Material modeling via Thermodynamics-based Artificial Neural Networks

Filippo Masi[a,b], Ioannis Stefanou[a], Paolo Vannucci[c], Victor Maffi-Berthier[b]

[a] *Institut de Recherche en Génie Civil et Mécanique, Ecole Centrale de Nantes*
[b] *Ingérop Conseil et Ingénierie*
[c] *LMV, Université de Versailles et Saint-Quentin*

From data-driven to physics-based neural networks material modeling

- Machine Learning methods and, in particular, **Artificial Neural Networks** (ANNs) have demonstrated promising capabilities in material constitutive modeling. One of the main drawbacks of such approaches is the **lack of a rigorous frame based on the laws of physics**.
- We propose a **new class of data-driven, physics-based, neural networks for constitutive modeling** of strain rate independent processes at the material point level, which we define as **Thermodynamics-based Artificial Neural Networks** (TANNs).

Thermodynamics principles

Under isothermal transformations and small strain regime, the **first law of thermodynamics** reads

$$\dot{\mathsf{F}} = \sigma \cdot \dot{\varepsilon} - \mathsf{D},$$

with F and $\dot{\mathsf{F}}$ the specific Helmholtz free energy (per unit volume) and its rate of change; σ the Cauchy stress tensor; ε and $\dot{\varepsilon}$ the small strain tensor and its rate of change; D the rate of mechanical dissipation. '$\cdot$' denotes contraction of adjacent indices.
From the **second law of thermodynamics** the rate of dissipation must be non-negative,

$$\mathsf{D} \geq 0.$$

For strain-rate independent materials

$$\mathsf{F} := \widetilde{\mathsf{F}}\left(\varepsilon, \mathcal{Z}\right) \quad \text{and} \quad \mathsf{D} := \widetilde{\mathsf{D}}\left(\varepsilon, \mathcal{Z}, \dot{\mathcal{Z}}\right),$$

where D is homogeneous of first-order (rate independency) in $\dot{\mathcal{Z}}$ and $\mathcal{Z} = (\zeta_i, \dots, \zeta_N)$ denotes a set of N (additional) internal state variables. It can be proven that

$$\sigma = \frac{\partial \mathsf{F}}{\partial \varepsilon}, \quad \mathsf{D} = -\sum_i \frac{\partial \mathsf{F}}{\partial \zeta_i} \cdot \dot{\zeta}_i = \sum_i \chi_i \cdot \dot{\zeta}_i.$$

Artificial Neural Networks

ANNs can be regarded as non-linear operators, composed of an assembly of mutually connected processing units−nodes−, which take an input signal $\mathcal{I}$ and return the output $\mathcal{O}$, namely

$$\mathcal{O} = \mathsf{ANN}@\mathcal{I}.$$

The signal flows from layer $(l-1)$ to layer (l) according to

$$p_k^{(l)} = \mathcal{A}^{(l)}(z_k^{(l)}),$$

with

$$z_k^{(l)} = \sum_s (w_{ks}^{(l)} p_s^{(l-1)}) + b_k^{(l)},$$

where $p_k^{(l)}$ is the output of node k, at layer (l); $\mathcal{A}^{(l)}$ is the activation function of layer (l); $w_{ks}^{(l)}$ is the *weight* between the s-th node in layer $(l-1)$ and the k-th node in layer (l); and $b_k^{(l)}$ are the *biases* of layer (l).

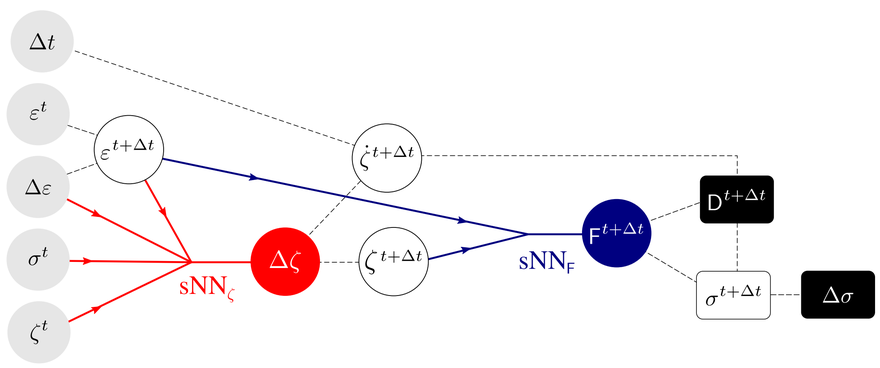
References

(+) Houlsby and Puzrin, 2007. Principles of Hyperplasticity. Springer Science Business Media.

(†) Masi et al., 2020. Thermodynamics-based Artificial Neural Networks for constitutive modeling. arXiv preprint arXiv:2005.12183.

(‡) Mozaffar et al., 2019. Deep learning predicts path-dependent plasticity. PNAS 116(52), 6414-26420.

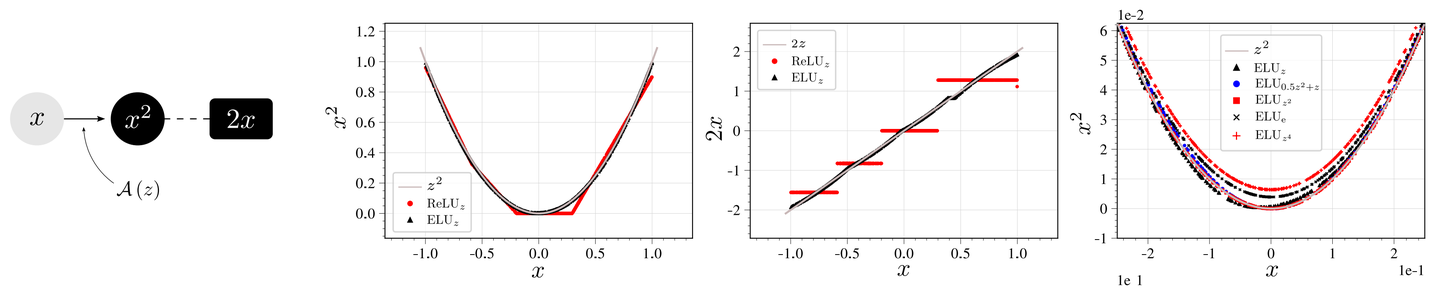
Architecture



1. $\varepsilon^{t+\Delta t} := \varepsilon^t + \Delta\varepsilon$
2. $\Delta\zeta = \mathsf{sNN}_\zeta@\left(\varepsilon^{t+\Delta t}, \Delta\varepsilon^t, \sigma^t, \zeta^t\right)$
3. $\dot{\zeta}^{t+1} \approx \frac{\Delta\zeta}{\Delta t}$, $\zeta^{t+1} := \zeta^t + \Delta\zeta^t$
4. $\mathsf{F}^{t+\Delta t} = \mathsf{sNN}_\mathsf{F}@(\varepsilon^{t+\Delta t}, \zeta^{t+\Delta t})$
5. $\mathsf{D}^{t+\Delta t} := -\dfrac{\partial \mathsf{F}^{t+\Delta t}}{\partial \zeta^{t+\Delta t}} \cdot \dot{\zeta}^{t+\Delta t}$
6. $\Delta\sigma := \dfrac{\partial \mathsf{F}^{t+\Delta t}}{\partial \varepsilon^{t+\Delta t}} - \sigma^t$

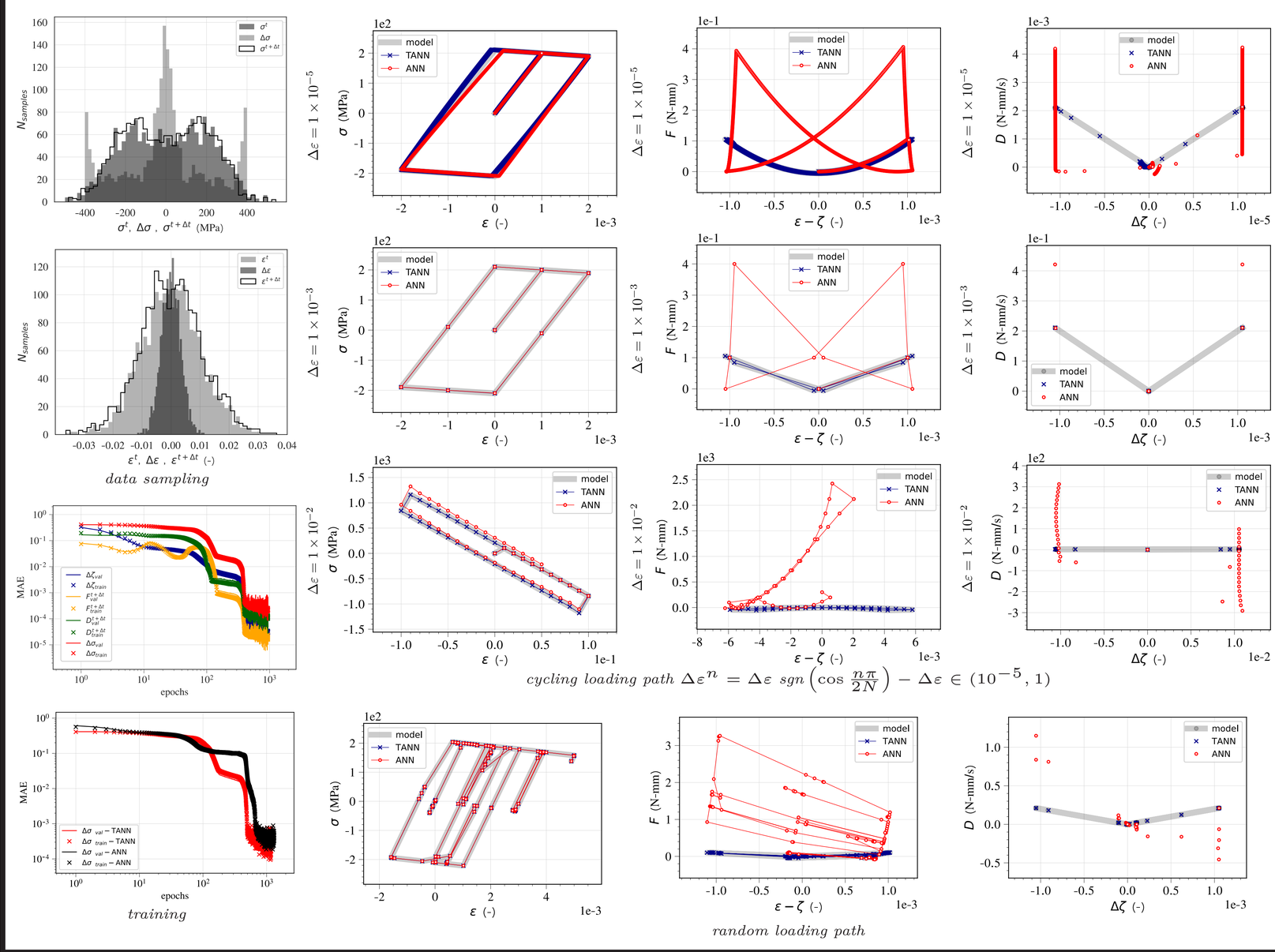
Second-order vanishing gradient

When dealing with TANNs, the issue of **second-order vanishing gradients** may arise. Similar to the first-order variant, it involves the **second derivatives of the activation functions**. Classical activation functions are not suitable choices.



Performance of TANNs vs standard ANNs

TANNs display **efficient and robust training**, and more **accurate predictions**. More important, TANNs predictions are always **thermodynamically consistent**, even for unseen data.



Concluding remarks

Thermodynamics-based Artificial Neural Networks (TANNs) do not have to identify the underlying pattern of thermodynamic laws during training, reducing the need of large data-sets and improving the **robustness** and the **performance** of **predictions**. The **predictions** remain **thermodynamically consistent**, even for unseen data. TANNs are excellent candidates for replacing constitutive calculations at Finite Element incremental formulations in solid mechanics.

Author I.S. would like to acknowledge the support of the European Research Council (ERC) under the EUH 2020 program (Grant ID 757848 CoQuake).