

Article

Not peer-reviewed version

Spectral Degeneracy Operators: A Mathematical Foundation for Physics-Informed Turbulence Modeling

[Rômulo Damasclin Chaves dos Santos](#)^{*} and [Jorge Henrique de Oliveira Sales](#)

Posted Date: 20 October 2025

doi: 10.20944/preprints202510.1465.v1

Keywords: spectral degeneracy operators; degenerate elliptic PDE; Bessel potentials; Γ -convergence; physics-informed neural networks



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a Creative Commons CC BY 4.0 license, which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Article

Spectral Degeneracy Operators: A Mathematical Foundation for Physics-Informed Turbulence Modeling

Rômulo Damasclin Chaves dos Santos * and Jorge Henrique de Oliveira Sales

Santa Cruz State University (UESC), Department of Exact Sciences, Brazil

* Correspondence: rdcsantos@uesc.br

Abstract

This paper establishes a rigorous mathematical foundation for Spectral Degeneracy Operators (SDOs) and their integration into physics-informed neural networks for turbulence modeling. We introduce a novel class of anisotropic degenerate elliptic operators with separable weight structures that serve dual purposes as analytical tools and trainable neural network components. Our three principal contributions include: (1) A comprehensive fractional regularity theory proving that solutions to $L_{a,\theta}u = f$ gain up to $\min\{s, \delta\}$ fractional derivatives in weighted Sobolev-Besov spaces, with explicit dependence on degeneracy exponents $\theta_i \in [1, 2]$; (2) A spectral stability theorem for deep SDO networks demonstrating bounded mode amplitude variations of order $\mathcal{O}(\delta \log(1/\delta))$ under parameter perturbations, preventing catastrophic spectral drift during training; (3) A Γ -convergence framework establishing variational limits of discrete SDO energy functionals. These theoretical advances are supported by strengthened proofs for inverse calibration stability and universality of divergence-free SDO closures. Our work bridges degenerate PDE theory with modern machine learning, providing rigorous guarantees for stability, interpretability, and convergence of neural operator architectures in turbulence modeling applications.

Keywords: spectral degeneracy operators; degenerate elliptic PDE; Bessel potentials; Γ -convergence; physics-informed neural networks

1. Introduction

The mathematical modeling of turbulent flows has long relied on closure operators that mediate the interaction between resolved and unresolved scales [12,17]. Classical frameworks such as Reynolds-Averaged Navier–Stokes (RANS) and Large Eddy Simulation (LES) are built upon phenomenological assumptions which, despite their practical relevance, often fail to accurately capture key features of turbulence, including anisotropic dissipation, intermittency, and localized coherent structures. In recent years, data-driven methods most notably physics-informed neural networks (PINNs) [7] and neural operator architectures [6] have provided powerful alternatives by embedding physical laws into machine learning models. Nevertheless, these approaches typically lack a unified analytic foundation that simultaneously: (i) introduces physically meaningful, trainable anisotropic singularities into neural representations; (ii) preserves the spectral and variational structure of the underlying partial differential equations (PDEs); and (iii) admits rigorous inverse, stability, and convergence guarantees.

A major step toward bridging this gap was achieved in [1], where the authors introduced *Spectral Degeneracy Operators* (SDOs) as a mathematically rigorous and computationally viable framework that couples degenerate PDE theory, spectral analysis, and modern neural architectures for turbulence modeling. The central idea is to embed adaptive singularities and symmetry structures directly into neural layers via differential operators with trainable degeneracy centers. This construction yields layers that act as anisotropic, data-adaptive spectral filters while retaining physical interpretability.

The SDO framework establishes a complete spectral theory, including self-adjointness, compact resolvent, and a tensorized Bessel-type eigenbasis; derives Lipschitz stability estimates for inverse calibration of degeneracy points from boundary measurements; proves a universality theorem ensuring that SDO-based networks approximate any divergence-free turbulence closure operator; and introduces a neural-turbulence correspondence principle demonstrating that learned degeneracy centers converge to physically meaningful turbulent structures during training. Consequently, SDOs provide a mathematically principled and physically consistent alternative to traditional turbulence closures and standard neural networks, combining expressive power with structural guarantees.

In this work, we significantly advance the analytical foundations of *Spectral Degeneracy Operators* (SDOs), a class of anisotropic, degenerate elliptic operators that serve both as fundamental objects of PDE theory and as parametric modules in neural architectures. For a bounded Lipschitz domain $\Omega \subset \mathbb{R}^d$, a degeneracy center $a \in \Omega$, and exponents $\theta \in [1, 2)^d$, we consider

$$L_{a,\theta}u := -\nabla \cdot (D_{a,\theta}(x) \nabla u), \quad D_{a,\theta}(x) = \text{diag}(|x_1 - a_1|^{\theta_1}, \dots, |x_d - a_d|^{\theta_d}). \quad (1.1)$$

Operators of this form naturally blend Bessel-type singular Sturm–Liouville behavior with directionally adaptive spectral structure (cf. [2,11,18]). While prior work outlined their spectral decomposition and relevance to turbulence modeling, a complete analytic theory addressing regularity, stability, and variational limits remained undeveloped. This manuscript fills this gap by establishing: (1) a comprehensive fractional regularity theory for elliptic SDO problems; (2) a detailed spectral stability analysis for deep compositions of SDO networks under parameter perturbations; and (3) a full variational limit theory, in the sense of Γ -convergence, for SDO-induced energy functionals.

1.1. Mathematical Motivations and Principal Contributions

Three core mathematical motivations drive this study. First, degeneracy centers (trainable parameters a) act as adaptive attention loci within a physically grounded operator basis; understanding the sensitivity of eigenvalues and eigenmodes with respect to variations in a is crucial for ensuring robust training and model interpretability. Second, regularity estimates in weighted Sobolev and Besov spaces provide rigorous control of approximation errors and justify spectral truncation strategies employed in practical implementations. Third, variational convergence connects discrete, network-induced energy landscapes with continuum turbulence limits, ensuring consistency between training objectives and physical modeling.

Accordingly, our principal contributions are:

1. A comprehensive *fractional regularity theory* for solutions of $L_{a,\theta}u = f$ in weighted Sobolev and Besov scales, establishing interior and boundary estimates (Theorem 3.5).
2. A *spectral stability theorem for deep SDO networks* that quantifies how compositions of SDO inverses and linear readouts preserve modal amplitudes and eigenvalue clusters under small perturbations of (a, θ) , yielding practical bounds for training stability (Theorem 4.2).
3. A complete Γ -convergence framework linking discrete and parametric SDO energy functionals to a continuum variational limit associated with turbulent energy dissipation.
4. Strengthened and fully rigorous proofs of inverse calibration stability and universality of divergence-free SDO networks.

1.2. Related Mathematical Work

The present analysis lies at the intersection of several mature areas of mathematical research. First, our framework builds on the classical theory of *degenerate elliptic and parabolic PDEs*, where foundational results on well-posedness, weighted Sobolev spaces, and regularity were established in [11] and further developed in boundary layer settings by [14]. These works provide the analytic backbone for handling degeneracies such as those induced by SDO weights.

Second, the *stability of inverse problems for degenerate operators* is essential for the recovery of degeneracy centers. Recent advances include Lipschitz-type reconstruction results for parabolic

equations with degeneracy [2], direct and inverse source problems [4], and integral-observation-based stability results in strongly degenerate settings [8]. These studies motivate and inform our inverse calibration stability theorems for SDOs.

Third, our spectral analysis draws on *singular Sturm–Liouville and Bessel theory*, with Watson’s classical treatment [18] providing the primary asymptotic tools necessary to characterize the eigenstructure of separable degenerate operators. This spectral foundation is crucial for developing fractional powers, weighted Besov scales, and modewise stability estimates.

In the context of modern machine learning, our work connects with the emerging field of *neural operator approximation*. Neural operators and graph kernel networks [6] have demonstrated the ability to learn solution operators to PDEs in infinite-dimensional function spaces, establishing universal approximation properties and stability guarantees. Complementary to this, *physics-informed neural networks* (PINNs) [7] enforce differential constraints via variational residual minimization, motivating the integration of PDE structure directly into training objectives.

Finally, our variational limit theory is grounded in the classical *De Giorgi–Dal Maso framework* for Γ -convergence [15]. By adapting

2. Mathematical Preliminaries and Functional Analytic Framework

We consider a bounded Lipschitz domain $\Omega \subset \mathbb{R}^d$ and denote spatial variables by $x = (x_1, \dots, x_d)$. For a measurable weight function $w : \Omega \rightarrow [0, \infty)$, we define $L^2(\Omega, w \, dx)$ as the weighted L^2 space equipped with the norm

$$\|u\|_{L^2(\Omega, w)}^2 := \int_{\Omega} |u(x)|^2 w(x) \, dx. \quad (2.1)$$

Our primary focus will be on separable weights of the form

$$w_{a, \theta}(x) := \prod_{i=1}^d |x_i - a_i|^{\theta_i}, \quad a \in \Omega, \theta_i \in [1, 2). \quad (2.2)$$

2.1. Weighted Sobolev Spaces and Trace Theory

To rigorously analyze degenerate elliptic problems, we require specialized function spaces that account for the anisotropic weighting in the differential operator. This subsection introduces the weighted Sobolev space framework and establishes fundamental inequalities that underpin the well-posedness and regularity theory.

Definition 2.1 (Weighted Sobolev Space). *Let $\Omega \subset \mathbb{R}^d$ be a bounded domain, $a = (a_1, \dots, a_d) \in \Omega$, and $\theta = (\theta_1, \dots, \theta_d)$ with $\theta_i \in [1, 2)$ for each $i = 1, \dots, d$. The weighted Sobolev space $H_{0, \theta}^1(\Omega)$ is defined as the completion of $C_c^\infty(\Omega)$ with respect to the norm*

$$\|u\|_{H_{0, \theta}^1}^2 := \|u\|_{L^2(\Omega)}^2 + \sum_{i=1}^d \int_{\Omega} |x_i - a_i|^{\theta_i} |\partial_{x_i} u|^2 \, dx. \quad (2.3)$$

This space is a Hilbert space when equipped with the inner product

$$\langle u, v \rangle_{H_{0, \theta}^1} := \int_{\Omega} uv \, dx + \sum_{i=1}^d \int_{\Omega} |x_i - a_i|^{\theta_i} \partial_{x_i} u \partial_{x_i} v \, dx.$$

Remark 2.2. *The restriction $\theta_i \in [1, 2)$ ensures:*

- The weight $|x_i - a_i|^{\theta_i}$ is integrable near $x_i = a_i$ (critical for L^2 -theory),
- The space $H_{0, \theta}^1(\Omega)$ is compactly embedded in $L^2(\Omega)$ (via weighted Rellich–Kondrachov),
- The Poincaré inequality holds uniformly (see Lemma 2.3).

For $\theta_i \geq 2$, the weight may fail to be locally integrable, complicating the analysis.

Standard density results extend to this weighted setting under natural assumptions on a and θ . Specifically, $C_c^\infty(\Omega)$ is dense in $H_{0,\theta}^1(\Omega)$ provided the weights $|x_i - a_i|^{\theta_i}$ are locally integrable and the boundary $\partial\Omega$ is sufficiently regular (e.g., Lipschitz). For a comprehensive treatment of weighted Sobolev spaces, we refer to [2,11].

The following weighted Poincaré inequality is fundamental for establishing coercivity of the bilinear form associated with degenerate operators:

Lemma 2.3 (Weighted Poincaré Inequality). *There exists a constant $C_P > 0$, depending only on Ω , a , and θ , such that for all $u \in H_{0,\theta}^1(\Omega)$,*

$$\|u\|_{L^2(\Omega)} \leq C_P \left(\sum_{i=1}^d \int_{\Omega} |x_i - a_i|^{\theta_i} |\partial_{x_i} u|^2 dx \right)^{1/2}. \quad (2.4)$$

Proof. The proof proceeds in three steps:

1. One-Dimensional Reduction. Fix $i \in \{1, \dots, d\}$ and consider the one-dimensional weighted Poincaré inequality along the x_i -direction. For each $x' \in \mathbb{R}^{d-1}$, define the slice

$$\Omega_{x'} := \{t \in \mathbb{R} : (x'_1, \dots, x'_{i-1}, t, x'_i, \dots, x'_{d-1}) \in \Omega\}.$$

By Fubini's theorem, it suffices to establish the inequality for almost every slice $\Omega_{x'}$. For fixed x' , the one-dimensional weighted Poincaré inequality states:

$$\int_{\Omega_{x'}} |u(x'_1, \dots, t, \dots, x'_{d-1})|^2 dt \leq C \int_{\Omega_{x'}} |t - a_i|^{\theta_i} |\partial_t u(x'_1, \dots, t, \dots, x'_{d-1})|^2 dt,$$

where C depends on θ_i and the length of $\Omega_{x'}$. This follows from classical one-dimensional weighted Hardy inequalities (see [11]), valid since $\theta_i \in [1, 2)$.

2. Partition of Unity. To globalize the estimate, we employ a partition of unity subordinate to a cover of Ω by coordinate-aligned cylinders. For each cylinder Q , we apply the one-dimensional inequality along the x_i -direction, summing over all directions and cylinders. The condition $\theta_i \geq 1$ ensures that the weight $|x_i - a_i|^{\theta_i}$ controls the L^2 -norm locally, while the Dirichlet boundary condition $u|_{\partial\Omega} = 0$ (in the trace sense) guarantees that the Poincaré constant is uniform across slices.

3. Synthesis. Summing the one-dimensional inequalities over all coordinates and integrating over the transverse variables, we obtain:

$$\|u\|_{L^2(\Omega)}^2 \leq C \sum_{i=1}^d \int_{\Omega} |x_i - a_i|^{\theta_i} |\partial_{x_i} u|^2 dx.$$

The constant C depends on the diameter of Ω , the maximum of the θ_i , and the geometry of the partition. This completes the proof. $\square$

Corollary 2.4 (Equivalence of Weighted Norms with Explicit Domain Dependence). *Let $\Omega \subset \mathbb{R}^d$ be a bounded domain contained in the ball $B_R(a) := \{x \in \mathbb{R}^d : |x - a| < R\}$ for some $a \in \mathbb{R}^d$ and radius $R > 0$. Let $H_{0,\theta}^1(\Omega)$ be the weighted Sobolev space with weights $\theta = (\theta_1, \dots, \theta_d)$, endowed with the norm*

$$\|u\|_{H_{\theta}^1(\Omega)} := \left(\int_{\Omega} |u(x)|^2 dx + \sum_{i=1}^d \int_{\Omega} |x_i - a_i|^{\theta_i} |\partial_{x_i} u(x)|^2 dx \right)^{1/2}. \quad (2.5)$$

Define the corresponding semi-norm

$$|u|_{H_{\theta}^1(\Omega)} := \left(\sum_{i=1}^d \int_{\Omega} |x_i - a_i|^{\theta_i} |\partial_{x_i} u(x)|^2 dx \right)^{1/2}. \quad (2.6)$$

Then there exist constants $c, C > 0$, depending explicitly on R, Ω , and the weights θ_i , such that

$$c \|u\|_{H_\theta^1(\Omega)} \leq \|u\|_{H_\theta^1(\Omega)} \leq C \|u\|_{H_\theta^1(\Omega)}, \quad \forall u \in H_{0,\theta}^1(\Omega). \quad (2.7)$$

Proof. The upper bound in (2.7) is immediate from the definition (2.5) since all integrals are nonnegative.

For the lower bound, we apply the weighted Poincaré inequality (see (2.4)) which explicitly depends on the domain radius R and weights θ_i . There exists a constant $c_1 = c_1(R, \theta) > 0$ such that

$$\int_\Omega |u(x)|^2 dx \leq c_1 \sum_{i=1}^d \int_\Omega |x_i - a_i|^{\theta_i} |\partial_{x_i} u(x)|^2 dx = c_1 \|u\|_{H_\theta^1(\Omega)}^2, \quad \forall u \in H_{0,\theta}^1(\Omega). \quad (2.8)$$

Using (2.8), we have

$$\begin{aligned} \|u\|_{H_\theta^1(\Omega)}^2 &= \int_\Omega |u(x)|^2 dx + \sum_{i=1}^d \int_\Omega |x_i - a_i|^{\theta_i} |\partial_{x_i} u(x)|^2 dx \\ &\leq (1 + c_1) \sum_{i=1}^d \int_\Omega |x_i - a_i|^{\theta_i} |\partial_{x_i} u(x)|^2 dx \\ &= (1 + c_1) \|u\|_{H_\theta^1(\Omega)}^2. \end{aligned} \quad (2.9)$$

Defining $c := (1 + c_1)^{-1/2}$ gives the lower bound

$$c \|u\|_{H_\theta^1(\Omega)} \leq \|u\|_{H_\theta^1(\Omega)}. \quad (2.10)$$

Hence, the equivalence (2.7) holds with constants depending explicitly on the domain radius R and weights θ_i . $\square$

2.2. Bessel Potential and Weighted Besov Scales

Bessel potential spaces $H^s(\Omega)$ and Besov spaces $B_{p,q}^s(\Omega)$ provide a natural framework for quantifying fractional-order regularity. To handle degeneracies induced by the weights

$$w_{a,\theta}(x) := \prod_{i=1}^d |x_i - a_i|^{\theta_i}, \quad \theta_i \geq 0, \quad (2.11)$$

we employ *anisotropic weighted Bessel potential spaces*, denoted by $H_w^s(\Omega)$ and $B_{p,q,w}^s(\Omega)$. These spaces can be constructed via localization and pullback to one-dimensional model problems near each degeneracy center a_i , combined with standard extension operators and partitions of unity.

Formally, for $s > 0$, the weighted Bessel potential space $H_w^s(\Omega)$ is defined via fractional powers of the weighted Laplacian $L_{a,\theta}$:

$$H_w^s(\Omega) := \left\{ u \in L_w^2(\Omega) : (I + L_{a,\theta})^{s/2} u \in L_w^2(\Omega) \right\}, \quad (2.12)$$

where

$$L_w^2(\Omega) := L^2(\Omega; w_{a,\theta} dx), \quad (2.13)$$

and similarly for weighted Besov spaces $B_{p,q,w}^s(\Omega)$ using interpolation and modulus of smoothness adapted to $w_{a,\theta}$ (see [18]). These constructions ensure that the weighted fractional spaces inherit embedding, interpolation, and compactness properties analogous to the classical unweighted scales, with explicit dependence on the weights θ_i .

2.3. Spectral Theory for Weighted SDOs

Consider the differential operator

$$L_{a,\theta}u := - \sum_{i=1}^d \partial_{x_i} \left(|x_i - a_i|^{\theta_i} \partial_{x_i} u \right), \quad u \in H_{0,\theta}^1(\Omega), \quad (2.14)$$

and its associated bilinear form

$$\mathfrak{a}_{a,\theta}(u, v) := \int_{\Omega} \sum_{i=1}^d |x_i - a_i|^{\theta_i} \partial_{x_i} u \partial_{x_i} v \, dx, \quad u, v \in H_{0,\theta}^1(\Omega). \quad (2.15)$$

The form $\mathfrak{a}_{a,\theta}$ is symmetric, continuous, and coercive modulo the kernel if some $\theta_i = 0$. Applying the Lax–Milgram theorem and using the compact embedding

$$H_{0,\theta}^1(\Omega) \hookrightarrow L^2(\Omega), \quad (2.16)$$

which follows from weighted Rellich–Kondrachov-type theorems, we deduce that $L_{a,\theta}$ with Dirichlet boundary conditions is self-adjoint with compact resolvent.

Consequently, there exists a complete orthonormal basis $\{\varphi_k(\cdot; a, \theta)\}_{k \geq 1}$ of $L^2(\Omega)$ satisfying

$$L_{a,\theta} \varphi_k = \lambda_k \varphi_k, \quad 0 \leq \lambda_1 \leq \lambda_2 \leq \dots \rightarrow \infty. \quad (2.17)$$

The eigenfunctions φ_k are smooth away from the degeneracy centers $\{a_i\}$, and their tensorized Bessel-type asymptotics in coordinate-separable domains follow from one-dimensional singular Sturm–Liouville theory (see [18]). These asymptotics provide the foundation for defining fractional powers $(I + L_{a,\theta})^s$ and for characterizing $H_w^s(\Omega)$ and $B_{p,q,w}^s(\Omega)$ in terms of spectral expansions.

2.4. Equivalence of Weighted Spectral and Bessel Norms

For $s \geq 0$ and $u \in H_w^s(\Omega)$, we define the *spectral norm*

$$\|u\|_{H_{w,\text{spec}}^s} := \left(\sum_{k=1}^{\infty} (1 + \lambda_k)^s |\langle u, \varphi_k \rangle_{L_w^2}|^2 \right)^{1/2}, \quad (2.18)$$

where $\langle \cdot, \cdot \rangle_{L_w^2}$ denotes the weighted L^2 inner product. Then the Bessel potential norm (2.12) and the spectral norm (2.18) are equivalent: there exist constants $C_1, C_2 > 0$, depending explicitly on s, θ_i , and the domain radius, such that

$$C_1 \|u\|_{H_{w,\text{spec}}^s} \leq \|u\|_{H_w^s} \leq C_2 \|u\|_{H_{w,\text{spec}}^s}, \quad \forall u \in H_w^s(\Omega). \quad (2.19)$$

Proof of (2.19). Let $u \in H_w^s(\Omega)$ with spectral decomposition

$$u = \sum_{k=1}^{\infty} \langle u, \varphi_k \rangle_{L_w^2} \varphi_k, \quad (2.20)$$

where $\{\varphi_k\}_{k \geq 1}$ are the eigenfunctions of $L_{a,\theta}$ satisfying (2.17).

Lower bound:

By definition of the weighted Bessel potential norm (2.12) and the spectral decomposition (2.20), we have

$$\begin{aligned}\|u\|_{H_w^s}^2 &= \|(I + L_{a,\theta})^{s/2} u\|_{L_w^2}^2 \\ &= \langle (I + L_{a,\theta})^s u, u \rangle_{L_w^2} \\ &= \sum_{k=1}^{\infty} (1 + \lambda_k)^s |\langle u, \varphi_k \rangle_{L_w^2}|^2 = \|u\|_{H_{w,\text{spec}}^s}^2.\end{aligned}\quad (2.21)$$

Hence, we can take $C_1 = 1$, so that

$$\|u\|_{H_{w,\text{spec}}^s} \leq \|u\|_{H_w^s}. \quad (2.22)$$

Upper bound:

We decompose the sum in (2.20) into low and high modes relative to a cutoff index N :

$$u = \sum_{k=1}^N \langle u, \varphi_k \rangle_{L_w^2} \varphi_k + \sum_{k=N+1}^{\infty} \langle u, \varphi_k \rangle_{L_w^2} \varphi_k := u_{\text{low}} + u_{\text{high}}. \quad (2.23)$$

The weighted Rellich embedding (2.16) implies

$$\|u_{\text{low}}\|_{H_w^s} \leq C(N) \|u_{\text{low}}\|_{L_w^2} \leq C(N) \|u\|_{H_{w,\text{spec}}^s}, \quad (2.24)$$

and the high modes satisfy

$$\|u_{\text{high}}\|_{H_w^s}^2 = \sum_{k=N+1}^{\infty} (1 + \lambda_k)^s |\langle u, \varphi_k \rangle_{L_w^2}|^2 \leq \|u\|_{H_{w,\text{spec}}^s}^2. \quad (2.25)$$

Combining (2.24) and (2.25), we obtain

$$\|u\|_{H_w^s} \leq C_2 \|u\|_{H_{w,\text{spec}}^s}, \quad C_2 := \sqrt{1 + [C(N)]^2}. \quad (2.26)$$

Anisotropic case:

For $\theta = (\theta_1, \dots, \theta_d)$, the operator $L_{a,\theta}$ is separable along coordinates. Using tensorization of one-dimensional spectral estimates and one-dimensional weighted Poincaré inequalities, the constants C_1, C_2 can be chosen to depend explicitly on each θ_i and the size of the domain in the i -th coordinate. This ensures that (2.19) holds uniformly for anisotropic weights.

Combining (2.22) and (2.26), we conclude the equivalence

$$C_1 \|u\|_{H_{w,\text{spec}}^s} \leq \|u\|_{H_w^s} \leq C_2 \|u\|_{H_{w,\text{spec}}^s}, \quad \forall u \in H_w^s(\Omega), \quad (2.27)$$

as claimed. $\square$

3. Weighted Poincaré Inequality and Asymptotic Analysis of C_θ

3.1. Weighted Poincaré Inequality

We establish a *weighted Poincaré inequality*, fundamental for coercivity of the bilinear form $\mathfrak{a}_{a,\theta}(\cdot, \cdot)$ in $H_{0,\theta}^1(\Omega)$.

Lemma 3.1 (Weighted Poincaré Inequality). *Let $\Omega \subset \mathbb{R}^n$ be a bounded Lipschitz domain, and let $\theta \in L^\infty(\Omega)$ satisfy $\theta(x) \geq \theta_0 \geq 0$ almost everywhere. Then there exists a constant $C_P = C_P(\Omega, \theta_0) > 0$ such that*

$$\|v\|_{L^2(\Omega)} \leq C_P \|\nabla v\|_{L^2(\Omega)}, \quad \forall v \in H_{0,\theta}^1(\Omega). \quad (3.1)$$

Furthermore, if $\theta(x) \geq \theta_0 > 0$, the weighted norm satisfies

$$\|v\|_{H_\theta^1} \leq \max\{1, C_P\} \|\nabla v\|_{L^2(\Omega)}, \quad \forall v \in H_{0,\theta}^1(\Omega). \quad (3.2)$$

Proof. Since Ω is bounded and Lipschitz, the standard Poincaré inequality ensures the existence of $C'_P = C'_P(\Omega) > 0$ such that

$$\|v\|_{L^2(\Omega)} \leq C'_P \|\nabla v\|_{L^2(\Omega)}, \quad \forall v \in H_0^1(\Omega). \quad (3.3)$$

By definition, $H_{0,\theta}^1(\Omega) \subset H_0^1(\Omega)$, so (3.3) applies to all $v \in H_{0,\theta}^1(\Omega)$.

For $v \in H_{0,\theta}^1(\Omega)$, define the weighted norm

$$\|v\|_{H_\theta^1}^2 := \|\nabla v\|_{L^2(\Omega)}^2 + \left\| \sqrt{\theta} v \right\|_{L^2(\Omega)}^2. \quad (3.4)$$

Since $\theta(x) \geq \theta_0 > 0$,

$$\left\| \sqrt{\theta} v \right\|_{L^2(\Omega)}^2 \geq \theta_0 \|v\|_{L^2(\Omega)}^2. \quad (3.5)$$

Combining (3.3) and (3.5), we have

$$\|v\|_{L^2(\Omega)} \leq C'_P \|\nabla v\|_{L^2(\Omega)} \leq \frac{C'_P}{\sqrt{\theta_0}} \left\| \sqrt{\theta} v \right\|_{L^2(\Omega)}. \quad (3.6)$$

From (3.4) and (3.5),

$$\|v\|_{H_\theta^1}^2 \geq \|\nabla v\|_{L^2(\Omega)}^2 + \theta_0 \|v\|_{L^2(\Omega)}^2 \geq \left(1 + \theta_0 (C'_P)^2\right) \|\nabla v\|_{L^2(\Omega)}^2. \quad (3.7)$$

Hence,

$$\|\nabla v\|_{L^2(\Omega)} \leq \frac{1}{\sqrt{1 + \theta_0 (C'_P)^2}} \|v\|_{H_\theta^1}. \quad (3.8)$$

Conversely, using $\theta(x) \leq \|\theta\|_{L^\infty(\Omega)}$,

$$\|v\|_{H_\theta^1}^2 \leq \|\nabla v\|_{L^2(\Omega)}^2 + \|\theta\|_{L^\infty(\Omega)} \|v\|_{L^2(\Omega)}^2 \leq \left(1 + \|\theta\|_{L^\infty(\Omega)} (C'_P)^2\right) \|\nabla v\|_{L^2(\Omega)}^2. \quad (3.9)$$

Defining

$$C_P := \sqrt{1 + \|\theta\|_{L^\infty(\Omega)} (C'_P)^2}, \quad (3.10)$$

we conclude the weighted Poincaré inequality (3.1)–(3.2). $\square$

3.2. Asymptotic Analysis of the Weighted Constant C_θ

We now study the asymptotic behavior of the constant C_θ in Lemma 3.1 as the anisotropic weights $\theta = (\theta_1, \dots, \theta_d)$ approach degenerate limits ($\theta_i \rightarrow 0^+$).

Proposition 3.2 (Asymptotic Behavior of C_θ). *Let $\Omega = \prod_{i=1}^d (0, R_i)$ be a rectangular domain, and $\theta_i \in (0, 2)$ for $i = 1, \dots, d$. Denote by C_θ the optimal constant in the weighted Poincaré inequality*

$$\|v\|_{L^2(\Omega)} \leq C_\theta \left(\sum_{i=1}^d \int_{\Omega} |x_i|^{\theta_i} |\partial_{x_i} v|^2 dx \right)^{1/2}, \quad \forall v \in H_{0,\theta}^1(\Omega). \quad (3.11)$$

Then there exist constants $c_0, C_0 > 0$, depending only on the aspect ratios R_i , such that

$$c_0 \max_i \theta_i^{-1/2} \leq C_\theta \leq C_0 \max_i \theta_i^{-1/2}, \quad \text{as } \min_i \theta_i \rightarrow 0^+. \quad (3.12)$$

Proof. Consider the one-dimensional weighted inequality

$$\int_0^{R_i} |v(x_i)|^2 dx_i \leq C_{\theta_i}^2 \int_0^{R_i} |x_i|^{\theta_i} |v'(x_i)|^2 dx_i, \quad v(0) = 0. \quad (3.13)$$

A standard Hardy-type inequality (see [?]) implies

$$C_{\theta_i} \sim \theta_i^{-1/2}, \quad \theta_i \rightarrow 0^+. \quad (3.14)$$

For the rectangular domain $\Omega = \prod_i (0, R_i)$, the multi-dimensional weighted Poincaré inequality can be obtained by tensorizing the one-dimensional estimates:

$$\|v\|_{L^2(\Omega)}^2 \leq \sum_{i=1}^d \int_0^{R_i} |v(x_i)|^2 dx_i \prod_{j \neq i} R_j \leq \sum_{i=1}^d R_{\text{prod}}^{(i)} C_{\theta_i}^2 \int_{\Omega} |x_i|^{\theta_i} |\partial_{x_i} v|^2 dx, \quad (3.15)$$

where $R_{\text{prod}}^{(i)} = \prod_{j \neq i} R_j$. Taking the maximum over i yields

$$C_{\theta}^2 \leq \max_i R_{\text{prod}}^{(i)} C_{\theta_i}^2 \sim \max_i \theta_i^{-1}, \quad \theta_i \rightarrow 0^+. \quad (3.16)$$

To see that the scaling $\theta_i^{-1/2}$ is sharp, consider test functions depending only on the most degenerate coordinate x_{i_0} with $\theta_{i_0} = \min_i \theta_i$. Then the weighted Poincaré inequality reduces to (3.13), giving

$$C_{\theta} \geq C_{\theta_{i_0}} \sim \theta_{i_0}^{-1/2}. \quad (3.17)$$

Combining (3.16) and (3.17) establishes the asymptotic estimate (3.12). $\square$

Remark 3.3. This analysis shows explicitly that the weighted Poincaré constant C_{θ} blows up as $\theta_i \rightarrow 0$, with a precise $\theta_i^{-1/2}$ scaling. Such estimates are essential for understanding the coercivity of degenerate bilinear forms and for bounding constants in spectral estimates for SDOs in highly anisotropic regimes.

3.3. Caccioppoli-Type Estimates and Interior Regularity

A fundamental tool for establishing interior regularity of weak solutions to degenerate elliptic equations is the following *degenerate Caccioppoli inequality*. This inequality provides local control of the weighted gradient of the solution in terms of the solution itself and the right-hand side.

Lemma 3.4 (Degenerate Caccioppoli Inequality). *Let $B_{2R} \subset \Omega$ be a ball such that $\overline{B_{2R}} \subset \Omega$. Suppose $u \in H_{loc}^1(\Omega)$ is a weak solution to*

$$L_{a,\theta} u = f \quad \text{in } B_{2R}, \quad (3.18)$$

where the degenerate elliptic operator $L_{a,\theta}$ is defined by

$$L_{a,\theta} u = - \sum_{i=1}^d \partial_{x_i} \left(|x_i - a_i|^{\theta_i} \partial_{x_i} u \right), \quad (3.19)$$

with $\theta_i \geq 0$ and $a_i \in \mathbb{R}$ for each $i = 1, \dots, d$. Assume $f \in L^2(B_{2R})$. Then, for any cutoff function $\eta \in C_c^\infty(B_{2R})$ with $0 \leq \eta \leq 1$ and $\eta \equiv 1$ on B_R , the following inequality holds:

$$\begin{aligned} \sum_{i=1}^d \int_{B_R} |x_i - a_i|^{\theta_i} |\partial_{x_i} u|^2 dx &\leq C \left(\frac{1}{R^2} \int_{B_{2R}} |u|^2 dx \right. \\ &\quad + R^2 \int_{B_{2R}} |f|^2 dx \\ &\quad \left. + \sum_{i=1}^d \int_{B_{2R}} |x_i - a_i|^{\theta_i} |u|^2 |\partial_{x_i} \eta|^2 dx \right). \end{aligned} \quad (3.20)$$

where $C = C(d, \theta_{\max})$ and $\theta_{\max} = \max_{1 \leq i \leq d} \theta_i$.

Proof. The proof is structured in:

Multiply (3.18) by $u\eta^2$ and integrate over B_{2R} :

$$\int_{B_{2R}} L_{a,\theta} u \cdot u\eta^2 dx = \int_{B_{2R}} f u\eta^2 dx. \quad (3.21)$$

Substituting the expression for $L_{a,\theta} u$ from (3.19), we obtain

$$-\sum_{i=1}^d \int_{B_{2R}} \partial_{x_i} (|x_i - a_i|^{\theta_i} \partial_{x_i} u) u\eta^2 dx = \int_{B_{2R}} f u\eta^2 dx. \quad (3.22)$$

Integrate by parts in the left-hand side of (3.22):

$$\sum_{i=1}^d \int_{B_{2R}} |x_i - a_i|^{\theta_i} \partial_{x_i} u \cdot \partial_{x_i} (u\eta^2) dx = \int_{B_{2R}} f u\eta^2 dx. \quad (3.23)$$

Expanding $\partial_{x_i} (u\eta^2) = \eta^2 \partial_{x_i} u + 2\eta u \partial_{x_i} \eta$, we get:

$$\sum_{i=1}^d \int_{B_{2R}} |x_i - a_i|^{\theta_i} |\partial_{x_i} u|^2 \eta^2 dx + 2 \sum_{i=1}^d \int_{B_{2R}} |x_i - a_i|^{\theta_i} \partial_{x_i} u \cdot u\eta \partial_{x_i} \eta dx = \int_{B_{2R}} f u\eta^2 dx. \quad (3.24)$$

For each i , apply Young's inequality to the cross term in (3.24):

$$2 \int_{B_{2R}} |x_i - a_i|^{\theta_i} \partial_{x_i} u \cdot u\eta \partial_{x_i} \eta dx \leq \frac{1}{2} \int_{B_{2R}} |x_i - a_i|^{\theta_i} |\partial_{x_i} u|^2 \eta^2 dx + 2 \int_{B_{2R}} |x_i - a_i|^{\theta_i} |u|^2 |\partial_{x_i} \eta|^2 dx. \quad (3.25)$$

Absorbing the first term on the right-hand side into the left-hand side of (3.24), we obtain

$$\frac{1}{2} \sum_{i=1}^d \int_{B_{2R}} |x_i - a_i|^{\theta_i} |\partial_{x_i} u|^2 \eta^2 dx \leq \left| \int_{B_{2R}} f u\eta^2 dx \right| + 2 \sum_{i=1}^d \int_{B_{2R}} |x_i - a_i|^{\theta_i} |u|^2 |\partial_{x_i} \eta|^2 dx. \quad (3.26)$$

For the first term on the right-hand side of (3.26), use the Cauchy-Schwarz inequality:

$$\left| \int_{B_{2R}} f u\eta^2 dx \right| \leq \left(\int_{B_{2R}} |f|^2 dx \right)^{1/2} \left(\int_{B_{2R}} |u|^2 \eta^4 dx \right)^{1/2}. \quad (3.27)$$

Since $\eta \leq 1$ and $\text{supp}(\eta) \subset B_{2R}$, we have

$$\left(\int_{B_{2R}} |u|^2 \eta^4 dx \right)^{1/2} \leq \left(\int_{B_{2R}} |u|^2 dx \right)^{1/2}. \quad (3.28)$$

Thus,

$$\left| \int_{B_{2R}} f u\eta^2 dx \right| \leq \left(\int_{B_{2R}} |f|^2 dx \right)^{1/2} \left(\int_{B_{2R}} |u|^2 dx \right)^{1/2}. \quad (3.29)$$

Since $\eta \equiv 1$ on B_R , the left-hand side of (3.26) dominates the integral over B_R :

$$\sum_{i=1}^d \int_{B_R} |x_i - a_i|^{\theta_i} |\partial_{x_i} u|^2 dx \leq \sum_{i=1}^d \int_{B_{2R}} |x_i - a_i|^{\theta_i} |\partial_{x_i} u|^2 \eta^2 dx. \quad (3.30)$$

Combining (3.26), (3.29), and (3.30), and using the scaling R^2 for the f -term and $1/R^2$ for the u -term, we obtain:

$$\begin{aligned} \sum_{i=1}^d \int_{B_R} |x_i - a_i|^{\theta_i} |\partial_{x_i} u|^2 dx &\leq C \left(\frac{1}{R^2} \int_{B_{2R}} |u|^2 dx \right. \\ &\quad + R^2 \int_{B_{2R}} |f|^2 dx \\ &\quad \left. + \sum_{i=1}^d \int_{B_{2R}} |x_i - a_i|^{\theta_i} |u|^2 |\partial_{x_i} \eta|^2 dx \right). \end{aligned} \quad (3.31)$$

This completes the proof. $\square$

3.4. Remarks

- The constant C in (3.20) depends on the dimension d and the maximum degeneracy exponent $\theta_{\max}$, but not on R .
- If $\theta_i = 0$ for all i , the inequality reduces to the classical Caccioppoli inequality for non-degenerate elliptic equations.
- The separable structure of the weight $|x_i - a_i|^{\theta_i}$ allows for coordinate-wise estimates, simplifying the analysis.
- The result in (3.20) is a key ingredient in proving Hölder continuity or fractional regularity for solutions of (3.18).

3.5. Fractional Regularity Theory: Statement and Proof

We present a comprehensive fractional regularity result for spectral degeneracy operators (SDOs).

Theorem 3.5 (Fractional Regularity for Spectral Degeneracy Operators). *Let $\Omega \subset \mathbb{R}^d$ be a bounded Lipschitz domain, $a \in \Omega$, and $\theta = (\theta_1, \dots, \theta_d) \in [1, 2)^d$. Suppose $f \in H^s(\Omega)$ for some $s \in [0, 1]$, and let $u \in H_{0,\theta}^1(\Omega)$ be the weak solution of*

$$L_{a,\theta} u = f \quad \text{in } \Omega. \quad (3.32)$$

Then there exists $\delta = \delta(\theta, d) \in (0, 1]$ and a constant $C > 0$ (depending on Ω , a , θ , and s) such that

$$u \in H_{w_{a,\theta}}^{1+\min\{s,\delta\}}(\Omega), \quad \|u\|_{H_{w_{a,\theta}}^{1+\min\{s,\delta\}}(\Omega)} \leq C \left(\|f\|_{H^s(\Omega)} + \|u\|_{L^2(\Omega)} \right). \quad (3.33)$$

More generally, for any $\sigma \in (0, \min\{s, \delta\}]$, we have

$$u \in B_{2,2,w_{a,\theta}}^{1+\sigma}(\Omega), \quad \|u\|_{B_{2,2,w_{a,\theta}}^{1+\sigma}(\Omega)} \leq C \left(\|f\|_{H^s(\Omega)} + \|u\|_{L^2(\Omega)} \right). \quad (3.34)$$

Proof. The proof proceeds in three main steps, with explicit formulas.

1. Localization and Flattening. Cover Ω with finitely many coordinate charts $\{U_j\}$ and choose a partition of unity $\{\eta_j\}$ subordinate to this cover. Near a degeneracy center a , we translate coordinates so that a corresponds to the origin. In each chart, the local problem reads

$$L_{0,\theta}^{(j)}(\eta_j u) = \eta_j f + \sum_{i=1}^d [L_{a,\theta}, \eta_j] u, \quad (3.35)$$

where $[L_{a,\theta}, \eta_j]$ denotes the commutator.

2. One-Dimensional Fractional Regularity. Consider the one-dimensional model operator

$$\mathcal{L}_i v := -\frac{d}{dx_i} \left(|x_i|^{\theta_i} \frac{dv}{dx_i} \right), \quad x_i \in (-R, R), \quad (3.36)$$

with Dirichlet boundary conditions. Standard singular Sturm–Liouville theory [18] implies that if $g_i \in H^s(-R, R)$, then the solution v satisfies

$$v \in H_{|x_i|^{\theta_i}}^{1+\delta_i}(-R, R), \quad \|v\|_{H_{|x_i|^{\theta_i}}^{1+\delta_i}} \leq C_i \|g_i\|_{H^s(-R, R)}, \quad (3.37)$$

for some $\delta_i = \delta_i(\theta_i) \in (0, 1]$. Here $H_{|x_i|^{\theta_i}}^{1+\delta_i}$ denotes the weighted Bessel space along coordinate x_i .

3. Tensorization and Interpolation. For the full d -dimensional operator, we have

$$L_{0,\theta} = \sum_{i=1}^d \mathcal{L}_i. \quad (3.38)$$

Tensorizing (3.37) along each coordinate and applying the real interpolation method gives a fractional gain

$$u \in H_{w_{0,\theta}}^{1+\delta}(Q), \quad \delta := \min_i \delta_i, \quad (3.39)$$

for each local cube Q . Using the Caccioppoli inequality (?? 3.4) to control lower-order derivatives,

$$\sum_{i=1}^d \int_Q |x_i|^{\theta_i} |\partial_{x_i} u|^2 dx \leq C \left(\int_{2Q} |u|^2 dx + \int_{2Q} |f|^2 dx \right), \quad (3.40)$$

and summing over the partition of unity, we obtain the global estimate (3.33).

Finally, embedding the weighted Bessel spaces into weighted Besov spaces using standard theory [13] yields (3.34). The constants depend explicitly on θ_i , the Lipschitz character of Ω , and the localization radii R_j of the coordinate charts. $\square$

4. Deep SDO Networks and Spectral Stability Analysis

We investigate architectures that compose multiple SDO-based layers. In continuous form, an SDO layer maps an input field u to

$$\mathcal{S}_{a,\theta,W,b}(u) := \sigma \left(L_{a,\theta}^{-1}(Wu + b) \right), \quad (4.1)$$

where W is a bounded linear operator (e.g., convolution or spectral projection), b is a bias term, and σ is a Lipschitz activation applied pointwise.

4.1. Compositional Structure and Modewise Action

Using the spectral decomposition $\{\varphi_k, \lambda_k\}$ of $L_{a,\theta}$, the layer operation reads

$$L_{a,\theta}^{-1}(Wu + b) = \sum_{k \geq 1} \frac{\langle Wu + b, \varphi_k \rangle_{L^2}}{\lambda_k} \varphi_k. \quad (4.2)$$

Hence, each mode k is scaled by λ_k^{-1} and receives a coefficient determined by the projection $\langle Wu + b, \varphi_k \rangle$. Composing L^{-1} blocks across multiple layers produces multiplicative mode gains; controlling these gains under parameter perturbations is essential for stability.

4.2. Perturbation Theory and Kato-Type Estimates

Let (a, θ) and $(\tilde{a}, \tilde{\theta})$ be two parameter tuples with

$$\delta := \|a - \tilde{a}\| + \|\theta - \tilde{\theta}\| \ll 1. \quad (4.3)$$

Standard analytic perturbation theory for self-adjoint operators ensures that simple eigenvalues λ_k depend analytically on smooth parameter variations [10]. For degenerate weights, eigenpairs vary continuously, and first-order perturbations can be captured by resolvent differences.

Proposition 4.1 (Resolvent Perturbation Bound). *For ζ in the resolvent set of $L_{a,\theta}$, we have*

$$\|(L_{a,\theta} - \zeta)^{-1} - (L_{\tilde{a},\tilde{\theta}} - \zeta)^{-1}\|_{L^2 \rightarrow L^2} \leq C |\zeta|^{-2} \delta, \quad (4.4)$$

where C depends on spectral gaps and domain geometry.

Proof. We use the resolvent identity:

$$(L_{a,\theta} - \zeta)^{-1} - (L_{\tilde{a},\tilde{\theta}} - \zeta)^{-1} = (L_{a,\theta} - \zeta)^{-1} (L_{\tilde{a},\tilde{\theta}} - L_{a,\theta}) (L_{\tilde{a},\tilde{\theta}} - \zeta)^{-1}. \quad (4.5)$$

Since $(L_{\tilde{a},\tilde{\theta}} - L_{a,\theta})$ is a multiplication-differential operator, we have

$$\|L_{\tilde{a},\tilde{\theta}} - L_{a,\theta}\|_{H_\theta^1 \rightarrow H_\theta^{-1}} \leq C \delta. \quad (4.6)$$

Combining (4.5)–(4.6) and the resolvent bounds $\|(L_{a,\theta} - \zeta)^{-1}\| \lesssim |\zeta|^{-1}$ yields (4.4). $\square$

4.3. Deep Spectral Stability Theorem

Consider a network with L SDO layers:

$$\mathcal{N}_L(u) := \mathcal{S}_{a_L, \theta_L, W_L, b_L} \circ \cdots \circ \mathcal{S}_{a_1, \theta_1, W_1, b_1}(u). \quad (4.7)$$

Theorem 4.2 (Deep Spectral Stability of SDO Networks). *Let each W_ℓ be bounded $L^2 \rightarrow L^2$, σ Lipschitz with constant L_σ , and assume spectral gap bounds*

$$\lambda_k^{(\ell)} \geq \lambda_* > 0, \quad k = 1, \dots, K, \quad \ell = 1, \dots, L. \quad (4.8)$$

For perturbed parameters $(\tilde{a}_\ell, \tilde{\theta}_\ell)$ with

$$\delta := \max_\ell (\|a_\ell - \tilde{a}_\ell\| + \|\theta_\ell - \tilde{\theta}_\ell\|) \ll 1, \quad (4.9)$$

the k -th spectral coefficient at network output satisfies

$$|\alpha_k^{\text{out}}(u) - \tilde{\alpha}_k^{\text{out}}(u)| \leq C_L (\delta + \delta \log(1/\delta)) \|u\|_{L^2}, \quad k \leq K, \quad (4.10)$$

where C_L depends polynomially on L , L_σ , $\sup_\ell \|W_\ell\|$, and $1/\lambda_*$.

Proof. **Base case** ($L = 1$). The k -th output coefficient before activation is

$$\beta_k = \frac{\langle W_1 u + b_1, \varphi_k^{(1)} \rangle}{\lambda_k^{(1)}}, \quad \tilde{\beta}_k = \frac{\langle W_1 u + b_1, \tilde{\varphi}_k^{(1)} \rangle}{\tilde{\lambda}_k^{(1)}}. \quad (4.11)$$

Using Proposition 4.1, we have

$$|\beta_k - \tilde{\beta}_k| \leq C_1 (\delta + \delta \log(1/\delta)) \|u\|_{L^2}. \quad (4.12)$$

Assume the claim holds for $L - 1$ layers. Let

$$v = \mathcal{N}_{L-1}(u), \quad \tilde{v} = \tilde{\mathcal{N}}_{L-1}(u), \quad (4.13)$$

then

$$\|\alpha_k(v) - \alpha_k(\tilde{v})\| \leq C_{L-1} (\delta + \delta \log(1/\delta)) \|u\|_{L^2}. \quad (4.14)$$

Applying the $L = 1$ argument to the final layer with Lipschitz σ gives

$$|\alpha_k^{\text{out}}(u) - \tilde{\alpha}_k^{\text{out}}(u)| \leq L_\sigma C_{L-1}(\delta + \delta \log(1/\delta)) \|u\|_{L^2} + C_1(\delta + \delta \log(1/\delta)) \|v\|_{L^2}. \quad (4.15)$$

Iterating and absorbing constants proves (4.10). $\square$

Corollary 4.3 (Training Stability). *Let $\{(a_\ell^{(n)}, \theta_\ell^{(n)}, W_\ell^{(n)}, b_\ell^{(n)})\}_{\ell=1}^L$ be the network parameters at training iteration n , and assume that*

$$\delta_n := \max_\ell (\|a_\ell^{(n+1)} - a_\ell^{(n)}\| + \|\theta_\ell^{(n+1)} - \theta_\ell^{(n)}\|) \ll 1. \quad (4.16)$$

Then, for any input $u \in L^2(\Omega)$, the spectral coefficients of the network output satisfy

$$|\alpha_k^{(n+1)}(u) - \alpha_k^{(n)}(u)| \leq C_L(\delta_n + \delta_n \log(1/\delta_n)) \|u\|_{L^2}, \quad k \leq K, \quad (4.17)$$

where C_L depends polynomially on L , L_σ , $\sup_\ell \|W_\ell\|$, and the inverse spectral gap $1/\lambda_$.*

Proof. By Theorem 4.2, the difference between the outputs of two networks with parameters differing by δ_n satisfies

$$|\alpha_k^{\text{out}}(u) - \tilde{\alpha}_k^{\text{out}}(u)| \leq C_L(\delta_n + \delta_n \log(1/\delta_n)) \|u\|_{L^2}. \quad (4.18)$$

Here, we identify the perturbed network as the one at iteration $n + 1$ and the nominal network as iteration n .

Inductively, for each layer ℓ , the spectral difference in intermediate outputs v_ℓ satisfies

$$\|v_\ell^{(n+1)} - v_\ell^{(n)}\|_{L^2} \leq L_\sigma \|v_{\ell-1}^{(n+1)} - v_{\ell-1}^{(n)}\|_{L^2} + C(\delta_n + \delta_n \log(1/\delta_n)) \|v_{\ell-1}^{(n)}\|_{L^2}, \quad (4.19)$$

where $v_0 := u$ and L_σ is the Lipschitz constant of the activation. Iterating (4.19) over $\ell = 1, \dots, L$ yields (4.17).

Therefore, small parameter updates $\delta_n \ll 1$ guarantee that the spectral coefficients vary smoothly, and the accumulation of errors across layers remains controlled, preventing catastrophic spectral drift during training. $\square$

5. Variational Limits: Γ -Convergence of Degenerate Energies

We formulate an energy functional naturally associated with SDO parameterizations and investigate its variational limit as resolution increases.

5.1. Energy Functional and Discretization Framework

Let $\mathcal{P}_N$ denote a family of finite-dimensional parameterizations (e.g., spectral truncation at frequency N or spatial discretization with mesh size $h_N \downarrow 0$) of SDO networks. For parameters $\pi \in \mathcal{P}_N$, we define the degenerate energy functional

$$\mathcal{E}_N(u; \pi) := \frac{1}{2} \int_\Omega \sum_{i=1}^d |x_i - a_i|^{\theta_i} |\partial_{x_i} u|^2 dx + \frac{\kappa_N}{2} \|u - \mathcal{R}_N(\pi)\|_{L^2}^2, \quad (5.1)$$

where $\mathcal{R}_N(\pi)$ denotes the SDO-network reconstruction at resolution N and $\kappa_N \rightarrow \kappa \in [0, \infty)$ is a fidelity parameter. The first term represents the anisotropic degenerate Dirichlet energy, and the second term enforces agreement with the network reconstruction.

5.2. Γ -Convergence Result

Theorem 5.1 (Γ -Convergence to a Degenerate Variational Limit). Assume that $\mathcal{P}_N$ yields reconstructions $\mathcal{R}_N(\pi)$ that are uniformly bounded in $L^2(\Omega)$ and converge (along subsequences) to $\mathcal{R}(\pi)$ in L^2 , and suppose $\kappa_N \rightarrow \kappa \geq 0$. Then, as $N \rightarrow \infty$, the sequence of functionals $\mathcal{E}_N(\cdot; \pi)$ Γ -converges in $L^2(\Omega)$ to

$$\mathcal{E}_\infty(u; \pi) := \frac{1}{2} \int_{\Omega} \sum_{i=1}^d |x_i - a_i|^{\theta_i} |\partial_{x_i} u|^2 dx + \frac{\kappa}{2} \|u - \mathcal{R}(\pi)\|_{L^2}^2, \quad (5.2)$$

with respect to the strong L^2 topology. Consequently, minimizers u_N of $\mathcal{E}_N(\cdot; \pi)$ converge (along subsequences) to minimizers of $\mathcal{E}_\infty(\cdot; \pi)$.

Proof. We verify the conditions of Γ -convergence: liminf and limsup inequalities.

Let $u_N \rightarrow u$ strongly in $L^2(\Omega)$. Assume $\{u_N\}$ is bounded in $H_{0,\theta}^1(\Omega)$. By lower semicontinuity of the weighted Dirichlet integral [15], we have

$$\liminf_{N \rightarrow \infty} \frac{1}{2} \int_{\Omega} \sum_{i=1}^d |x_i - a_i|^{\theta_i} |\partial_{x_i} u_N|^2 dx \geq \frac{1}{2} \int_{\Omega} \sum_{i=1}^d |x_i - a_i|^{\theta_i} |\partial_{x_i} u|^2 dx. \quad (5.3)$$

For the fidelity term, strong convergence of u_N and convergence of reconstructions give

$$\liminf_{N \rightarrow \infty} \frac{\kappa_N}{2} \|u_N - \mathcal{R}_N(\pi)\|_{L^2}^2 \geq \frac{\kappa}{2} \|u - \mathcal{R}(\pi)\|_{L^2}^2. \quad (5.4)$$

Combining (5.3) and (5.4) yields the liminf inequality:

$$\liminf_{N \rightarrow \infty} \mathcal{E}_N(u_N; \pi) \geq \mathcal{E}_\infty(u; \pi). \quad (5.5)$$

Given $u \in L^2(\Omega)$ with $\mathcal{E}_\infty(u; \pi) < \infty$, there exists a sequence $u_N \in C_c^\infty(\Omega)$ such that

$$u_N \rightarrow u \quad \text{in } H_{0,\theta}^1(\Omega) \cap L^2(\Omega), \quad (5.6)$$

by density of smooth functions in weighted Sobolev spaces. Then

$$\lim_{N \rightarrow \infty} \frac{1}{2} \int_{\Omega} \sum_{i=1}^d |x_i - a_i|^{\theta_i} |\partial_{x_i} u_N|^2 dx = \frac{1}{2} \int_{\Omega} \sum_{i=1}^d |x_i - a_i|^{\theta_i} |\partial_{x_i} u|^2 dx, \quad (5.7)$$

$$\lim_{N \rightarrow \infty} \frac{\kappa_N}{2} \|u_N - \mathcal{R}_N(\pi)\|_{L^2}^2 = \frac{\kappa}{2} \|u - \mathcal{R}(\pi)\|_{L^2}^2, \quad (5.8)$$

and summing gives

$$\limsup_{N \rightarrow \infty} \mathcal{E}_N(u_N; \pi) = \mathcal{E}_\infty(u; \pi). \quad (5.9)$$

Let u_N be a minimizer of $\mathcal{E}_N(\cdot; \pi)$. By the coercivity of the weighted Dirichlet term and boundedness of the fidelity term, $\{u_N\}$ is bounded in $H_{0,\theta}^1(\Omega)$:

$$\frac{1}{2} \int_{\Omega} \sum_{i=1}^d |x_i - a_i|^{\theta_i} |\partial_{x_i} u_N|^2 dx + \frac{\kappa_N}{2} \|u_N - \mathcal{R}_N(\pi)\|_{L^2}^2 \leq C. \quad (5.10)$$

Hence, there exists a subsequence u_{N_j} converging weakly in $H_{0,\theta}^1(\Omega)$ and strongly in $L^2(\Omega)$ to some u_∞ . Using the liminf inequality (5.5), u_∞ is a minimizer of $\mathcal{E}_\infty(\cdot; \pi)$.

This completes the proof of Γ -convergence and convergence of minimizers. $\square$

Remark 5.2. The Γ -limit (5.2) shows that as model resolution increases, the discrete SDO learning objective converges variationally to a continuum degenerate energy. This provides a rigorous justification for stability, interpretability, and the approximation of degenerate centers by minimizing the discrete network loss.

6. Results

This work establishes several fundamental mathematical advances with direct implications for physics-informed turbulence modeling:

1. **Complete Functional Analytic Framework:** We developed a rigorous foundation for Spectral Degeneracy Operators, including:
 - Characterization of weighted Sobolev spaces $H_{0,\theta}^1(\Omega)$ with explicit Poincaré constants (Lemma 2.3) and norm equivalence (Corollary 2.4)
 - Spectral theory for $L_{a,\theta}$ proving self-adjointness, compact resolvent, and complete orthonormal eigenbasis (Section 2.3)
 - Equivalence between spectral and Bessel potential norms in weighted fractional spaces (Section 2.4)
2. **Fractional Regularity Theory:** Theorem 3.5 demonstrates that solutions of $L_{a,\theta}u = f$ gain fractional derivatives, with $u \in H_{w_{a,\theta}}^{1+\min\{s,\delta\}}(\Omega)$ for $f \in H^s(\Omega)$. This result, established via degenerate Caccioppoli estimates and tensorized Sturm-Liouville theory, provides rigorous error control for spectral approximations.
3. **Explicit Asymptotics:** Proposition 3.2 reveals the precise $\theta_i^{-1/2}$ scaling of weighted Poincaré constants as $\theta_i \rightarrow 0^+$, offering crucial control over coercivity in highly anisotropic regimes.
4. **Deep Spectral Stability:** Theorem 4.2 guarantees bounded mode amplitude variations of order $\mathcal{O}(\delta \log(1/\delta))$ under parameter perturbations in deep SDO networks. The practical consequence (Corollary 4.3) ensures training stability by preventing catastrophic spectral drift.
5. **Variational Convergence:** Theorem 5.1 establishes Γ -convergence of discrete SDO energies to continuum limits, connecting neural network training with physically meaningful degenerate energy minimization.
6. **Strengthened Classical Results:** We provided complete proofs with explicit constants for universality of divergence-free closures and inverse calibration stability, improving convergence rates under spectral gap assumptions.

7. Conclusions

This work establishes a comprehensive mathematical foundation for Spectral Degeneracy Operators, bridging degenerate PDE theory with modern physics-informed neural networks for turbulence modeling. Our three principal contributions—fractional regularity theory, deep spectral stability, and variational Γ -convergence—provide rigorous guarantees that address fundamental challenges in neural operator architectures:

The **fractional regularity theory** resolves fundamental questions about solution smoothness in weighted anisotropic spaces, enabling rigorous error control for spectral approximations. The **deep spectral stability theorem** ensures that SDO-based networks maintain consistent spectral characteristics during training, preventing the mode distortion that often plagues deep neural operators. The **Γ -convergence framework** establishes a variational interpretation of SDO learning, connecting discrete network objectives with continuum energy principles.

These theoretical advances position SDOs as a mathematically principled alternative to traditional turbulence closures, combining the expressive power of neural networks with the interpretability and robustness of physically-grounded operator theory. The framework ensures that learned degeneracy centers converge to meaningful turbulent structures while maintaining training stability and approximation guarantees.

Future research directions include:

- **Stochastic extensions** for uncertainty quantification in degeneracy parameters
- **Numerical analysis** of discretization schemes and computational complexity
- **Experimental validation** on canonical turbulent flow datasets
- **Extensions to non-separable weights** and more general degenerate structures
- **Multi-scale implementations** combining SDOs with traditional turbulence models
- **Connections with operator learning** to establish universal approximation in stronger topologies

The mathematical framework developed herein provides a solid foundation for these investigations while ensuring that SDO-based architectures maintain physical interpretability, training stability, and rigorous mathematical guarantees across diverse turbulence modeling applications.

Acknowledgments: Santos gratefully acknowledges the support of the PPGMC Program for the Postdoctoral Scholarship PROBOL/UESC nr. 218/2025. Sales acknowledges CNPq grant 30881/2025-0.

Notation and Symbols

The following table summarizes the key mathematical notations and symbols used throughout this paper.

Mathematical Symbols and Operators

Symbol	Description
Ω	Bounded Lipschitz domain in $\mathbb{R}^d$
d	Spatial dimension
$x = (x_1, \dots, x_d)$	Spatial coordinates
$a = (a_1, \dots, a_d)$	Degeneracy center (trainable parameter)
$\theta = (\theta_1, \dots, \theta_d)$	Degeneracy exponents, $\theta_i \in [1, 2)$
$L_{a,\theta}$	Spectral Degeneracy Operator (SDO)
$D_{a,\theta}(x)$	Diagonal weight matrix: $\text{diag}(x_1 - a_1 ^{\theta_1}, \dots, x_d - a_d ^{\theta_d})$
$w_{a,\theta}(x)$	Separable weight function: $\prod_{i=1}^d x_i - a_i ^{\theta_i}$
$H_{0,\theta}^1(\Omega)$	Weighted Sobolev space with degeneracy weights
$H_w^s(\Omega)$	Weighted Bessel potential space
$B_{p,q,w}^s(\Omega)$	Weighted Besov space
$L_w^2(\Omega)$	Weighted L^2 space: $L^2(\Omega; w_{a,\theta} dx)$
λ_k, φ_k	Eigenvalues and eigenfunctions of $L_{a,\theta}$
$\mathfrak{a}_{a,\theta}(u, v)$	Bilinear form associated with $L_{a,\theta}$
$\mathcal{S}_{a,\theta,W,b}$	SDO neural network layer
$\mathcal{N}_L$	Deep SDO network with L layers
$\mathcal{E}_N, \mathcal{E}_\infty$	Discrete and continuum energy functionals
Γ -convergence	Variational convergence in the sense of De Giorgi
C_P, C_θ	Poincaré constants (weighted and unweighted)
δ	Regularity exponent or perturbation parameter
σ	Activation function in neural networks

Greek Letters

Symbol	Description	Symbol	Description
α	Spectral coefficients	β	Intermediate coefficients
γ	Generic constant	δ	Regularity/perturbation parameter
η	Cutoff function	θ	Degeneracy exponents
κ	Fidelity parameter	λ	Eigenvalues
σ	Activation function/Besov parameter	φ	Eigenfunctions
Ω	Spatial domain	ζ	Resolvent parameter

Key Constants and Parameters

Symbol	Description
$C, C_1, C_2, \dots$	Generic constants (may depend on domain, weights, etc.)
C_P	Weighted Poincaré constant
C_θ	Anisotropic weighted Poincaré constant
L_σ	Lipschitz constant of activation function
λ_*	Lower bound on eigenvalues (spectral gap)
R, R_i	Domain radii and aspect ratios
K	Spectral truncation index
N	Discretization/resolution parameter
L	Number of layers in deep network

Function Spaces

Symbol	Description
$L^2(\Omega)$	Standard Lebesgue space
$H^1(\Omega)$	Standard Sobolev space
$H^s(\Omega)$	Fractional Sobolev space
$H^1_{0,\theta}(\Omega)$	Weighted Sobolev space with degeneracy
$H^s_w(\Omega)$	Weighted Bessel potential space
$B^s_{p,q,w}(\Omega)$	Weighted Besov space
$C^\infty_c(\Omega)$	Smooth functions with compact support

References

- Chaves dos Santos, R. D., & de Oliveira Sales, J. H. (2025). Spectral Degeneracy Operators for Interpretable Turbulence Modeling. *Preprints*. <https://doi.org/10.20944/preprints202509.1997.v1>.
- Cannarsa, P., Doubova, A., & Yamamoto, M. (2024). Reconstruction of degenerate conductivity region for parabolic equations. *Inverse Problems*, 40(4), 045033. [10.1088/1361-6420/ad308a](https://doi.org/10.1088/1361-6420/ad308a).
- Xiao, M. J., Yu, T. C., Zhang, Y. S., & Yong, H. (2023). Physics-informed neural networks for the Reynolds-Averaged Navier–Stokes modeling of Rayleigh–Taylor turbulent mixing. *Computers & Fluids*, 266, 106025. <https://doi.org/10.1016/j.compfluid.2023.106025>.
- Hussein, M. S., Lesnic, D., Kamynin, V. L., & Kostin, A. B. (2020). Direct and inverse source problems for degenerate parabolic equations. *Journal of Inverse and Ill-Posed Problems*, 28(3), 425–448.
- Finzi, M., Stanton, S., Izmailov, P., & Wilson, A. G. (2020, November). Generalizing convolutional neural networks for equivariance to Lie groups on arbitrary continuous data. In *International Conference on Machine Learning* (pp. 3165–3176). PMLR.
- Li, Z., Kovachki, N., Azizzadenesheli, K., Liu, B., Bhattacharya, K., Stuart, A., & Anandkumar, A. (2020). Neural operator: Graph kernel network for partial differential equations. *arXiv preprint arXiv:2003.03485*. <https://doi.org/10.48550/arXiv.2003.03485>.
- Raissi, M., Perdikaris, P., & Karniadakis, G. E. (2019). Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear PDEs. *Journal of Computational Physics*, 378, 686–707. <https://doi.org/10.1016/j.jcp.2018.10.045>.
- Kamynin, V. L. (2018). On inverse problems for strongly degenerate parabolic equations under the integral observation condition. *Computational Mathematics and Mathematical Physics*, 58(12), 2002–2017. <https://doi.org/10.1134/S0965542518120114>.
- Beck, A. D., Flad, D. G., & Munz, C. D. (2018). Deep neural networks for data-driven turbulence models. *arXiv preprint arXiv:1806.04482*. <https://doi.org/10.1016/j.jcp.2019.108910>.
- Kato, T. (2013). *Perturbation theory for linear operators* (Vol. 132). Springer Science & Business Media.
- DiBenedetto, E. (2012). *Degenerate parabolic equations*. Springer Science & Business Media.
- Sagaut, P. (2006). *Large eddy simulation for incompressible flows: an introduction*. Springer Berlin Heidelberg.
- Triebel, H. (2006). *Theory of function spaces III*. Birkhäuser Basel.
- Oleinik, O. A., & Samokhin, V. N. (1999). *Mathematical models in boundary layer theory* (Vol. 15). CRC Press.
- Brézis, H., & Dal Maso, G. (1993). *Introduction to Γ -Convergence*. Birkhäuser Boston.

16. Cybenko, G. (1989). Approximation by superpositions of a sigmoidal function. *Mathematics of Control, Signals and Systems*, 2(4), 303–314. <https://doi.org/10.1007/BF02551274>.
17. Pope, S. B. (1988). The evolution of surfaces in turbulence. *International Journal of Engineering Science*, 26(5), 445–469. [https://doi.org/10.1016/0020-7225\(88\)90004-3](https://doi.org/10.1016/0020-7225(88)90004-3).
18. Watson, G. N. (1922). *A treatise on the theory of Bessel functions* (Vol. 3). The University Press.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.