POLITECNICO
MILANO 1863

SCUOLA DI INGEGNERIA INDUSTRIALE
E DELL'INFORMAZIONE

Executive Summary of the Thesis

On the Choice of the Entropy Function and Iterative Solver in Structure-Preserving Discontinuous Methods for the Fisher-Kolmogorov-Petrovsky-Piskunov Equation

Laurea Magistrale in Mathematical Engineering - Ingegneria Matematica

Author: Nicola Ciliberti

Advisor: Prof. Paola F. Antonietti

Co-advisor: Prof. Ilaria Perugia, Dr. Mattia Corti

Academic year: 2025-2026

1. Introduction

Neurodegenerative diseases such as Alzheimer's and Parkinson's disease are characterized by the accumulation propagation of misfolded proteins – Amyloid-β and α-Synuclein, respectively – through brain tissue [5]. A widely used mathematical model to describe such phenomenon is the **Fisher–Kolmogorov–Petrovsky–Piskunov (FKPP)** equation, a nonlinear reaction–diffusion partial differential equation (PDE) that naturally reproduces the traveling wavefront patterns observed clinically in disease progression [4].

A central challenge in the numerical approximation of the FKPP equation is the preservation of the physical bounds the relative concentration of pathological proteins. Standard discretisation methods do not guarantee in general to preserve such physical bounds, particularly in regions with sharp solution fronts.

To address this issue, Antonietti et al. [1, 2] recently introduced a **structure-preserving Local Discontinuous Galerkin (LDG)** framework based on an *entropy variable* reformulation of the mathematical model (see [3]). The key idea is to rewrite the relative unknown concentration, $c(\cdot)$ as $c = u(w)$ for a suitable transformation u, so that the discrete solution automatically satisfies the physical constraints at all times.

This thesis extends the framework of [1] by addressing two open questions left in the original work:

1. **Choice of entropy transformation.** While any transformation satisfying certain theoretical properties is admissible, different choices may exhibit substantially different numerical behaviour. We test three alternatives to the standard Fermi-Dirac function considered in [3].
2. **Robustness of the iterative solver.** Newton's method, employed to solve the nonlinear system at each time step, is sensitive to the initial guess, which can be problematic in challenging clinical settings. We propose and test hybrid iterative strategies to improve robustness.

2. Mathematical and Numerical Framework

Let $\Omega \subset \mathbb{R}^2$ be a polygonal domain with Lipschitz boundary. For a given time T, we seek $c : \Omega \times (0,T) \to \mathbb{R}$ satisfying

$$\begin{aligned} \partial_t c - \nabla \cdot (\mathbf{D}\nabla c) &= \alpha\, c(1-c) && \text{in } \Omega \times (0,T), && \text{(1a)} \\ (\mathbf{D}\nabla c) \cdot \mathbf{n} &= 0 && \text{on } \partial\Omega \times (0,T), && \text{(1b)} \\ c(\cdot,0) &= c_0 && \text{in } \Omega, && \text{(1c)} \end{aligned}$$

where $\alpha \in L^\infty(\Omega)$ is a positive reaction coefficient and $\mathbf{D}$ is a symmetric positive-definite diffusion tensor. Under the assumption $0 \le c_0 \le 1$, the solution satisfies $c \in [0,1]$ for all time, see [4].

2.1. Entropy Variable Reformulation

An *entropy function* $s : [0,1] \to \mathbb{R}$ with $s' > 0$, $s'' > 0$ leads to the transformation $u = (s')^{-1}$, defining the *entropy variable* $w = u^{-1}(c) \in \mathbb{R}$. Rewriting (1) in terms of w, the physical bounds are encoded implicitly, i.e., $c = u(w) \in [0,1]$ for all $w \in \mathbb{R}$ whenever u maps $\mathbb{R}$ into $(0,1)$.
The standard choice of [1] is the **Fermi-Dirac** function, i.e.,

$$u(w) = \frac{e^w}{1+e^w},$$

for which $s''(c) = 1/(c(1-c))$ and $u'(w) = u(w)(1-u(w))$. Beyond Fermi-Dirac, we test three alternative entropy transformations:
(1) Shannon-Boltzmann:

$$u_{\mathrm{SB}}(w) = e^w.$$

Unlike Fermi-Dirac, this is unbounded above, but $c = 1$ being a stable equilibrium of (1) prevents blow-up.
(2) Hellinger:

$$u_{\mathrm{H}}(w) = \frac{1}{2}\left(1 + \frac{w}{\sqrt{1+w^2}}\right).$$

This function maps $\mathbb{R} \to (0,1)$ and satisfies all theoretical requirements, but has $s''(c) \to \infty$ as $c \to 0$ more rapidly than Fermi-Dirac.
(3) k-Fermi-Dirac:

$$u_{k\text{-FD}}(w) = \frac{k^w}{1+k^w}, \quad k > 0,$$

a one-parameter family generalising Fermi-Dirac (recovered at $k = e$).

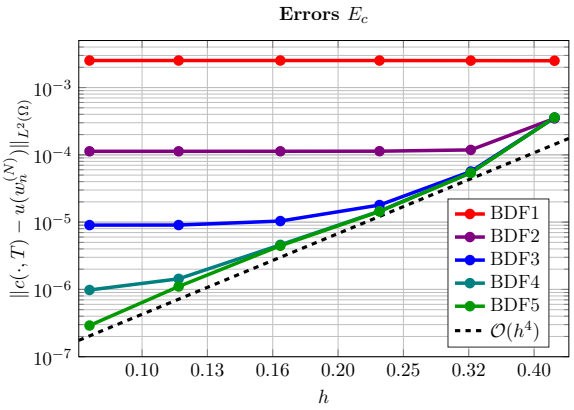

Figure 1: Fermi-Dirac error L^2

2.2. Structure-Preserving BDF–LDG Scheme

The spatial discretisation uses polygonal meshes $\mathcal{T}_h$ with piecewise polynomial spaces $\mathcal{W}_\ell(\mathcal{T}_h)$ of degree ℓ. The LDG method introduces three auxiliary flux variables $\boldsymbol{z}_h, \boldsymbol{\sigma}_h, \boldsymbol{r}_h$ to rewrite the problem 1 as a first-order system, allowing purely *volumetric* evaluation of the nonlinear operator (no surface integrals involving $u(w)$). Time integration uses Backward Differentiation Formulas (BDF) of order $\nu \in \{1,\dots,6\}$. The resulting nonlinear system at each time step is

$$\mathbf{F}(\mathbf{W}_h^{(n+1)}) = \mathbf{0}, \qquad (2)$$

where $\mathbf{W}_h^{(n+1)}$ collects the degrees of freedom of $w_h^{(n+1)}$. The penalty parameter is chosen on each space as $\eta_F = \eta_0(\ell+1)^2 h_F^{-1}$, with $\eta_0 = 1$. The time interval $(0,T)$ is also partitioned by a mesh $\mathcal{T}_\tau$ defined by the points $0 = t_0 < t_1 < \cdots < t_N = T$. For $n = 1,\dots,N$, we define the time intervals $I_n = (t_{n-1}, t_n)$.

3. Numerical Results

We consider the test on a traveling wave solution:

$$c_{\mathrm{ex}}(x,t) = \frac{1}{4}\left(1 + \tanh\left(8 - \sqrt{\tfrac{\alpha}{24 d_{\mathrm{ext}}}}\,\xi\right)\right)^2,$$
$$\xi = x - vt,$$

with $\mathbf{D} = 10^{-3}\mathbf{I}$, on $(0,3) \times (0,1)$. The plots of the L^2 error for both the Fermi-Dirac transformation and the Shannon-Boltzmann transformation are plotted in Figures 1 and 2, respectively. This solution exibits a sharp front where c transitions from 1 to 0.

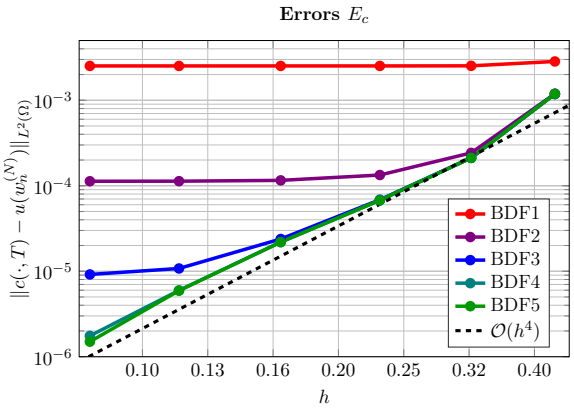

Figure 2: Shannon-Boltzmann error L^2

The Hellinger transformation fails consistently on this test. The scheme reconstructs the wavefront less steeply than the exact solution – a *numerical diffusion* artefact – causing error accumulation and eventual divergence. The failure mechanism is due to the second derivative of the entropy:

$$s''_{\mathrm{H}}(c) = \frac{1}{2c^{3/2}(1-c)^{3/2}} \xrightarrow{c\to 0} +\infty$$

more rapidly than Fermi-Dirac's $s''_{\mathrm{FD}}(c) = 1/(c(1-c))$. As $c \to 0$, the large s'' forces $u'(w) \to 0$, making the Jacobian of $\mathcal{F}$ nearly singular in the front region and degrading scheme sensitivity exactly where the solution is most critical. The **Shannon-Boltzmann** transformation, on the other hand, remains convergent on traveling wave solutions at the expected rate, confirming its theoretical validity, albeit with a constant-factor accuracy penalty relative to Fermi-Dirac. Lastly, the k**-Fermi-Dirac** transformation. For $k > e$, the transformation is steeper near $w = 0$, improving accuracy for fine meshes but inducing instability on coarse ones. No single k dominates uniformly, confirming that performance depends weakly on the slope at $w = 0$ and more on the asymptotic behaviour of s''.

4. Performance of Iterative Methods

Newton's method applied to (2) has quadratic convergence of the form:

$$\|e^{n+1}\|_{L^2} \le C\|e^n\|_{L^2}^2,$$

where $\|\cdot\|_{L^2}$ is the L^2 norm on the domain Ω and $e^n = w_h^{(n)} - w_{ex}$, where $w_h^{(n)}$ is the output of the n-th iteration and w_{ex} is the exact solution. On the other hand, its performance degrades when the initial guess $w_h^{(n)}$ (the solution from the previous time step) is far from the solution at the new step. This occurs for large time steps, low BDF orders, or complex geometries.

4.1. Warm-Start Method

We propose a **"Warm-Start"** (WS) strategy that precedes Newton iterations with a few *Richardson* (fixed-point) steps, see 1:

Algorithm 1 Warm-Start

Estimate $\delta \approx 2/\lambda_{\max}$ via Power Method
while not converged **do**
  if $n_{\mathrm{it}} < n_{\mathrm{FP}}$ (Richardson phase) **then**
    $w^{n+1} = w^n - \delta\,\mathcal{F}(w^n)$
  else if $n_{\mathrm{it}} < n_{\mathrm{FP}} + n_{\mathrm{rel}}$ (relaxed phase) **then**
    $w^{n+1} = \alpha_r\, w^n + (1-\alpha_r)\big(w^n - J^{-1}\mathcal{F}(w^n)\big)$
  else
    $w^{n+1} = w^n - J(\mathcal{F}(w^n))^{-1}\mathcal{F}(w^n)$
  end if
end while

The WS method proved itself to be more robust than the regular Newton's method, being able to deal with simulations with a large time step Δt and a perturbed initial datum.

4.2. Alternating Newton

The Alternating Newton method interleaves k Richardson iterations with one Newton step throughout the entire solve. While theoretically appealing, it incurs approximately $2\times$ the computational cost of Newton for negligible robustness gains.

5. Computational Cost and Robustness

We first investigate the computational costs. Table 1 reports the average time per time step across three methods on Test Case 1 with 300 elements, $\ell = 3$, $\Delta t = 0.5$. Warm-Start requires $\approx 2\%$ overhead over Newton (12.87 vs 12.64 s), while Alternating Newton costs $\approx 2.4\times$ more (30.01 s).
Next, we address robustness with respect to the time step Δt Table 2 shows the convergence

	Newton	WS	Alt.
Avg. it.	8.0	12.5	18.25
Avg. time (s)	12.64	12.87	30.01
Time/iter (s)	1.578	1.098	1.261

Table 1: Computational cost comparison, $\Delta t = 0.5$.

methods, Newton and Warm-Start on the traveling wave test with increasing Δt and varying polynomial degrees ℓ. The presence of "–" implies failure in reaching convergence. Warm-Start recovers convergence in almost all cases where Newton fails, demonstrating its practical value.

		Newton		Warm-Start	
Δt	ℓ	BDF1	BDF2	BDF1	BDF2
0.25	1	✓	✓	✓	✓
0.5	1	✓	✓	✓	✓
0.25	2	—	✓	✓	✓
0.5	2	—	—	✓	—

Table 2: Convergence results for large time step test.

6. Brain Mesh Simulation

A 2D brain cross-section mesh (534 elements) is used to test the scheme in an application setting, with $\ell = 1$, $\tau = 2.5 \times 10^{-2}$, and BDF order 6. Both Newton and Warm-Start propagate the protein front correctly up to $t \approx 3.95$, where the solver diverges – a sign that the failure is caused by error accumulation over long time horizons rather than by a poor initial guess. In this regime, Warm-Start provides no advantage, confirming that its benefit is localized to the initial transient of each time step.

7. Conclusions

This thesis has extended the structure-preserving BDF–LDG framework of [1] for the FKPP equation by investigating the impact of the entropy transformation and the iterative solver.

Entropy transformations. Theoretical admissibility is a necessary but *not sufficient* condition for practical performance. The Shannon-Boltzmann function yields optimal convergence rates but larger errors, attributable to the unboundedness of the exponential map. The Hellinger transformation fails systematically on sharp-front solutions due to excessive numerical diffusion, caused by the rapid divergence of $s''(c)$ as $c \to 0$. The k-Fermi-Dirac family shows that no single parameter value k performs uniformly well: the optimal choice depends on mesh size, and the asymptotic properties of s'' play a critical role. Overall, the **Fermi-Dirac transformation seems to be the most robust and accurate choice** within this framework.

Iterative methods. The Warm-Start approach significantly improves solver robustness in challenging regimes – particularly for large time steps and low polynomial degrees – at a modest additional cost of approximately 2% over Newton. The relaxed transition phase is essential: without it, Newton's method discards the benefit of the Richardson pre-iterations. The Alternating Newton method is impractical due to its $\approx 2.4\times$ overhead with only marginal gains. Neither method overcomes divergence caused by long-term error accumulation, which remains an open challenge.

Future directions. Promising extensions include: the application of the framework to three-dimensional brain meshes using the DOLFINx/Vulpes framework for tetrahedral geometries; adaptive strategies for the Richardson parameter δ; and a theoretical analysis of the basin of convergence of the Warm-Start method.

Acknowledgements

I want to thank sincerely thanks Prof. Paola F. Antonietti for proposing this project and for her guidance throughout; Prof. Ilaria Perugia for her generous hospitality in Vienna; and Dr. Mattia Corti for his invaluable day-to-day support and encouragement. The brain MRI images were provided by OASIS-3: Longitudinal Multimodal Neuroimaging: Principal Investigators: T. Benzinger, D. Marcus, J. Morris; NIH P30 AG066444, P50 AG00561, P30 NS09857781, P01 AG026276, P01 AG003991, R01 AG043434, UL1 TR000448, R01 EB009352. AV-45 doses were provided by Avid Radiopharmaceuticals, a wholly-owned subsidiary of Eli Lilly.

References

[1] Paola F. Antonietti, Mattia Corti, Sergio Gómez, and Ilaria Perugia. A structure-

preserving ldg discretization of the fisher–kolmogorov equation for modeling neurodegenerative diseases. *Mathematics and Computers in Simulation*, 241:351–366, 2026.

[2] Sergio Gómez, Ansgar Jüngel, and Ilaria Perugia. Structure-preserving local discontinuous galerkin method for nonlinear cross-diffusion systems. *arXiv preprint*, 2025.

[3] Ansgar Jüngel. The boundedness-by-entropy method for cross-diffusion systems. *Nonlinearity*, 28(6):1963–1984, 2015.

[4] S. Salsa and G. Verzini. *Partial Differential Equations in Action: From Modelling to Theory*, volume 147 of *Unitext*. Springer, Cham, fourth updated edition edition, 2022.

[5] L. Stefanis. α-Synuclein in Parkinson's Disease. *Cold Spring Harbor Perspectives in Medicine*, 2(2):a009399, 2012.