
Learning Closure Relations using Differentiable Programming: An Example in Radiation Transport

A. J. Crilly

ac116@ic.ac.uk

Centre for Inertial Fusion Studies
The Blackett Laboratory, Imperial College London
London SW7 2AZ, United Kingdom

I-X Centre for AI In Science
Imperial College London, White City Campus
84 Wood Lane, London W12 0BZ, United Kingdom

B. Duhig

Centre for Inertial Fusion Studies
The Blackett Laboratory, Imperial College London
London SW7 2AZ, United Kingdom

N. Bouziani

Department of Mathematics
Imperial College London
London SW7 2AZ, United Kingdom

I-X Centre for AI In Science
Imperial College London, White City Campus
84 Wood Lane, London W12 0BZ, United Kingdom

Abstract

The continuous flow or ‘transport’ of a macroscopic system of particles is a high dimensional problem and therefore often solved using reduced order models. This necessarily introduces unknown closure relations into these models. In this work, we present a machine learning approach to finding accurate closure relations utilising differentiable programming. As a case study, we consider the transport of photons and use a literature radiation transport test problem as a training dataset. We present novel ML closures for a number of reduced order models which outperform their literature counterparts in both trained and unseen problems.

1 Introduction

Reduced order models provide a method to model physical systems efficiently but approximately when more accurate models may prove prohibitively expensive. These models often have ‘closure relations’, where in order to close the system of equations, one needs to encode higher order behaviour in a function of the variables which are being solved for. The primary objective of this work is to utilise differentiable programming to provide a proof of principle machine learning (ML) method to the task of deriving equation closures. By using a physics-driven model enhanced with ML we can obtain better performance than both pure data-driven and traditional numerical physics techniques. This methodology has cross-disciplinary application as the closure problem is ubiquitous in the

physical sciences. In our example, we will show that state-of-the-art reduced order models from the literature can be improved in accuracy by an order of magnitude using ML closures.

Machine learning is apt in the case of finding closures as these act as unknowns in otherwise physically motivated reduced models. We will investigate ML closures in reduced order solutions to a particular analytic result in radiation transport from Su and Olson[1]. Radiation transport is an important phenomenon in many fields of physics including astrophysics, solar physics, plasma physics and nuclear fusion. The Su & Olson problem considers a homogeneous medium which isotropically absorbs and emits radiation. Embedded within this medium is a transient and localised radiation source, Q . The solution describes how the radiation, W , and material, V , energy densities evolve over space, x , and time, τ . Further details can be found in the appendix. The data set considered is small within a ML context and therefore we need strong inductive bias in the models we train. These constraints will come from radiation transport theory and will improve the ML closures' ability to model unseen problems.

2 Learning closure relations using differentiable programming

The problem of learning closure relations can be specified as an optimisation problem. We have data we wish to match e.g. an analytical solution or experimental data. We also have a model with a number of trainable parameters, i.e. our reduced order system of equations, and a parametrised closure relation. Finally, we must define a loss between the reduced order model and analytic solution.

With these three components, one can utilise differentiable programming to compute gradients of the 'loss' with respect to our closure parameters. This involves back-propagating through the discretised form of the PDE model solution[2], which amounts to solving the adjoint equation of the PDE. As a result, we obtain an end-to-end differentiable architecture that can be used to learn closure relations, thanks to automatic gradient calculations. In this work, we utilise the ML framework *JAX*[3] and libraries *optax*[4] and *diffjax*[2]. These libraries perform the optimisation task with modern ML techniques, within a very small number of lines of user code. The main challenges lie in creating a numerically-stable spatially-discretised version of the PDEs and specifying an effective parametrisation of the closure which enforces the physical constraints.

2.1 Variable Eddington factor closure, $p(f)$

The first reduced order model we will consider is variable Eddington factor (VEF):

$$\frac{\partial W}{\partial \tau} + \frac{\partial F}{\partial x} + W = V + Q, \quad (1a)$$

$$\frac{\partial V}{\partial \tau} = W - V \quad (1b)$$

$$\frac{\partial F}{\partial \tau} + \frac{\partial}{\partial x}(pW) + F = 0, \text{ with closure } p \left(f = \frac{|F|}{W} \right). \quad (1c)$$

These equations appear in the radiation transport literature with various proposed closure relations $p(f)$, e.g. the Kershaw[5] and P1[6] closures. However, the literature closures cannot reproduce the Su Olson problem, as shown in Fig. 1 (a). We will use our ML approach to find an optimal closure and improve the VEF equations' prediction.

For optimisation purposes, we define the best solution as one which minimises the squared errors between the model solution and the Su Olson radiation energy density, W , giving the following loss function:

$$\mathcal{L} = \sum_i \sum_j (W_{\text{SO}}(x_j, \tau_i) - W_{\text{model}}(x_j, \tau_i; \theta))^2, \quad (2)$$

We also restrict to the time points for which the external source, Q , is on, $\tau \leq 10$. For these times, transport behaviour beyond diffusion is most significant, growing to a peak around $\tau = 10$. This also allows a smaller domain size and therefore fewer numerical time steps. The correct late time diffusion behaviour, $F \rightarrow -(1/3)(\partial W/\partial x)$, will be enforced by constraints on the closures. Shortening the training time further to exclude the $\tau = 10$ data caused a precipitous drop in accuracy.

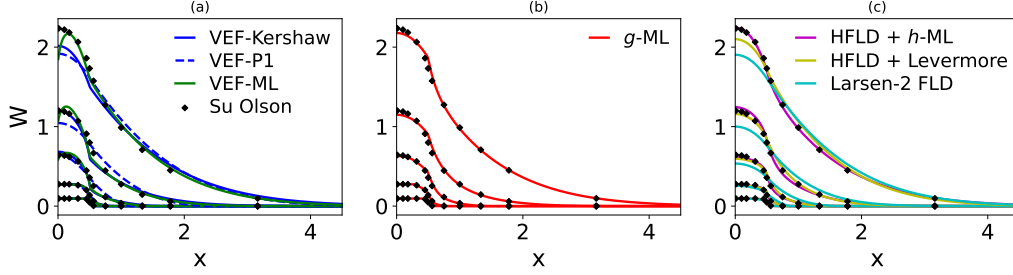


Figure 1: A comparison of energy densities (W) calculated with various reduced models and Su & Olson's analytic results. These are shown as function of position (x) at various times ($\tau = 0.1, 0.316, 1.0, 3.16, 10.0$), note that W monotonically increases in time.

Secondly, we must define a suitable parameterisation of the closure which respects the radiation transport limiting behaviour. This is done through a constrained Padé approximant:

$$p(f) = \frac{\sum_{n=0}^N a_n (1-f)^n}{b_0 + \sum_{m=1}^M b_m (1-f)^m}, \quad a_0 = 1, b_0 = 1 \text{ such that } p(1) = 1, \quad (3)$$

$$a_1 = b_1 - 2 \text{ such that } \frac{dp}{df}(1) = 2, \quad a_N = \frac{1}{3} \sum_{m=0}^M b_m - \sum_{n=0}^{N-1} a_n \text{ such that } p(0) = \frac{1}{3},$$

where the a_n and b_m are the trainable parameters we will optimise. The constrained Padé approximant form was chosen as it required a low number of trainable parameters, allowed easy implementation of physical constraints and is a more flexible function approximator than a truncated power series.

Finally, our end-to-end differentiable PDE model produces AD gradients which can be utilised by mainstream optimisation methods. The small amount of data also means that no batching is required, thus every space and time point is evaluated for each new updated parameter set. We solve the transport equations explicitly using literature, finite differencing methods; for details see the code in the supplementary material or code repository. The explicit time stepping means the speed of light restricts the numerical CFL time step. For a spatially resolution dx , $N_t = \frac{10}{C dx}$ time steps are needed to reach $\tau = 10$, where C is the Courant number. For our training we chose $C = 0.1$ and $dx = 0.015$, thus $N_t \sim 7000$ time steps of our reduced models must be taken in order to evaluate the loss function once. This highlights the need for differentiable programming as having efficient access to accurate gradient information will greatly reduced the number of loss function evaluations needed. For optimisation, we use the Adam optimiser with a decaying, piecewise-constant learning rate. The results from the optimal closure found using this ML method described above are shown in Fig. 1 (a).

These results point to an issue with VEF methods in this case. Indeed, in the Su Olson problem it is known that p is less than $1/3$ at the origin[6] (where $f = 0$ by symmetry). Improved agreement could be found by allowing $p(f = 0)$ to deviate away from $1/3$. This however removes the models ability to model the diffusive limit, occurring when $f \rightarrow 0$. These deficiencies point to the fact we are exiting the models range of validity and the accuracy of the model is no longer limited by the closure.

2.2 Third angular moment closure, $g(p)$

The third angular moment equations extend the VEF equations with the following additional equation:

$$\frac{\partial P}{\partial \tau} + \frac{\partial}{\partial x}(gF) + P = \frac{1}{3}(V + Q), \text{ with closure } g\left(p = \frac{P}{W}\right), \quad (4)$$

The third angular moment equations are a much less common approach than the VEF equations and no mainstream literature closures exist. Therefore, the presented ML methodology is vital in deriving a viable third order moment model. From its definition, we can enforce the following restrictions on the closure, $g(0) = 0$, $g(1) = 1$ and $dg/dp(p = 1) = 1$, using a constrained Padé approximant. Then we use the same optimisation method as for the VEF results to find the optimal values of a_n

and b_m for the novel $g(p)$ closure to reproduce the Su Olson solution. The results are summarised in Fig. 1 (b). As shown, the agreement between the model and analytic solution is improved for this higher angular moment scheme and ML closure.

2.3 Consequences for flux limited diffusion

Given the results from the higher order schemes, it is obligatory to investigate the consequences for arguably the most popular closure model - Flux Limited Diffusion[6] (FLD). The FLD model resembles the VEF model (Eq. (1)) but with the flux, F , approximated using the following equation:

$$F = -\lambda \frac{\partial W}{\partial x}, \quad (5)$$

where λ is known as the flux limiter; when $\lambda \rightarrow 1/3$ pure diffusion is obtained. Various literature closures for λ exist, such as Larsen[7], but as with the VEF equations no single FLD closure will be able to capture the required transport behaviour in the Su & Olson problem. Instead, we propose a novel higher order flux limited diffusion (HFLD) model using the third order moment equations, but with the radiative flux PDE replaced with Eq. (5) and use the following flux limiter:

$$\lambda = \frac{-1 + \sqrt{1 + 4R^2 h(p)}}{2R^2} \quad \text{where } R \equiv \frac{1}{W} \frac{\partial W}{\partial x}. \quad (6)$$

This HFLD model is enabled by the closed form of the third order moment model we were able to obtain using ML methods. We have also introduced a new flux limiter closure term, $h(p)$. From the literature, Levermore[8] proposes $h(p) = p$ but then uses other literature VEF closures to relate p to R rather than explicitly solving Eq. (4) for p . For our ML model, we again choose $h(p)$ to have a constrained Padè form, with $h(0) = 0$, $h(1) = 1$ and $h(1/3) = 1/3$ to ensure the correct limiting behaviours. We also enforce $h'(0) > 0$ to (non-rigorously) prevent negative values which would allow complex λ values. Training of the HFLD $h(p)$ closure using the Su Olson results proceeds in a similar fashion to both the VEF and third order moment models, albeit with a more restrictive CFL condition on the numerical time step.

3 Application of the ML-closures in a modified problem

As the ML-closures discussed above are specifically trained on the Su-Olson problem, one might question their generic applicability in other problems. Here, we consider a modification to the Su-Olson test problem in which there is an additional reflective symmetry about $x = 2$. This introduces another region of zero flux but from symmetric converging flow ($p > 1/3$) rather than symmetric diverging flow ($p < 1/3$) – this acts to test a closure’s ability to capture anisotropy in two ‘fluxless’ configurations, one of which is absent from the training data set. We also numerically solved the full transport problem using discrete ordinates[6], S_N , with ordinate order $N = 32$. We compare the reduced order models and S_{32} results in Fig. 2 (a) & (b), and the reduced model losses in Fig. 2 (c).

We see that the g-ML model generalises well as the moment equations allow the radiation pressure to respond to flux gradients. The VEF-ML model struggles in both ‘fluxless’ ($f = 0$) regions as the requirement needed for capturing diffusive behaviour ($p(f \rightarrow 0) \rightarrow 1/3$) creates large inaccuracy, as discussed in Section 2.1. For HFLD, the novel ML $h(p)$ closure captures the source regions correctly in both problems. It performs similarly to the HFLD Levermore closure ($h(p) = p$) in the converging flow region (about $x = 2$) of the modified problem. This is a strong indication that the ML closure has not been over-fit to the original Su Olson training data. It was found that the diffusive limit, $h(1/3) = 1/3$, constraint is particularly important in preventing over-fitting.

4 Discussion

In this work, we describe a methodology for using accurate data to learn closure relations in reduced order models using differentiable programming. We use the example of the Su Olson radiation transport problem[1] to find novel closures to the variable Eddington factor, third order angular moment and flux limited diffusion models. All machine learnt models exhibit a lower mean-squared distance to the Su Olson analytic results than their respective literature counterparts, see Fig. 2 (c).

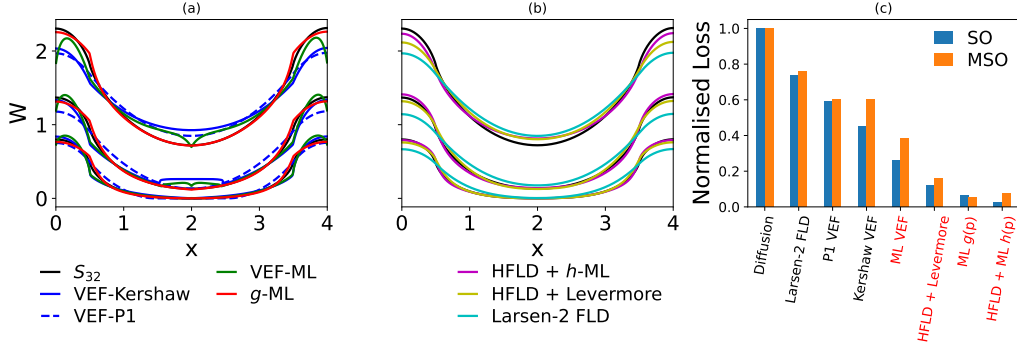


Figure 2: (a) & (b) The energy density in the modified Su Olson problem. The numerical results from the reduced order equations and discrete ordinates are shown at various times ($\tau = 1.5, 4.0, 10.0$). (c) Plot of the loss normalised to the pure diffusion result for the various models considered in the original (SO) and modified (MSO) Su Olson problem. The loss for the modified Su Olson problem used the S_{32} results as the ground truth values. The models introduced in this work have their labels highlighted in red.

The performance of trained ML models did not suffer when a modification was made to the Su Olson problem - still outperforming literature counterparts on an unseen problem. The third order angular moment performed best, proving the most consistent across problem types. The higher order flux-limited diffusion models still proved an improvement. This shows that the ML methods presented can produce more accurate closures as part of reduced models which respect physical laws and generalise to unseen problems. In the case of VEF after the ML closure is trained, we find the accuracy is not limited by the closure but instead by the assumptions of the reduced model itself.

Closure problems are ubiquitous in the physical sciences and machine learning methods have been employed to address the closure problem in the literature. Trained neural networks have proven successful closure models, as long as physical constraints are addressed. Huang *et al.* directly trained the network to match the higher order solution's closure term[9, 10] rather than training to match the dynamics at a reduced order. Han *et al.*[11] took a similar approach but found that an end-to-end learning approach was beneficial. In contrast, we learnt the closures in an end-to-end approach using only the dynamics of the lowest order terms in the loss, i.e. the energy density, W . We chose not to use neural networks to specify the closure for two reasons. Firstly, enforcing hard inductive biases on the closure is important for ensuring accurate physical behaviour. Including these biases in a loss term would not strictly enforce these. While neural network architecture choices could be used to include these biases, we opted for using a simpler parametrisation of the closures. Secondly, a low number of trainable parameters was preferred due to the sparsity of training data, therefore favouring our Padé approximant approach. With more training data, a neural network could be used within the differentiable programming framework described here.

In this work, we trained closures with PDE solutions using selected numerical methods. A different choice of methods might alter the numerical solutions obtained from the same underlying PDE, affecting the optimal parameter values. However, with sufficiently accurate methods the dynamics of the numerical solution are the same. Therefore, the optimal closure relation parameters found using the differentiable programming methodology are applicable if a different numerical method is utilised, up to fine-tuning. Finally, it should be noted that the closure relation problem is not unique to radiation transport and therefore a similar approach could be considered in other transport problems throughout the physical sciences.

Acknowledgments and Disclosure of Funding

The authors would like to thank Prof. S. J. Rose and Dr. A. Joglekar for many useful discussions and helpful comments on this work.

This project is supported by the Eric and Wendy Schmidt AI in Science Postdoctoral Fellowship, a Schmidt Futures program.

Appendix

4.1 The Su & Olson radiation transport problem

We summarise the relevant details on the Su & Olson test problem in the following. The coupled system of equations to be solved are the following:

$$\left(\epsilon \frac{\partial}{\partial \tau} + \mu \frac{\partial}{\partial x} + 1 \right) U(x, \mu, \tau) = \frac{1}{2} V(x, \tau) + \frac{1}{2} Q(x, \tau), \quad (7)$$

$$\frac{\partial V}{\partial \tau} = W(x, \tau) - V(x, \tau), \quad (8)$$

where:

$$W(x, \tau) \equiv \int_{-1}^{+1} U(x, \mu, \tau) d\mu. \quad (9)$$

Here U , W , V and Q are the scaled radiation intensity, radiation energy density, material energy density and external radiation source respectively. The coordinates x , μ , and τ are the normalised distance, cosine of angle of propagation and the normalised time respectively.

The analytic solutions given are performed for a source term of the following form:

$$Q(x, \tau) = \frac{1}{2x_0} [\Theta(x + x_0) - \Theta(x - x_0)] [\Theta(\tau) - \Theta(\tau - \tau_0)] \quad (10)$$

where Θ is the Heaviside function.

Su and Olson provide analytic solutions at various x and τ values for $\epsilon = 1$, $x_0 = 0.5$ and $\tau_0 = 10$.

4.2 ML closure coefficient tables

Here we give tables of the Padé coefficients for each ML closure. For completeness, every coefficient is given, even those set by physical constraints.

Table 1: Coefficient tables for VEF-ML, g-ML and HFLD-ML closures respectively. Both VEF-ML and g-ML use Padé forms $f(x) = A(1-x)/B(1-x)$ while HFLD-ML uses $f(x) = A(x)/B(x)$, where $A()$ and $B()$ are power series with coefficients a_n and b_m respectively.

a_0	1.0	b_0	1.0	a_0	1.0	b_0	1.0	a_0	0.0	b_0	1.0
a_1	-1.97	b_1	3.28e-2	a_1	-8.97e-1	b_1	1.03e-1	a_1	1.12e-1	b_1	-3.85
a_2	2.03	b_2	-3.29e-1	a_2	1.18e-1	b_2	-1.30e-1	a_2	2.68e-1	b_2	5.16
a_3	-1.30	b_3	-6.00e-1	a_3	1.99e-1	b_3	-2.84e-1	a_3	-8.36e-2	b_3	1.66
a_4	3.89e-2			a_4	2.76e-1			a_4	4.61	b_4	9.39e-1
a_5	2.37e-1			a_5	-6.96e-1						

References

- [1] Bingjing Su and Gordon L Olson. An analytical benchmark for non-equilibrium radiative transfer in an isotropically scattering medium. *Annals of Nuclear Energy*, 24(13):1035–1055, 1997.
- [2] Patrick Kidger. *On Neural Differential Equations*. PhD thesis, University of Oxford, 2021.

- [3] James Bradbury, Roy Frostig, Peter Hawkins, Matthew James Johnson, Chris Leary, Dougal Maclaurin, George Necula, Adam Paszke, Jake VanderPlas, Skye Wanderman-Milne, and Qiao Zhang. JAX: composable transformations of Python+NumPy programs, 2018.
- [4] Igor Babuschkin, Kate Baumli, Alison Bell, Surya Bhupatiraju, Jake Bruce, Peter Buchlovsky, David Budden, Trevor Cai, Aidan Clark, Ivo Danihelka, Antoine Dedieu, Claudio Fantacci, Jonathan Godwin, Chris Jones, Ross Hemsley, Tom Hennigan, Matteo Hessel, Shaobo Hou, Steven Kapturowski, Thomas Keck, Iurii Kemaev, Michael King, Markus Kunesch, Lena Martens, Hamza Merzic, Vladimir Mikulik, Tamara Norman, George Papamakarios, John Quan, Roman Ring, Francisco Ruiz, Alvaro Sanchez, Rosalia Schneider, Eren Sezener, Stephen Spencer, Srivatsan Srinivasan, Wojciech Stokowiec, Luyu Wang, Guangyao Zhou, and Fabio Viola. The DeepMind JAX Ecosystem, 2020.
- [5] David S Kershaw. Flux limiting natures own way—a new method for numerical solution of the transport equation. Technical report, Lawrence Livermore National Lab.(LLNL), Livermore, CA (United States), 1976.
- [6] Gordon L Olson, Lawrence H Auer, and Michael L Hall. Diffusion, p1, and other approximate forms of radiation transport. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 64(6):619–634, 2000.
- [7] Bingjing Su. Variable eddington factors and flux limiters in radiative transfer. *Nuclear science and engineering*, 137(3):281–297, 2001.
- [8] C David Levermore. Relating eddington factors to flux limiters. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 31(2):149–160, 1984.
- [9] Juntao Huang, Yingda Cheng, Andrew J Christlieb, and Luke F Roberts. Machine learning moment closure models for the radiative transfer equation i: directly learning a gradient based closure. *Journal of Computational Physics*, 453:110941, 2022.
- [10] Juntao Huang, Yingda Cheng, Andrew J Christlieb, Luke F Roberts, and Wen-An Yong. Machine learning moment closure models for the radiative transfer equation ii: Enforcing global hyperbolicity in gradient-based closures. *Multiscale Modeling & Simulation*, 21(2):489–512, 2023.
- [11] Jiequn Han, Chao Ma, Zheng Ma, and Weinan E. Uniformly accurate machine learning-based hydrodynamic models for kinetic equations. *Proceedings of the National Academy of Sciences*, 116(44):21983–21991, 2019.