
Multi-Constitutive Neural Network for Large Deformation Poromechanics Problem

Qi Zhang^{1†}, Yilin Chen^{1†}, Ziyi Yang^{2†}, Eric Darve^{2,3}

¹ Department of Civil and Environmental Engineering, Stanford University

² Department of Mechanical Engineering, Stanford University

³ Institute for Computational and Mathematical Engineering, Stanford University

† Equal Contribution

{qzhang94, yilinc2, ziyi.yang, darve}@stanford.edu

Abstract

In this paper, we study the problem of large-strain consolidation in poromechanics with deep neural networks (DNN). Given different material properties and different loading conditions, the goal is to predict pore pressure and settlement. We propose a novel method “multi-constitutive neural network” (MCNN) such that one model can solve several different constitutive laws. We introduce a one-hot encoding vector as an additional input vector, which is used to label the constitutive law we wish to solve. Then we build a DNN which takes (\hat{X}, \hat{t}) as input along with a constitutive law label and outputs the corresponding solution. It is the first time, to our knowledge, that we can evaluate multi-constitutive laws through only one training process while still obtaining good accuracies. We found that MCNN trained to solve multiple PDEs outperforms individual neural network solvers trained with PDE in some cases.

1 Introduction

In the last decade, machine learning, especially deep neural networks (DNN) have yielded revolutionary results across diverse disciplines, and a fair amount of research has also been done related to differential equations that appear in many engineering fields. In the work done by Raissi et al. [15, 16, 17], the physics-informed neural network (PINN) is proposed to solve a given physics problem described by a specific partial differential equation (PDE). However, there might exist multiple PDEs for a given problem, corresponding to choosing different constitutive laws. Traditionally, multiple models need to be built and trained to solve them separately, which might be computationally expensive and time-consuming. The multi-constitutive problem is especially important in large deformation poromechanics where many different constitutive laws which are derived from distinct forms of strain-energy density functions [10] can exist for one application [22]. In this paper, we propose a multi-constitutive neural network (MCNN); it uses an input vector which encodes the choice of constitutive law as an extra input to the network. Formally, let us index each constitutive law by some integer i . Denote $J_i(x, t)$ the solution to the PDE (or the sought after function) with constitutive law i , as a function of x and t . Our neural network is a function $\tilde{J}(x, t, i; \theta)$, where θ are the usual DNN parameters (weights and biases), i is the index of the constitutive law, and $\tilde{J}(x, t, i; \theta)$ is an approximation of the solution $J_i(x, t)$ for law i .

By solving the large-strain consolidation problem [6, 14, 18, 19, 22], we have demonstrated the generalization capability of our model when incorporating multiple constitutive laws; our approach also requires fewer model parameters. The novel contributions of this paper are twofold: (1) A general deep learning framework MCNN that ensembles multi-constitutive laws and solves the resulting governing equations “essentially simultaneously” (simply by changing the input i , encoded as a

one-hot vector); (2) Simultaneous training is more efficient and provides better generalization than individual training. In our benchmarks, for the second and the third cases, MCNN is more accurate than individual PINN models which are trained independently for each constitutive law.

2 Multi-Constitutive Laws in Large Deformation Poromechanics

We briefly review the multi-constitutive laws used in the large-strain consolidation problem. In order to eliminate the effect of choice of units in data, we employ a non-dimensionalization technique, which means variables are described in their dimensionless form. In this paper, three constitutive laws of hyper-elasticity are investigated [10], and here we give their forms after incorporating the momentum balance equation and effective stress principle [3, 4, 5, 21, 22]:

(i) Saint-Venant Kirchhoff law (law 1):

$$\frac{\partial \hat{p}}{\partial \hat{X}} = \frac{3J^2 - 1}{2} \frac{\partial J}{\partial \hat{X}} \quad (1)$$

(ii) Modified Saint-Venant Kirchhoff law (law 2):

$$\frac{\partial \hat{p}}{\partial \hat{X}} = \left[\hat{\gamma} \frac{1 - \log J}{J^2} + \hat{\mu} (3J^2 - 1) \right] \frac{\partial J}{\partial \hat{X}} \quad (2)$$

(iii) Neo-Hookean law (law 3):

$$\frac{\partial \hat{p}}{\partial \hat{X}} = \left[\hat{\gamma} \frac{1 - \log J}{J^2} + \hat{\mu} \left(1 + \frac{1}{J^2} \right) \right] \frac{\partial J}{\partial \hat{X}} \quad (3)$$

where in above three equations, J is a measure of the deformation gradient, \hat{p} is the dimensionless fluid (pore) pressure, $\hat{X} \in [0, 1]$ is the dimensionless coordinate, $\hat{\gamma}$ and $\hat{\mu}$ are the dimensionless Lamé constants. The mass balance equation is independent of multi-constitutive laws and is given by

$$\frac{\partial J}{\partial \hat{t}} - \frac{1}{\varphi_0^3} \frac{\partial}{\partial \hat{X}} \left[\frac{(J - 1 + \varphi_0)^3}{J^2} \frac{\partial \hat{p}}{\partial \hat{X}} \right] = 0 \quad (4)$$

where \hat{t} is the dimensionless time and φ_0 is the reference porosity. In Eq. (4), we have considered the dynamic change of permeability through the Kozeny-Carman equation [3, 22]. For the initial conditions (ICs) and boundary conditions (BCs), on the top surface ($\hat{X} = 0$), the surface overburden leads to a Dirichlet boundary condition $J = \bar{J}$ [22], while at the bottom ($\hat{X} = 1$), the no-flow condition leads to $\partial J / \partial \hat{X} = 0$. The initial condition is $J = 1$ when $\hat{t} = 0$.

To summarize, we have one governing Eq. (4) and three constitutive laws (with proper ICs & BCs). Let's index each constitutive law by i ($i = 1, 2, \text{ or } 3$). We want to build a model that takes as input i and outputs the solution of Eq. (1), (2), or (3) (coupled with Eq. (4) which is common to all). We discuss our approach in the next section.

3 Multi-Constitutive Neural Network

We propose the Multi-Constitutive Neural Network (MCNN) as a novel modeling approach for problems with multiple constitutive laws. As illustrated in Figure 1, we introduce a one-hot constitutive law encoding vector \vec{e} as an extra input to a PINN. Concretely, we use (1, 0, 0), (0, 1, 0), (0, 0, 1) to represent Eqns. (1), (2), and (3), respectively. Therefore, the dimension of each data sample is 5 (\hat{X} , \hat{t} , and the one-hot encoding vector \vec{e} of length 3). The network output J (after possible final transformation) can be substituted into the three PDEs to calculate the 3×1 residual vector through automatic differentiation (AD) [1, 11, 20]. The final scalar PDE loss for one sample is the square of inner product of this residual vector with the same one-hot encoding for this sample. The other loss terms, including boundary and initial conditions closely follow standard PINN (see [2, 9] for details). To be more specific, the scalar loss \mathcal{L} of one training sample could be mathematically expressed as:

$$\mathcal{L} = \left\{ \sum_{i=1}^3 e_i f_i \left[\frac{\partial}{\partial \hat{t}} J(\hat{X}, \hat{t}, \vec{e}), \frac{\partial}{\partial \hat{X}} J(\hat{X}, \hat{t}, \vec{e}), \frac{\partial^2}{\partial \hat{X}^2} J(\hat{X}, \hat{t}, \vec{e}) \right] \right\}^2 + \mathcal{L}_{BI} \quad (5)$$

$$\mathcal{L}_{BI} = \begin{cases} [J(0, \hat{t}, \vec{e}) - \bar{J}]^2 & \hat{X} = 0, \hat{t} > 0 \\ \left[\frac{\partial}{\partial \hat{X}} J(\hat{X}, \hat{t}, \vec{e}) \Big|_{\hat{X}=1} \right]^2 & \hat{X} = 1, \hat{t} > 0 \\ [J(\hat{X}, 0, \vec{e}) - 1]^2 & \hat{t} = 0, 0 < \hat{X} < 1 \\ 0 & \hat{t} > 0, 0 < \hat{X} < 1 \end{cases} \quad (6)$$

where \mathcal{L}_{BI} refers to the loss from the boundary and initial conditions, e_i is the i^{th} component of vector \vec{e} , and $f_i[\cdot, \cdot, \cdot]$ represents the final PDE form for the i^{th} constitutive law, which is obtained by combining Eq. (1), (2), or (3) with Eq. (4).

The material properties and network architectures are presented in Table 1 and 2. By leveraging the constitutive law encoding vector, we are able to assimilate multiple constitutive laws and solve all the resulting governing equations through simultaneous training.

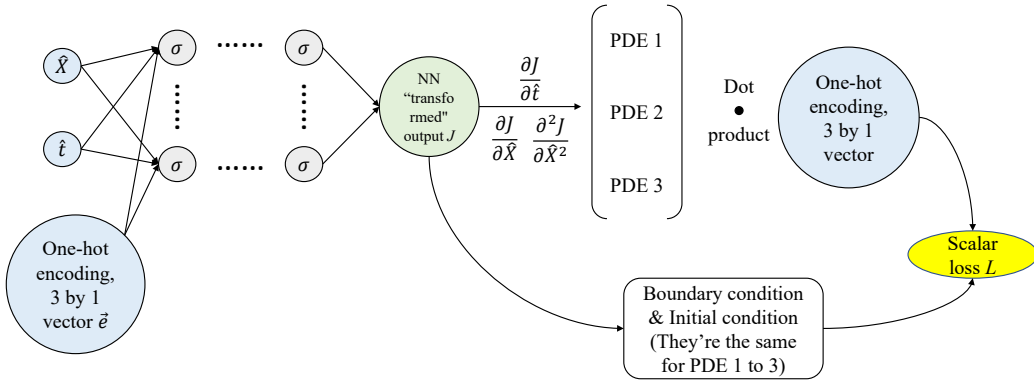


Figure 1: A schematic diagram of our proposed MCNN. The inputs (left) include \hat{X} , t and the law encoding vector \vec{e} . The σ in the neural network represents the activation function; weights and biases are not shown for simplicity. Time and spatial derivatives of J for the PDE formulations are computed through AD [1, 11, 20].

Table 1: Material properties (same for all three PDEs).

$\hat{\gamma}$	$\hat{\mu}$	φ_0	\bar{J}	\hat{t}
1/3	1/3	0.3	0.8	$0 \leq \hat{t} \leq 1$

Table 2: Model parameters of the neural networks. The total size of the training set for MCNN is the same as each independent PINN. All architectures use 5 (hidden layers) \times 50 (neurons) with \tanh as the activation function. The test sets are always equispaced within the domain. Due to the issue of numerical instability of the Saint-Venant Kirchhoff law [22], we adopt a larger number of epochs for law 1 to make the optimization process more stable. When the number of epochs is fixed, we tune the learning rate and find that the best value is $5e-4$. These learning rates and numbers of epochs are also typical values used in [7, 8, 9].

DNN	Training set	Test set	Optim.	Rate	Epochs
MCNN (ours)	1000 (per law)	10^4 (per law)	Adam [12]	$5e-4$	10^5
PINN of law 1	3000	10^4	Adam	$5e-4$	5×10^4
PINN of law 2	3000	10^4	Adam	$5e-4$	2×10^4
PINN of law 3	3000	10^4	Adam	$5e-4$	10^4

4 Results and Discussions

In this section we evaluate the inference accuracy of MCNN by simultaneously training with three different constitutive laws. To confirm the effectiveness of MCNN, for baseline models, we train an independent PINN respectively for each constitutive law and also calculate the reference solution. To build our DNN structure, we multiply the network output (see Figure 1) with \hat{X} and add \bar{J} ; this becomes our final output J . In this way, our output automatically satisfies the Dirichlet boundary condition at $\hat{X} = 0$. The reference prediction J is computed using a conditionally stable finite difference (FD) algorithm [18, 22] with $\Delta \hat{t} = 10^{-5}$ and $\Delta \hat{X} = 0.02$ (second-order accurate in space and first-order accurate in time).

We compute the relative errors to the reference predictions for MCNN and individual PINNs using the formula $\sqrt{\text{Ave} [(J/J_{\text{ref}} - 1)^2]}$, as shown in Table 3. Surprisingly, MCNN achieves lower relative error than individual PINN on the cases of law 2 and law 3, even with smaller number of training points for each constitutive law. This shows the data efficiency of MCNN from simultaneous training. Since the trend of J with \hat{X} and \hat{t} is fixed (similar to the trend of excess pore pressure in Terzaghi's consolidation equation [5]), the training points for different one-hot encoding vectors can "communicate" useful information and benefit from each other. This might explain why we observe better generalization for our method.

In Figure 2, we plot the prediction of J for MCNN and FD versus \hat{X} and \hat{t} , from which we can see that MCNN is able to distinguish among three different encoding vectors. In other words, MCNN successfully learns different constitutive laws. In Figure 3, we visualize the dimensionless settlement \hat{U} for each constitutive law. Here \hat{U} is obtained from a trapezoidal numerical integration of J since $J = 1 + \partial \hat{U} / \partial \hat{X}$ and $\hat{U}(\hat{X} = 1) = 0$. MCNN achieves accurate predictions for all three constitutive laws, which again demonstrates the powerful capacity of MCNN in modeling the multi-constitutive problem.

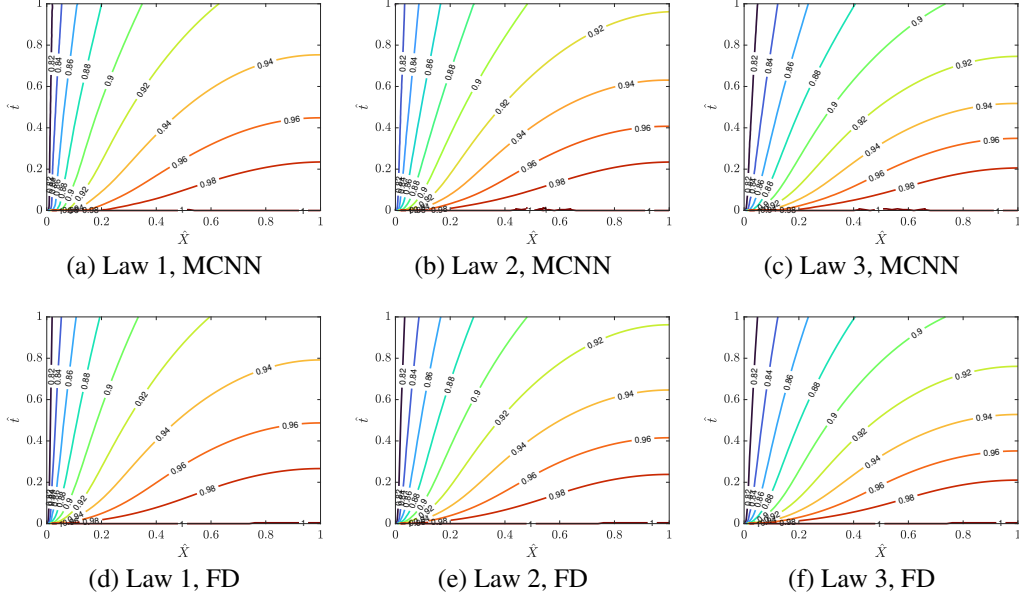


Figure 2: Plots of predicted J with MCNN and FD. MCNN is able to make an accurate inference for each constitutive law.

5 Conclusion

In this paper, we propose a novel modeling technique called Multi-Constitutive Neural Network (MCNN) to study multi-constitutive problems. Different from classical PINN, by using a law-

Table 3: Relative error of the proposed MCNN and PINNs for different constitutive laws. Although MCNN is learning to solve 3 PDEs simultaneously, it outperforms PINNs optimized independently on the cases of law 2 and law 3.

Method	Law 1	Law 2	Law 3
MCNN (Ours)	0.3864%	0.1316%	0.1658%
Independent PINN	0.2888%	0.2392%	0.4089%

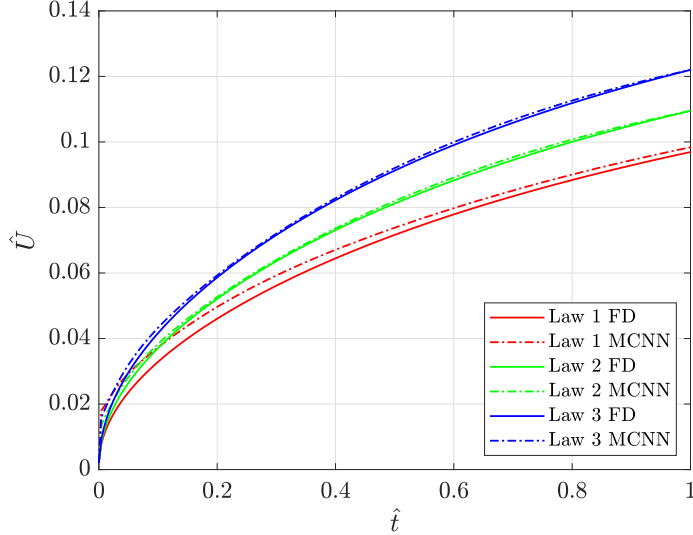


Figure 3: A visualization of MCNN predictions and reference solutions for the dimensionless settlement \hat{U} . Predictions from MCNN, even though they are generated by the same model, are very accurate for all three constitutive laws.

encoding input vector, MCNN is able to use a single model to solve for different constitutive laws. Using applications from nonlinear large deformation poromechanics problem, we demonstrated that MCNN could achieve good accuracies. For future work, we plan on studying encoding vectors of the form (a, b, c) with $a + b + c = 1$, $a \geq 0$, $b \geq 0$, $c \geq 0$ (instead of $(1,0,0)$ for example). This is akin to the idea of fractional derivatives [13].

Broader impact

As described and benchmarked in this paper, our approach opens the path for endowing deep learning with powerful capacity to study poromechanics. Poromechanics deals with the basic interactions between solid and fluid, which is fundamental in reservoir simulations, geothermal engineering, and oil and gas exploration. As a result, we expect MCNN to have an impact in these fields in the future. For example, in reservoir simulation, by using MCNN, we could compare the differences between Darcy and non-Darcy flows using a single training process. This research may have some negative impacts on society because it may encourage over-exploitation of oil and gas sites, pollution of sites (e.g., fracking), and may lead to an overall increase in production and consumption of non-renewable energy sources, contributing for example to global warming.

References

- [1] A. G. Baydin, B. A. Pearlmutter, A. A. Radul, J. M. Siskind, Automatic differentiation in machine learning: a survey, *Journal of Machine Learning Research* 18 (2017).
- [2] J. Berg, K. Nyström, A unified deep artificial neural network approach to partial differential equations in complex geometries, *Neurocomputing* 317 (2018).

- [3] J. Choo, Large deformation poromechanics with local mass conservation: An enriched Galerkin finite element framework, *International Journal for Numerical Methods in Engineering* 116 (2018) 66–90.
- [4] O. Coussy, Poromechanics, John Wiley and Sons (2004).
- [5] E. Detournay, A. H. D. Cheng, Fundamentals of poroelasticity, *Analysis and Design Methods: Principles, Practice and Projects* (1993) 113–171.
- [6] R. E. Gibson, G. L. England, M. J. L. Hussey, The theory of one-dimensional consolidation of saturated clays: I. Finite non-linear consolidation of thin homogeneous layers, *Géotechnique* 17 (1967) 261–273.
- [7] Y. Guo, X. Cao, B. Liu, M. Gao, Solving partial differential equations using deep learning and physical constraints, *Applied Sciences* 10 (2020).
- [8] E. Haghighat, M. Raissi, A. Moure, H. Gomez, R. Juanes, A deep learning framework for solution and discovery in solid mechanics, <https://arxiv.org/abs/2003.02751> (2020).
- [9] Q. He, D. Barajas-Solano, G. Tartakovsky, A. M. Tartakovsky, Physics-informed neural networks for multi-physics data assimilation with application to subsurface transport, *Advances in Water Resources* (2020).
- [10] G. A. Holzapfel, Nonlinear solid mechanics, a continuum approach for engineering, John Wiley and Sons (2000).
- [11] D. Z. Huang, K. Xu, C. Farhat, E. Darve, Learning constitutive relations from indirect observations using deep neural networks, *Journal of Computational Physics* 416 (2020).
- [12] D. P. Kingma, J. Ba, Adam: A method for stochastic optimization, published at International Conference on Learning Representations (ICLR) 2015, <https://arxiv.org/abs/1412.6980> (2014).
- [13] C. Li, A. Chen, Numerical methods for fractional partial differential equations, *International Journal of Computer Mathematics* 95 (2018) 1048–1099.
- [14] C. W. MacMinn, E. R. Dufresne, J. S. Wettlaufer, Large deformation of a soft porous material, *Physical Review Applied* 5 (2016).
- [15] M. Raissi, P. Perdikaris, G. E. Karniadakis, Physics informed deep learning (part i): data-driven solutions of nonlinear partial differential equations, <https://arxiv.org/abs/1711.10561> (2017).
- [16] M. Raissi, P. Perdikaris, G. E. Karniadakis, Physics informed deep learning (part ii): data-driven discovery of nonlinear partial differential equations, <https://arxiv.org/abs/1711.10566> (2017).
- [17] M. Raissi, P. Perdikaris, G. E. Karniadakis, Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations, *Journal of Computational Physics* 378 (2019) 686–707.
- [18] A. P. S. Selvadurai, A. P. Suvorov, Coupled hydro-mechanical effects in a poro-hyperelastic material, *Journal of the Mechanics and Physics of Solids* 91 (2016) 311–333.
- [19] K. H. Xie, C. J. Leo, Analytical solutions of one-dimensional large strain consolidation of saturated and homogeneous clays, *Computers and Geotechnics* 31 (2004) 301–314.
- [20] K. Xu, E. Darve, Learning constitutive relations using symmetric positive definite neural networks, <https://arxiv.org/abs/2004.00265> (2020).
- [21] Q. Zhang, Hydromechanical modeling of solid deformation and fluid flow in the transversely isotropic fissured rocks, *Computers and Geotechnics* 128 (2020).
- [22] P. Zheng, K. Zhang, On the effective stress law and its application to finite deformation problems in a poroelastic solid, *International Journal of Mechanical Sciences* 161&162 (2019).