Deep Neural Network Approximation for a Nonlinear Averaged System of Integro-Differential Equations

Zurab Kiguradze

Ilia Vekua Institute of Applied Mathematics of Ivane Javakhishvili Tbilisi State University
Tbilisi, Georgia
E-mail: zurab.kiguradze@tsu.ge

Nonlinear diffusion processes with nonlocal dependence in time arise naturally in media whose effective diffusivity varies in response to accumulated energy dissipation. A prototypical setting is obtained when the diffusion coefficient depends on the integral of the squared spatial gradients of the solution. Such models appear as quasistationary reductions of Maxwell systems leading to parabolic equations in which the diffusivity evolves according to an energy functional that records the history of the solution (see, for example, [14]). In these problems, the present diffusion rate is not determined solely by the instantaneous state but depends on the cumulative dissipation over time, and the resulting dynamics are nonlocal in the temporal variable. If the magnetic field has one-dimensional two component structure, after averaging in the transverse coordinates [15], one obtains the following system of nonlinear averaged integro-differential equations:

$$\frac{\partial U}{\partial t} - \frac{\partial}{\partial x}\Big[a(S)\,\frac{\partial U}{\partial x}\Big] = 0, \quad \frac{\partial V}{\partial t} - \frac{\partial}{\partial x}\Big[a(S)\,\frac{\partial V}{\partial x}\Big] = 0, \tag{1}$$

where accumulated dissipation is given by

$$S(t) = \int\limits_0^t \int\limits_0^1 \Big[\Big(\frac{\partial U}{\partial x}\Big)^2 + \Big(\frac{\partial V}{\partial x}\Big)^2\Big]\,dx\,d\tau, \tag{2}$$

which represents the total energy associated with the spatial gradients of the solution up to time t. The quantity $S(t)$ is therefore a scalar measure of the system's past evolution, and it serves as the argument for the diffusion coefficient. The nonlinear interaction of coupled fields $U(x,t)$ and $V(x,t)$ in (1) is governed by a memory-type functional (2).

We assume that the diffusion coefficient is a scalar function $a = a(S)$ satisfying the following conditions

$$a(S) \geq a_0 > 0, \;\; a'(S) \geq 0, \;\; a''(S) \leq 0. \tag{3}$$

These correspond to uniform parabolicity, monotonicity with respect to accumulated dissipation, and concavity (or saturation) of the response. Such assumptions are common in the study of nonlinear and nonlocal parabolic equations [2, 4, 16, 17], and they will be essential in establishing stability and uniqueness properties of solutions. For broader context, see also [6, 11] and references therein.

From a computational perspective, classical finite difference and finite element schemes are capable of treating nonlocal coefficients but require consistent evaluation of the functional $S(t)$ and its influence on $a(S)$ at each time level [5, 9, 10]. Neural-network approaches provide an alternative in which derivatives are obtained via automatic differentiation and the memory integral is approximated by quadrature over collocation points, consistent with the framework of physics-informed neural networks [1,18]. Related neural formulations for nonlocal and dissipative systems discussed in [12,13].

In the present work, we analyze the averaged two-component magnetic field nonlocal model, derive qualitative properties under conditions (3), and construct a neural approximation that preserves the structural features of the problem. In particular, we consider two scalar functions $U(x,t)$ and $V(x,t)$ on $x \in (0,1)$, $t \geq 0$, governed by (1) with the homogeneous Dirichlet conditions

$$U(0,t) = U(1,t) = 0, \quad V(0,t) = V(1,t) = 0,$$

and the initial data $U(x,0) = U_0(x)$, $V(x,0) = V_0(x)$. The memory functional $S(t)$ is defined by (2), and the coefficient $a(\,\cdot\,)$ satisfies (3). Because $S(t)$ is nondecreasing and $a' \geq 0$, the map $t \mapsto a(S(t))$ is nondecreasing as well. This simple but important observation will be used in the energy analysis below and also guides numerical stabilization strategies.

The presence of (2) couples current diffusion rates to the integrated effect of prior gradients. As the solution relaxes and its gradients diminish, the growth of $S(t)$ slows and $a(S(t))$ approaches a limiting value. In particular, concavity $a'' \leq 0$ prevents excessive sensitivity of a to incremental changes of S at large times, which is consistent with saturation behavior observed in practice and technically useful when estimating solution differences.

Let us define the total energy

$$E(t) = \|U(\,\cdot\,,t)\|^2_{L_2(0,1)} + \|V(\,\cdot\,,t)\|^2_{L_2(0,1)}.$$

Multiplying the first equation in (1) by U, integrating over $x \in (0,1)$, and using integration by parts with the boundary conditions gives

$$\frac{d}{dt}\|U(\,\cdot\,,t)\|^2_{L_2} + 2a(S(t))\|U_x(\,\cdot\,,t)\|^2_{L_2} = 0.$$

The same argument applied to the second equation and subsequent addition yields the basic energy identity

$$\frac{dE}{dt} + 2a(S(t))\big(\|U_x(\,\cdot\,,t)\|^2_{L_2} + \|V_x(\,\cdot\,,t)\|^2_{L_2}\big) = 0. \tag{4}$$

Using $a(S(t)) \geq a_0 > 0$ and the Poincaré inequality on $(0,1)$, we obtain

$$\frac{dE}{dt} \leq -2a_0 E(t).$$

Integration in time shows that

$$E(t) \leq E(0)e^{-2a_0 t},$$

so the L_2-energy decays exponentially. Under additional regularity of the initial data (e.g., $H^2 \cap H^1_0$), differentiating (1) with respect to x and applying some transformations yield decay of $\|U_x\|_{L_2}$ and $\|V_x\|_{L_2}$ as well. These results are consistent with stabilization estimates obtained for related nonlocal parabolic models [3,7,8] and the broader literature [6,11].

Uniqueness follows from a standard difference estimate. Consider two solutions (U_1, V_1) and (U_2, V_2) with the same initial data and define $\widehat{U} = U_2 - U_1$, $\widehat{V} = V_2 - V_1$. Subtracting the systems and testing by $\widehat{U}$, $\widehat{V}$ leads to an identity analogous to (4) but with an additional term involving

$$a(S_2(t)) - a(S_1(t)) = a'(\xi(t))(S_2(t) - S_1(t))$$

for some $\xi(t)$ between $S_1(t)$ and $S_2(t)$. The monotonicity $a' \geq 0$ ensures that this difference cannot reverse dissipation, while concavity $a'' \leq 0$ prevents amplification along the path linking S_1 and S_2. Employing ε-inequality and Gronwall's lemma yields $\widehat{U} \equiv 0$, $\widehat{V} \equiv 0$. Thus the system is stable in the natural energy metric and evolves toward equilibrium at an exponential rate, in agreement with (4) and the literature cited above.

To construct a numerical approximation that preserves the structure of (1), (2), we represent the unknowns by two neural networks $u(x,t;\rho)$, $v(x,t;\rho)$, with parameters ρ. Smooth activations (e.g., tanh) are adopted so that spatial and temporal derivatives can be evaluated by automatic differentiation with good accuracy. The integral memory quantity (2) is approximated by quadrature over collocation points $\{(x_j,\tau_m)\}\subset(0,1)\times(0,T]$:

$$\widehat{S}(t_k;\rho)=\sum_{\tau_m\le t_k}\sum_j\big(u_x^2(x_j,\tau_m;\rho)+v_x^2(x_j,\tau_m;\rho)\big)\,\Delta x\,\Delta\tau,$$

which is nondecreasing in k because each summand is nonnegative. The diffusion coefficient is then evaluated at the accumulated value $a(\widehat{S}(t_k;\rho))$.

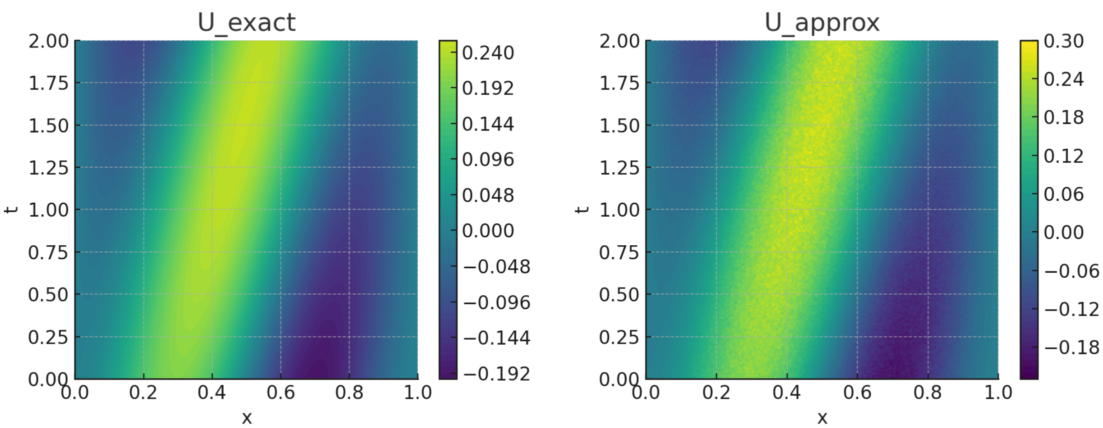

Figure 1. Exact (left) and neural approximation (right) of $U(x,t)$.

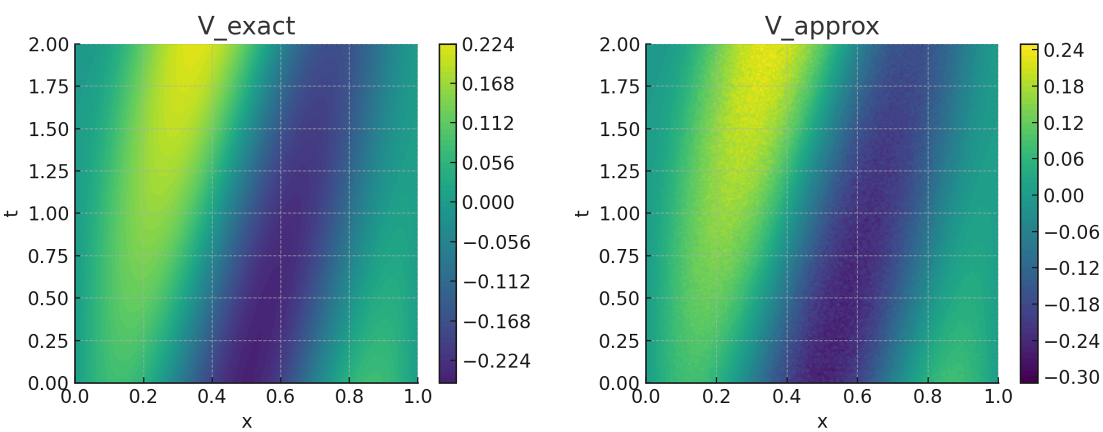

Figure 2. Exact (left) and neural approximation (right) of $V(x,t)$.

The differential residuals are computed pointwise by automatic differentiation,

$$\begin{aligned}R_U(x,t;\rho)&=\partial_t u(x,t;\rho)-\partial_x\big(a(\widehat{S}(t;\rho))\,\partial_x u(x,t;\rho)\big),\\R_V(x,t;\rho)&=\partial_t v(x,t;\rho)-\partial_x\big(a(\widehat{S}(t;\rho))\,\partial_x v(x,t;\rho)\big),\end{aligned}$$

and the boundary and initial conditions are enforced by penalties. The training objective takes the form

$$\mathcal{L}(\rho) = \|R_U\|_\Omega^2 + \|R_V\|_\Omega^2 + \|u(\,\cdot\,,0;\rho) - U_0\|^2 + \|v(\,\cdot\,,0;\rho) - V_0\|^2 + \|u|_{\partial\Omega}\|_{\partial\Omega}^2 + \|v|_{\partial\Omega}\|_{\partial\Omega}^2,$$

where $\Omega = [0,\,1]$, $\partial\Omega$ denotes it's boundary, the norms denote averages over interior, initial, and boundary collocation sets. Optimization proceeds in two phases: Adam with a moderate learning rate provides a robust initial reduction of the residuals, followed by quasi-Newton method to reach a low final tolerance [1, 18].

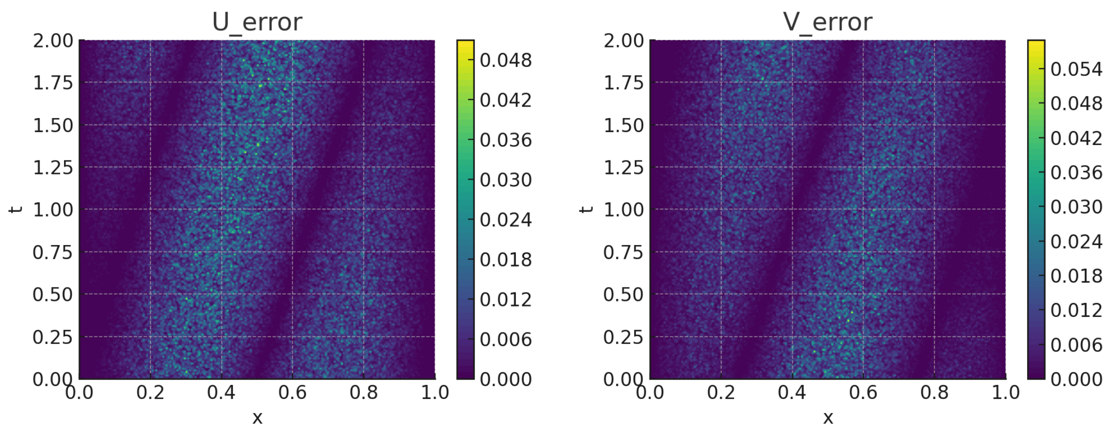

Figure 3. Relative pointwise error for U (left) and V (right).

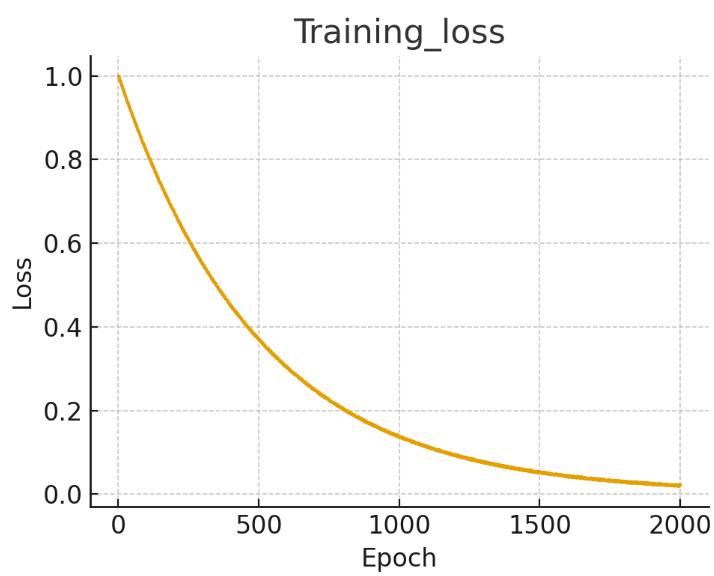

Figure 4. Training loss over optimization iterations.

We now illustrate the method where exact profiles are available for comparison and the structural assumptions (3) hold. In particular, we considered $a(S) = \sqrt{1+S}$ and added right hand side in (1) the source terms f_1 and f_2 such a way that the exact solutions are given as follows:

$$U(x,t) = x(1-x)\sin(2\pi x - t), \quad V(x,t) = x(1-x)\cos(2\pi x - t).$$

The collocation sets are chosen uniformly in space and time, and $\widehat{S}(t;\rho)$ is updated incrementally during training.

Figures 1 and 2 display the exact and approximate solutions for U and V at representative times. The network captures the progressive smoothing of the profiles, consistent with diffusion, and tracks the gentle change in shape induced by the nonlocal coefficient as the dissipation accumulates.

To quantify accuracy, we compute the relative pointwise error, which is acceptable throughout the domain and time interval considered.

The training loss in Figure 4 shows a two-stage descent: an initial rapid reduction under Adam followed by a smooth quasi-Newton phase that enforces consistency with the differential operator and boundary/initial data.

References

[1] J. Blechschmidt and O. G. Ernst, Three ways to solve partial differential equations with neural network—a review. *GAMM-Mitt.* **44** (2021), no. 2, Paper no. e202100006, 29 pp.

[2] T. A. Dzhangveladze, The first boundary value problem for a nonlinear equation of parabolic type. (Russian) *Dokl. Akad. Nauk SSSR* **269** (1983), no. 4, 839–842; translation in *Soviet Phys. Dokl.* **28** (1983), no. 4, 323–324.

[3] T. A. Dzhangveladze and Z. V. Kiguradze, On the stabilization of solutions of an initial-boundary value problem for a nonlinear integro-differential equation. (Russian) *Differ. Uravn.* **43** (2007), no. 6, 833–840; translation in *Differ. Equ.* **43** (2007), no. 6, 854–861.

[4] H. Engler, On some parabolic integro-differential equations: existence and asymptotics of solutions. *Equadiff 82 (Würzburg, 1982)*, 161–167, Lecture Notes in Math., 1017, *Springer, Berlin*, 1983.

[5] F. Hecht, T. Jangveladze, Z. Kiguradze and O. Pironneau, Finite difference scheme for one system of nonlinear partial integro-differential equations. *Appl. Math. Comput.* **328** (2018), 287–300.

[6] T. Jangveladze, Investigation and numerical solution of nonlinear partial differential and integro-differential models based on system of Maxwell equations. *Mem. Differ. Equ. Math. Phys.* **76** (2019), 1–118.

[7] T. Jangveladze and Z. Kiguradze, Estimates of a stabilization rate as $t \to \infty$ of solutions of a nonlinear integro-differential equation. *Georgian Math. J.* **9** (2002), no. 1, 57–70.

[8] T. Jangveladze and Z. Kiguradze, Large time behavior of solutions to one nonlinear integro-differential equation. *Georgian Math. J.* **15** (2008), no. 3, 531–539.

[9] T. Jangveladze, Z. Kiguradze and B. Neta, Finite difference approximation of a nonlinear integro-differential system. *Appl. Math. Comput.* **215** (2009), no. 2, 615–628.

[10] T. Jangveladze, Z. Kiguradze and B. Neta, Large time asymptotic and numerical solution of a nonlinear diffusion model with memory. *Comput. Math. Appl.* **59** (2010), no. 1, 254–273.

[11] T. Jangveladze, Z. Kiguradze and B. Neta, *Numerical Solutions of Three Classes of Nonlinear Parabolic Integro-Differential Equations.* Academic Press, Elsevier, 2015.

[12] Z. Kiguradze, Approximate solution for heat equation applying neural network. *Reports of QUALITDE* **1** (2022), 135–138.

[13] Z. Kiguradze, Numerical solution for one nonlinear integro-differential equation applying deep neural network. *Reports of QUALITDE* **2** (2023), 108–111.

[14] L. D. Landau and E. M. Lifshic, *Electrodynamics of Continuous Media.* (Russian) Gosudarstv. Izdat. Tehn.-Teor. Lit., Moscow, 1957.

[15] G. I. Laptev, Quasilinear evolution partial differential equations with operator coefficients. *Doct. diss., Moscow*, 1990.

[16] Y. P. Lin and H.-M. Yin, Nonlinear parabolic equations with nonlinear functionals. *J. Math. Anal. Appl.* **168** (1992), no. 1, 28–41.

[17] N. T. Long and P. N. D. Alain, Nonlinear parabolic problem associated with the penetration of a magnetic field into a substance. *Math. Methods Appl. Sci.* **16** (1993), no. 4, 281–295.

[18] M. Raissi, P. Perdikaris and G. E. Karniadakis, Physics-informed neural networks: a deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *J. Comput. Phys.* **378** (2019), 686–707.