

Guaranteeing Conservation of Integrals with Projection in Physics-Informed Neural Networks

Anonymous submission

Abstract

We propose a novel projection method that guarantees the conservation of integral quantities in Physics-Informed Neural Networks (PINNs) (Raissi, Perdikaris, and Karniadakis 2019). While the soft constraint PINNs use to enforce the structure of partial differential equations (PDEs) enables necessary flexibility during training, it also allows for the violation of these physical laws. We demonstrate that our projection method guarantees the conservation of the linear and quadratic integrals, both separately and jointly. We derived the projections used in these settings from constrained non-linear optimization problems and found that our PINN modified with the projection, which we call PINN-Proj, reduced the error in the conservation of these quantities by three to four orders of magnitude compared to the soft constraint and marginally reduced the error in the PDE solution. We also found that the projection improved convergence and the conditioning of the loss landscape. Our method holds promise as a general framework to guarantee the conservation of any integral quantity in a PINN if a projection is found and is tractable in neural network training.

Introduction

Recently, neural network-based approaches have emerged as powerful methods to solve and model PDEs. Among these are Physics-Informed Neural Networks (PINNs) (Raissi, Perdikaris, and Karniadakis 2019), which can parameterize the solution of a partial differential equation (PDE) with little to no training data and without requiring a mesh. The structure of the PDE and the initial and boundary conditions of the solution are enforced by the loss function. This creates a soft constraint on the physics equation, where the PINN is penalized for predictions that deviate from it. The soft constraint allows flexibility during training so that the PINN learns a solution that fits both the training data and the governing PDE.

This soft constraint, however, permits the solution learned by the PINN to violate conservation laws, which are fundamental principles of physics. Many conservation laws can be expressed as integral quantities that must remain constant or evolve according to changes in the solution domain. If the solutions that the PINN learned exactly obeyed conservation laws, they would be more physically accurate and more trustworthy when applied to model real-world physical systems. We cannot exactly enforce conservation laws using

another soft constraint, as it could still be violated in a valid solution. Instead, we must impose a hard constraint that ensures that the PINN always obeys the laws of conservation regardless of the nature of the data.

In this work, we propose a novel framework that can guarantee the conservation of some integral quantity in the PINN’s solution by projecting its prediction into a functional space where the conservation is not violated. This is done by framing the projection as a constrained non-linear optimization problem where we are minimizing the distance between the solution and a hyperplane where the integral is conserved. We then use Lagrange multipliers to find the unique solution to this problem. Given that the solution is a projection that is differentiable and tractable in training, it will ensure that the predictions from the PINN obey conservation exactly. We demonstrate that this projection method guarantees global conservation of the linear and quadratic integral, both separately and at the same time, works on one-dimensional and two-dimensional PDEs, and on conserved and non-conserved systems. We define *guarantee* as ensuring that the only sources of error are from the data and machine precision and not from the method.

We refer to our PINN model modified with the projection method as PINN-Proj. We compare its performance to an unmodified PINN and a PINN model modified with a soft constraint on the conservation of the same quantity or quantities, referred to as PINN-SC. The models were trained on small subsets of PDE solutions and evaluated on the error of their predictions of the solution of the PDE u and the conservation of the chosen integral quantity. We refer to these quantities as *conserved quantities* c . Our main contributions in this paper are as follows:

- We propose a novel projection method and show that it guarantees the conservation of the linear integral, quadratic integral, and both simultaneously
- We derive the projections for our conserved quantities by solving non-linear constrained optimization problems
- We show that PINN-Proj successfully guarantees the conservation of the integral quantities while marginally improving the accuracy of the PDE solution
- We analyze the convergence and loss landscape of PINN and PINN-Proj using spectral analysis of the Hessian of the loss

Related Work

Soft Constraints on PINNs A soft constraint can be used to enforce the conservation law in a PINN by adding a loss term that penalizes the difference between the conserved quantity in the current prediction and the ground truth. Training can either be done entirely with this modified loss function (Fang et al. 2022), or in two stages where the modified loss function is used in a second stage of training (Lin and Chen 2022). Creating soft constraints in this way can also be used to enforce conservation of flux between two neighboring discrete subdomains (Jagtap, Kharazmi, and Karniadakis 2020).

Hard Constraints on PINNs Hard constraints have been used in PINNs to strictly enforce initial and boundary conditions (Xiao et al. 2024). This was done by modifying the PINN solution with a distance function that vanishes at the boundaries and a particular solution that automatically satisfies the boundary and initial conditions. KKT-hPINN uses KKT conditions, which describe optimality conditions in constrained optimization, to create an untrainable projection on the output of a PINN that conserves a desired quantity, but is limited to linear constraints due to relying on the closed-form analytical solution of a linear system to create the projection (Chen, Flores, and Li 2024). Projection on the output of a PINN has also been used to guarantee conservation by applying multiple Newton iterations after each training step to move the solution back into the conservation manifold (Cardoso-Bihlo and Bihlo 2025). This projection method, however, relies on making iterative soft updates on the PINN output and does not incorporate the conservation constraint directly as a differentiable layer into the PINN. Numerical error values are also not reported, so it is difficult to quantify the performance of this hard constraint.

Exactly Learning Conservation Laws Conservation laws can be directly incorporated into the structure of neural networks to create hard constraints as well. The Hamiltonian (Greydanus, Dzamba, and Yosinski 2019) or Lagrangian (Cranmer et al. 2019), which describe how total energy is conserved in a system, can be parametrized by a neural network so that the quantities are always conserved in any solution it learns. Similarly, hard constraints can also be created by construction in a neural network on conservation laws by parameterizing a divergence-free vector field that corresponds to the continuity equation (Richter-Powell, Lipman, and Chen 2022). Due to the divergence-free assumption, however, these neural conservation laws are limited to forms of conserved systems, whereas PINNs can solve and model non-conservative systems that have sources, sinks, or boundary flux. PINNs are also not limited to PDEs that are forms of the continuity equation. Enforcing conservation using the continuity equation modeled as a divergence-free field has also been applied to neural operators, which instead learn the governing PDE directly from solution data (Liu et al. 2023). A differential projection layer can also be incorporated after a neural network to find the optimal linear combinations of basis functions to learn the governing PDE and enforce it as a hard constraint (Négiar, Mahoney, and Krishnapriyan 2022).

Method

Physics-Informed Neural Network

We define the general form of a partial differential equation (PDE) as

$$0 = u_t + \mathcal{N}[u], \quad (1)$$

where $u(x, t)$ is the solution and a function of spatial coordinate x and time t . $\mathcal{N}[\cdot]$ is a differentiable non-linear operator that acts on $u(x, t)$. The physics-informed neural network (PINN) is defined as

$$f := u_t + \mathcal{N}[u], \quad (2)$$

where $f(x, t)$ is a function and the residual between u_t and $\mathcal{N}[u]$. The PINN learns a parameterized solution $u_\theta(x, t)$ when $\mathcal{N}[\cdot]$ is known by using the loss function, \mathcal{L} , defined as

$$\mathcal{L} = \frac{1}{N_u} \sum_{i=1}^{N_u} |u_\theta(x_u^i, t_u^i) - u^i|^2 + \frac{1}{N_f} \sum_{i=1}^{N_f} |f(x_f^i, t_f^i)|^2. \quad (3)$$

Here, $\{x_u^i, t_u^i, u^i\}$ are the initial and boundary data on $u(x, t)$, while $\{x_f^i, t_f^i\}$ are collocation points randomly sampled from the domain independently. The first term of \mathcal{L} is the mean squared error between the PINN prediction $u_\theta(x, t)$ and ground truth u^i . The second term of \mathcal{L} is the mean squared error of $f(x, t)$, as the residual of the PDE is dependent on the learned $u_\theta(x, t)$ and is penalized to enforce the PDE's structure on the parameterized solution.

Physics-Informed Neural Network with Soft Constraint

We include PINN-SC to act as a benchmark for a soft constraint used to enforce conservation. This constraint was incorporated by adding a term to the loss function that penalized the difference between the predicted conserved quantity \hat{c} and ground truth c

$$\mathcal{L}_{SC} = \mathcal{L} + \lambda \frac{1}{T} \sum_{t=1}^T |c(t) - \hat{c}(t)|^2. \quad (4)$$

We found that adding a weight term λ to the conservation loss improved the performance of PINN-SC, establishing it as a more rigorous benchmark.

Conservation Laws

Many PDEs' governing physical systems are naturally expressed as conservation laws. A canonical form of such equation is

$$\frac{\partial u}{\partial t} + \nabla \cdot \mathbf{F}(u) = 0, \quad (5)$$

where the flux $\mathbf{F}(u)$ denotes the flow of a quantity u across a given spatial domain. This equation illustrates that the change in the conserved quantity over time is related to the divergence of its flux. To analyze the global behavior, we integrate the conservation law over a spatial domain \mathcal{X} ,

$$\frac{d}{dt} \int_{\mathcal{X}} u \, dx = - \int_{\mathcal{X}} \nabla \cdot \mathbf{F}(u) \, dx = - \oint_{\partial \mathcal{X}} \mathbf{F}(u) \cdot d\mathbf{S}. \quad (6)$$

The second equality is from applying the divergence theorem, which relates the volume integral of the divergence of a vector field to a surface integral over the boundary of the domain. This formulation identifies how the net change of the conserved quantity only depends on the flux across the boundary. If the flux through the boundary is zero, then the conserved quantity remains constant within the domain \mathcal{X} . If there exists flux across the boundary, the changes in the conserved quantity can be explicitly calculated,

$$c(t) = \int_{\mathcal{X}} u(x, t) dx, \quad (7)$$

where \mathcal{X} represents the computational domain and $c(t)$ denotes the known total conserved quantity. The changes in the conserved quantity can also be calculated if there is a source or sink term in the domain. Under periodic, reflective, or homogeneous Neumann boundary conditions, $c(t)$ remains constant. Under other prescribed boundary conditions, $c(t)$ can be determined through temporal integration.

Conservation of the Linear Integral

We define the linear integral of the parameterized solution of the PDE, $u_{\theta}(x, t)$ over the spatial domain \mathcal{X} as

$$c(t) = \int_{\mathcal{X}} u_{\theta}(x, t) dx. \quad (8)$$

If we select a PDE where u describes the velocity of a fluid, then this conserved quantity can be physically interpreted as the total momentum of the system.

Conservation of the Quadratic Integral

We define the quadratic integral of the parameterized solution of the PDE, $u_{\theta}(x, t)$ over the spatial domain \mathcal{X} as

$$c(t) = \int_{\mathcal{X}} u_{\theta}(x, t)^2 dx. \quad (9)$$

If we select a PDE where u describes the velocity of a fluid, then this conserved quantity can be physically interpreted as total energy of the system.

Projection Method for Conservation of the Linear Integral

Because these integral constraints are intractable under a discretized PDE solution and neural network parameterized output $u_{\theta}(x, t)$, we instead approximate the integral with discretization and numerical integration over the neural network outputs,

$$\hat{c}(t) = \int_{\mathcal{X}} u_{\theta}(x, t) dx \approx \sum_{i=1}^n u_{\theta}(x_i, t) \Delta x = \mathbf{1}_n^{\top} u_{\theta}(x, t) \Delta x, \quad (10)$$

where $\mathbf{1}_n$ is the all-ones vector, Δx represents the uniform discretization size, and n is the total number of discretization points. To force the neural network output to respect the linear integral constraint, we consider the following optimization problem

$$\min_y \|\bar{u}_{\theta}(x, t) - y\|^2 \text{ s.t. } \Delta x \mathbf{1}_n^{\top} y = c(t), \quad (11)$$

where $\bar{u}_{\theta}(x, t) = [u_{\theta}(x_1, t), u_{\theta}(x_2, t), \dots, u_{\theta}(x_n, t)]^{\top}$ is a vector that represents the discretized ground truth values of $u_{\theta}(x, t)$ arranged as a column vector, and y is our prediction from our projection method. We solve this constrained optimization problem by Lagrange multiplier. This problem has unique global minimum because, geometrically, we are finding the point on a hyperplane that has the smallest distance from a given point. We find the solution to be

$$\tilde{u}_{\theta}(x, t) = u_{\theta}(x, t) + \frac{c(t)}{n \Delta x} - \sum_{i=1}^n \frac{\bar{u}_{\theta}(x_i, t)}{n} \quad (12)$$

where $u_{\theta}(x, t)$ is the unprojected output of the neural network and $\tilde{u}_{\theta}(x, t)$ is the projected output of the neural network.

Projection Method for Conservation of the Quadratic Integral

We can now use the same method to find a projection for a physical system that is governed by quadratic summation constraints. The conservation principle can be expressed as

$$\hat{c}(t) = \int_{\mathcal{X}} u_{\theta}(x, t)^2 dx \approx \sum_{i=1}^n u_{\theta}(x_i, t)^2 \Delta x = y^{\top} y \Delta x. \quad (13)$$

We then consider the following optimization problem,

$$\min_y \|\bar{u}_{\theta}(x, t) - y\|^2, \text{ s.t. } \Delta x y^{\top} y = c(t), \quad (14)$$

and find the solution to be

$$\tilde{u}_{\theta}(x, t) = u_{\theta}(x, t) \sqrt{\frac{c(t)}{\sum_{i=1}^n \bar{u}_{\theta}(x_i, t)^2}}. \quad (15)$$

Combined Projection Method for Linear and Quadratic Integral

For a physical system that is governed by both linear and quadratic summation constraints, we can also find a conservation principle that obeys both constraints. We apply both previously incorporated constraints in the same manner as before,

$$\min_y \|\bar{u}_{\theta}(x, t) - y\|^2 \text{ s.t. } \Delta x y^{\top} y = c_1(t), \Delta x \mathbf{1}_n^{\top} y = c_2(t) \quad (16)$$

and find the solution to be

$$\tilde{u}_{\theta}(x, t) = (u_{\theta}(x_i, t) - \sum_{i=1}^n \frac{u_{\theta}(x_i, t)}{n}) \sqrt{C} + \frac{c_2(t)}{n \Delta x} \quad (17)$$

$$C = \frac{c_1(t)/\Delta x - c_2^2(t)/(n \Delta x^2)}{\sum_{i=1}^n (u_{\theta}(x_i, t) - \sum_{i=1}^n \frac{u_{\theta}(x_i, t)}{n})^2}.$$

The complete derivations of all projections can be found in the Appendix.

Projection Method for Non-Conserved Systems

Our projection method allows $c(t)$ to either be constant, as in a conserved system, or changing, as in a non-conserved system where there is boundary flux or the addition or removal of mass. In the latter case, the projection method requires an additional step to fully characterize $c(t)$. $c(t)$ is calculated from a discretized PDE solution $u(x, t)$, so $c(t)$ is only defined at the t values in this solution. However, since the collocation points used to calculate $f(x, t)$ are generated using Latin Hypercube Sampling over the computational domain, some of these t values will not have defined $c(t)$ values. This does not affect a conserved system, as $c(t)$ is assumed to be the same for all times. However, for a non-conserved system, we estimate these missing $c(t)$ values by linear interpolation using the second-order accurate gradient from the closest defined $c(t)$ value. We avoided using more complex approximation techniques for simplicity and efficiency.

Projection Method for 2D Systems

We can extend the projection method to two dimensional PDEs by integrating over both spatial dimensions in the calculation of $c(t)$. For spatial dimensions x and y , the linear integral is

$$\hat{c}(t) = \int_{\mathcal{Y}} \int_{\mathcal{X}} u_{\theta}(x, y, t) dx dy \approx \sum_{j=1}^n \sum_{i=1}^n u_{\theta}(x_i, y_j, t) \Delta x \Delta y, \quad (18)$$

and we use a similar formula for other conserved quantities. We can extend the projection method into an arbitrary number of dimensions by integrating over all spatial dimensions of the solution to find the conserved quantity.

Experimental Setup

We trained three different models for evaluation: PINN, PINN-SC, and PINN-Proj. The training setup for all models is the same as the original PINN method (Raissi, Perdikaris, and Karniadakis 2019) unless otherwise mentioned. We used five PDE datasets generated from the Reaction-Diffusion Equation, 1D and 2D Advection Equation, Wave Equation, and Korteweg-De Vries Equation. These were selected to include PDEs of different orders and applications and because they all conserve both the linear and quadratic integral (with the exception of Reaction-Diffusion). All PDEs were one-dimensional unless otherwise noted.

The computational domain for all 1D PDEs was discretized into 256 spatial points and 100 temporal points. For the 2D advection equation, we used a 32x32 spatial grid with 100 time steps. All solutions were generated using various FDM solvers and had homogeneous Neumann boundary conditions. The Advection Equation, Wave Equation, and Korteweg-De Vries Equation solutions were conserved systems with no net change in mass. The Reaction-Diffusion Equation included a reaction term that added mass to the system over time, so the modified projection method for non-conserved systems was used in this case.

The training set contained 100 points for the 1D PDEs and 3000 points for the 2D PDEs that were randomly sam-

pled from each PDE dataset, and contained 10,000 collocation points that were generated using Latin Hypercube Sampling. Each test set was composed of all spatial and temporal points in the domain to assess the PINN’s ability to model the full PDE solution. Training was done until the convergence criterion of the L-BFGS optimizer was reached. More details can be found in the Appendix.

Evaluation

The models were evaluated on the error in predicting the PDE solution, called Error u , and on the error in the conservation of the conserved quantity, called Error c , in the solution. Error c_L refers to the error of conservation of the linear integral and Error c_Q refers to the error of conservation of the quadratic integral. Relative \mathcal{L}_2 error was used for Error u , and absolute \mathcal{L}_1 error was used for Error c . We also report the duration and number of epochs of training. We averaged all values over 10 trials.

Results

Table 1 shows the error on solution prediction and conservation when the linear integral is the conserved quantity, Table 2 shows the error on solution prediction and conservation when the quadratic integral is the conserved quantity, and Table 3 shows the error on solution prediction and conservation when both the linear and quadratic integrals are conserved quantities. The relative performance of the models was very similar across the three settings. On solution prediction, PINN-Proj achieved the lowest Error u on 13 of the 15 trials, and PINN achieved the lowest Error u on two of the 15 trials. These two trials were both on the Advection Equation. It is notable, however, that the margin between Error u values for all models was always less than one order of magnitude. Regarding conservation, PINN-Proj achieved the lowest Error c_L and Error c_Q on all datasets by three to four orders of magnitude compared to the next best model, which was always PINN-SC.

Between different conserved quantities, PINN-SC always improved on PINN in Error c , but did not consistently improve Error u over PINN for the linear integral, and more often worsened Error u for the quadratic integral and when both integrals were conserved. Error c_L and Error c_Q are also slightly higher for each PDE when both integrals are conserved compared to when only one is conserved.

Table 4 shows the average duration and number of epochs of training for PINN, PINN-SC, and PINN-Proj. The average was taken across trials with different conserved quantities, so there is a single value for each model and PDE. The soft constraint increased training time on average by 11x on the 1D PDEs and by a substantial 90x on the 2D PDE. The projection method increased training time on average by 6x on the 1D PDEs and by 30x on the 2D PDE. PINN-Proj always took the fewest epochs to converge, and PINN-SC always took the most epochs to converge across all trials.

Figure 1 visualizes the evolution of $c(t)$ over time for the 1D conserved system PDEs. The ground truth $c(t)$ is the black dashed line. We can observe that these predictions are consistent with the error values, as the predicted $c(t)$ values

PDE	PINN		PINN-SC		PINN-Proj	
	Error u	Error c_L	Error u	Error c_L	Error u	Error c_L
1D Advection Eq.	2.61E-02	3.80E-01	2.93E-02	7.87E-02	2.28E-02	6.63E-06
2D Advection Eq.	5.43E-02	2.29E-01	5.35E-02	1.05E-02	5.22E-02	9.85E-06
Wave Eq.	3.74E-02	3.40E+00	2.56E-02	4.05E-01	2.06E-02	1.06E-05
KdV Eq.	6.87E-03	1.76E-01	8.54E-03	2.21E-02	2.47E-03	1.81E-05
React-Diff Eq.	2.29E-03	9.00E-02	2.21E-03	5.30E-02	1.56E-03	2.57E-05

Table 1: Average error for solution prediction (Error u) and conservation (Error c_L) when linear integral is conserved

PDE	PINN		PINN-SC		PINN-Proj	
	Error u	Error c_Q	Error u	Error c_Q	Error u	Error c_Q
1D Advection Eq.	2.96E-02	6.22E-02	2.65E-02	6.07E-03	2.24E-02	3.22E-06
2D Advection Eq.	5.36E-02	1.34E-01	5.49E-02	6.07E-03	5.65E-02	1.96E-06
Wave Eq.	3.28E-02	4.99E+00	7.61E-02	5.25E-01	2.27E-02	9.38E-06
KdV Eq.	2.22E-03	3.87E-02	9.93E-03	1.42E-02	1.77E-03	9.94E-06
React-Diff Eq.	2.13E-03	1.94E-01	3.50E-03	4.06E-02	2.05E-03	1.51E-05

Table 2: Average error for solution prediction (Error u) and conservation (Error c_Q) when quadratic integral is conserved

PDE	PINN			PINN-SC			PINN-Proj		
	Error u	Error c_L	Error c_Q	Error u	Error c_L	Error c_Q	Error u	Error c_L	Error c_Q
1D Adv. Eq.	2.23E-02	4.47E-01	8.80E-01	3.02E-02	7.89E-02	1.93E-02	3.22E-02	6.29E-06	4.94E-06
2D Adv. Eq.	5.34E-02	1.06E-01	1.33E-01	5.31E-02	1.21E-02	5.73E-03	5.27E-02	1.09E-05	2.68E-06
Wave Eq.	3.38E-02	2.41E+00	3.93E+00	4.20E-02	3.61E-01	7.11E-01	1.97E-02	9.72E-06	1.34E-05
KdV Eq.	4.19E-03	7.94E-02	6.14E-02	9.04E-03	4.69E-02	3.01E-02	1.91E-03	2.73E-05	2.80E-05
React-Diff Eq.	2.01E-03	1.26E-01	2.16E-01	2.50E-03	5.17E-02	5.21E-02	1.31E-03	4.43E-05	4.95E-05

Table 3: Average error for solution prediction (Error u) and conservation (Error c_L and Error c_Q) when linear and quadratic integrals are conserved

PDE	Metric	PINN	PINN-SC	PINN-Proj
1D Advection Eq.	Duration (s)	5.66	57.02	35.47
	Epochs	1235	1970	928
2D Advection Eq.	Duration (s)	2.08	193.00	65.30
	Epochs	421	898	234
Wave Eq.	Duration (s)	1.29	27.16	7.10
	Epochs	187	901	172
Korteweg-De Vries Eq.	Duration (s)	11.42	49.96	44.16
	Epochs	1307	1565	1048
Reaction-Diffusion Eq.	Duration (s)	7.24	61.80	52.54
	Epochs	1336	2253	1187

Table 4: Average duration of training and epochs across models.

from PINN diverge most from the ground truth, those from PINN-SC diverge less, and those from PINN-Proj are difficult to discern from the ground truth on scales where all curves are visible.

Discussion

On solution prediction, PINN-Proj achieved the best performance on 13 of the 15 systems tested. However, we note that the margin between PINN-Proj and the next best model

was always less than one order of magnitude. On conservation of quantities, PINN-Proj was always the best performing model, outperforming the soft constraint by three to four orders of magnitude. Our results thus demonstrate that the projection method was able to reliably and significantly improve the conservation of the linear and quadratic integral, both together and separately. Additionally, the projection method also marginally improved the accuracy of the PDE solution.

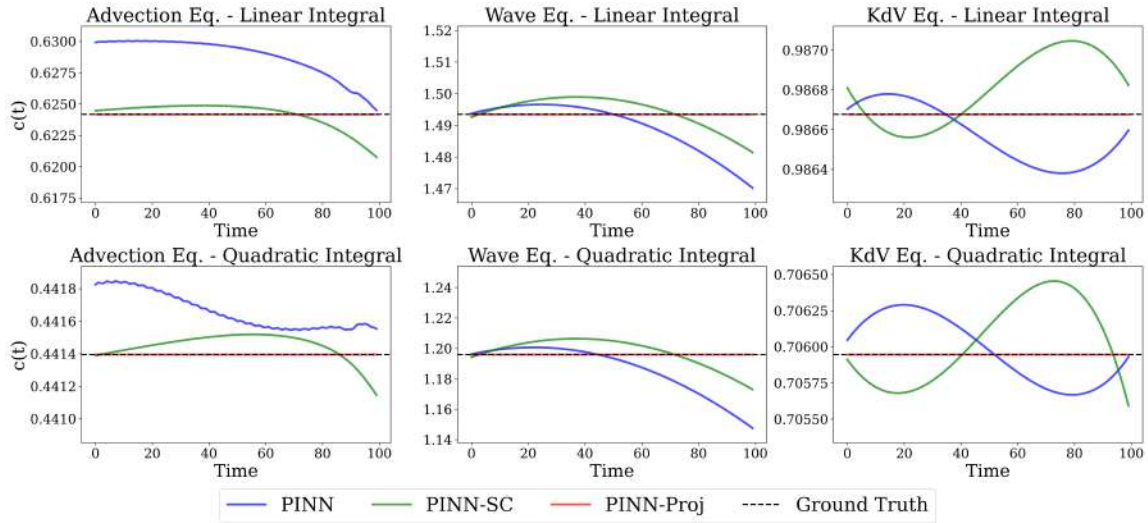


Figure 1: $c(t)$ values over time for the 1D conserved system PDEs with one conserved quantity

Error c was consistently the lowest on both the 1D and 2D Advection Equation, which is a first-order PDE, and higher on the Wave and Korteweg-De Vries Equations, which are second-order and third-order PDEs, respectively. This indicates that Error c for PINN-Proj increased as the order of the PDE increased. Despite the Reaction-Diffusion Equation being a first order PDE, it still had Error c values comparable to the higher order PDEs. This is likely due to the error of the linear approximation of $c(t)$, and could be reduced using higher order approximation methods. Error c_L and c_Q were also similar whether the linear and quadratic integrals were conserved at the same time or not, so conserving both quantities did not seem to significantly degrade performance.

Errors and Guarantee Across all datasets and conserved quantities, the lowest Error c values for PINN-Proj reached approximately $3E-6$. This is slightly over one order of magnitude away from machine precision for 32-bit floating point, at approximately $1E-7$. However, we claim that our guarantee is validated, and that the discrepancy between these values originates from approximation error that we subsequently discuss in the solution data and model architecture. All numerically computed PDE solutions contain discretization error from the approximation of continuous derivatives and integrals on a finite computational grid. This error then accumulates at the integration step in the projection method and limits its accuracy. There is also numerical error from the solver method. We can estimate its contribution to Error c by the change in $c(t)$ over time. Since $c(t)$ should remain constant in a conserved system, any variation must come from the numerical error. To address this, the mean $c(t)$ value was used as the ground truth of $c(t)$. Across all PDEs, the variation in $c(t)$ was less than $1E-4$ for the linear integral, and less than $5E-4$ for the quadratic integral.

There is also approximation error inherent in the PINN architecture from the neural network and the auto-differentiation step in the PDE structure enforcement. These methods can have difficulty accurately representing solu-

tions with sharp features and higher-order derivatives. The PINN is also limited by the set of collocation points that enforce the PDE structure. We observed that increasing the number of collocation points decreased Error c , as it allows the PINN to increase the resolution of its approximation of the PDE residual loss. Though these errors limit the application of the projection method, one could apply more sophisticated solver methods, increase the spatial resolution of the PDE solution and the collocation points, and increase the size of the neural network to further reduce conservation error and fully realize the benefits of the projection method.

PINN Training and Convergence Both the addition of the soft constraint in PINN-SC and the projection method in PINN-Proj significantly increased training time. This is because the projection method required the neural network $u_\theta(x, t)$ to be evaluated at all x and t values in the computational domain at each training step. Additionally, both the soft constraint and the projection layer also require an additional summation calculation over the spatial points in the forward pass, which greatly increases the complexity of gradient computations during backpropagation. This caused a significant increase in training time for the 2D Advection Equation, as it contained four times as many grid points as the 1D equations. However, through this increased cost, we were able to guarantee the conservation of a chosen quantity and even improve the accuracy of the solution. For applications where safety is paramount, this increased cost could likely be justified. The cost is a one-time investment at training time, which can be amortized over all future inference uses. Additionally, in some applications such as in chemical reactors (Chen, Flores, and Li 2024) or inverse design (Lu et al. 2021), hard constraints are necessary for safe and accurate modeling.

Though PINN-SC and PINN-Proj both added to training time, PINN-Proj always took the fewest epochs to converge. Combined with the improved Error u , the results suggest that the projection improved convergence. Since an accurate

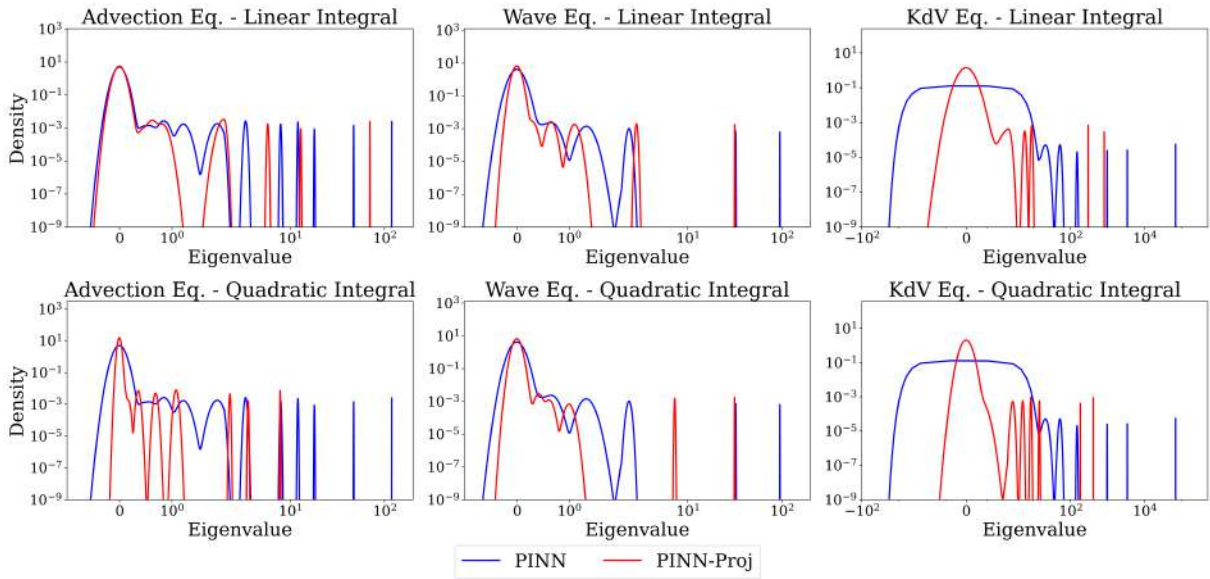


Figure 2: Spectral densities of Hessian for the 1D conserved system PDEs with one conserved quantity

reconstruction of the solution would maintain the conserved quantity, the projection could have constrained the search space and guided gradient descent to a better solution more rapidly. We can verify this by analyzing the loss landscape using the spectral density of the loss function’s Hessian matrix (Rathore et al. 2024). These spectral densities describe the distribution of the eigenvalues of the Hessian to give information on the local curvature of the loss function around the optimum found by the models. When the eigenvalues of the Hessian are far apart in magnitude, the loss landscape is very steep in some directions and flat in others, so the loss is ill-conditioned and difficult for gradient descent to minimize. When eigenvalues are more similar in magnitude, there is more uniform change in all directions, so the loss landscape is better conditioned. This results in more uniform sensitivity in all directions and makes gradient descent more effective.

Figure 2 contains the spectral densities of the Hessians of PINN and PINN-Proj for the 1D conserved systems. The spectral densities for PINN-Proj are always more tightly clustered than PINN, which means that its eigenvalues are closer together in magnitude. Also, the greatest eigenvalue of PINN-Proj always has a lower value than that of PINN. This shows that the projection method improved the conditioning of the loss landscape and improved convergence.

The spectral densities also help explain why PINN only outperformed PINN-Proj on Error u only on the Advection Equation. By looking at how much more tightly clustered the eigenvalues are, we see that PINN-Proj improved the conditioning of the Korteweg-De Vries Equation the most, as it lowered the greatest eigenvalue by around a factor of 100, and improved the conditioning of the Advection Equation when conserving the linear integral the least, as it lowered the greatest eigenvalue by significantly less than a factor of 10. The Advection Equation was the simplest PDE tested, so

its loss landscape would be relatively simple. The projection could, therefore, have had less of a benefit on the conditioning, as it was already well-conditioned. We discussed how the projection method also worsened some approximation errors, and this could have harmed the solution reconstruction. The other PDEs were, therefore, sufficiently complex such that the projection improved accuracy despite the accumulation of approximation errors.

Future work could involve extending the projection method to other integral quantities beyond the linear and quadratic integrals, which would involve solving a new non-linear optimization problem that is differentiable and tractable in neural network training. This direction is promising owing to the demonstrated effectiveness of the projection method. Future work could also include evaluating the projection method on more complex PDEs and initial conditions to further quantify its performance. The training cost of PINN-Proj could also be improved through better memory management and application of other hardware-aware techniques.

Conclusion

We proposed a novel projection method to guarantee conservation of integral quantities in a PINN. To determine its efficacy, we compared a PINN that uses the projection method, PINN-Proj, to an unmodified PINN and a PINN with a soft constraint, PINN-SC, on 1D, 2D, conserved, and non-conserved PDE systems. We found that while adding computational cost, the addition of the projection significantly reduced the violation of conservation laws and guaranteed the conservation of the linear and quadratic integrals, both when applied separately and jointly, over PINN and PINN-SC. We also found that the projection improved the convergence of PINN using spectral analysis and marginally lowered the error in solution reconstruction.

References

- Cardoso-Bihlo, E.; and Bihlo, A. 2025. Exactly conservative physics-informed neural networks and deep operator networks for dynamical systems. *Neural Networks*, 181: 106826.
- Chen, H.; Flores, G. E. C.; and Li, C. 2024. Physics-informed neural networks with hard linear equality constraints. *Computers & Chemical Engineering*, 189: 108764.
- Cranmer, M.; Greydanus, S.; Hoyer, S.; Battaglia, P.; Spergel, D.; and Ho, S. 2019. Lagrangian Neural Networks. In *ICLR 2020 Workshop on Integration of Deep Neural Models and Differential Equations*.
- Fang, Y.; Wu, G.-Z.; Kudryashov, N. A.; Wang, Y.-Y.; and Dai, C.-Q. 2022. Data-driven soliton solutions and model parameters of nonlinear wave models via the conservation-law constrained neural network method. *Chaos, Solitons & Fractals*, 158: 112118.
- Greydanus, S.; Dzamba, M.; and Yosinski, J. 2019. Hamiltonian neural networks. *Advances in neural information processing systems*, 32.
- Jagtap, A. D.; Kharazmi, E.; and Karniadakis, G. E. 2020. Conservative physics-informed neural networks on discrete domains for conservation laws: Applications to forward and inverse problems. *Computer Methods in Applied Mechanics and Engineering*, 365: 113028.
- Lin, S.; and Chen, Y. 2022. A two-stage physics-informed neural network method based on conserved quantities and applications in localized wave solutions. *Journal of Computational Physics*, 457: 111053.
- Liu, N.; Fan, Y.; Zeng, X.; Klöwer, M.; Zhang, L.; and Yu, Y. 2023. Harnessing the power of neural operators with automatically encoded conservation laws. *arXiv preprint arXiv:2312.11176*.
- Lu, L.; Pestourie, R.; Yao, W.; Wang, Z.; Verdugo, F.; and Johnson, S. G. 2021. Physics-informed neural networks with hard constraints for inverse design. *SIAM Journal on Scientific Computing*, 43(6): B1105–B1132.
- Négiar, G.; Mahoney, M. W.; and Krishnapriyan, A. S. 2022. Learning differentiable solvers for systems with hard constraints. *arXiv preprint arXiv:2207.08675*.
- Raissi, M.; Perdikaris, P.; and Karniadakis, G. E. 2019. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational physics*, 378: 686–707.
- Rathore, P.; Lei, W.; Frangella, Z.; Lu, L.; and Udell, M. 2024. Challenges in training pinns: A loss landscape perspective. *arXiv preprint arXiv:2402.01868*.
- Richter-Powell, J.; Lipman, Y.; and Chen, R. T. 2022. Neural conservation laws: A divergence-free perspective. *Advances in Neural Information Processing Systems*, 35: 38075–38088.
- Xiao, Z.; Ju, Y.; Li, Z.; Zhang, J.; and Zhang, C. 2024. On the Hard Boundary Constraint Method for Fluid Flow Prediction based on the Physics-Informed Neural Network. *Applied Sciences*, 14(2): 859.
- Yao, Z.; Gholami, A.; Keutzer, K.; and Mahoney, M. W. 2020. Pyhessian: Neural networks through the lens of the hessian. In *2020 IEEE international conference on big data (Big data)*, 581–590. IEEE.

Appendix

Training Setup

The PINN models contained 9 hidden linear layers of 20 neurons. The weights of the neural network were initialized using Xavier normal initialization. All activation functions were hyperbolic tangent. The optimizer was L-BFGS. The L-BFGS optimizer used the change in the loss function was less than machine epsilon, which was approximately $1\text{E-}7$. Thus, the number of epochs of training varied between trials and between datasets. We used a weight of $\lambda = 10$ on PINN-SC across all trials. All training was done on an Nvidia RTX 4090.

PDE Setup

Each 1D PDE dataset had parameters $\Delta x = 1/128$, $\mathcal{X} = [0, 2]$, $\Delta t = 0.01$, $\mathcal{T} = [0, 0.99]$, $n_x = 256$, and $n_t = 100$, so each had 25,600 total state values u . The 2D Advection dataset had parameters $\Delta x = 1/16$, $(\mathcal{X}, \mathcal{Y}) = [0, 2] \times [0, 2]$, $\Delta t = 0.01$, $\mathcal{T} = [0, 0.99]$, $(n_x, n_y) = (32, 32)$, and $n_t = 100$, so there were 102,400 total state values u . We used figures for the 1D conserved system PDEs to isolate the effects of the projection method without confounding factors from higher dimensionality or non-conservative terms.

The solutions for the Advection and Wave Equation were generated using a Weighted Essentially Non-Oscillatory (WENO) solver method. The solutions for the Korteweg-De Vries and Reaction-Diffusion Equation were generated using a Crank-Nicolson solver method. The Reaction-Diffusion equation had a diffusion coefficient of 0.1 and a reaction coefficient of 0.5. The Advection equation had a transport coefficient 0.1. The Wave equation had a diffusion coefficient of 0.1. The Korteweg-De Vries had a non-linear coefficient of 1 and a dispersive coefficient of 0.0025.

Hessian Spectrum Setup

The Hessian spectrum was generated by stochastic Lanczos quadrature (SLQ) using PyHessian (Yao et al. 2020). SLQ estimates the Hessian spectrum by using the Lanczos algorithm, which transforms a large matrix into a much smaller tridiagonal matrix whose eigenvalues approximate the original matrix’s eigenvalues, with random probe vectors to iteratively construct a tridiagonal matrix whose eigenvalues approximate those of the full Hessian. The spectral densities were smoothed using Gaussian kernels with a variance of $5\text{E-}5$.

Advection Equation

The Advection Equation is a PDE that describes the transport of a quantity $u(x, t)$, such as mass, heat, or momentum, within a fluid due to its bulk motion.

1D:

$$\frac{\partial u}{\partial t} + c \frac{\partial u}{\partial x} = 0$$

Initial Condition:

$$u_0 = \exp\left(-\frac{1}{2} \frac{(x-1)^2}{(0.25)^2}\right)$$

Boundary Condition:

$$\frac{\partial u}{\partial x}(0, t) = 0, \quad \frac{\partial u}{\partial x}(2, t) = 0$$

2D:

$$\frac{\partial u}{\partial t} + c \frac{\partial u}{\partial x} + c \frac{\partial u}{\partial y} = 0$$

Initial Condition:

$$u_0 = \exp\left(-\frac{1}{2} \frac{(x-1)^2 + (y-1)^2}{(0.25)^2}\right)$$

Boundary Condition:

$$\frac{\partial u}{\partial x}(0, t) = 0, \quad \frac{\partial u}{\partial x}(2, t) = 0$$

$$\frac{\partial u}{\partial y}(0, t) = 0, \quad \frac{\partial u}{\partial y}(2, t) = 0$$

Wave Equation

The Wave Equation is a second-order PDE that describes the propagation of waves where a quantity $u(x, t)$ is the wave's displacement.

$$\frac{\partial^2 u}{\partial t^2} = c^2 \frac{\partial^2 u}{\partial x^2}$$

Initial Condition:

$$u_0 = \exp(-(x-1)^2)$$

Boundary Condition:

$$\frac{\partial u}{\partial x}(0, t) = 0, \quad \frac{\partial u}{\partial x}(2, t) = 0$$

Korteweg-De Vries Equation

The Korteweg-de Vries equation is a third-order non-linear PDE that describes the propagation of shallow water waves where a quantity $u(x, t)$ represents the wave's amplitude.

$$\frac{\partial u}{\partial t} + au \frac{\partial u}{\partial x} + b \frac{\partial^3 u}{\partial x^3} = 0$$

Initial Condition:

$$u_0 = \exp\left(-\frac{(x-1)^2}{2}\right)$$

Boundary Condition:

$$\frac{\partial u}{\partial x}(0, t) = 0, \quad \frac{\partial u}{\partial x}(2, t) = 0$$

Reaction-Diffusion Equation

The Reaction-Diffusion Equation models the interaction between diffusion, which spreads a quantity over space, and reaction, which governs the flux into the system where a quantity $u(x, t)$ is the concentration of a substance.

$$\frac{\partial u}{\partial t} = D \frac{\partial^2 u}{\partial x^2} + R(u)$$

In this case,

$$R(u) = ku.$$

Initial Condition:

$$u_0 = \exp\left(-\frac{(x-1)^2}{(0.5)^2}\right)$$

Boundary Condition:

$$\frac{\partial u}{\partial x}(0, t) = 0, \quad \frac{\partial u}{\partial x}(2, t) = 0$$

Linear Approximation of $c(t)$

Second-order central difference:

$$c'(t_i) = \frac{c(t_{i+1}) - c(t_{i-1}))}{t_{i+1} - t_{i-1}}, \quad (19)$$

with forward differences at the left boundary:

$$c'(t_0) = \frac{-3c(t_0) + 4c(t_1) - c(t_2))}{t_2 - t_0}, \quad (20)$$

and backward differences at the right boundary:

$$c'(t_n) = \frac{3c(t_n) - 4c(t_{n-1}) + c(t_{n-2}))}{t_n - t_{n-2}}. \quad (21)$$

We approximate $c(t_f)$ using gradient-based interpolation, where t^* is the closest value to t_f where $c(t)$ is defined:

$$c(t_f) = c(t^*) + c'(t^*)(t_f - t^*). \quad (22)$$

Derivation of Conservation of Linear Integral

$$\min_y ||\bar{u}_\theta(x, t) - y||^2, \text{ s.t. } \Delta x \mathbf{1}_n^\top y = c(t) \quad (23)$$

The Lagrangian has the form:

$$\mathcal{L}(y, \lambda) = (\bar{u}_\theta(x, t) - y)^\top (\bar{u}_\theta(x, t) - y) + \lambda (\Delta x \mathbf{1}_n^\top y - c(t)). \quad (24)$$

Take partial derivatives and set to zero:

$$\partial \mathcal{L} / \partial y_i = -2(\bar{u}_\theta(x_i, t) - y_i) + \lambda \Delta x = 0 \quad (25)$$

$$\partial \mathcal{L} / \partial \lambda = \Delta x \mathbf{1}_n^\top y - c(t) = 0 \quad (26)$$

From the first equation, we get:

$$y_i = \bar{u}_\theta(x_i, t) - \frac{\lambda \Delta x}{2} \quad (27)$$

Substitute this into the second equation:

$$c(t) = \sum_{i=1}^n y_i \Delta x = \sum_{i=1}^n (\bar{u}_\theta(x_i, t) - \frac{\lambda \Delta x}{2}) \Delta x \quad (28)$$

$$\frac{\lambda \Delta x}{2} = \sum_{i=1}^n \frac{\bar{u}_\theta(x_i, t)}{n} - \frac{c(t)}{n \Delta x} \quad (29)$$

Substitute back into the equation for y_i to find the projection solution on the discretized point:

$$\tilde{u}_\theta(x_i, t) = y_i = \bar{u}_\theta(x_i, t) + \frac{c(t)}{n\Delta x} - \sum_{i=1}^n \frac{\bar{u}_\theta(x_i, t)}{n} \quad (30)$$

$$= \bar{u}_\theta(x_i, t) + \frac{c(t)}{n\Delta x} - \mathbf{1}_n^\top \frac{\bar{u}_\theta(x_i, t)}{n} \quad (31)$$

We can verify that this solution satisfies the constraint:

$$\begin{aligned} \sum_{i=1}^n \tilde{u}_\theta(x_i, t) \Delta x &= \sum_{i=1}^n \left(\bar{u}_\theta(x_i, t) + \frac{c(t)}{n\Delta x} \right. \\ &\quad \left. - \sum_{i=1}^n \frac{\bar{u}_\theta(x_i, t)}{n} \right) \Delta x = c(t) \end{aligned} \quad (32)$$

Extending the projection to continuous space yields:

$$\tilde{u}_\theta(x, t) = u_\theta(x, t) + \frac{c(t)}{n\Delta x} - \sum_{i=1}^n \frac{\bar{u}_\theta(x_i, t)}{n} \quad (33)$$

Derivation of Conservation of Quadratic Integral

$$\min_y ||\bar{u}_\theta(x, t) - y||^2, \text{ s.t. } \Delta x y^\top y = c(t) \quad (34)$$

The Lagrangian has the form:

$$\begin{aligned} \mathcal{L}(y, \lambda) &= (\bar{u}_\theta(x, t) - y)^\top (\bar{u}_\theta(x, t) - y) \\ &\quad + \lambda (\Delta x y^\top y - c(t)) \end{aligned} \quad (35)$$

Take partial derivatives and set to zero:

$$\partial \mathcal{L} / \partial y_i = -2(\bar{u}_\theta(x_i, t) - y_i) + 2\lambda \Delta x y_i = 0 \quad (36)$$

$$\partial \mathcal{L} / \partial \lambda = \Delta x y^\top y - c(t) = 0 \quad (37)$$

From the first equation, we get:

$$y_i = \frac{\bar{u}_\theta(x_i, t)}{1 + \lambda \Delta x} \quad (38)$$

Substitute this into the second equation:

$$c(t) = \sum_{i=1}^n y_i^2 \Delta x = \sum_{i=1}^n \frac{\bar{u}_\theta(x_i, t)^2 \Delta x}{(1 + \lambda \Delta x)^2} \quad (39)$$

$$\frac{1}{1 + \lambda \Delta x} = \sqrt{\frac{c(t)}{\sum_{i=1}^n \bar{u}_\theta(x_i, t)^2}} \quad (40)$$

On the dicretized point, the projected solution that satisfies the constraint is:

$$\tilde{u}_\theta(x_i, t) = y_i = \bar{u}_\theta(x_i, t) \sqrt{\frac{c(t)}{\sum_{i=1}^n \bar{u}_\theta(x_i, t)^2}} \quad (41)$$

Extending the projection to continuous space yields:

$$\tilde{u}_\theta(x, t) = u_\theta(x, t) \sqrt{\frac{c(t)}{\sum_{i=1}^n \bar{u}_\theta(x_i, t)^2}} \quad (42)$$

Derivation of Conservation of Combined Linear and Quadratic Integral

$$\min_y ||\bar{u}_\theta(x, t) - y||^2 \quad (43)$$

$$\text{s.t. } \Delta x y^\top y = c_1(t) \quad (44)$$

$$\Delta x \mathbf{1}_n^\top y = c_2(t)$$

The Lagrangian has the form:

$$\mathcal{L}(y, \lambda) = (\bar{u}_\theta(x, t) - y)^\top (\bar{u}_\theta(x, t) - y) \quad (45)$$

$$+ \lambda_1 (\Delta x y^\top y - c_1(t))$$

$$+ \lambda_2 (\Delta x \mathbf{1}_n^\top y - c_2(t))$$

Take partial derivatives and set to zero:

$$\partial \mathcal{L} / \partial y_i = -2(u_\theta(x_i, t) - y_i) \quad (46)$$

$$+ 2\lambda_1 \Delta x y_i + \lambda_2 \Delta x = 0$$

$$\partial \mathcal{L} / \partial \lambda_1 = \Delta x y^\top y - c_1(t) = 0 \quad (47)$$

$$\partial \mathcal{L} / \partial \lambda_2 = \Delta x \mathbf{1}_n^\top y - c_2(t) = 0 \quad (48)$$

From the first equation, we get:

$$y_i = \frac{2u_\theta(x_i, t) - \lambda_2 \Delta x}{2 - 2\lambda_1 \Delta x} \quad (49)$$

Substitute into constraint:

$$c_1(t) = \Delta x \sum_{i=1}^n \frac{(2u_\theta(x_i, t) - \lambda_2 \Delta x)^2}{(2 - 2\lambda_1 \Delta x)^2} \quad (50)$$

$$c_2(t) = \Delta x \sum_{i=1}^n \frac{2u_\theta(x_i, t) - \lambda_2 \Delta x}{2 - 2\lambda_1 \Delta x} \quad (51)$$

From second equation:

$$\lambda_2 \Delta x = \frac{-(2 - 2\lambda_1 \Delta x)c_2(t)}{n\Delta x} + \sum_{i=1}^n \frac{2u_\theta(x_i, t)}{n} \quad (52)$$

$$\begin{aligned} y_i &= \frac{2u_\theta(x_i, t) - \lambda_2 \Delta x}{2 - 2\lambda_1 \Delta x} \\ &= \frac{u_\theta(x_i, t) - \sum_{i=1}^n \frac{u_\theta(x_i, t)}{n}}{1 - \lambda_1 \Delta x} + \frac{c_2(t)}{n\Delta x} \end{aligned} \quad (53)$$

Let $\bar{u}_\theta = \sum_{i=1}^n \frac{u_\theta(x_i, t)}{n}$ and substitute the above into the first equation:

$$\begin{aligned} &= \Delta x \sum_{i=1}^n \left(\frac{(u_\theta(x_i, t) - \bar{u}_\theta)^2}{(1 - \lambda_1 \Delta x)^2} + \right. \\ &\quad \left. 2 \frac{u_\theta(x_i, t) - \bar{u}_\theta}{1 - \lambda_1 \Delta x} \frac{c_2(t)}{n\Delta x} + \frac{c_2^2(t)}{(n\Delta x)^2} \right) \end{aligned} \quad (54)$$

$u_\theta(x_i, t) - \bar{u}_\theta$ cancels out after the summation, so the middle

term is 0

$$c_1(t) = \Delta x \sum_{i=1}^n \left(\frac{(u_\theta(x_i, t) - \bar{u}_\theta)^2}{(1 - \lambda_1 \Delta x)^2} + \frac{c_2^2(t)}{(n \Delta x)^2} \right) \quad (55)$$

$$\frac{c_1(t)}{\Delta x} - \frac{c_2^2(t)}{n \Delta x^2} = \frac{1}{(1 - \lambda_1 \Delta x)^2} \sum_{i=1}^n (u_\theta(x_i, t) - \bar{u}_\theta)^2 \quad (56)$$

$$\frac{1}{1 - \lambda_1 \Delta x} = \sqrt{\frac{c_1(t)/\Delta x - c_2^2(t)/(n \Delta x^2)}{\sum_{i=1}^n (u_\theta(x_i, t) - \bar{u}_\theta)^2}} \quad (57)$$

Substitute back to the expression of y_i

$$y_i = (u_\theta(x_i, t) - \bar{u}_\theta) \sqrt{\frac{c_1(t)/\Delta x - c_2^2(t)/(n \Delta x^2)}{\sum_{i=1}^n (u_\theta(x_i, t) - \bar{u}_\theta)^2}} + \frac{c_2(t)}{n \Delta x} \quad (58)$$