operators on specific fractals, study sparse or random self-similar networks in the spirit of [17], or consider higher-dimensional reference spaces Q and anisotropic kernels. From the particle-system viewpoint, it would also be interesting to understand systematically which qualitative properties of the reference PDE survive after transport and Galerkin discretization: dissipation for diffusion, conservative or one-sided structures for transport and scalar conservation laws, maximum principles, entropy inequalities, or synchronization thresholds in oscillator-type models. Clarifying these correspondences would significantly strengthen the conceptual content of the framework.

Another promising direction is optimal transport on fractals. The map Φ identifies probability measures on K that are absolutely continuous with respect to ν with probability measures on Q that are absolutely continuous with respect to λ . If the cost is defined through the symbolic coordinate, the problem is reduced to a classical one-dimensional optimal-transport problem; if the cost is intrinsic on K (for example Euclidean or resistance distance), its pullback becomes a nonlocal cost on Q . This gives a flexible framework in which one can separate measure-theoretic transport, intrinsic geometry, and numerical discretization.

Acknowledgements. This work was partially supported by the National Science Foundation through grant DMS-2406941 (to GSM).

References

- [1] B. Bárány, K. Simon, and B. Solomyak. *Self-similar and self-affine sets and measures*, volume 276 of *Math. Surv. Monogr.* Providence, RI: American Mathematical Society (AMS), 2023.
- [2] F. Baudoin and C. Lacaux. Fractional gaussian fields on the Sierpiński gasket and related fractals. *J. Anal. Math.*, 146:719–739, 2022.
- [3] C. J. Bishop and Y. Peres. *Fractals in probability and analysis*, volume 162 of *Camb. Stud. Adv. Math.* Cambridge: Cambridge University Press, 2017.
- [4] A. M. Caetano, S. N. Chandler-Wilde, X. Claeys, A. Gibbs, D. P. Hewett, and A. Moiola. Integral equation methods for acoustic scattering by fractals. *Proceedings of the Royal Society A*, 481(2306):20230650, 2025.
- [5] J. P. Chen, M. Hinz, and A. Teplyaev. From non-symmetric particle systems to non-linear PDEs on fractals. In *Stochastic partial differential equations and related fields*, volume 229 of *Springer Proc. Math. Stat.*, pages 503–513. Springer, Cham, 2018.
- [6] F. Cipriani and J.-L. Sauvageot. Derivations as square roots of Dirichlet forms. *Journal of Functional Analysis*, 201(1):78–120, 2003.
- [7] M. G. Crandall and A. Majda. Monotone difference approximations for scalar conservation laws. *Math. Comp.*, 34(149):1–21, 1980.
- [8] C. M. Dafermos. *Hyperbolic conservation laws in continuum physics*, volume 325 of *Grundlehren der mathematischen Wissenschaften*. Springer, Berlin, fourth edition, 2016.
- [9] M. Fukushima and T. Shima. On a spectral analysis for the Sierpiński gasket. *Potential Analysis*, 1(1):1–35, 1992.
- [10] A. Grigor’yan and M. Yang. Local and non-local Dirichlet forms on the Sierpiński carpet. *Transactions of the American Mathematical Society*, 372(6):3985–4030, 2019.
- [11] M. Hinz, D. Koch, and M. Meinert. Sobolev spaces and calculus of variations on fractals. In *Analysis, probability and mathematical physics on fractals. Based on the presentations at the 6th conference, Cornell University, Ithaca, NY, USA, June 2017*, pages 419–450. Hackensack, NJ: World Scientific, 2020.
- [12] M. Hinz and M. Meinert. Approximation of partial differential equations on compact resistance spaces. *Calculus of Variations and Partial Differential Equations*, 61(1):19, 2022.
- [13] M. Hinz, M. Röckner, and A. Teplyaev. Vector analysis for Dirichlet forms and quasilinear PDE and SPDE on metric measure spaces. *Stochastic Processes and their Applications*, 123(12):4373–4406, 2013.

- [14] M. Hinz and W. Schefer. On equations of continuity and transport type on metric graphs and fractals. *arXiv preprint arXiv:2412.07988*, 2024.
- [15] M. Hinz and A. Teplyaev. Vector analysis on fractals and applications. In *Fractal geometry and dynamical systems in pure and applied mathematics II: Fractals in applied mathematics.*, pages 147–163. Providence, RI: American Mathematical Society (AMS), 2013.
- [16] J. E. Hutchinson. Fractals and self-similarity. *Indiana Univ. Math. J.*, 30(5):713–747, 1981.
- [17] D. Kaliuzhnyi-Verbovetskyi and G. S. Medvedev. Sparse Monte Carlo method for nonlocal diffusion problems. *SIAM J. Numer. Anal.*, 60(6):3001–3028, 2022.
- [18] J. Kigami. *Analysis on fractals*, volume 143 of *Cambridge Tracts in Mathematics*. Cambridge University Press, Cambridge, 2001.
- [19] J. Kigami. Measurable Riemannian geometry on the Sierpinski gasket: the Kusuoka measure and the Gaussian heat kernel estimate. *Mathematische Annalen*, 340(4):781–804, 2008.
- [20] G. S. Medvedev. Interacting dynamical systems on networks and fractals: discrete and continuous models, mean-field limit, and convergence rates. Preprint, arXiv:2601.23175 [math.DS] (2026), 2026.
- [21] G. S. Medvedev and M. S. Mizuhara. The Kuramoto model on the Sierpinski gasket II: twisted states. *arXiv preprint arXiv:2506.12940*, 2025.
- [22] G. S. Medvedev and M. S. Mizuhara. Kuramoto model on sierpinski gasket i: Harmonic maps, 2026.
- [23] U. Mosco. Analysis and numerics of some fractal boundary value problems. *Boll. Unione Mat. Ital. (9)*, 6(1):53–73, 2013.
- [24] T. Paul and E. Trélat. Universal approximations of quasilinear PDEs by finite distinguishable particle systems. Preprint, arXiv:2501.11387 [math.AP] (2025), 2025.
- [25] R. S. Strichartz. *Differential equations on fractals*. Princeton University Press, Princeton, NJ, 2006. A tutorial.
- [26] D. W. Stroock. *Mathematics of Probability*, volume 149 of *Grad. Stud. Math.* Providence, RI: American Mathematical Society (AMS), 2013.
- [27] A. Teplyaev. Harmonic coordinates on fractals with finitely ramified cell structure. *Canadian Journal of Mathematics*, 60(2):457–480, 2008.
- [28] M. Yang. Equivalent semi-norms of non-local Dirichlet forms on the Sierpiński gasket and applications. *Potential Analysis*, 49(2):287–308, 2018.