Doctoral School 2023

Machine Learning in Geomechanics

ALERT Geomaterials

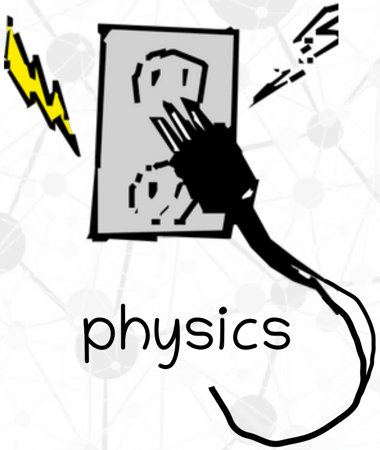
Physics-informed and thermodynamics-based artificial neural nets

Filippo Masi

The University of Sydney

Ioannis Stefanou

Ecole Centrale de Nantes



THE UNIVERSITY OF SYDNEY

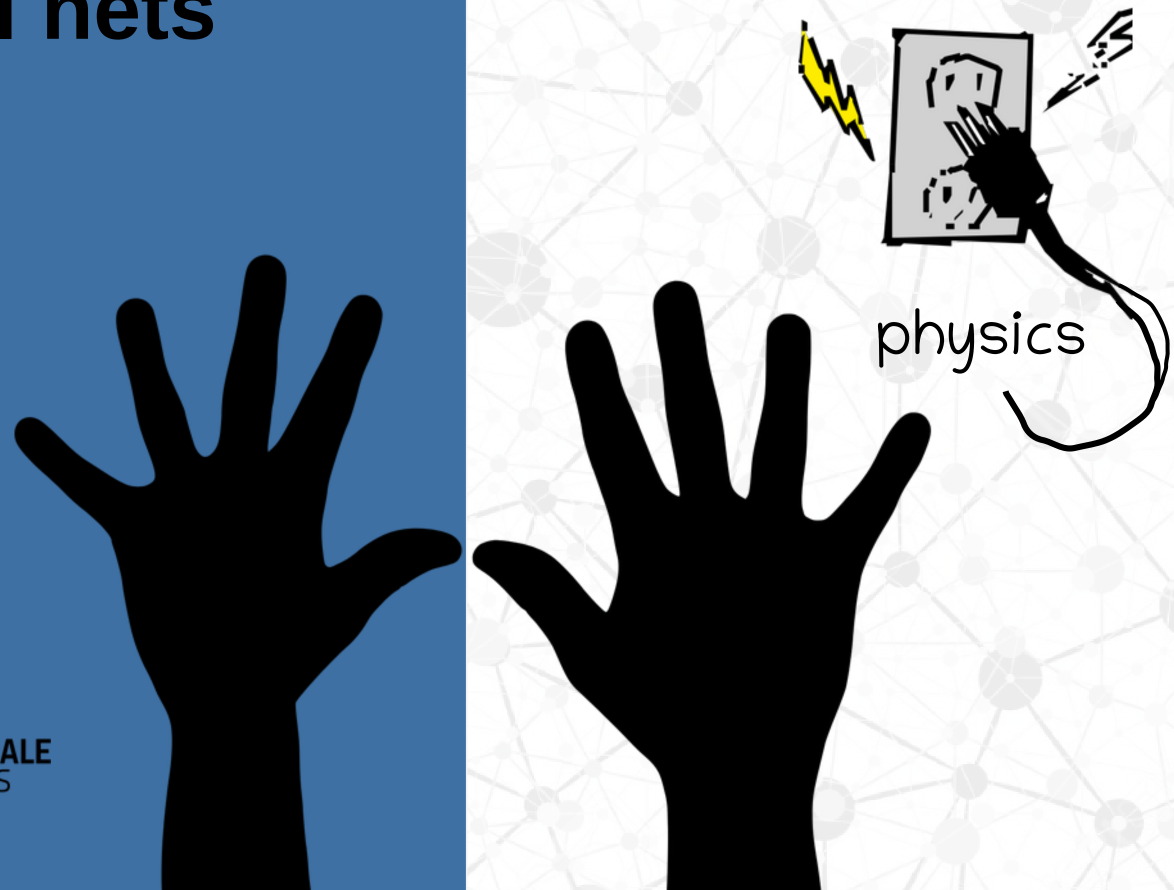
CENTRALE NANTES

Doctoral School 2023
Machine Learning in Geomechanics

ALERT Geomaterials

Physics-informed and thermodynamics-based artificial neural nets

Filippo Masi
The University of Sydney
Ioannis Stefanou
Ecole Centrale de Nantes



THE UNIVERSITY OF SYDNEY

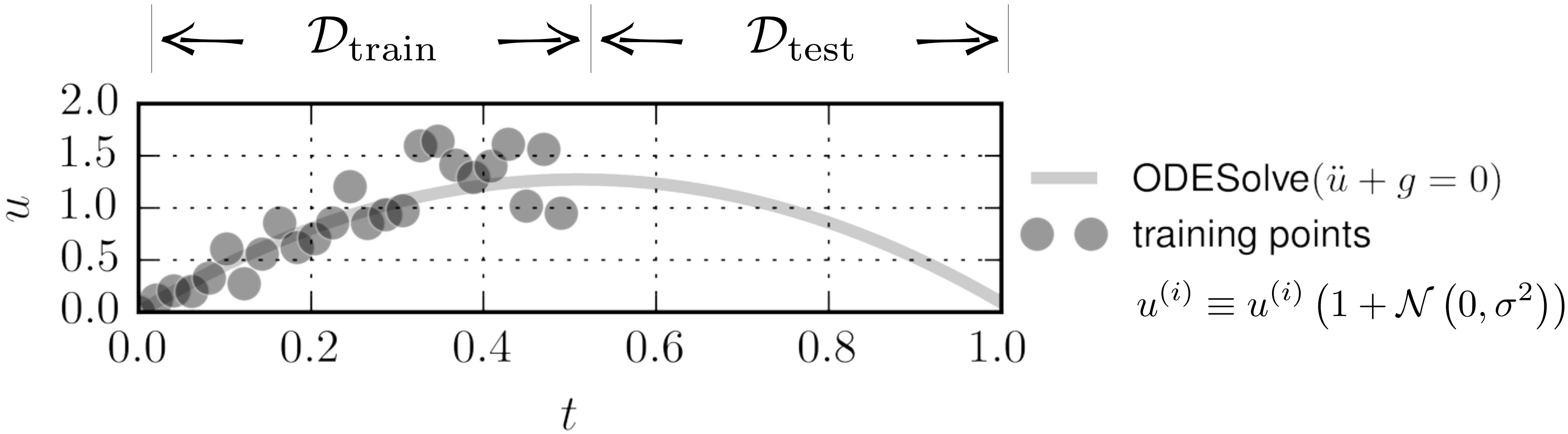
CENTRALE NANTES

Re-discover ballistics: PINN vs NN

PINN: Physics-informed NN

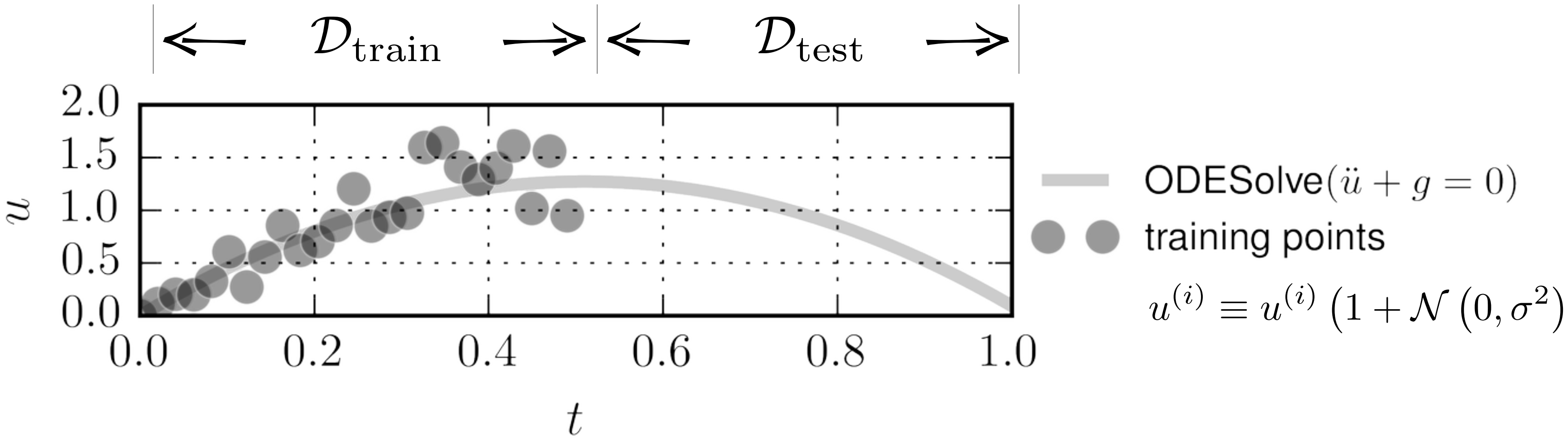
Re-discover ballistics: PINN vs NN

PINN: Physics-informed NN

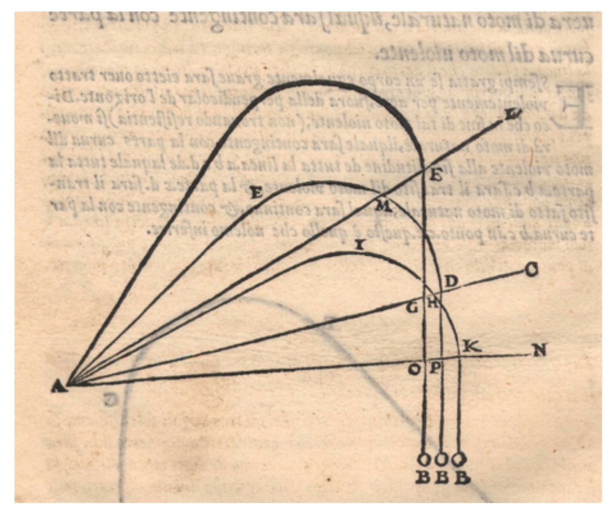


Re-discover ballistics: PINN vs NN

PINN: Physics-informed NN

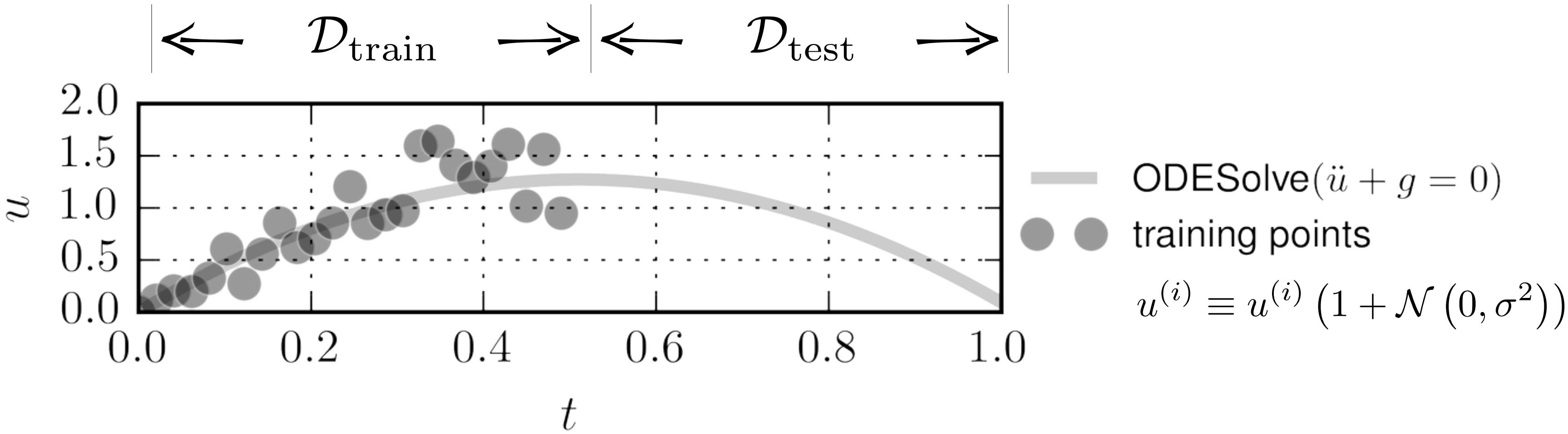


Nicolò Tartaglia *(1558)* *La nova scientia*

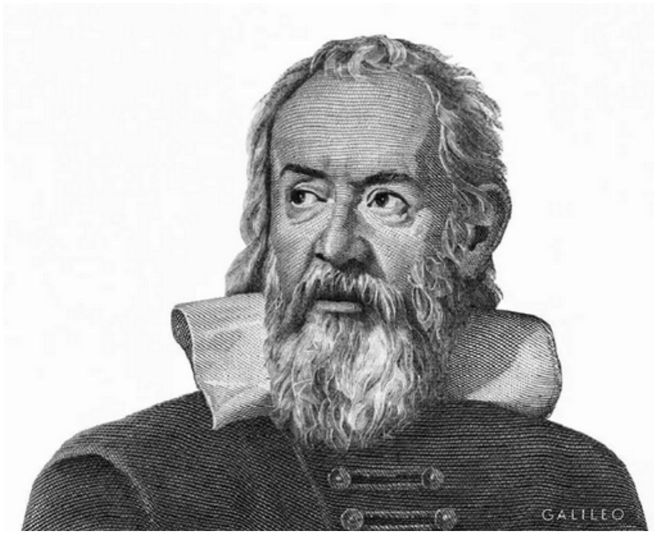


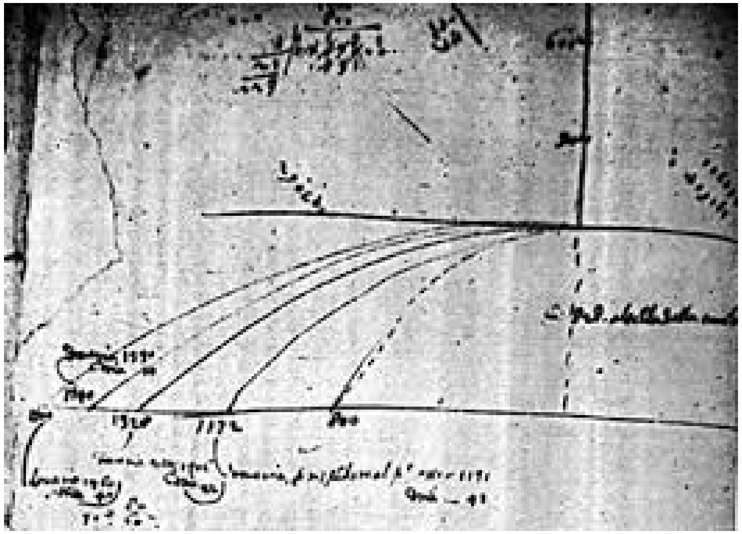
Re-discover ballistics: PINN vs NN

PINN: Physics-informed NN



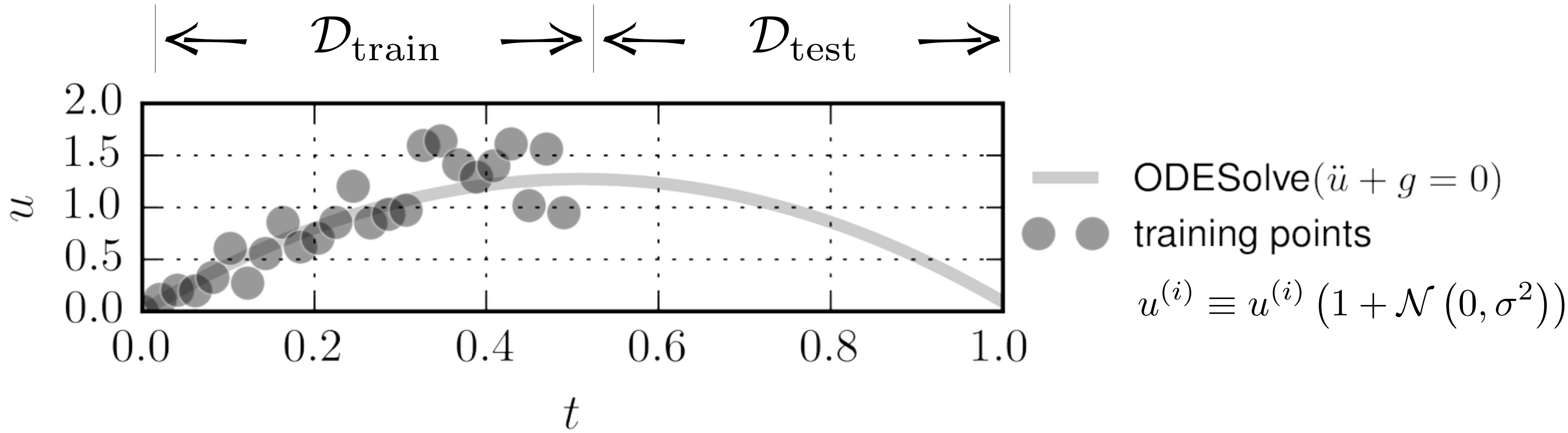
Galileo Galilei (1589) *De motu antiquiora*



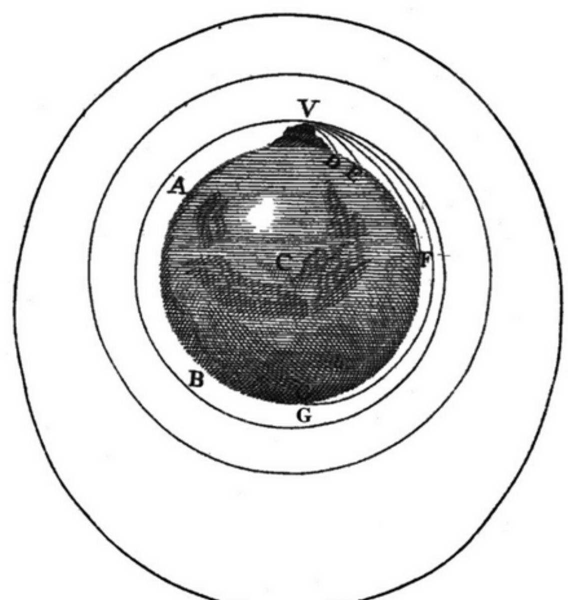


Re-discover ballistics: PINN vs NN

PINN: Physics-informed NN

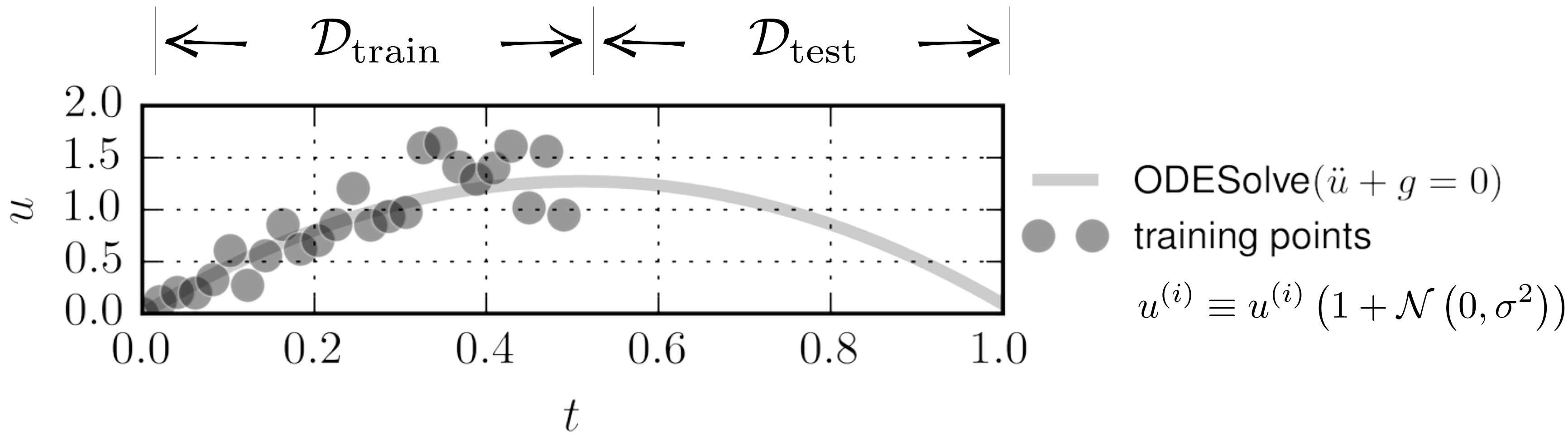


Isaac Newton (1731) *On the system of the world*



Re-discover ballistics: PINN vs NN

PINN: Physics-informed NN



GOALS 1) train a NN such that $u^{(i)} = \underbrace{\hat{u}_{\boldsymbol{\theta}}\left(t^{(i)}\right)}_{\text{neural net}} \; \forall\, i \in [0, N]\,,\; t^N = 0.5\ \mathrm{s}$

2) train its physics-informed twin (*Galilean!*) $\left(\ddot{u}^{(i)} + \lambda = 0\right)$

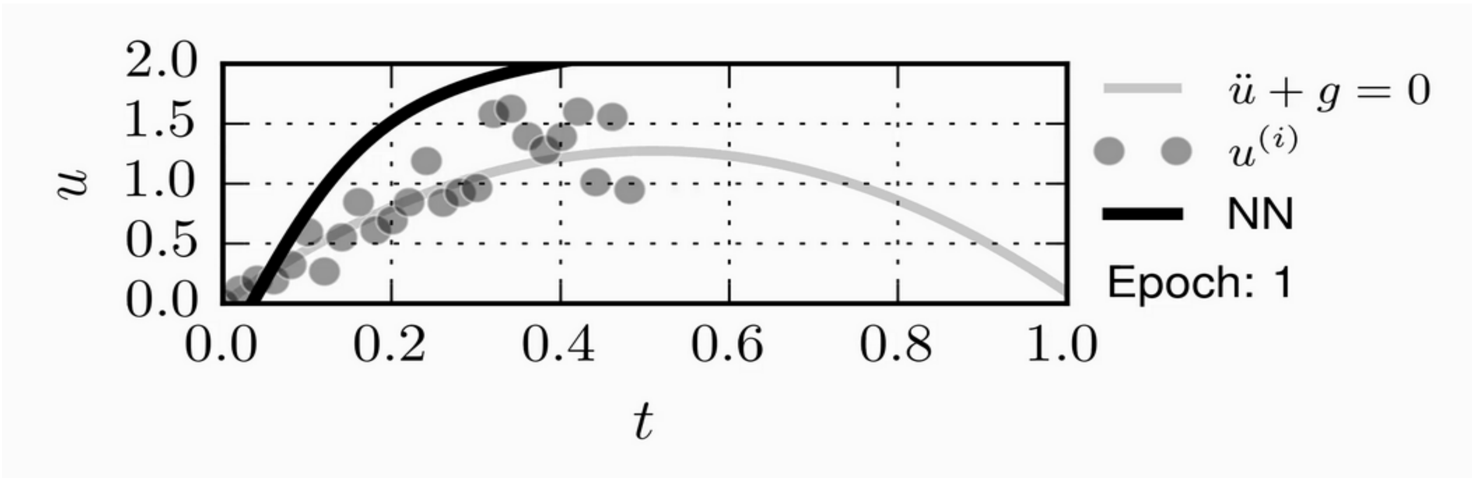
3) evaluate the predictions $\forall\, i \in [0, 2N]\,,\; t^{2N} = 1\ \mathrm{s}$

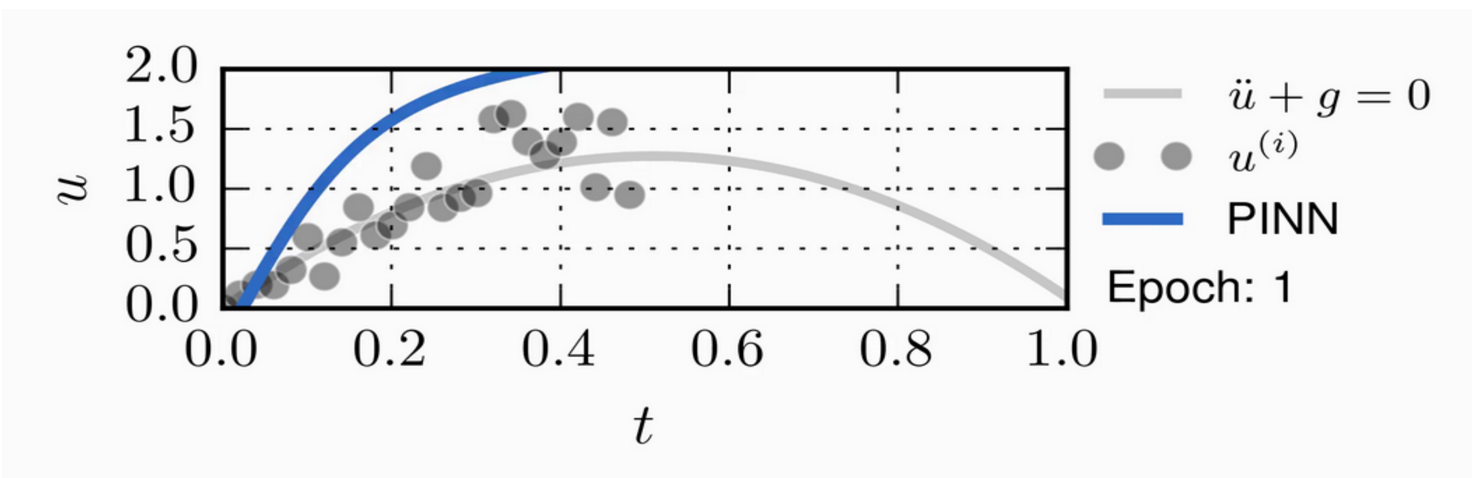
4) (+) repeat using ReLU activations (instead of tanh)

PINN+TANN / PINN (hands-on)

https://qrco.de/piml

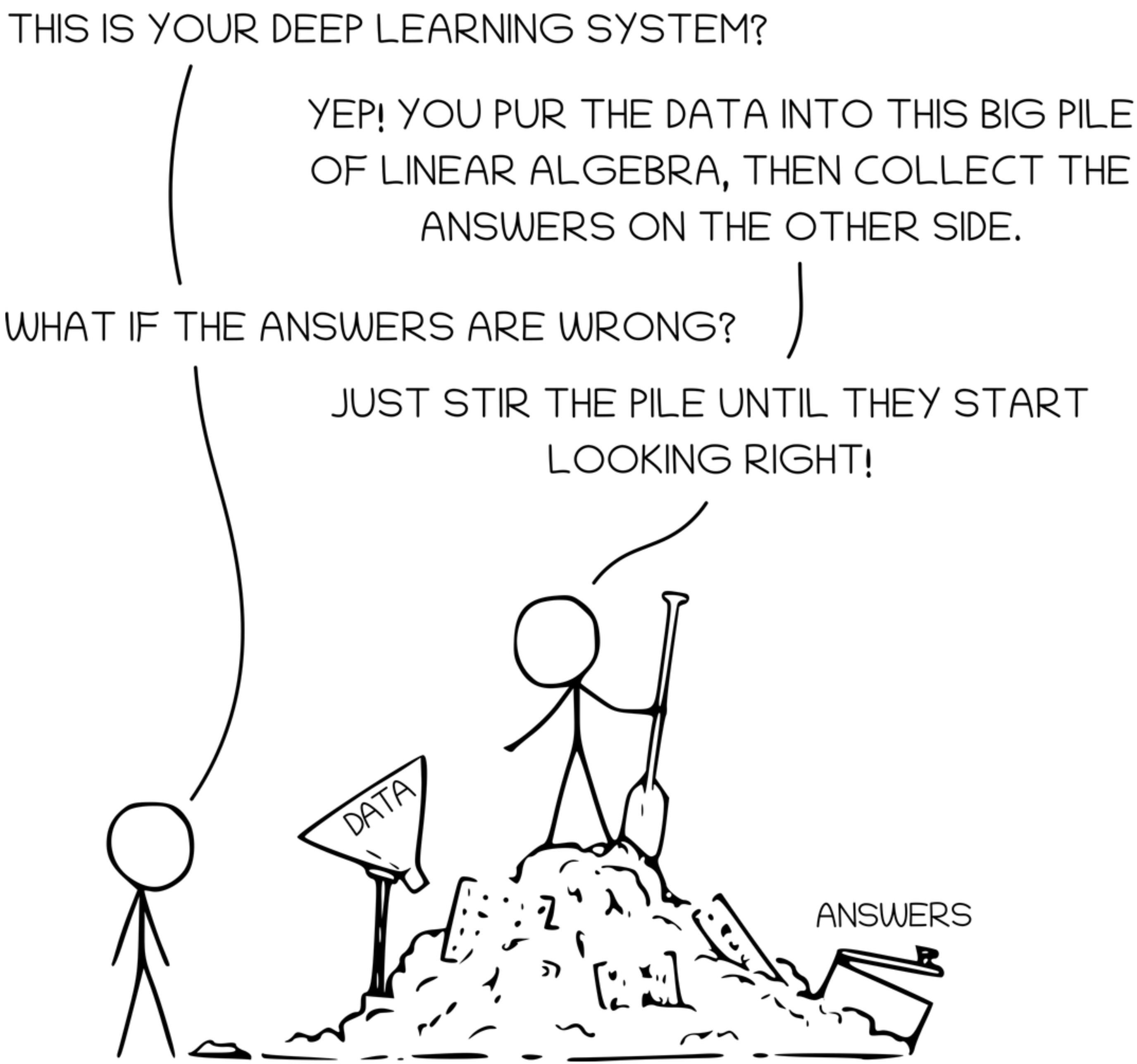
Re-discover ballistics: PINN vs NN



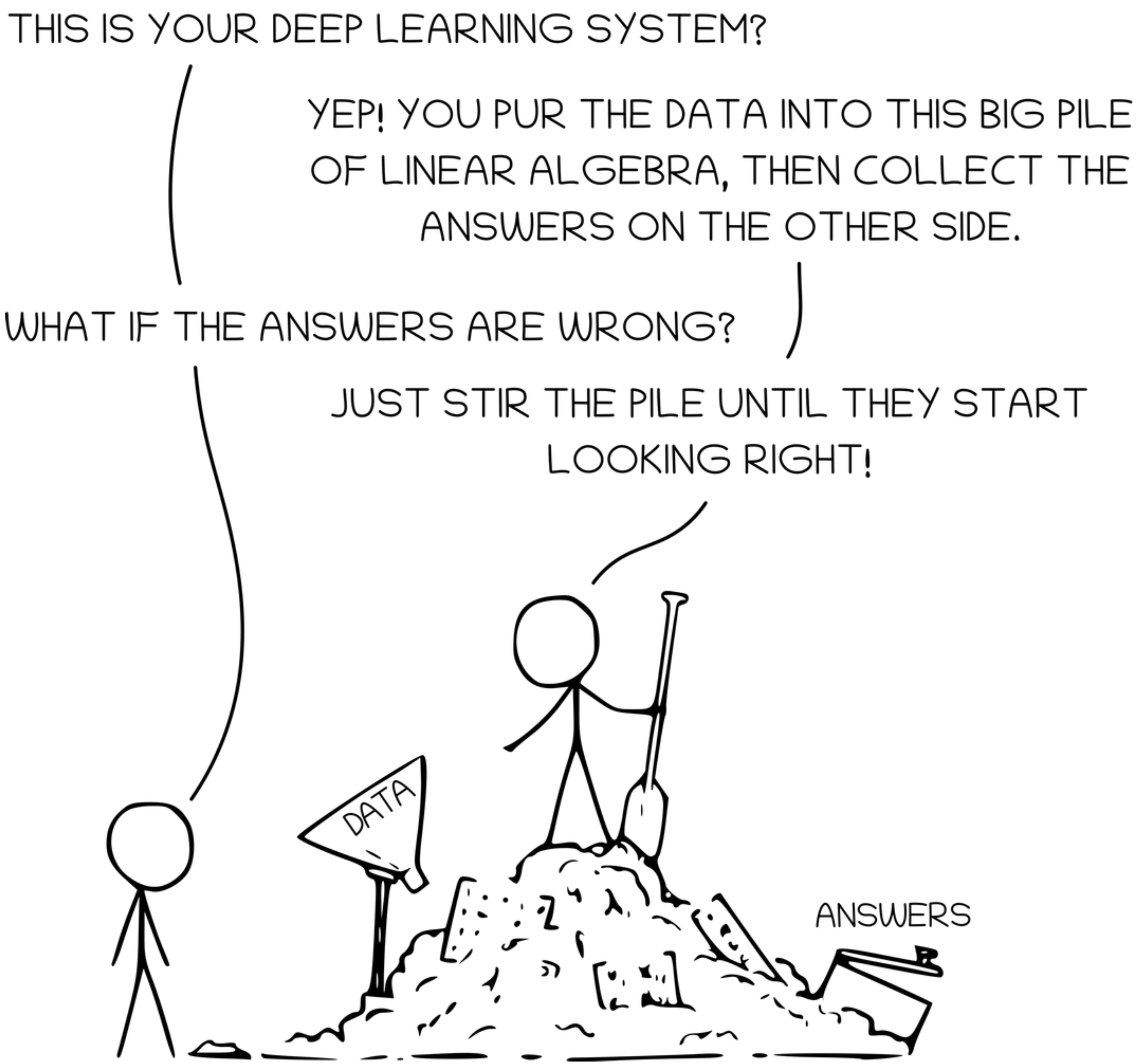


If Galileo had used PINN, he would have discovered the equations of motion for ballistics 140 years before Isaac did.

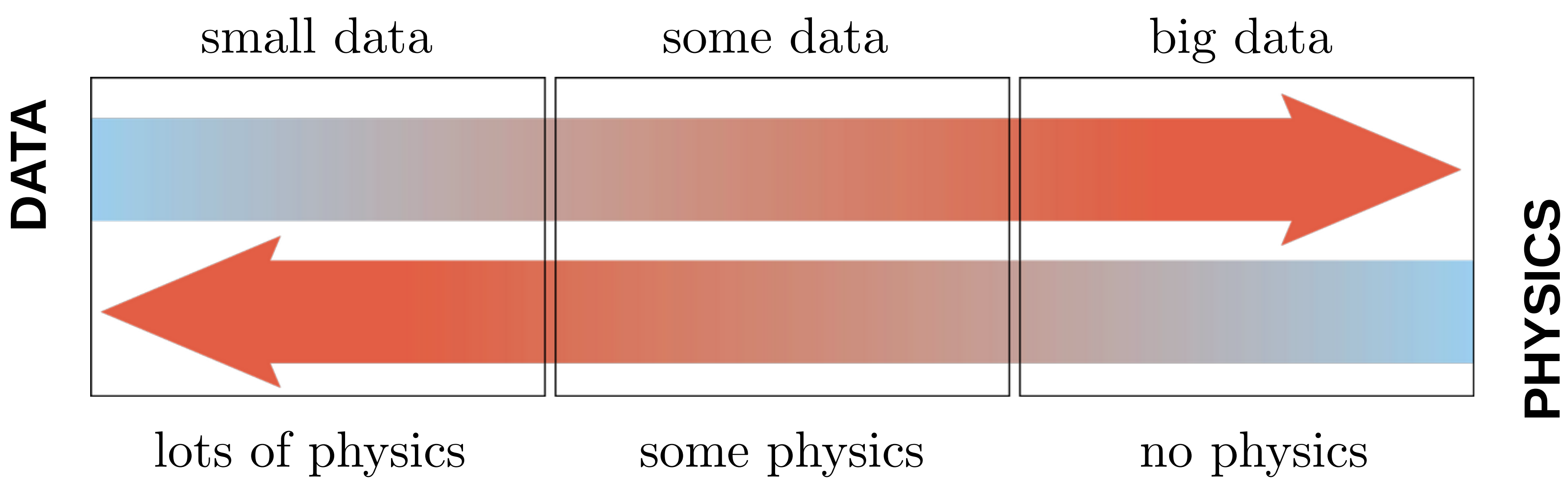
“*Physics-informed*” machine learning


THIS IS YOUR DEEP LEARNING SYSTEM?
YEP! YOU PUR THE DATA INTO THIS BIG PILE OF LINEAR ALGEBRA, THEN COLLECT THE ANSWERS ON THE OTHER SIDE.
WHAT IF THE ANSWERS ARE WRONG?
JUST STIR THE PILE UNTIL THEY START LOOKING RIGHT!
DATA
ANSWERS

The flaws of data-intensive models

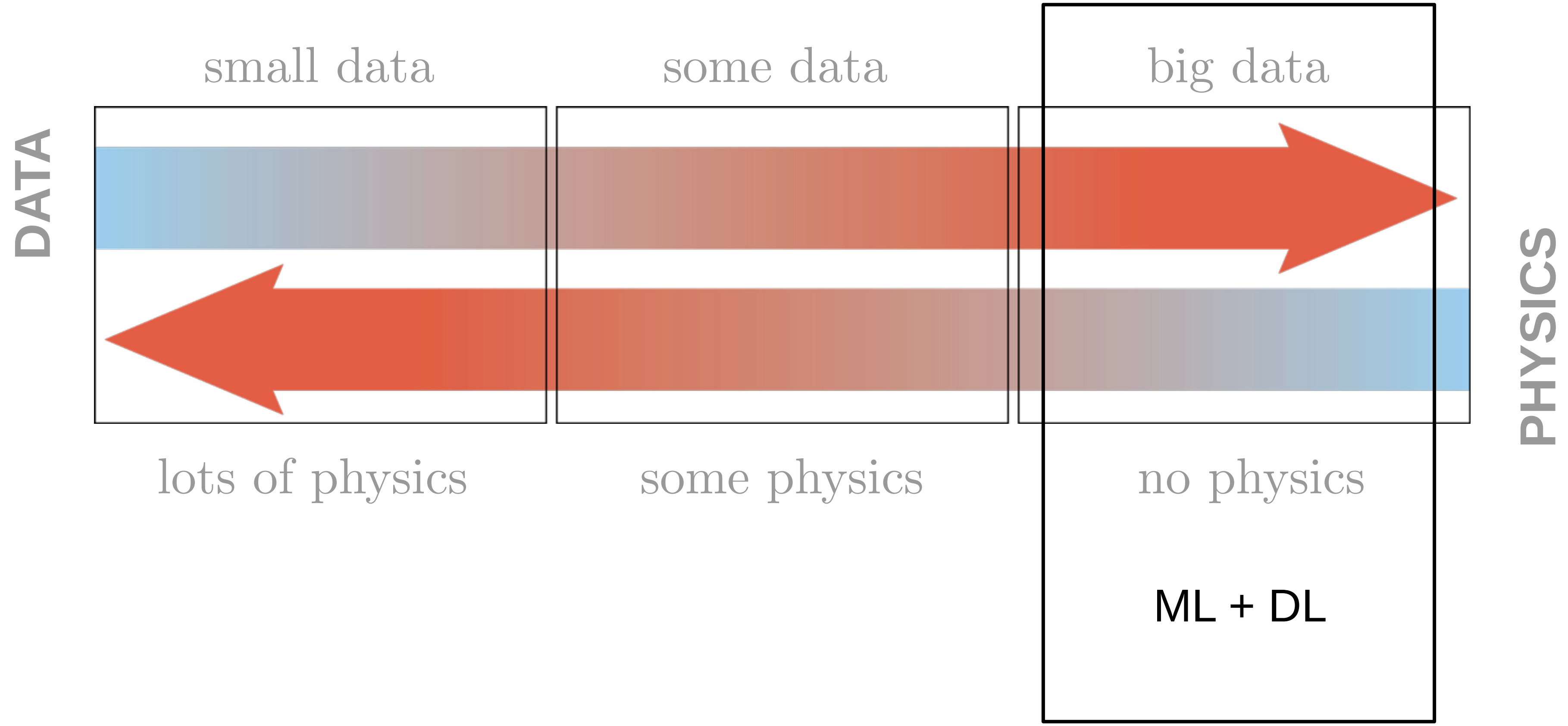


The flaws of data-intensive models



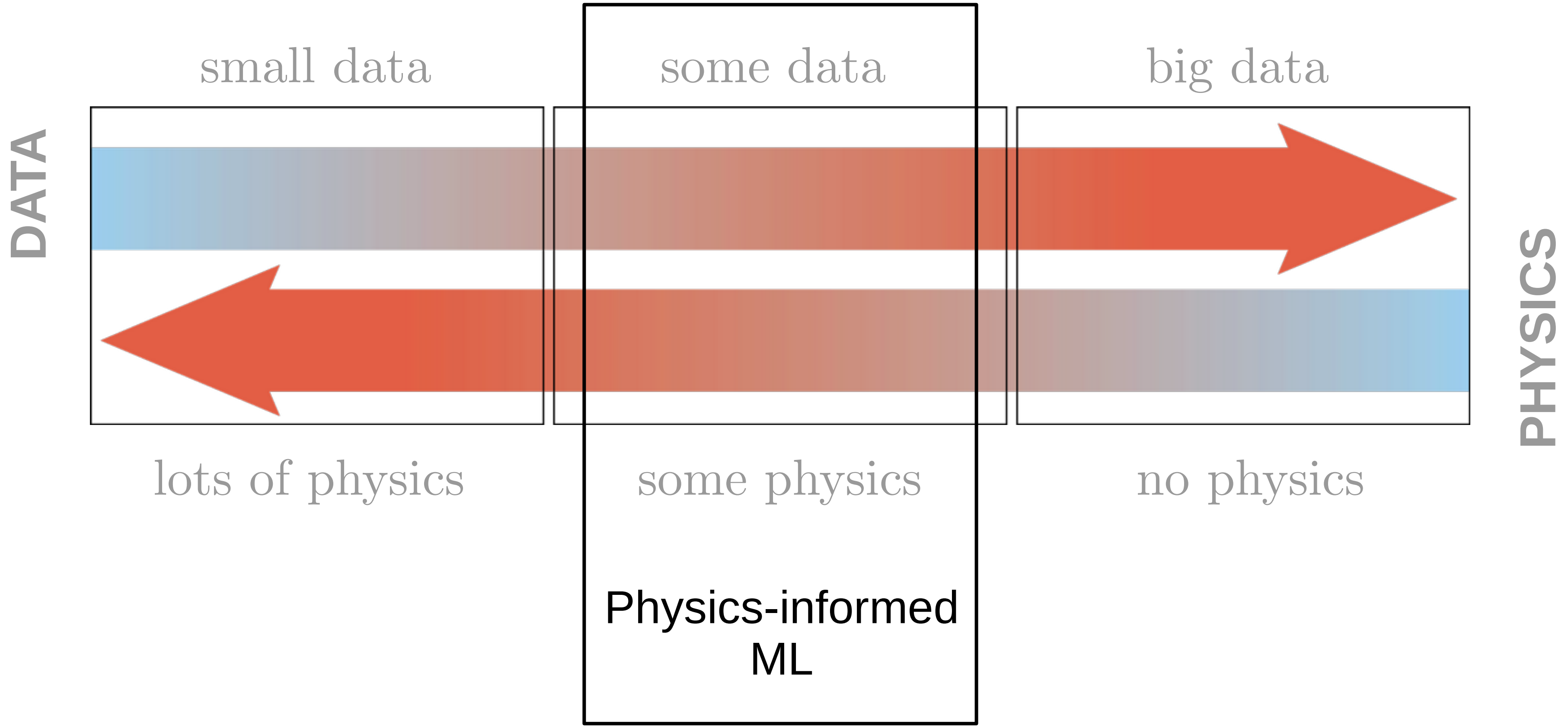
Karniadakis et al, 2021. *Nature Reviews Physics* 3(6)

The flaws of data-intensive models



Karniadakis et al, 2021. *Nature Reviews Physics* 3(6)

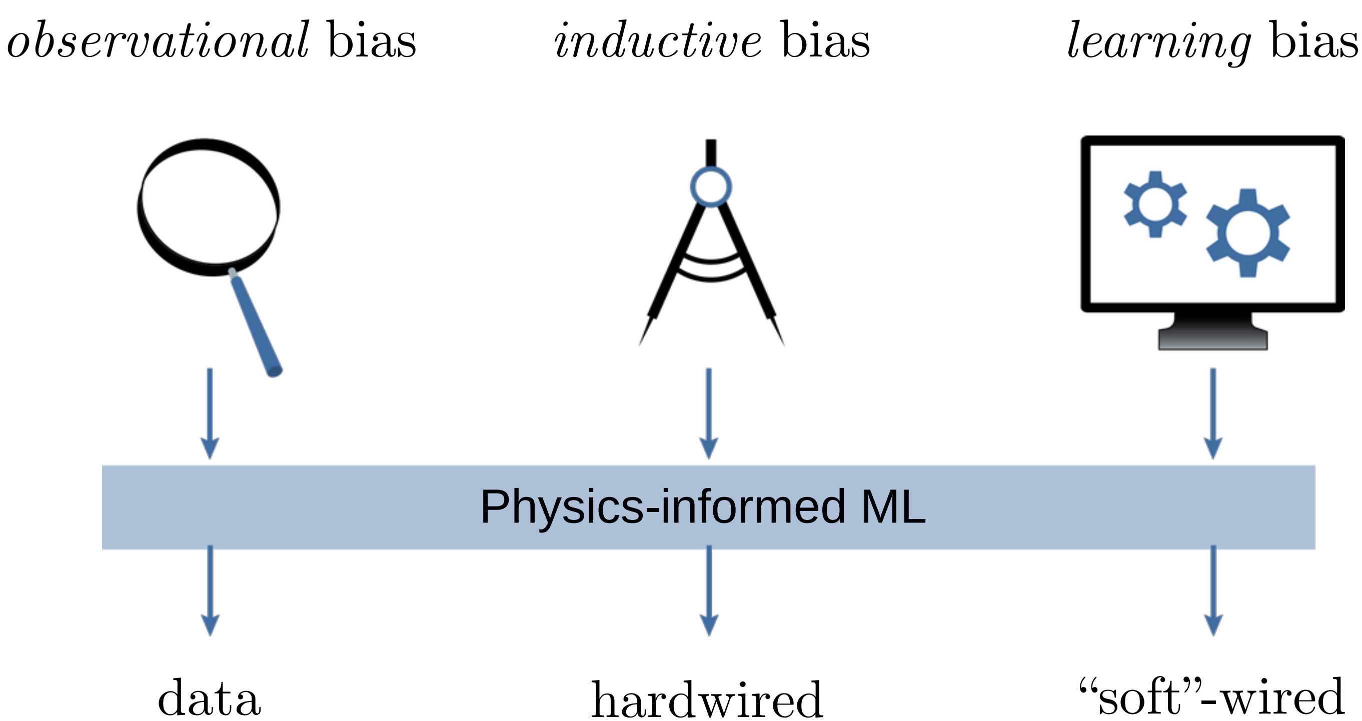
The flaws of data-intensive models



Karniadakis et al, 2021. *Nature Reviews Physics* 3(6)

What is physics-informed ML?

ML models that embed prior knowledge stemming from our physical understanding of the world (+ performance, reliability, interpretability)



Karniadakis et al, 2021. *Nature Reviews Physics* 3(6)

Physics-informed neural nets

PINN: Physics-informed neural nets

Supervised DL algorithms encoding physical laws governing data whose evolution can be described by partial differential equations (PDE)

$$\begin{gathered}\dot{\boldsymbol{u}} + \mathcal{N}_\gamma[\boldsymbol{u}] = 0 \\ + \\ \text{boundary conditions}\end{gathered} \tag{1}$$

$\boldsymbol{u} = \hat{\boldsymbol{u}}(\boldsymbol{x}, t)$ latent (hidden) solution

$\mathcal{N}_\gamma[\,\cdot\,]$ nonlinear differential operator parametrized by γ

$\boldsymbol{x} \in \Omega \subset \mathbb{R}^n, t \in \mathbb{R}^+$ spatial coordinates and time

Raissi et al, 2019. *Journal of Computational physics*, 378

PINN: Physics-informed neural nets

Supervised DL algorithms encoding physical laws governing data whose evolution can be described by partial differential equations (PDE)

$$\begin{gathered}\dot{\boldsymbol{u}} + \mathcal{N}_\gamma[\boldsymbol{u}] = 0 \\ + \\ \text{boundary conditions}\end{gathered} \tag{1}$$

$\boldsymbol{u} = \hat{\boldsymbol{u}}(\boldsymbol{x}, t)$ latent (hidden) solution

$\mathcal{N}_\gamma[\,\cdot\,]$ nonlinear differential operator parametrized by γ

$\boldsymbol{x} \in \Omega \subset \mathbb{R}^n, t \in \mathbb{R}^+$ spatial coordinates and time

GOALS 1) Solution of PDEs

2) Discovery of PDEs

Raissi et al, 2019. *Journal of Computational physics*, 378

1 – Solution of PDEs

1 – Solution of PDEs

Find an approximation of $\hat{\boldsymbol{u}}$ such that (1) holds true for fixed γ, by means of a deep neural network $\boldsymbol{u}_{\boldsymbol{\theta}}$, with parameters $\boldsymbol{\theta}$

1 – Solution of PDEs

Find an approximation of $\hat{\boldsymbol{u}}$ such that (1) holds true for fixed γ, by means of a deep neural network $\boldsymbol{u}_{\boldsymbol{\theta}}$, with parameters $\boldsymbol{\theta}$

$$\varrho(\boldsymbol{x}, t) \equiv \dot{\boldsymbol{u}}_{\boldsymbol{\theta}} + \mathcal{N}[\boldsymbol{u}_{\boldsymbol{\theta}}] \quad \text{residual}$$

$$\mathcal{L} = \lambda_u \underbrace{\frac{1}{N_u} \sum_{i=1}^{N_u} \left\| \boldsymbol{u}_{\boldsymbol{\theta}} \left(\boldsymbol{x}^{(i)}, t^{(i)} \right) - \boldsymbol{u}^{(i)} \right\|}_{\mathcal{L}_u} + \lambda_{\mathrm{PDE}} \underbrace{\frac{1}{N_{\mathrm{PDE}}} \sum_{k=1}^{N_{\mathrm{PDE}}} \left\| \varrho \left(\boldsymbol{x}^{(k)}, t^{(k)} \right) \right\|}_{\mathcal{L}_{\mathrm{PDE}}}$$

supervised loss (data)

unsupervised loss (physics)

1 – Solution of PDEs

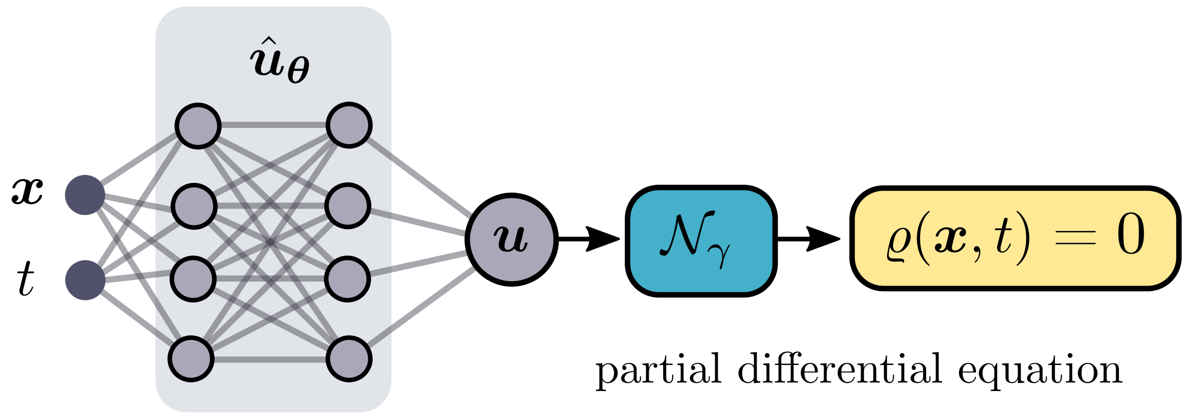
Find an approximation of $\hat{\boldsymbol{u}}$ such that (1) holds true for fixed γ, by means of a deep neural network $\boldsymbol{u}_{\boldsymbol{\theta}}$, with parameters $\boldsymbol{\theta}$

$$\varrho(\boldsymbol{x}, t) \equiv \dot{\boldsymbol{u}}_{\boldsymbol{\theta}} + \mathcal{N}[\boldsymbol{u}_{\boldsymbol{\theta}}] \quad \text{residual}$$

$$\mathcal{L} = \lambda_u \underbrace{\frac{1}{N_u} \sum_{i=1}^{N_u} \left\| \boldsymbol{u}_{\boldsymbol{\theta}} \left(\boldsymbol{x}^{(i)}, t^{(i)} \right) - \boldsymbol{u}^{(i)} \right\|}_{\mathcal{L}_u} + \lambda_{\mathrm{PDE}} \underbrace{\frac{1}{N_{\mathrm{PDE}}} \sum_{k=1}^{N_{\mathrm{PDE}}} \left\| \varrho \left(\boldsymbol{x}^{(k)}, t^{(k)} \right) \right\|}_{\mathcal{L}_{\mathrm{PDE}}}$$

supervised loss (data)

unsupervised loss (physics)



1 – Solution of PDEs

Find an approximation of $\hat{\boldsymbol{u}}$ such that (1) holds true for fixed γ, by means of a deep neural network $\boldsymbol{u}_{\boldsymbol{\theta}}$, with parameters $\boldsymbol{\theta}$

$$\varrho(\boldsymbol{x},t) \equiv \dot{\boldsymbol{u}}_{\boldsymbol{\theta}} + \mathcal{N}[\boldsymbol{u}_{\boldsymbol{\theta}}] \quad \text{residual}$$

$$\mathcal{L} = \underbrace{\lambda_u \frac{1}{N_u}\sum_{i=1}^{N_u} \left\| \boldsymbol{u}_{\boldsymbol{\theta}}\left(\boldsymbol{x}^{(i)}, t^{(i)}\right) - \boldsymbol{u}^{(i)} \right\|}_{\mathcal{L}_u} + \underbrace{\lambda_{\mathrm{PDE}} \frac{1}{N_{\mathrm{PDE}}}\sum_{k=1}^{N_{\mathrm{PDE}}} \left\| \varrho\left(\boldsymbol{x}^{(k)}, t^{(k)}\right) \right\|}_{\mathcal{L}_{\mathrm{PDE}}}$$

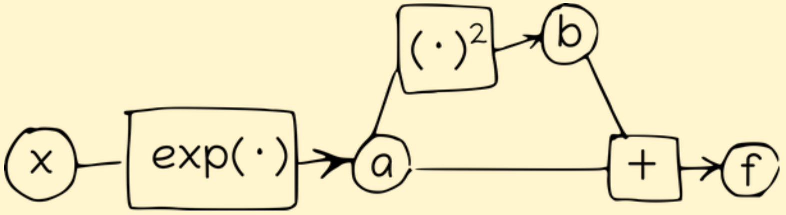
supervised loss (data)

unsupervised loss (physics)

Automatic differentiation

$$f = \exp(x) + \exp(x)^2 = a + b$$

$$\frac{df}{dx} = \frac{df}{da}\frac{da}{dx} + \frac{df}{db}\frac{db}{dx} = \exp(x) + 2\exp(x)$$



2 – Discovery of PDEs

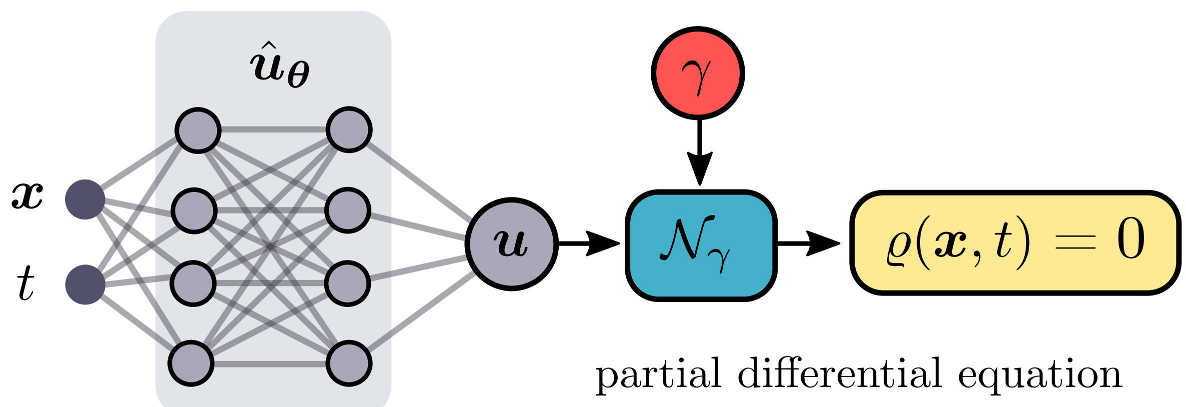
The same but now γ is unknown!

$\varrho(\boldsymbol{x}, t) \equiv \dot{\boldsymbol{u}}_{\boldsymbol{\theta}} + \mathcal{N}_\gamma[\boldsymbol{u}_{\boldsymbol{\theta}}]$ residual

$$\mathcal{L} = \lambda_u \underbrace{\frac{1}{N_u} \sum_{i=1}^{N_u} \left\| \boldsymbol{u}_{\boldsymbol{\theta}} \left(\boldsymbol{x}^{(i)}, t^{(i)} \right) - \boldsymbol{u}^{(i)} \right\|}_{\mathcal{L}_u} + \lambda_{\mathrm{PDE}} \underbrace{\frac{1}{N_{\mathrm{PDE}}} \sum_{k=1}^{N_{\mathrm{PDE}}} \left\| \varrho \left(\boldsymbol{x}^{(k)}, t^{(k)} \right) \right\|}_{\mathcal{L}_{\mathrm{PDE}}}$$

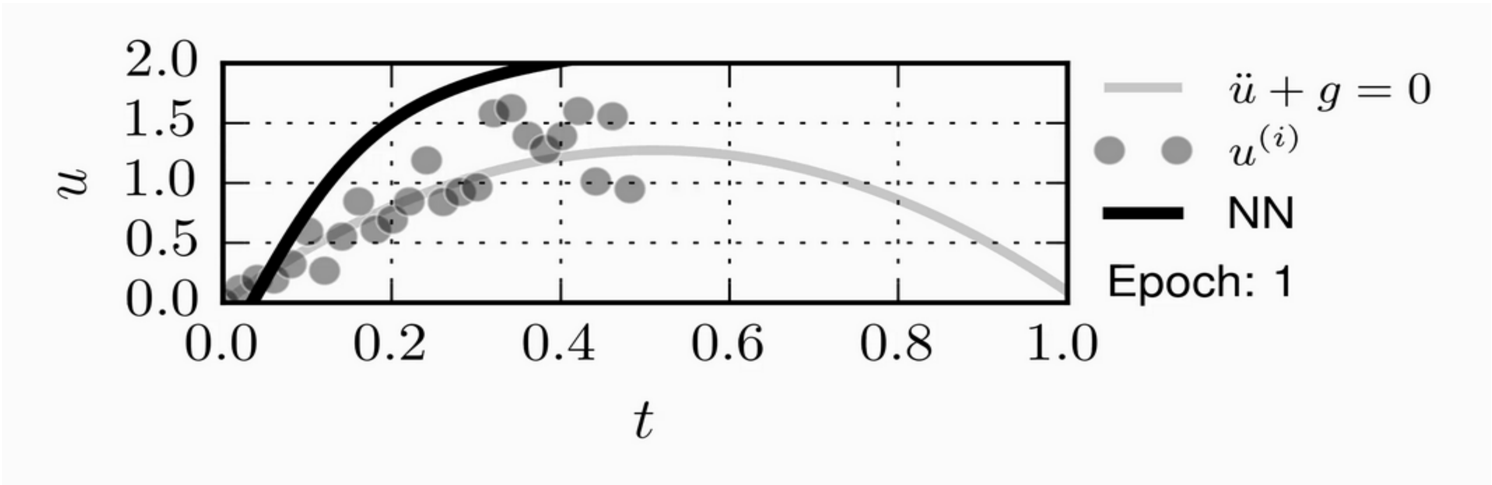
supervised loss (data)

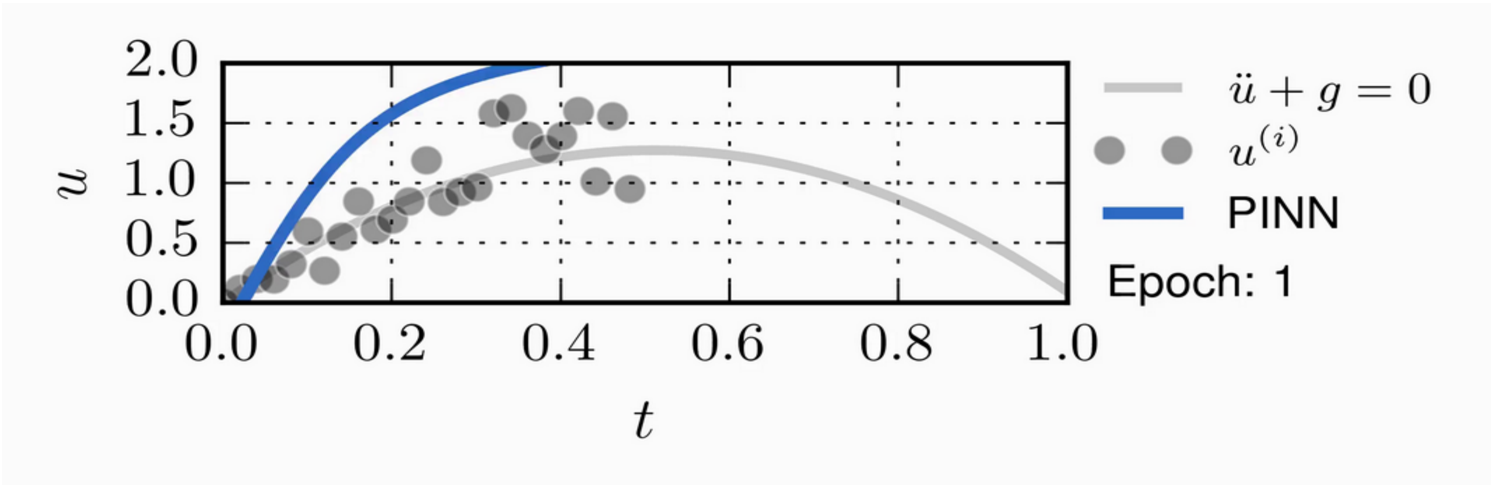
unsupervised loss (physics)



Decoding the hands-on

Data-driven discovery of the governing equations for ballistics

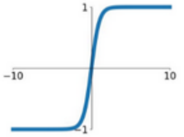




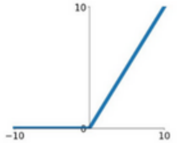
...

using ReLU activations (instead of tanh), we obtain

tanh
$\tanh(x)$



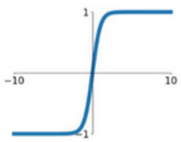
ReLU
$\max(0, x)$



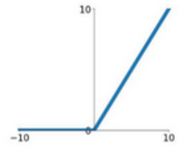
...

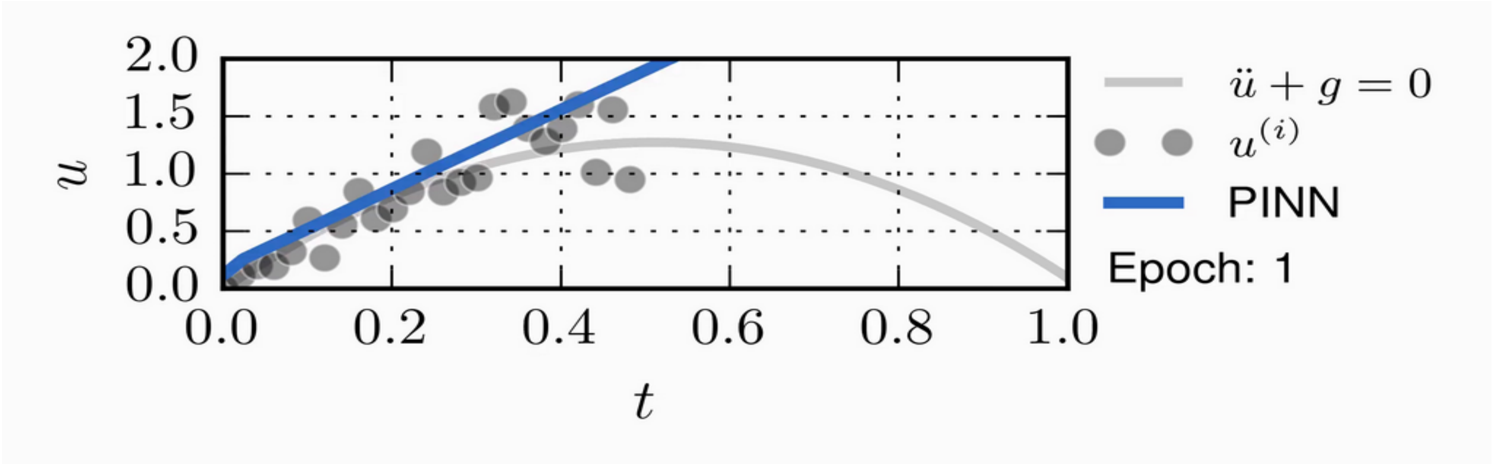
using ReLU activations (instead of tanh), we obtain

tanh
$\tanh(x)$



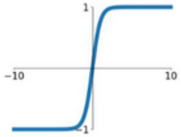
ReLU
$\max(0, x)$



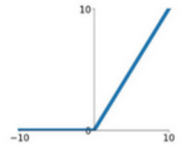


… Higher-order vanishing gradients

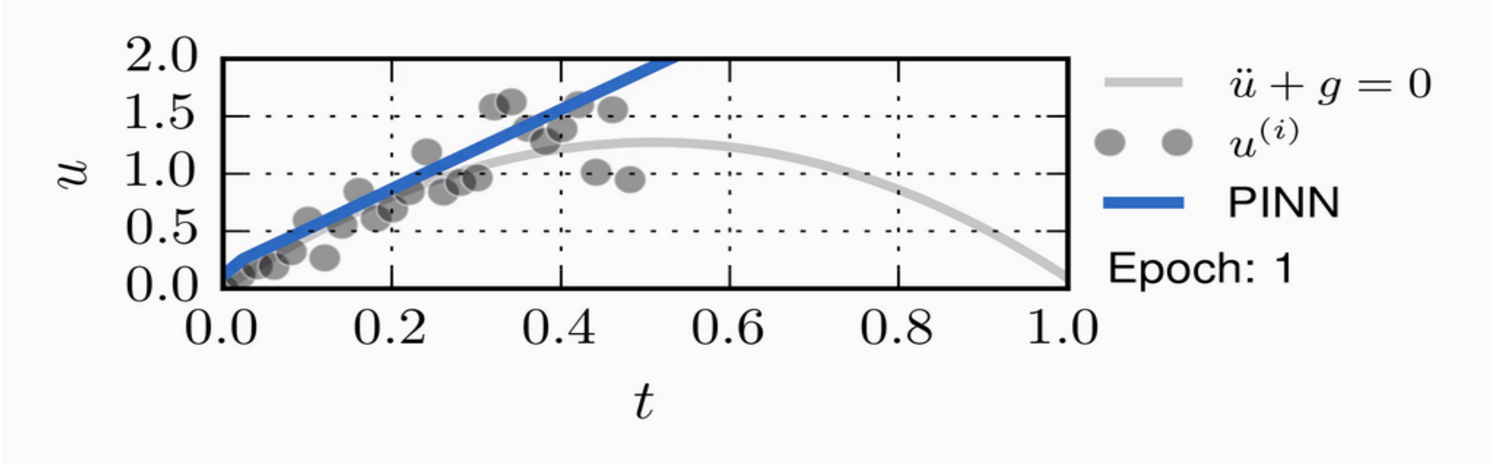
tanh
$\tanh(x)$



ReLU
$\max(0, x)$



using ReLU activations (instead of tanh), we obtain



$$\left.\begin{aligned} \boldsymbol{u}_{\boldsymbol{\theta}}(t) &= \boldsymbol{g}(\boldsymbol{\theta} \odot t) \\ \dot{\boldsymbol{u}}_{\boldsymbol{\theta}}(t) &= \boldsymbol{g}'(\cdot) \odot \boldsymbol{\theta} \\ \ddot{\boldsymbol{u}}_{\boldsymbol{\theta}}(t) &= \boldsymbol{g}''(\cdot) \odot \boldsymbol{\theta} \odot \boldsymbol{\theta} \end{aligned}\right\} \quad \frac{\partial \mathcal{L}_{\mathrm{PDE}}}{\partial \boldsymbol{\theta}} \backsim \boldsymbol{g}''(\cdot) \odot \boldsymbol{\theta} + \mathcal{O}(\boldsymbol{g}'''(\cdot))$$

if $\boldsymbol{g}''(\cdot) = 0 \Rightarrow \dfrac{\partial \mathcal{L}_{\mathrm{PDE}}}{\partial \boldsymbol{\theta}} = 0$

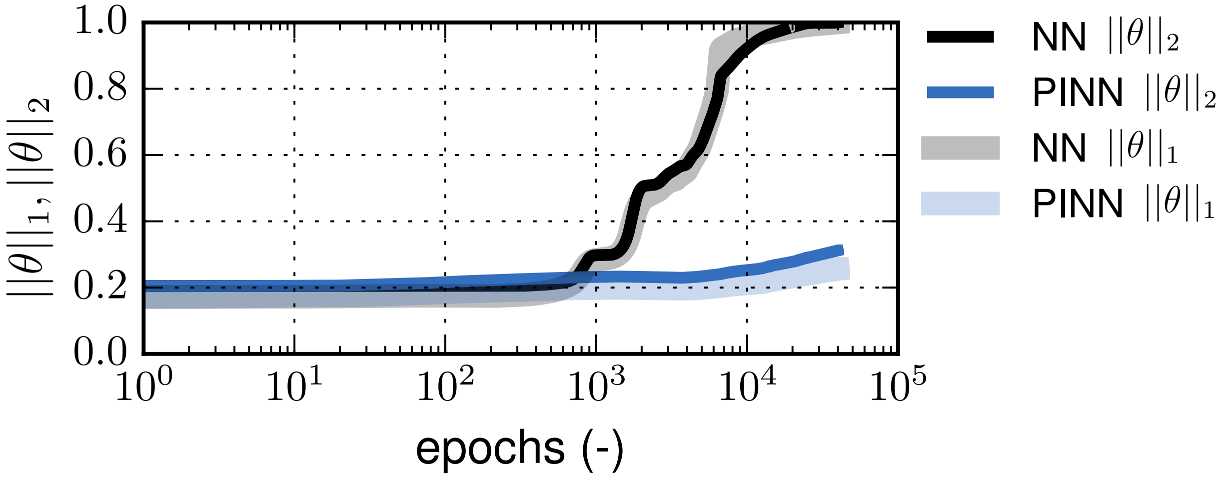
Parsimony or physics?

Decoding the influence of physical constraints as learning biases

Parsimony or physics?

Decoding the influence of physical constraints as learning biases

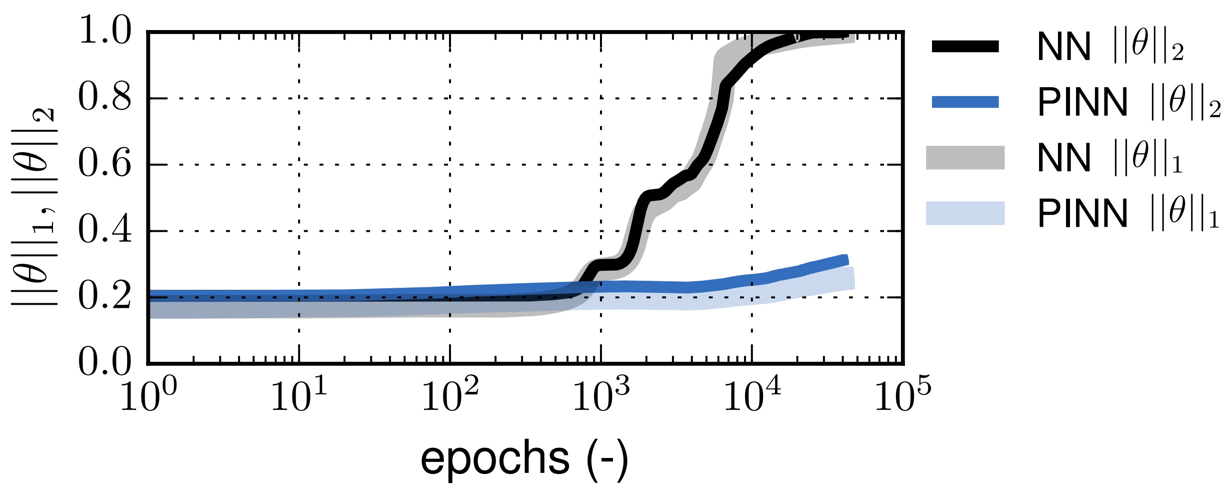
$\tanh(x)$



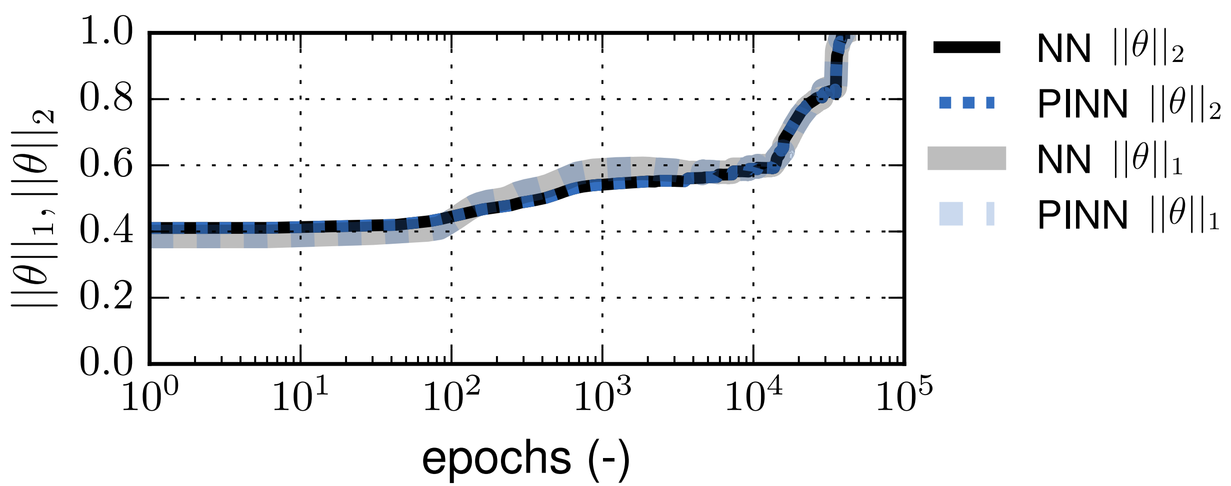
Parsimony or physics?

Decoding the influence of physical constraints as learning biases

$\tanh(x)$

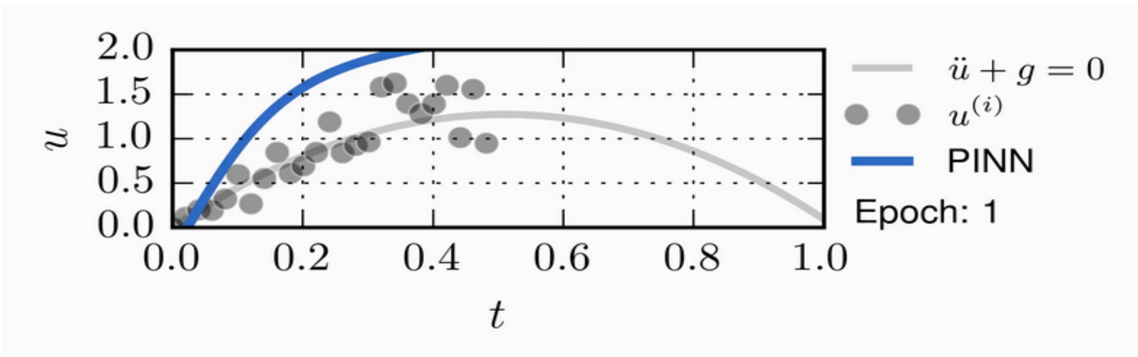


$\mathrm{ReLU}(x)$

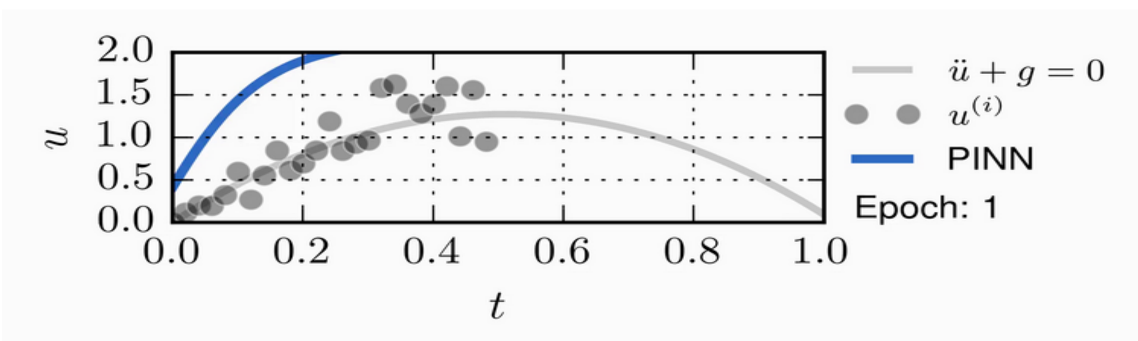


The flaws of solving IVP/BVP

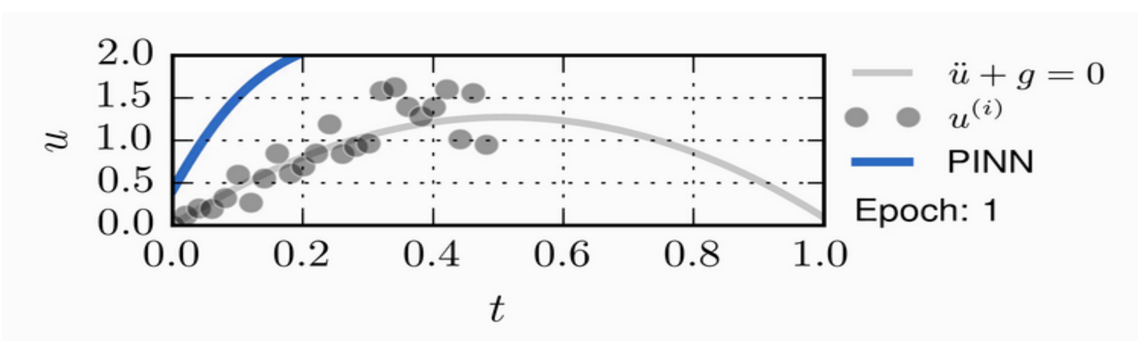
Dependence on the sampling of the collocation points (but also BC/IV)



$N_{\text{PDE}} = 20$



$N_{\text{PDE}} = 2$



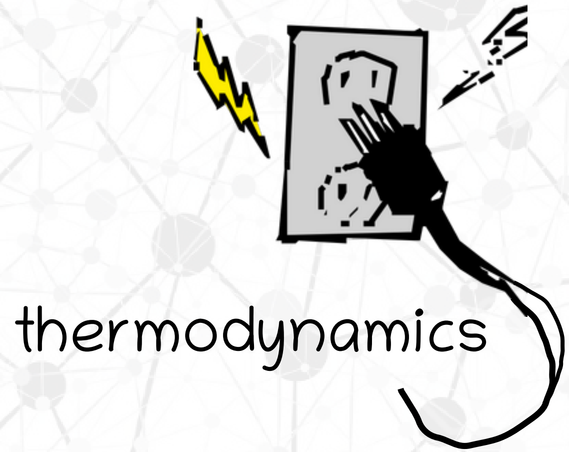
$N_{\text{PDE}} = 1$

Doctoral School 2023
Machine Learning in Geomechanics

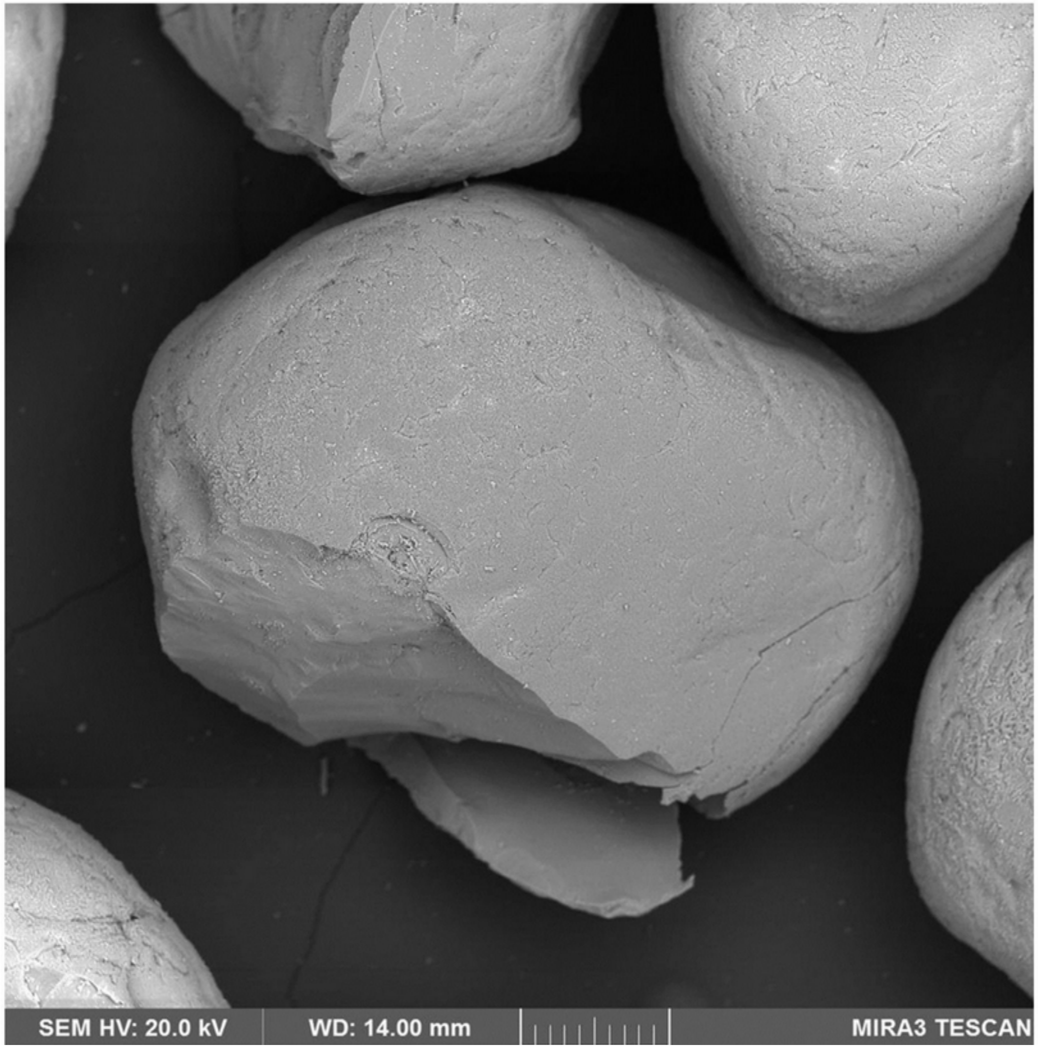
ALERT Geomaterials

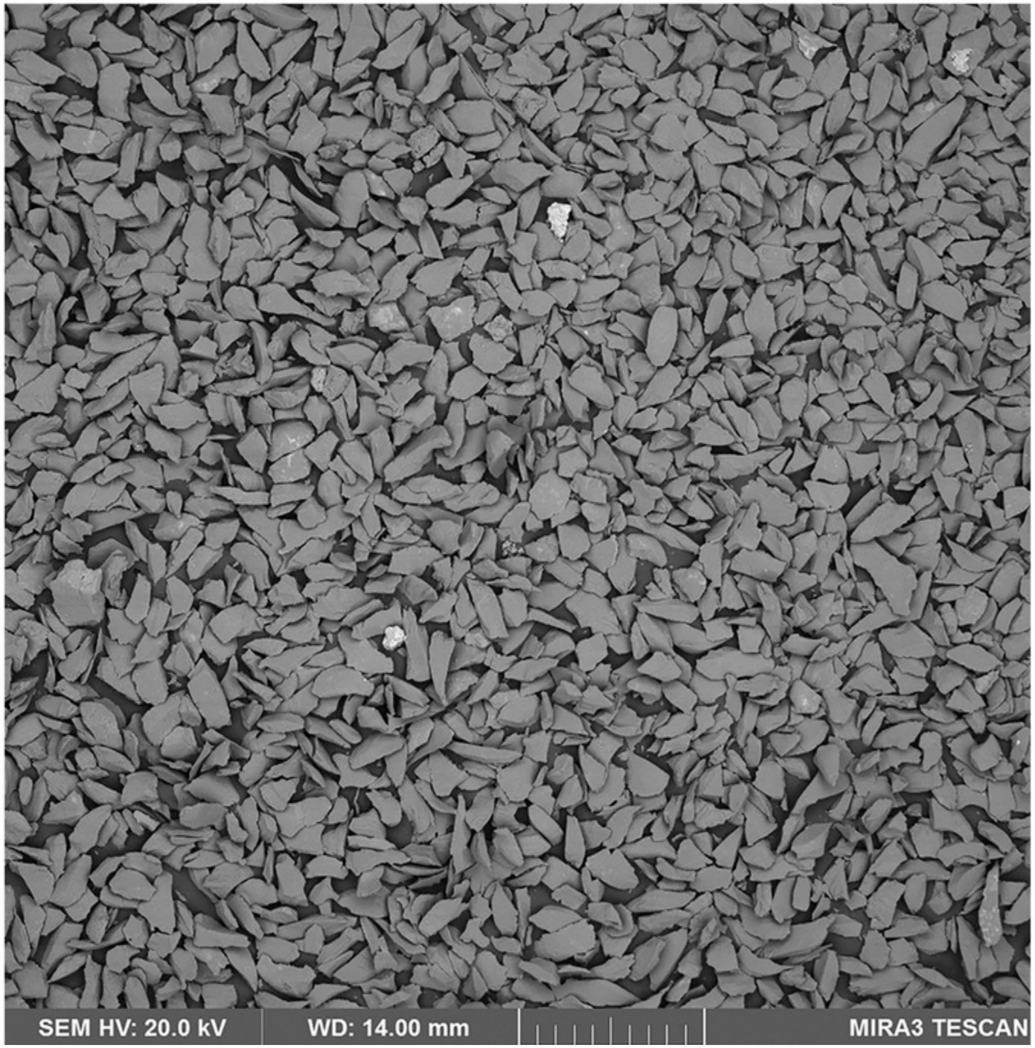
Physics-informed and thermodynamics-based artificial neural nets

Filippo Masi
The University of Sydney
Ioannis Stefanou
Ecole Centrale de Nantes



Complex materials





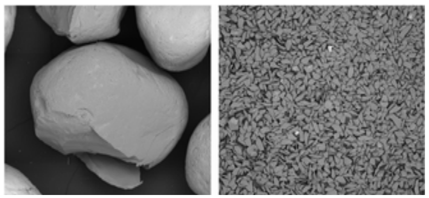
Zheng et al. *Int J Geomech*, 2020

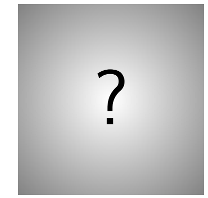
Bridging across the scales

Bridging across the scales

microscale

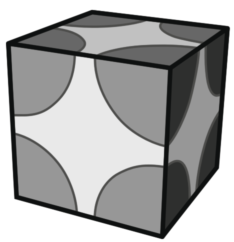
Macroscale





auxiliary problem

Boundary Value Problem

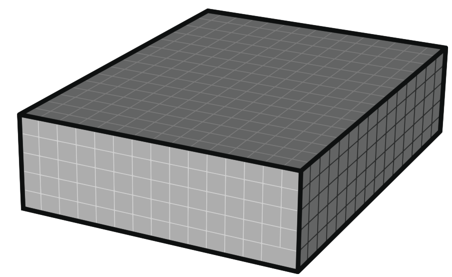


unit cell

$\leftarrow \quad \Delta\varepsilon \quad \leftarrow$

$\rightarrow \quad \Delta\boldsymbol{\sigma} \quad \rightarrow$

FE^2
FEM×DEM
FEM×...

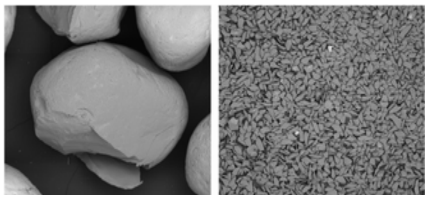


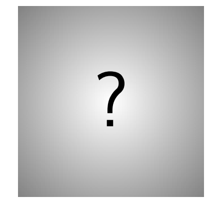
$$\begin{cases} \operatorname{div}\boldsymbol{\sigma} + \boldsymbol{f} = 0 \quad \text{in } \mathcal{V} \\ \boldsymbol{u} = \bar{\boldsymbol{u}} \quad \text{on } \partial\mathcal{V}_u \\ \boldsymbol{\sigma}\cdot n = \bar{\boldsymbol{t}} \quad \text{on } \partial\mathcal{V}_t \end{cases}$$

Pinho-da Cruz et al, 2009; Geers et al, 2010; Lloberas Valls et al., 2019; Nitka et al., 2011

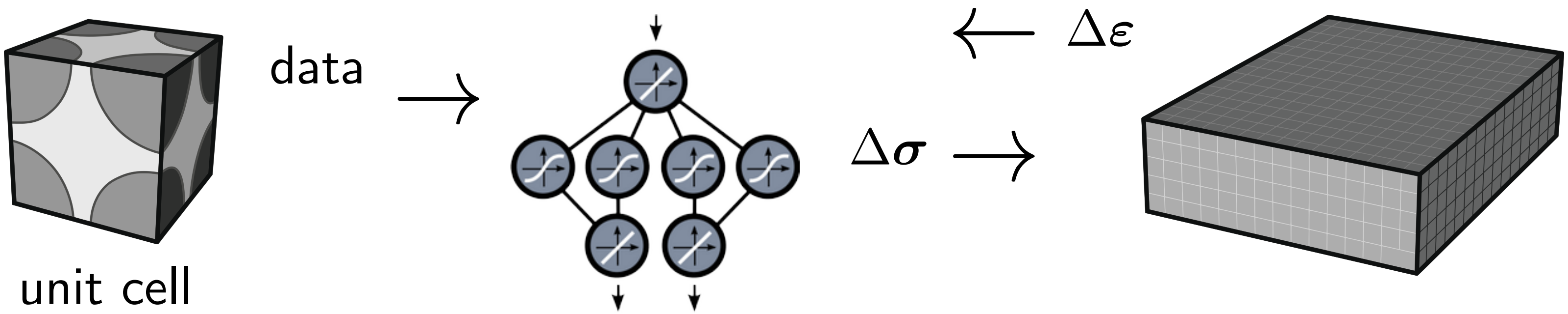
Bridging across the scales

microscale — Macroscale





auxiliary problem — *artificial neural nets* — Boundary Value Problem



unit cell

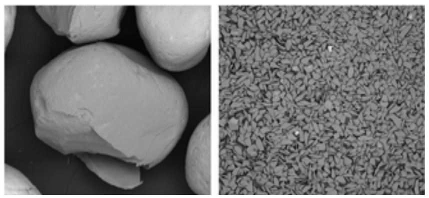
$$\begin{cases} \operatorname{div}\boldsymbol{\sigma} + \boldsymbol{f} = 0 \quad \text{in } \mathcal{V} \\ \boldsymbol{u} = \bar{\boldsymbol{u}} \quad \text{on } \partial\mathcal{V}_u \\ \boldsymbol{\sigma}\cdot n = \bar{\boldsymbol{t}} \quad \text{on } \partial\mathcal{V}_t \end{cases}$$

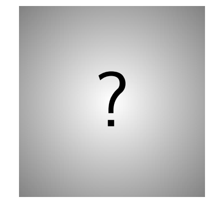
Ghaboussi et al, 1991; Lefik and Schrefler, 2003; Mozaffar et al, 2019; Mianroodi et al, 2021; etc.

Bridging across the scales

microscale

Macroscale

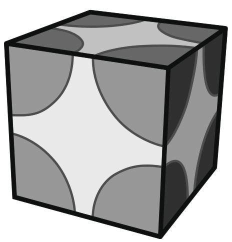




auxiliary problem

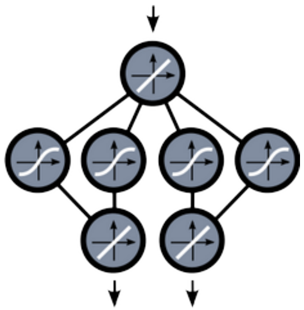
artificial neural nets

Boundary Value Problem



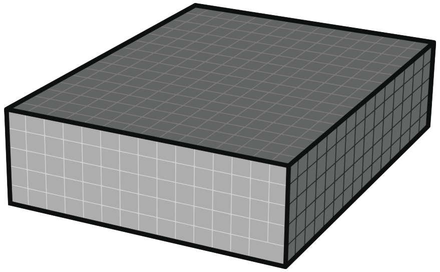
data $\longrightarrow$

unit cell



$\longleftarrow \Delta\varepsilon$

$\Delta\boldsymbol{\sigma} \longrightarrow$



poor reliability
poor generalization
huge amount of data

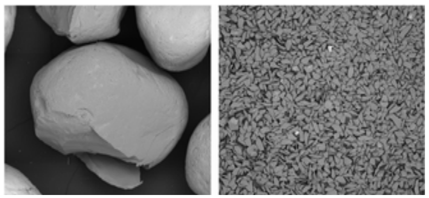
$$\begin{cases} \operatorname{div}\boldsymbol{\sigma} + \boldsymbol{f} = 0 \quad \text{in } \mathcal{V} \\ \boldsymbol{u} = \bar{\boldsymbol{u}} \quad \text{on } \partial\mathcal{V}_u \\ \boldsymbol{\sigma}\cdot n = \bar{\boldsymbol{t}} \quad \text{on } \partial\mathcal{V}_t \end{cases}$$

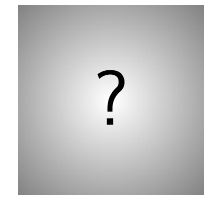
Ghaboussi et al, 1991; Lefik and Schrefler, 2003; Mozaffar et al, 2019; Mianroodi et al, 2021; etc.

Bridging across the scales

microscale

Macroscale

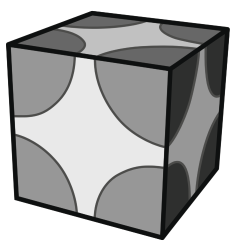




auxiliary problem

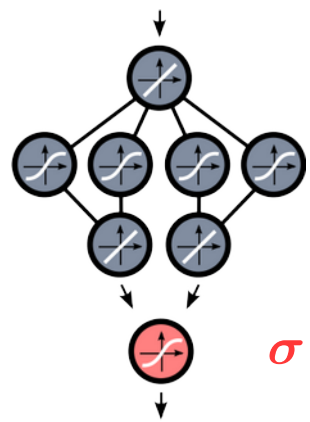
artificial neural nets

Boundary Value Problem



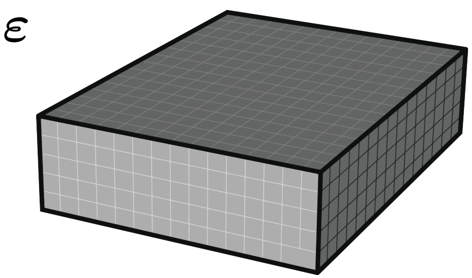
data $\longrightarrow$

unit cell



$\longleftarrow \Delta\varepsilon$

$\Delta\boldsymbol{\sigma} \longrightarrow$



$\boldsymbol{\sigma} : \dot{\boldsymbol{\varepsilon}} = \dot{\psi} + d$

$$\begin{cases} \operatorname{div} \boldsymbol{\sigma} + \boldsymbol{f} = 0 & \text{in } \mathcal{V} \\ \boldsymbol{u} = \bar{\boldsymbol{u}} & \text{on } \partial\mathcal{V}_u \\ \boldsymbol{\sigma} \cdot n = \bar{\boldsymbol{t}} & \text{on } \partial\mathcal{V}_t \end{cases}$$

thermodynamics/physics

PINN Karniadakis et al, 2019; SPNN Hernandez et al, 2021;
TANN Masi et al, 2021; Masi and Stefanou, 2022,2023

Thermodynamics

Thermodynamics

Consider a solid with reference configuration $\mathcal{B}$

$$\underbrace{\dot{e}}_{\substack{\text{int. energy}\\\text{rate}}} = \underbrace{\boldsymbol{P}}_{\substack{\text{1st Piola-}\\\text{Kirchoff stress}}} : \underbrace{\dot{\boldsymbol{F}}}_{\substack{\text{def. grad.}\\\text{rate}}} - \nabla \cdot \underbrace{\boldsymbol{q}}_{\substack{\text{ref. heat}\\\text{flux}}} + \underbrace{r}_{\substack{\text{heat}\\\text{source}}}$$

$$\underbrace{\gamma}_{\substack{\text{entropy}\\\text{product. rate}}} = \underbrace{\dot{\eta}}_{\substack{\text{entropy}\\\text{rate}}} - \nabla \cdot \left(\frac{\boldsymbol{q}}{T}\right) - \frac{r}{T} \geq 0$$

Coleman and Gurtin, 1967

Thermodynamics

Consider a solid with reference configuration $\mathcal{B}$

$$\underbrace{\dot{e}}_{\substack{\text{int. energy}\\ \text{rate}}} = \underbrace{\boldsymbol{P}}_{\substack{\text{1st Piola-}\\ \text{Kirchoff stress}}} : \underbrace{\dot{\boldsymbol{F}}}_{\substack{\text{def. grad.}\\ \text{rate}}} - \nabla \cdot \underbrace{\boldsymbol{q}}_{\substack{\text{ref. heat}\\ \text{flux}}} + \underbrace{r}_{\substack{\text{heat}\\ \text{source}}}$$

$$\underbrace{\gamma}_{\substack{\text{entropy}\\ \text{product. rate}}} = \underbrace{\dot{\eta}}_{\substack{\text{entropy}\\ \text{rate}}} - \nabla \cdot \left(\frac{\boldsymbol{q}}{T}\right) - \frac{r}{T} \geq 0$$

Coleman and Gurtin, 1967

Clausius-Duhem inequality $\gamma = \boldsymbol{P} : \dot{\boldsymbol{F}} - \left(\dot{\psi} + \eta \dot{T}\right) - \boldsymbol{q} \cdot \frac{\nabla T}{T} \geq 0$

Thermodynamics

Consider a solid with reference configuration $\mathcal{B}$

$$\underbrace{\dot{e}}_{\substack{\text{int. energy}\\ \text{rate}}} = \underbrace{\boldsymbol{P}}_{\substack{\text{1st Piola-}\\ \text{Kirchoff stress}}} : \underbrace{\dot{\boldsymbol{F}}}_{\substack{\text{def. grad.}\\ \text{rate}}} - \nabla \cdot \underbrace{\boldsymbol{q}}_{\substack{\text{ref. heat}\\ \text{flux}}} + \underbrace{r}_{\substack{\text{heat}\\ \text{source}}}$$

$$\underbrace{\gamma}_{\substack{\text{entropy}\\ \text{product. rate}}} = \underbrace{\dot{\eta}}_{\substack{\text{entropy}\\ \text{rate}}} - \nabla \cdot \left(\frac{\boldsymbol{q}}{T}\right) - \frac{r}{T} \geq 0$$

Coleman and Gurtin, 1967

Clausius-Duhem inequality

$$\gamma = \boldsymbol{P} : \dot{\boldsymbol{F}} - \left(\dot{\psi} + \eta \dot{T}\right) - \boldsymbol{q} \cdot \frac{\nabla T}{T} \geq 0$$

$$d = \boldsymbol{P} : \dot{\boldsymbol{F}} - \left(\dot{\psi} + \eta \dot{T}\right) \geq 0$$

Thermodynamics

Consider a solid with reference configuration $\mathcal{B}$

$$\underbrace{\dot{e}}_{\substack{\text{int. energy}\\ \text{rate}}} = \underbrace{\boldsymbol{P}}_{\substack{\text{1st Piola-}\\ \text{Kirchoff stress}}} : \underbrace{\dot{\boldsymbol{F}}}_{\substack{\text{def. grad.}\\ \text{rate}}} - \nabla \cdot \underbrace{\boldsymbol{q}}_{\substack{\text{ref. heat}\\ \text{flux}}} + \underbrace{r}_{\substack{\text{heat}\\ \text{source}}}$$

$$\underbrace{\gamma}_{\substack{\text{entropy}\\ \text{product. rate}}} = \underbrace{\dot{\eta}}_{\substack{\text{entropy}\\ \text{rate}}} - \nabla \cdot \left(\frac{\boldsymbol{q}}{T}\right) - \frac{r}{T} \geq 0$$

Coleman and Gurtin, 1967

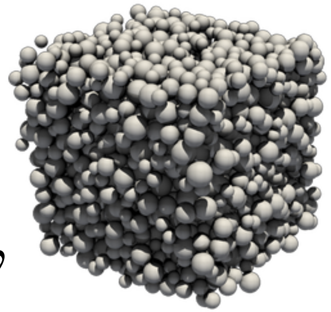
Clausius-Duhem inequality $\quad \gamma = \boldsymbol{P} : \dot{\boldsymbol{F}} - \left(\dot{\psi} + \eta \dot{T}\right) - \boldsymbol{q} \cdot \frac{\nabla T}{T} \geq 0$

$$d = \boldsymbol{P} : \dot{\boldsymbol{F}} - \left(\dot{\psi} + \eta \dot{T}\right) \geq 0$$

Volume average

$$d^{(Y)} = \boldsymbol{P}^{(Y)} : \dot{\boldsymbol{F}}^{(Y)} - \left(\dot{\psi}^{(Y)} + \eta^{(Y)} \dot{T}^{(Y)}\right) \geq 0$$

volume average operator $\quad \psi^{(Y)} = \langle\, \psi \,\rangle = \frac{1}{|\mathcal{V}|} \int_{\mathcal{V}} \psi \, dy$



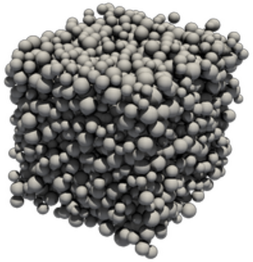
$\mathcal{V}$

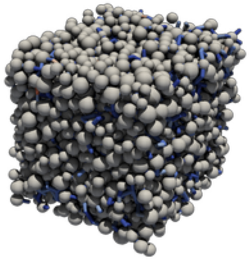
unit cell

Thermodynamics

State functions and variables

$$\psi = \hat{\psi}\left(\boldsymbol{F}, T, \boldsymbol{z}\right)$$





$\boldsymbol{z}$ **internal state variables**
internal material structure and mechanisms

evolution equations

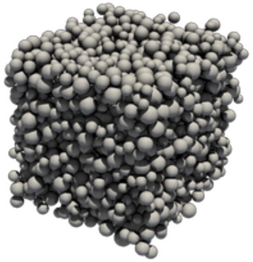
$$\dot{\boldsymbol{z}} = f\left(\boldsymbol{F}, T, \boldsymbol{z}\right)$$

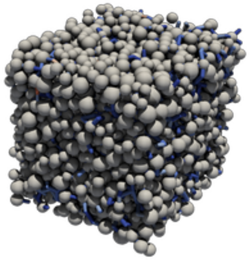
Coleman and Gurtin, 1967

Thermodynamics

State functions and variables

$$\psi = \hat{\psi}\left(\boldsymbol{F}, T, \boldsymbol{z}\right)$$





$\boldsymbol{z}$ **internal state variables**
internal material structure and mechanisms

evolution equations

$$\dot{\boldsymbol{z}} = f\left(\boldsymbol{F}, T, \boldsymbol{z}\right)$$

Coleman and Gurtin, 1967

Time differentiation

$$\dot{\psi} = \left(\partial_{\boldsymbol{F}}\hat{\psi}\right) \cdot \dot{\boldsymbol{F}} + \left(\partial_{T}\hat{\psi}\right)\dot{T} + \left(\partial_{\boldsymbol{z}}\hat{\psi}\right) \cdot \dot{\boldsymbol{z}}$$

Thermodynamics

State functions and variables

$$\psi = \hat{\psi}\left(\boldsymbol{F}, T, \boldsymbol{z}\right)$$


for more, see ALERT DS 2018, 2021
(I Einav, E Gerolymatou, C Tamagnini, D Mašin)

$\boldsymbol{z}$ **internal state variables**
internal material structure and mechanisms

evolution equations

$$\dot{\boldsymbol{z}} = f\left(\boldsymbol{F}, T, \boldsymbol{z}\right)$$

Coleman and Gurtin, 1967

Time differentiation

$$\dot{\psi} = \left(\partial_{\boldsymbol{F}}\hat{\psi}\right) \cdot \dot{\boldsymbol{F}} + \left(\partial_{T}\hat{\psi}\right)\dot{T} + \left(\partial_{\boldsymbol{z}}\hat{\psi}\right) \cdot \dot{\boldsymbol{z}}$$

Thermodynamic admissible processes

$$d = \boldsymbol{P} : \dot{\boldsymbol{F}} - \left(\dot{\psi} + \eta\dot{T}\right)$$

stress function $\boldsymbol{P} = \partial_{\boldsymbol{F}}\hat{\psi}\left(\boldsymbol{F}, T, \boldsymbol{z}\right)$

entropy function $\eta = -\partial_{T}\hat{\psi}\left(\boldsymbol{F}, T, \boldsymbol{z}\right)$

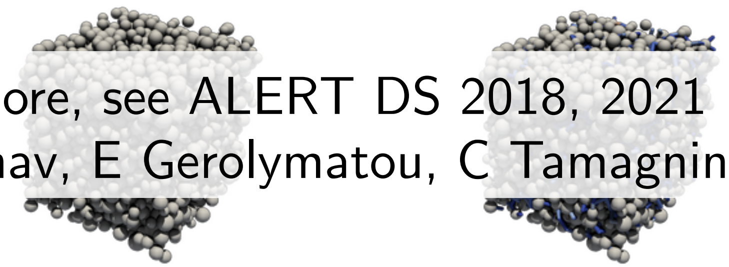
internal dissipation rate $d = \boldsymbol{\Pi}\left(\boldsymbol{F}, T, \boldsymbol{z}\right) \cdot \dot{\boldsymbol{z}}$

$$\boldsymbol{\Pi} \equiv -\partial_{\boldsymbol{z}}\hat{\psi}\left(\boldsymbol{F}, T, \boldsymbol{z}\right)$$

Thermodynamics

State functions and variables

$$\psi = \hat{\psi}\left(\boldsymbol{F}, T, \boldsymbol{z}\right)$$



for more, see ALERT DS 2018, 2021
(I Einav, E Gerolymatou, C Tamagnini, D Mašin)

$\boldsymbol{z}$ **internal state variables**
internal material structure and mechanisms

evolution equations

$$\dot{\boldsymbol{z}} = f\left(\boldsymbol{F}, T, \boldsymbol{z}\right)$$

Coleman and Gurtin, 1967

Time differentiation

$$\dot{\psi} = \left(\partial_{\boldsymbol{F}}\hat{\psi}\right)\cdot\dot{\boldsymbol{F}} + \left(\partial_{T}\hat{\psi}\right)\dot{T} + \left(\partial_{\boldsymbol{z}}\hat{\psi}\right)\cdot\dot{\boldsymbol{z}}$$

Thermodynamic admissible processes

$$d = \boldsymbol{P} : \dot{\boldsymbol{F}} - \left(\dot{\psi} + \eta\dot{T}\right) \geq 0$$

stress function $\boldsymbol{P} = \partial_{\boldsymbol{F}}\hat{\psi}\left(\boldsymbol{F}, T, \boldsymbol{z}\right)$

entropy function $\eta = -\partial_{T}\hat{\psi}\left(\boldsymbol{F}, T, \boldsymbol{z}\right)$

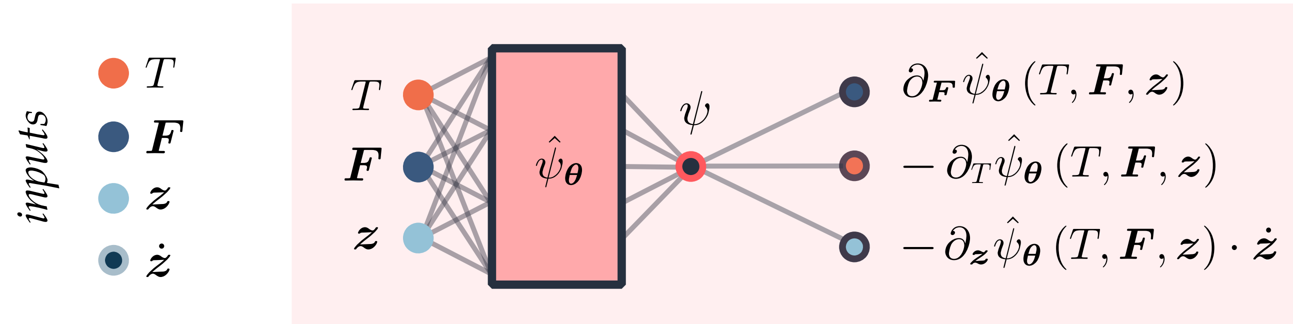
internal dissipation rate $d = \boldsymbol{\Pi}\left(\boldsymbol{F}, T, \boldsymbol{z}\right)\cdot\dot{\boldsymbol{z}} \geq 0$

$$\boldsymbol{\Pi} \equiv -\partial_{\boldsymbol{z}}\hat{\psi}\left(\boldsymbol{F}, T, \boldsymbol{z}\right)$$

Thermodynamics-based Artificial Neural Networks

TANN: Thermodynamics-based Artificial Neural Networks

Free energy density network



$$\text{minimize} \qquad \mathcal{L}_{\psi}(\boldsymbol{s}_{\hat{\psi}}) + \mathcal{L}_{\mathrm{grad}\psi}(\boldsymbol{s}_{\hat{\psi}}) + \lambda_{\mathrm{reg}}\mathcal{L}_{\mathrm{reg}}(\boldsymbol{s}_{\hat{\psi}})$$

$$\mathcal{L}_{\psi} = \left\|\psi - \hat{\psi}\left(T,\boldsymbol{F},\boldsymbol{z}\right)\right\|$$

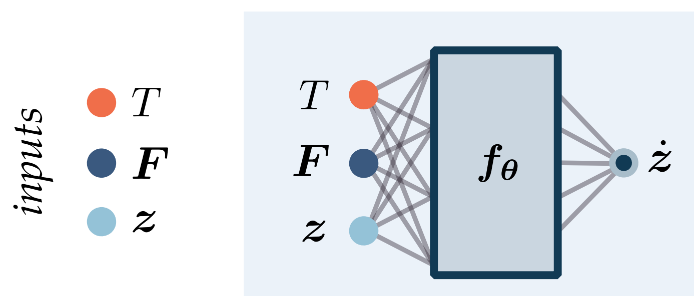
$$\begin{aligned}\mathcal{L}_{\mathrm{grad}\psi} = &\left\|\boldsymbol{P} - \partial_{\boldsymbol{F}}\hat{\psi}\left(T,\boldsymbol{F},\boldsymbol{z}\right)\right\| + \\ &\left\|\eta + \partial_{T}\hat{\psi}\left(T,\boldsymbol{F},\boldsymbol{z}\right)\right\| + \\ &\left\|d + \partial_{\boldsymbol{z}}\hat{\psi}\left(T,\boldsymbol{F},\boldsymbol{z}\right)\cdot\dot{\boldsymbol{z}}\right\|\end{aligned}$$

$$\mathcal{L}_{\mathrm{reg}} = \left\|\left\{\partial_{\boldsymbol{z}}\hat{\psi}\left(T,\boldsymbol{F},\boldsymbol{z}\right)\cdot\dot{\boldsymbol{z}}\right\}\right\|$$

$$\{x\} = \max(x,0)$$

Evolution equation network

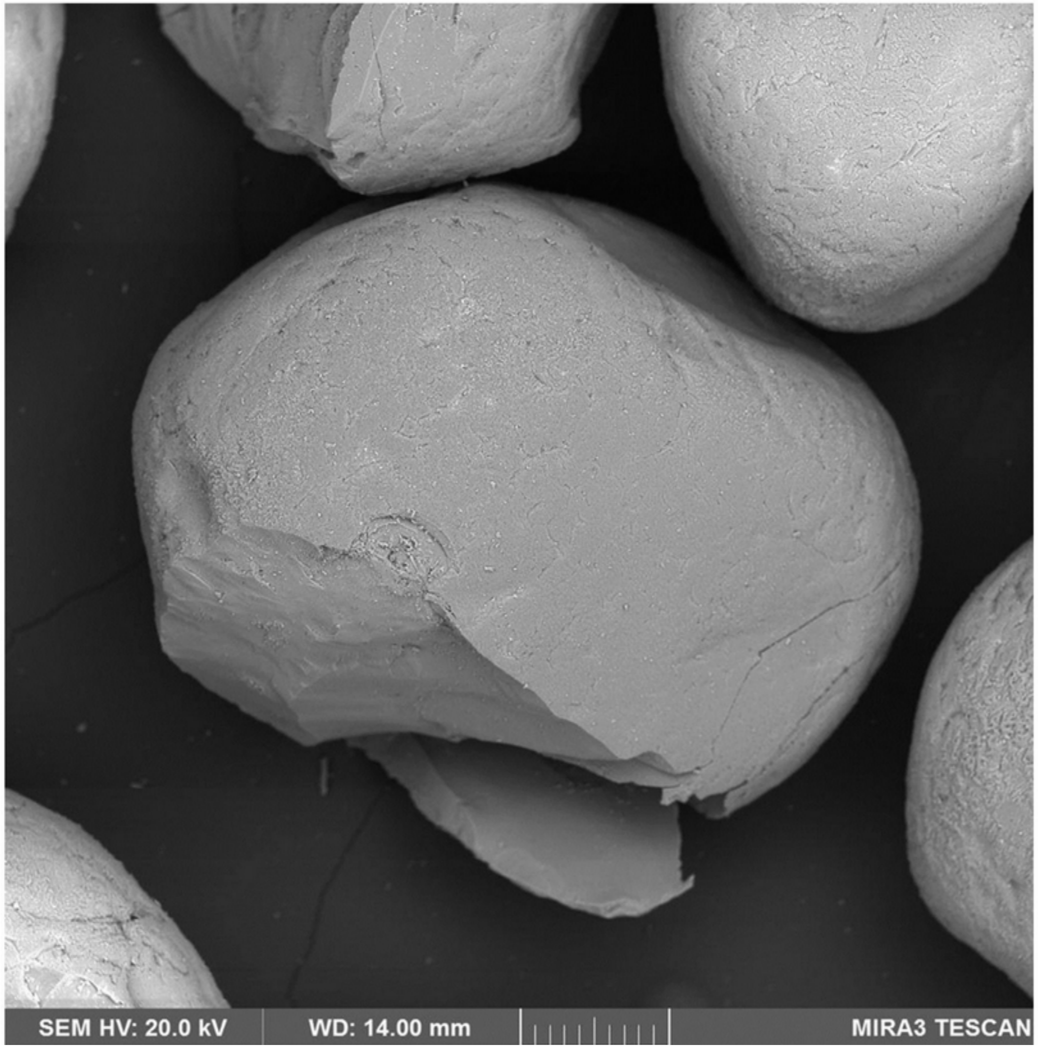
Masi, Stefanou. *J Mech Phys Solids*, 2023

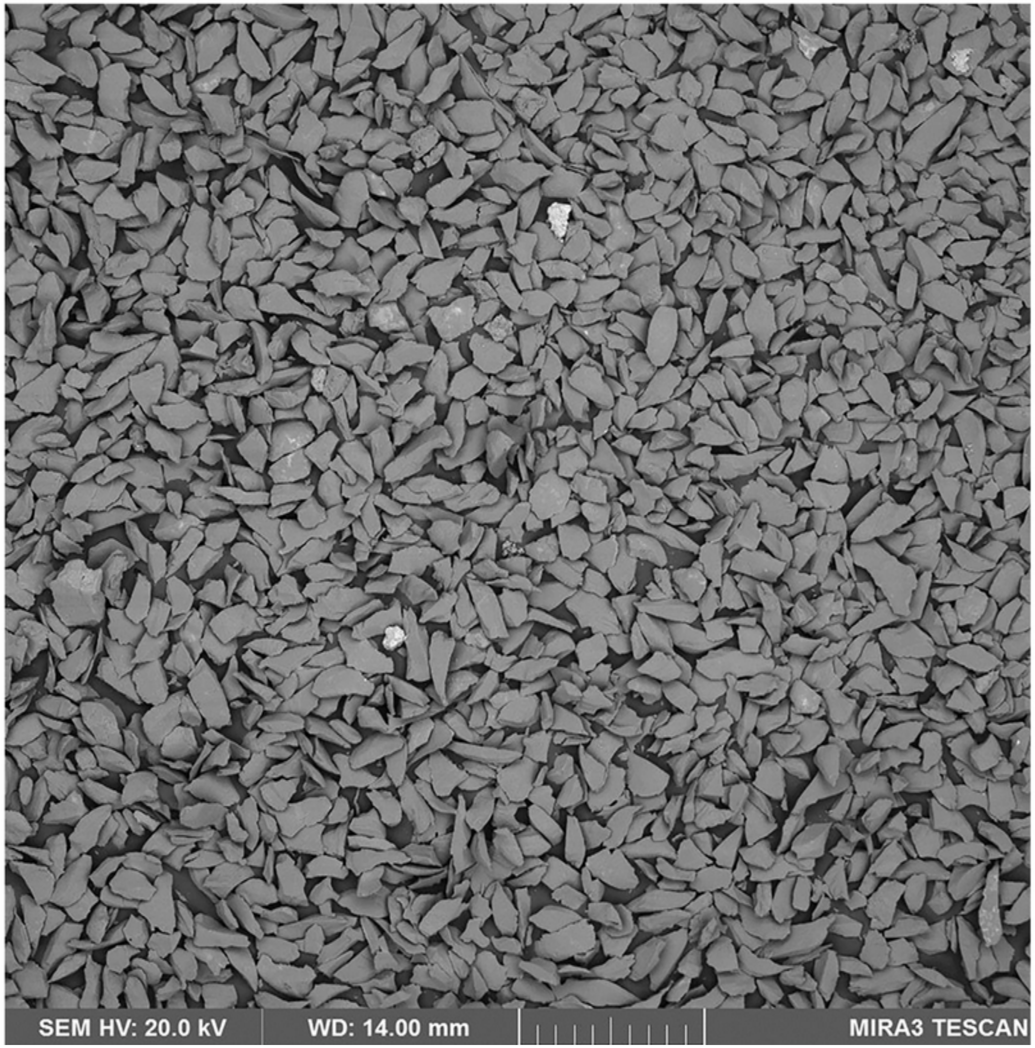


$$\mathcal{L}_{\dot{\boldsymbol{z}}} = \left\|\dot{\boldsymbol{z}} - \boldsymbol{f}\left(T,\boldsymbol{F},\boldsymbol{z}\right)\right\| \qquad \text{minimize} \qquad \mathcal{L}_{\dot{\boldsymbol{z}}}(\boldsymbol{s}_{\boldsymbol{f}})$$

Discovery of internal variables

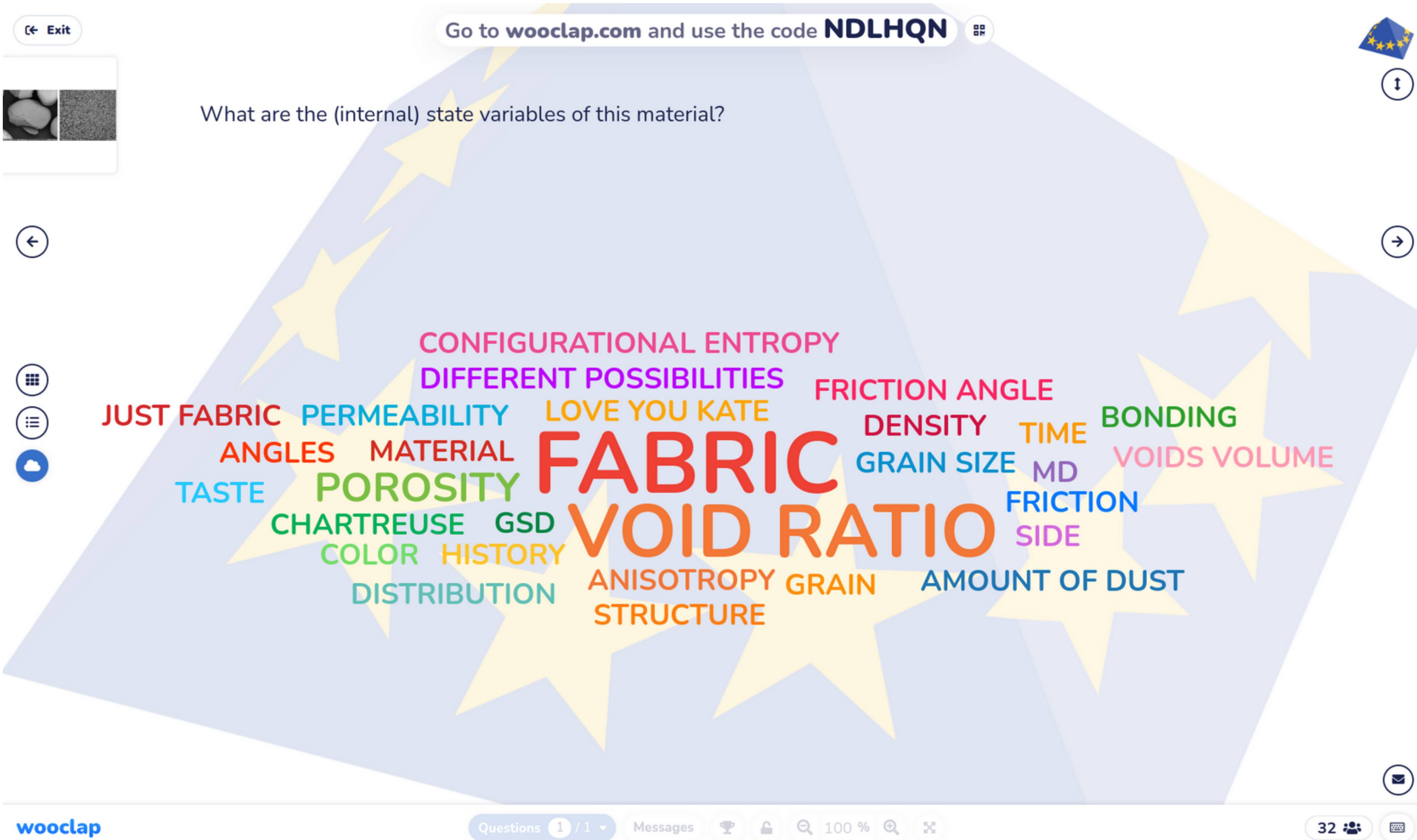
The quest for internal variables





Zheng et al. *Int J Geomech*, 2020

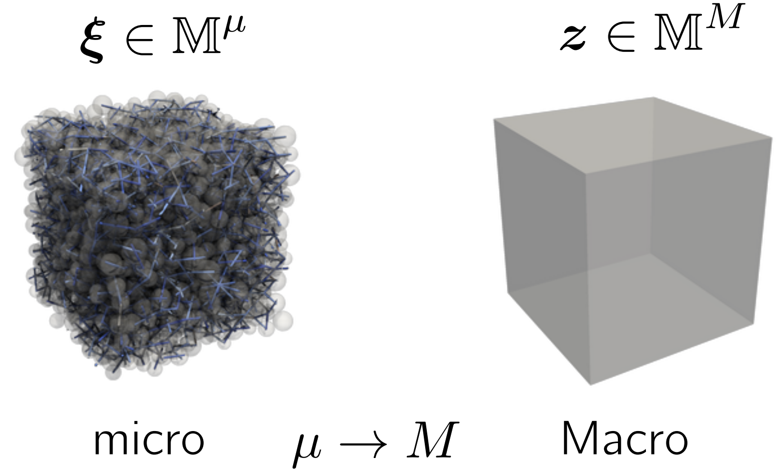
The quest for internal variables



The quest for internal variables

def. **internal coordinates**

those microscopic quantities
describing internal mechanisms



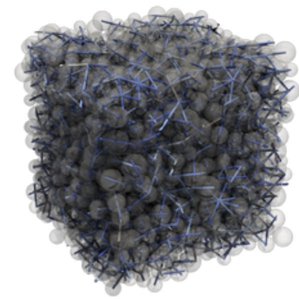
Masi et Stefanou. *Comput Methods Appl Mech Eng.* 2022

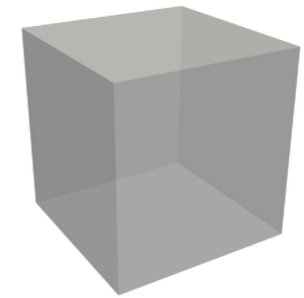
The quest for internal variables

def. **internal coordinates**

those microscopic quantities
describing internal mechanisms

$\boldsymbol{\xi} \in \mathbb{M}^{\mu}$ $\qquad$ $\boldsymbol{z} \in \mathbb{M}^{M}$





micro $\quad \mu \rightarrow M \quad$ Macro

internal variables

latent representations of the internal coordinates

$\boldsymbol{z} = \boldsymbol{h}\left(\boldsymbol{\xi}\right)$ s.t. $\psi = \hat{\psi}\left(T, \boldsymbol{F}, \boldsymbol{z}\right)$ $\qquad$ $\tilde{\boldsymbol{\xi}} = \boldsymbol{g}\left(\boldsymbol{z}\right)$ pseudoinverse of $\boldsymbol{h}(\cdot)$

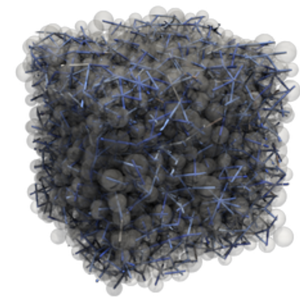
Masi et Stefanou. *Comput Methods Appl Mech Eng.* 2022

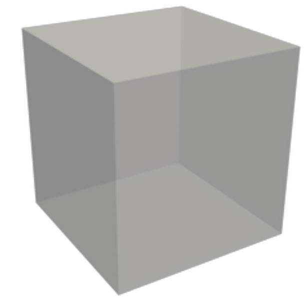
The quest for internal variables

def. **internal coordinates**

those microscopic quantities
describing internal mechanisms

$\boldsymbol{\xi} \in \mathbb{M}^{\mu}$ $\boldsymbol{z} \in \mathbb{M}^{M}$





micro $\mu \to M$ Macro

internal variables

latent representations of the internal coordinates

$$\boldsymbol{z} = \boldsymbol{h}\left(\boldsymbol{\xi}\right) \qquad \text{s.t.} \qquad \psi = \hat{\psi}\left(T, \boldsymbol{F}, \boldsymbol{z}\right)$$

$$\tilde{\boldsymbol{\xi}} = \boldsymbol{g}\left(\boldsymbol{z}\right)$$ pseudoinverse of $\boldsymbol{h}(\cdot)$

evolution equations

from the evolution of the internal coordinates

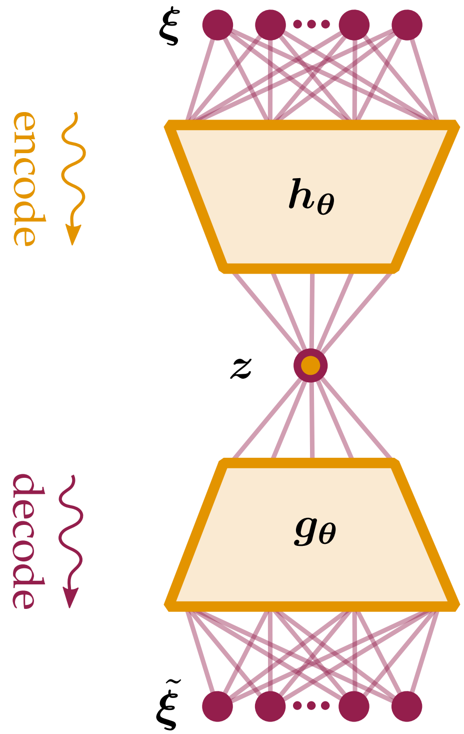
$$\dot{\boldsymbol{z}} = \dot{\boldsymbol{h}}(\boldsymbol{\xi}) = \partial_{\boldsymbol{\xi}} \boldsymbol{h}(\boldsymbol{\xi}) \cdot \dot{\boldsymbol{\xi}} \qquad \dot{\boldsymbol{z}} = \boldsymbol{f}\left(\boldsymbol{F}, T, \boldsymbol{z}\right)$$

and viceversa

$$\dot{\tilde{\boldsymbol{\xi}}} = \dot{\boldsymbol{g}}(\boldsymbol{z}) = \partial_{\boldsymbol{z}} \boldsymbol{g}(\boldsymbol{z}) \cdot \dot{\boldsymbol{z}}$$

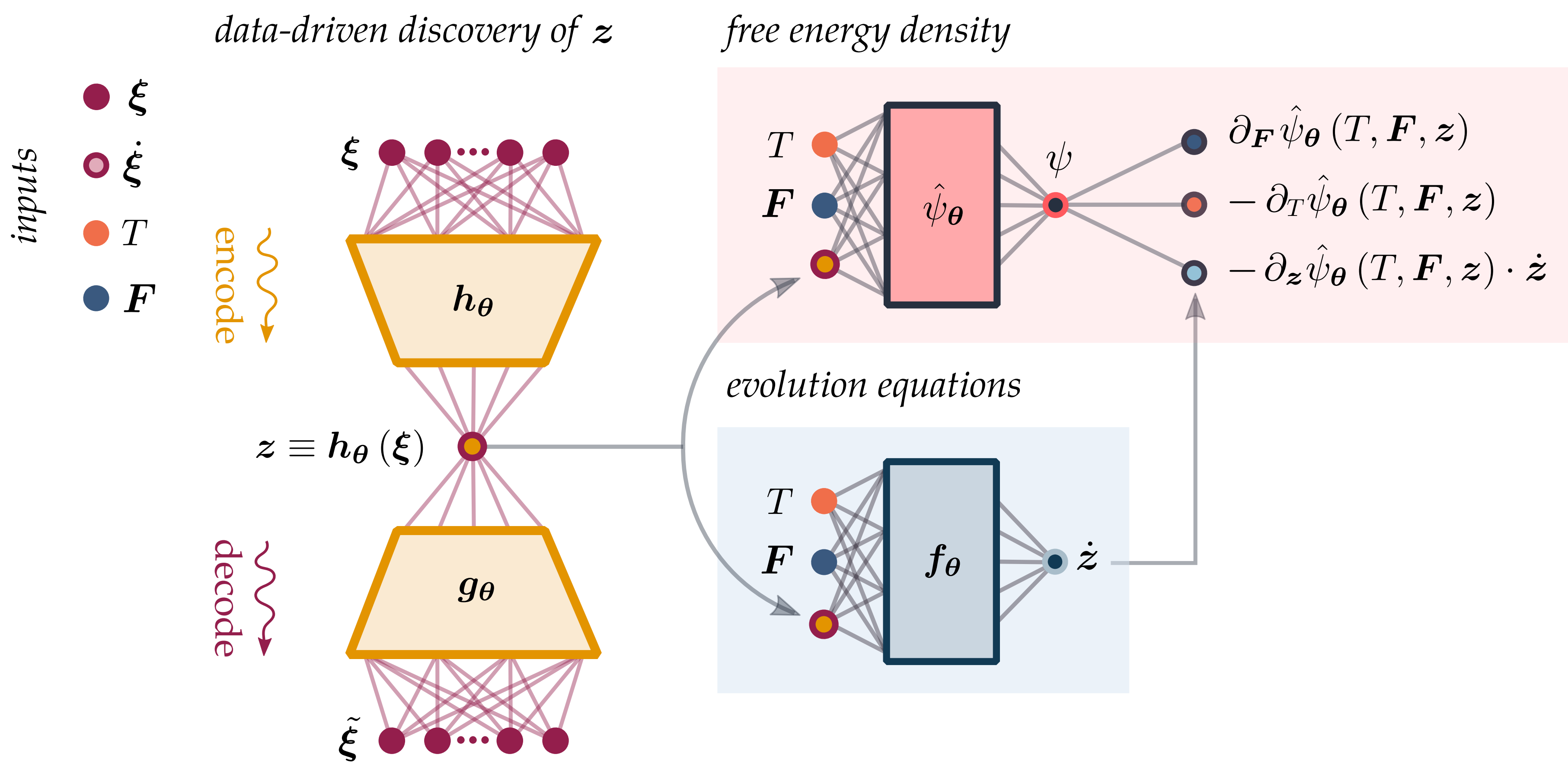
Masi et Stefanou. *Comput Methods Appl Mech Eng.* 2022

TANN and the discovery of internal variables



$$\underbrace{\left\|\boldsymbol{\xi} - \boldsymbol{g}\left(\boldsymbol{h}\left(\boldsymbol{\xi}\right)\right)\right\|}_{\text{reconstruction loss}}$$

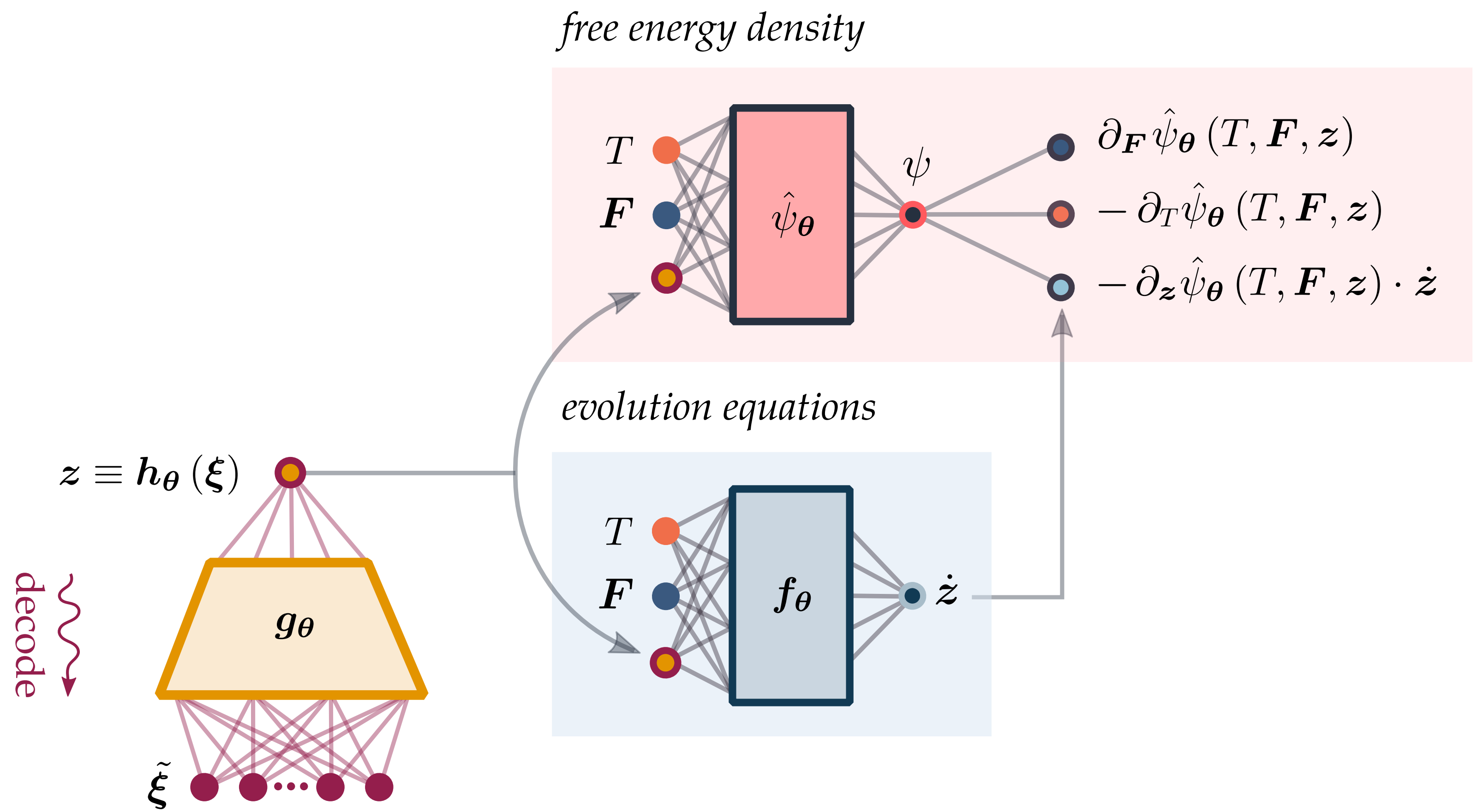
TANN and the discovery of internal variables



$$\underbrace{\left\|\boldsymbol{\xi} - \boldsymbol{g}\left(\boldsymbol{h}\left(\boldsymbol{\xi}\right)\right)\right\|}_{\text{reconstruction loss}} + \underbrace{\left\|\dot{\boldsymbol{\xi}} - \partial_{\boldsymbol{z}}\boldsymbol{g}\left(\boldsymbol{z}\right) \cdot \dot{\boldsymbol{z}}\right\|}_{\text{loss in } \dot{\boldsymbol{\xi}}} + \underbrace{\left\|\dot{\boldsymbol{z}} - \boldsymbol{f}\left(\theta, \boldsymbol{F}, \boldsymbol{z}\right)\right\|}_{\text{loss in } \dot{\boldsymbol{z}}} + \underbrace{\left\|\psi - \hat{\psi}\left(\theta, \boldsymbol{F}, \boldsymbol{z}\right)\right\|}_{\text{loss in } \psi} +$$

$$\underbrace{\left\|\boldsymbol{P} - \partial_{\boldsymbol{F}}\hat{\psi}\left(\theta, \boldsymbol{F}, \boldsymbol{z}\right)\right\| + \left\|\eta + \partial_\theta\hat{\psi}\left(\theta, \boldsymbol{F}, \boldsymbol{z}\right)\right\| + \left\|d + \partial_{\boldsymbol{z}}\hat{\psi}\left(\theta, \boldsymbol{F}, \boldsymbol{z}\right) \cdot \dot{\boldsymbol{z}}\right\|}_{\text{loss in grad}\,\psi} + \lambda_{\text{reg}} \underbrace{\left\|\left\{\partial_{\boldsymbol{z}}\hat{\psi}\left(\theta, \boldsymbol{F}, \boldsymbol{z}\right) \cdot \dot{\boldsymbol{z}}\right\}\right\|}_{\text{regularization}}$$

Inference



Constitutive equations $\quad \dot{\boldsymbol{z}}(t), \boldsymbol{P}(t) = \boldsymbol{e}\mathbf{TANN}\big(T(t), \boldsymbol{F}(t), \boldsymbol{z}(t)\big) \quad \forall\, t$

Initial Value Problem $\quad \dot{\boldsymbol{z}}(t) = \boldsymbol{f}\big(T(t), \boldsymbol{F}(t), \boldsymbol{z}(t)\big), \quad \boldsymbol{z}(t_0) = \boldsymbol{z}_0.$

Downscaling – "*localization*" $\quad \boldsymbol{\xi}(t) = \boldsymbol{g}\big(\boldsymbol{z}(t)\big)$

Constitutive modeling at the material point level

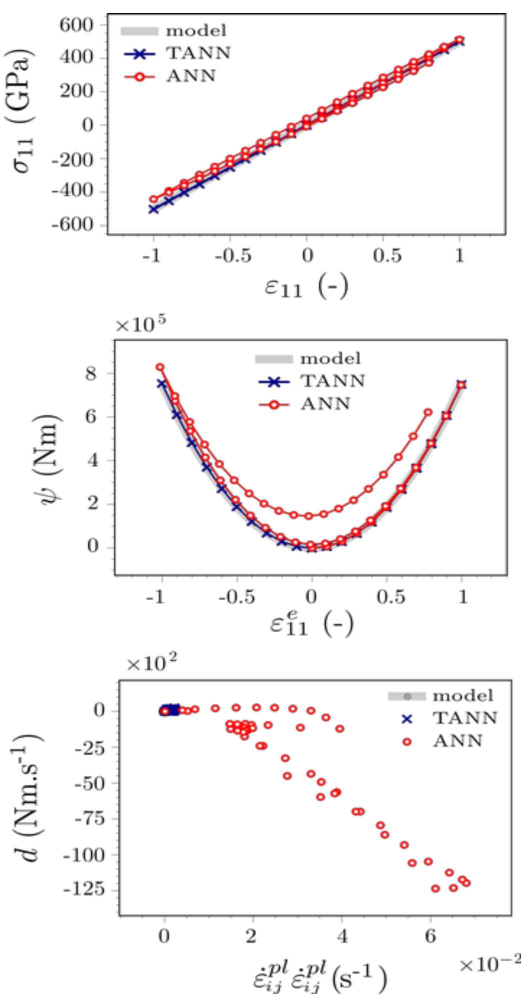
Why TANN and not ANN?

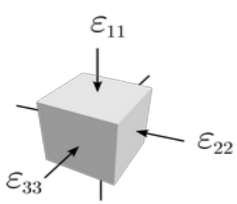
Masi et al, *J Mech Phys Solids.* 2021

hp. isothermal material processes and small-strain regime

$$\theta(t) = \theta_0 \; \forall \, t \qquad \boldsymbol{\sigma} \approx \boldsymbol{P} \qquad \boldsymbol{\varepsilon} \approx \boldsymbol{F}^{\mathrm{Sym}} + \boldsymbol{I}$$

Hyper-plasticity

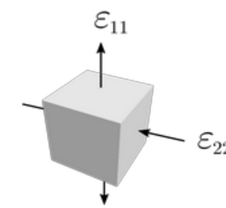


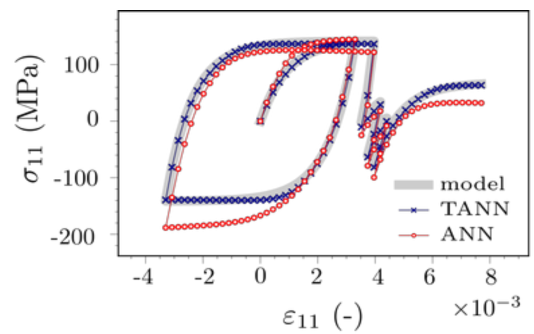


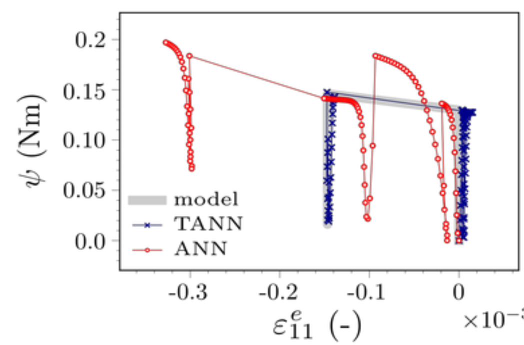
ANN

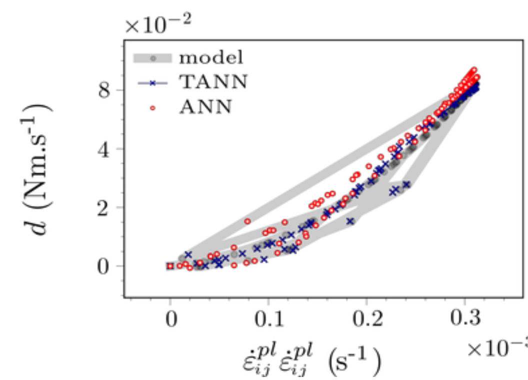
TANN

Hypo-plasticity



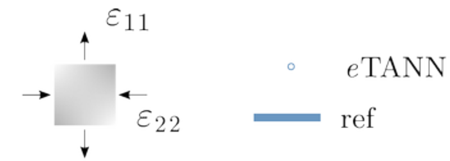




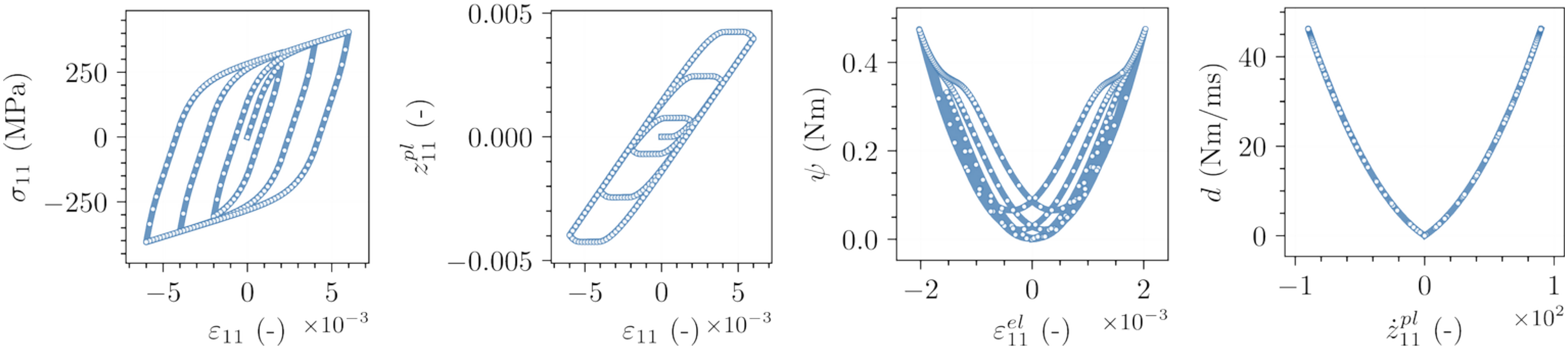


Rate-dependent materials

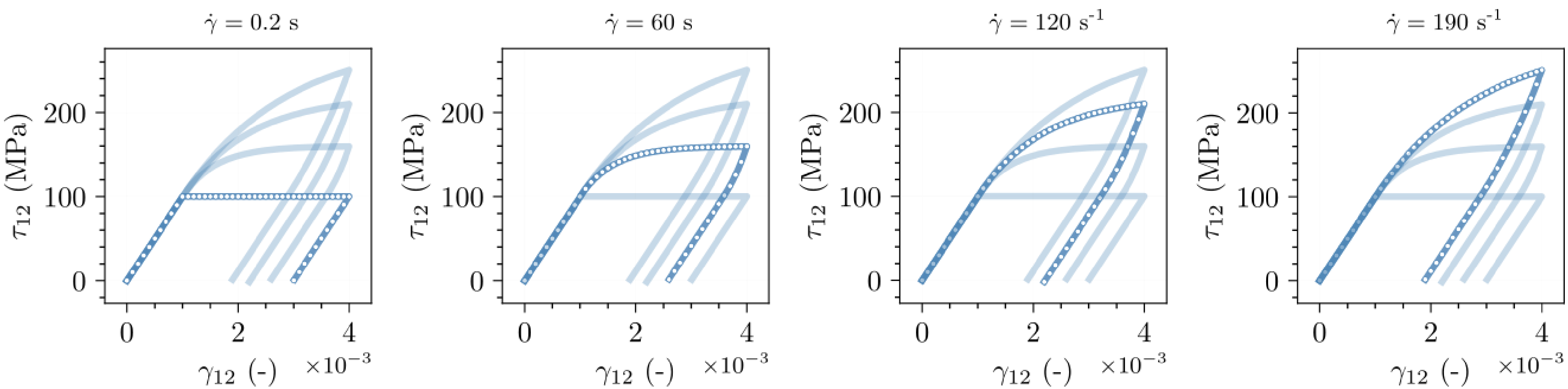
$$K = 167\ \text{GPa},\ G = 77\ \text{GPa},\ c = 100\ \text{MPa},\ \mu = 1\ \text{s}$$



- Inference mode (continuous-time)



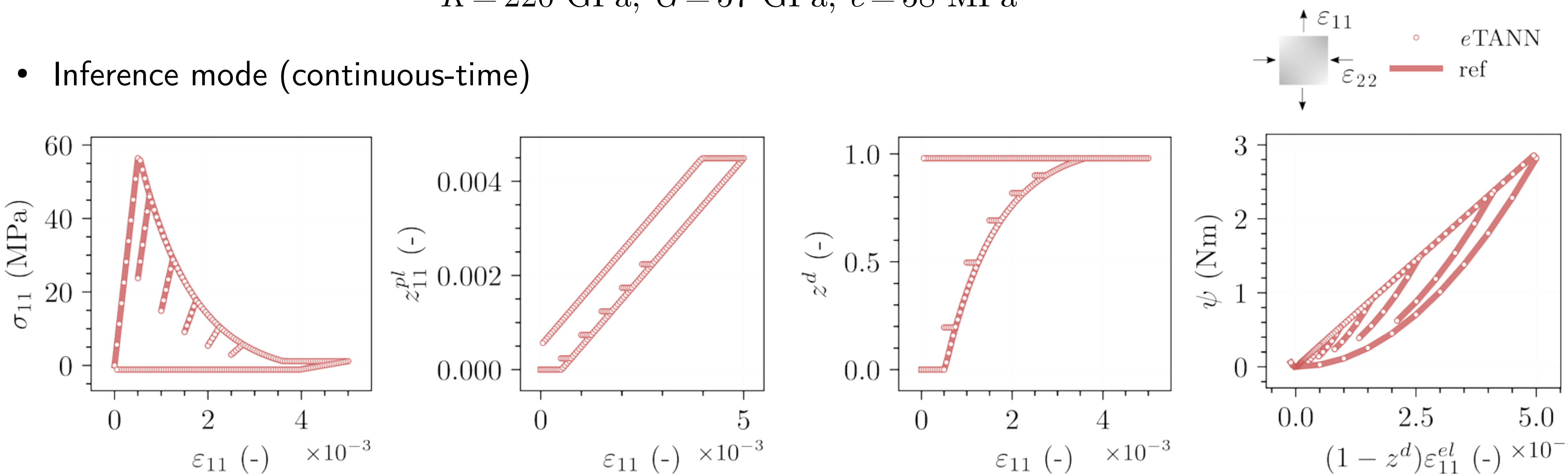
- Rate effects



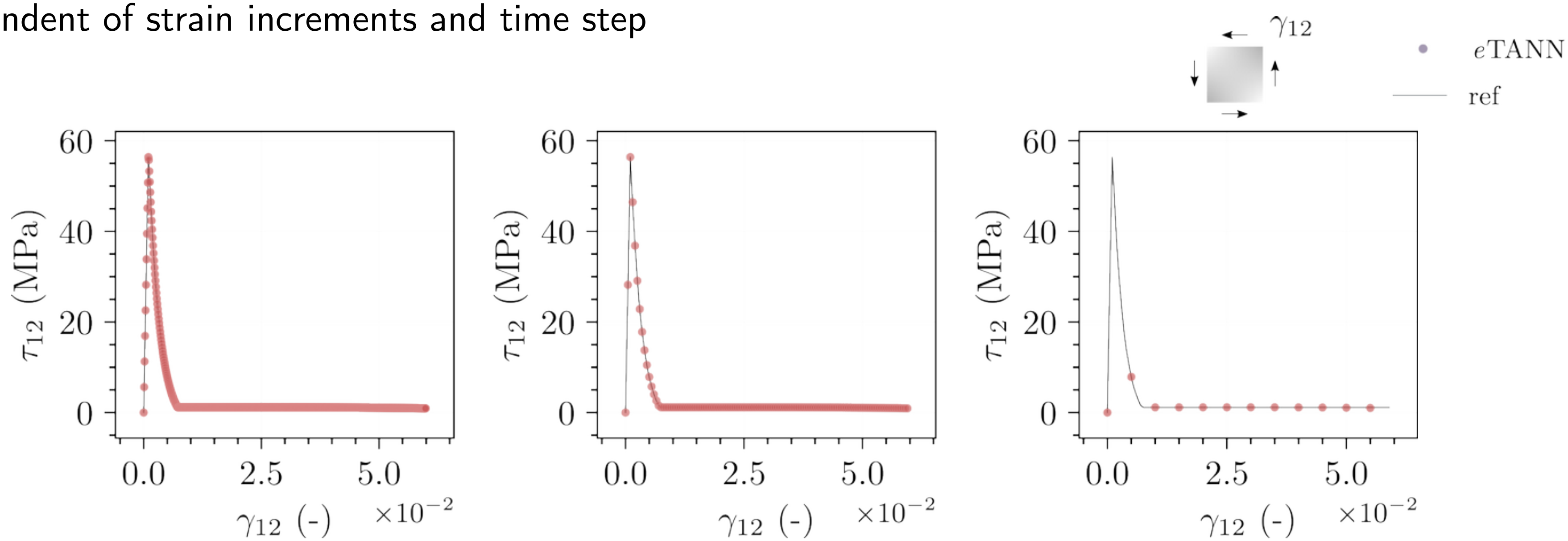
Elasto-plastic materials with damage

$$K = 220 \text{ GPa}, \; G = 57 \text{ GPa}, \; c = 58 \text{ MPa}$$

- Inference mode (continuous-time)



- Independent of strain increments and time step

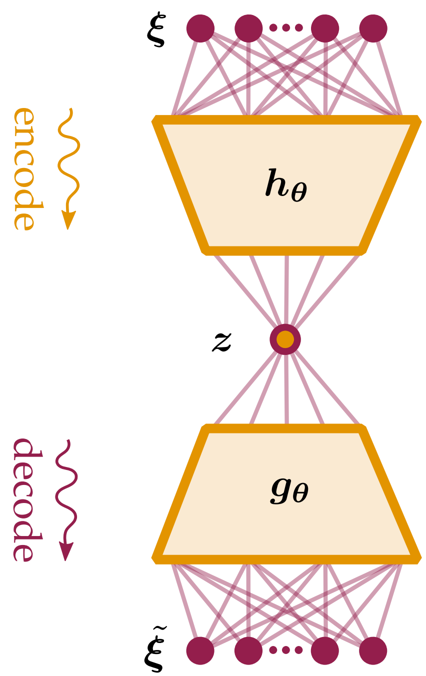


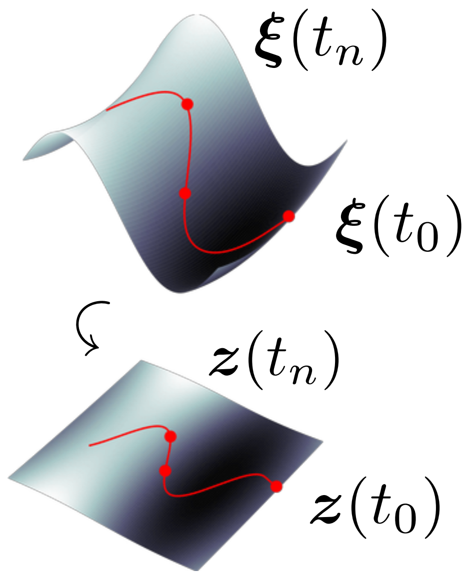
PINN+TANN / TANN (hands-on)

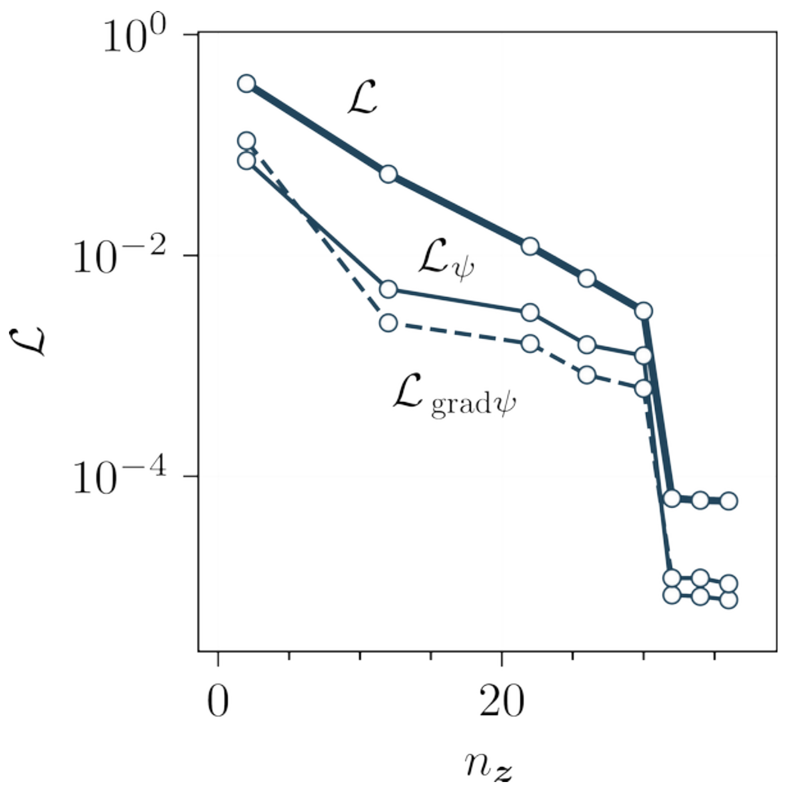
https://qrco.de/piml

Heterogeneous materials

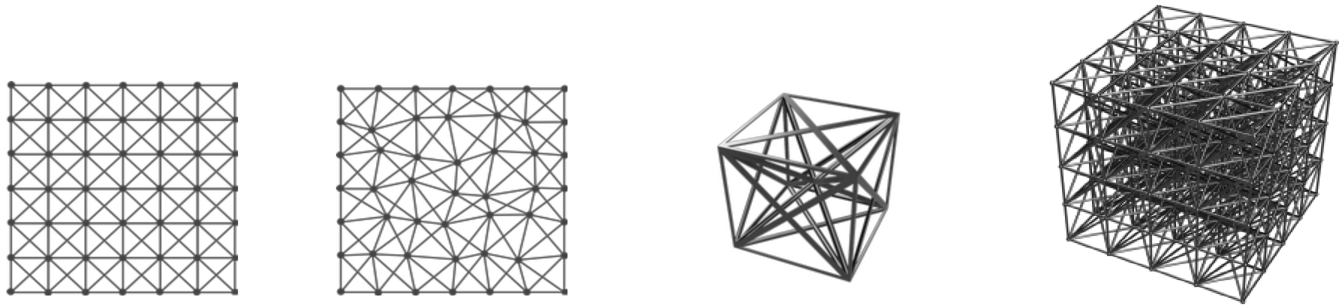
Internal variables discovery: how many internal variables?



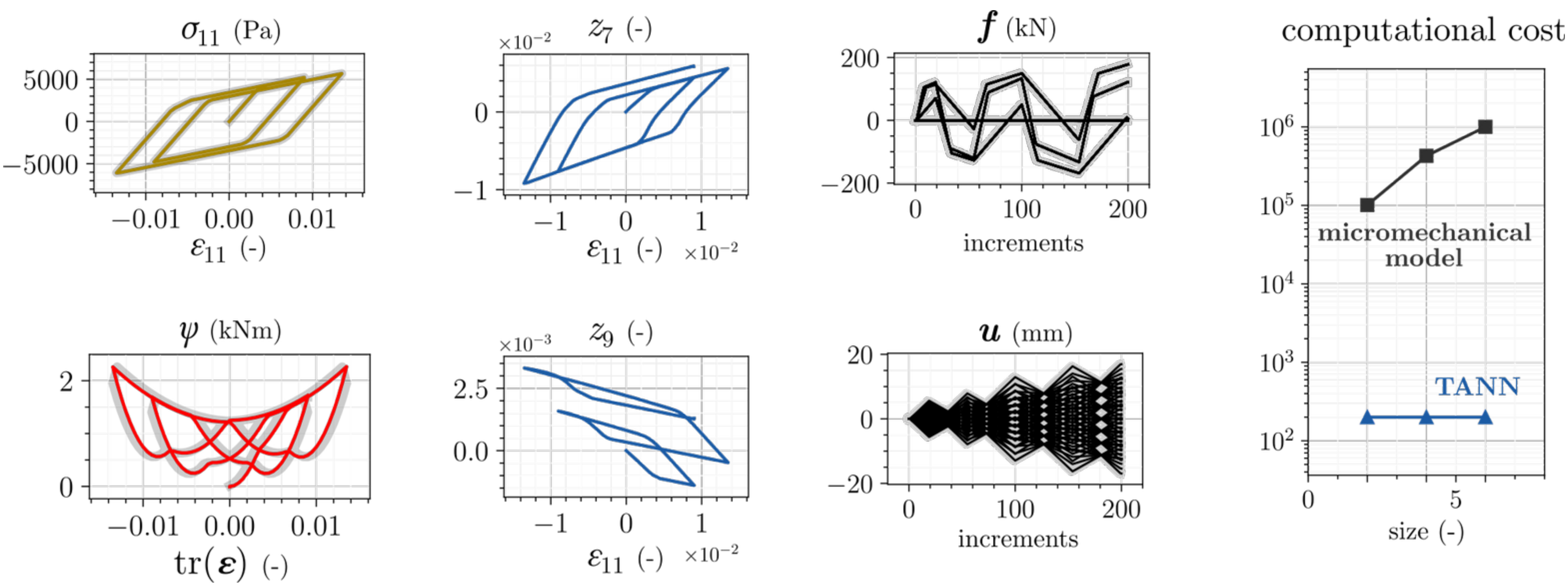




Lattice materials



2D regular　2D irregular　3D cell　3D regular

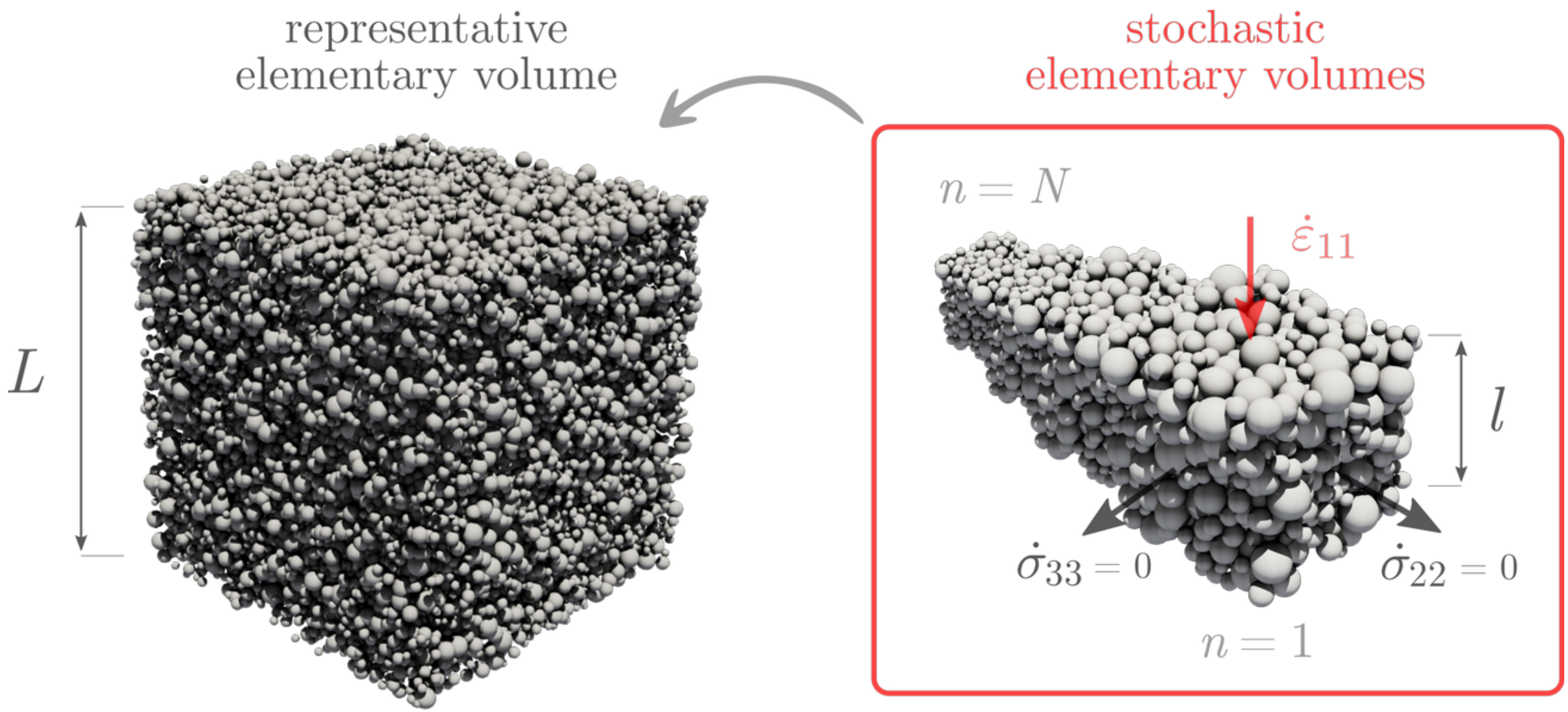


Digital twins of granular materials: a pedagogic example

Digital twins of granular materials: a pedagogic example

Virtual twin: DEM – stochastic representation (statistical ensemble)

Multiple, small periodic stochastic elementary volumes (SEVs) lead, after averaging, to the identification of the representative elementary volume (REV)

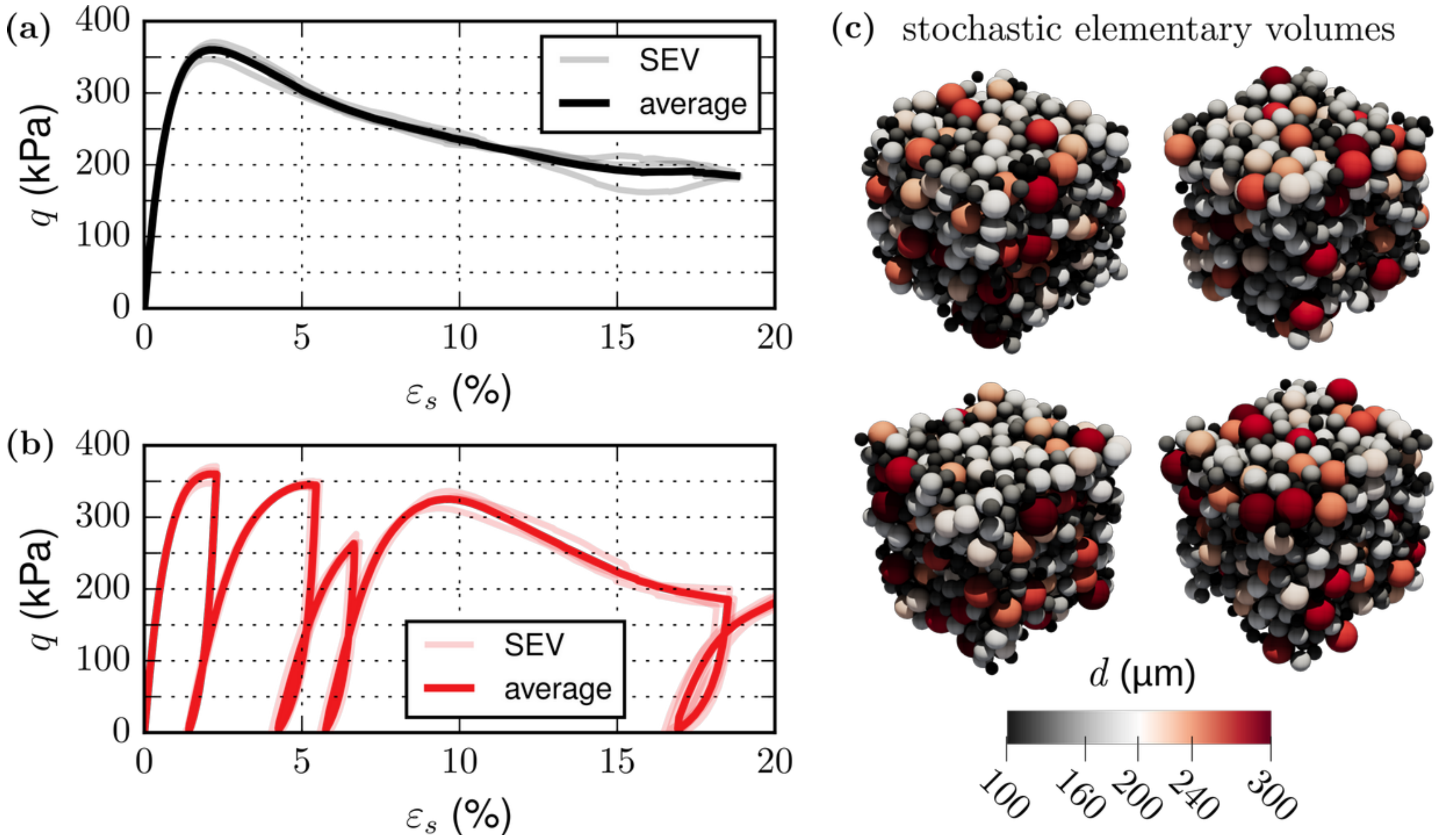


Papachristos et al., 2022, JGR: Solid Earth 128 (1)

Digital twins of granular materials

Virtual twin: DEM – stochastic representation (statistical ensemble)

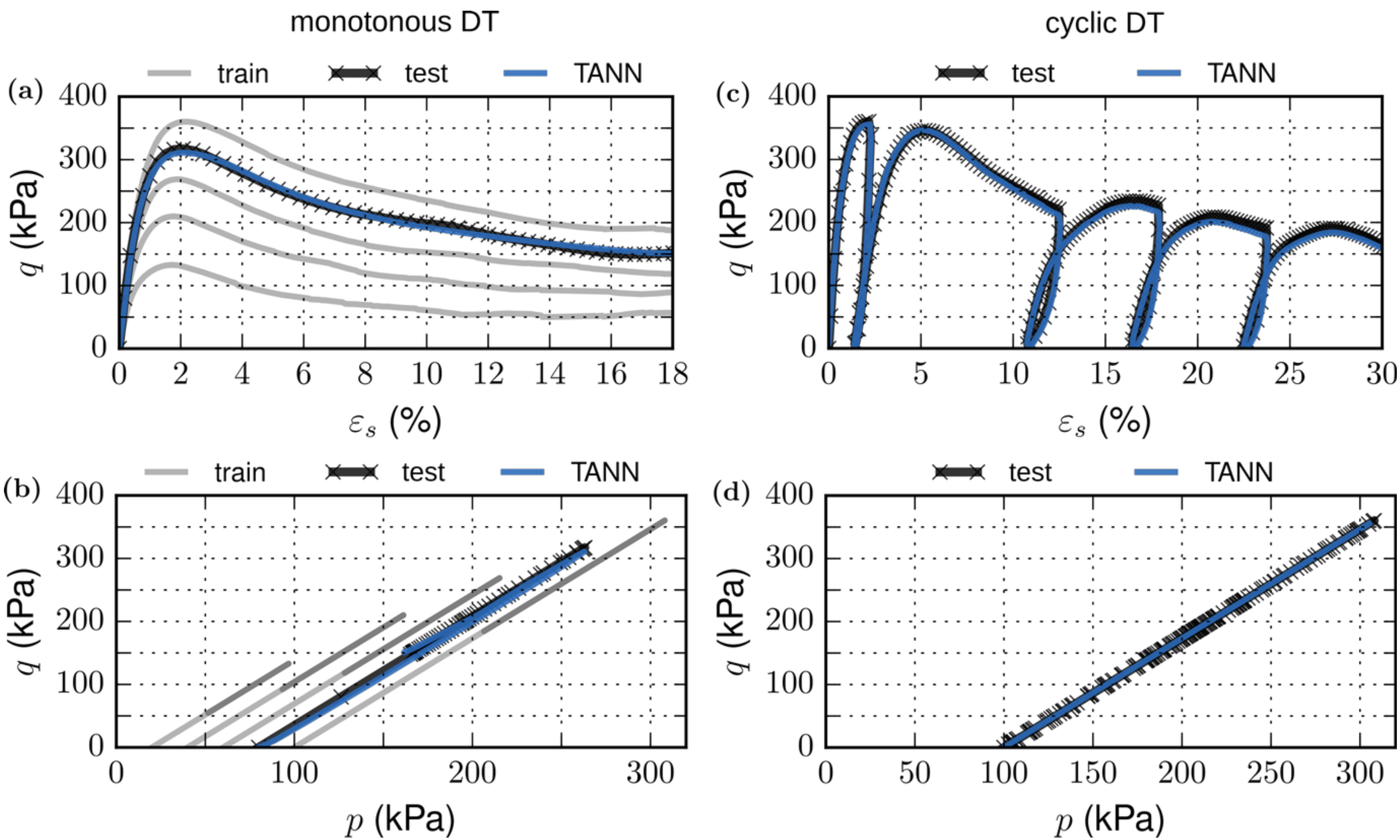
Multiple, small periodic stochastic elementary volumes (SEVs) lead, after averaging, to the identification of the representative elementary volume (REV)



Papachristos et al., 2022, JGR: Solid Earth 128 (1)

Digital twins of granular materials

Digital twin: TANN inference

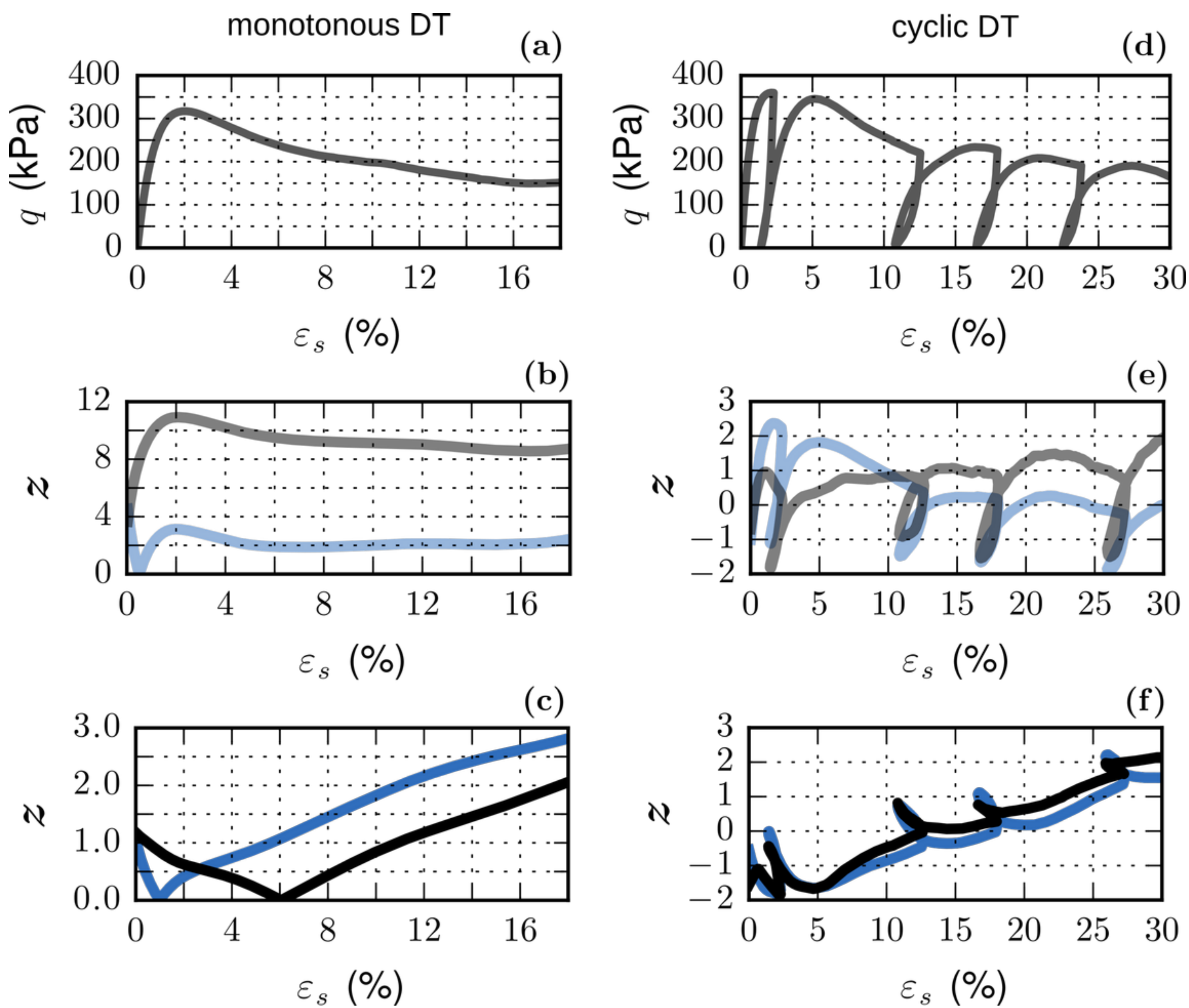


Digital twins of granular materials

Discovered internal state variables...!

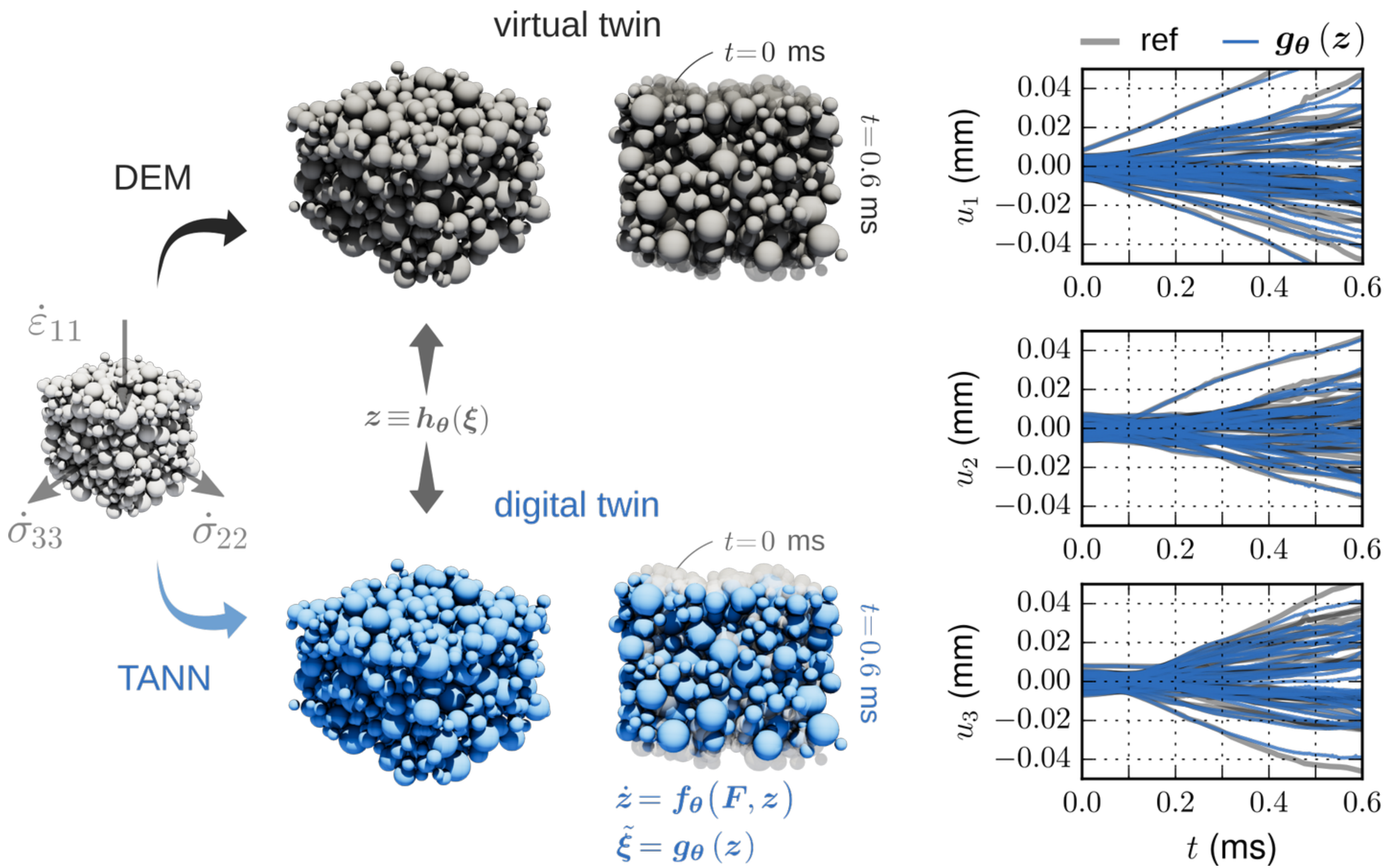
Digital twins of granular materials

Discovered internal state variables...!



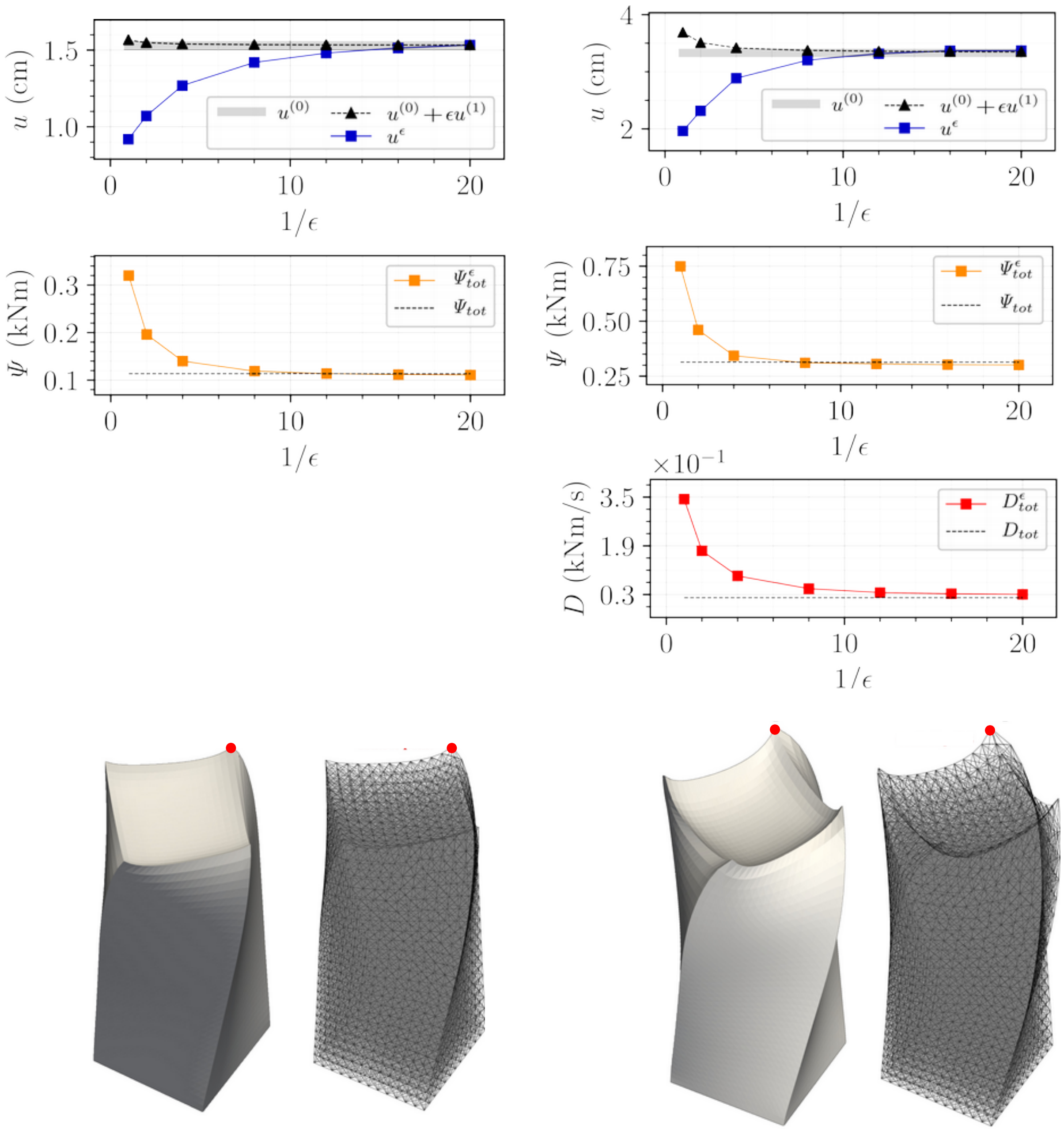
Digital twins of granular materials

From micro to Macro (on-the-fly reconstruction)



Bridging the multiple scales

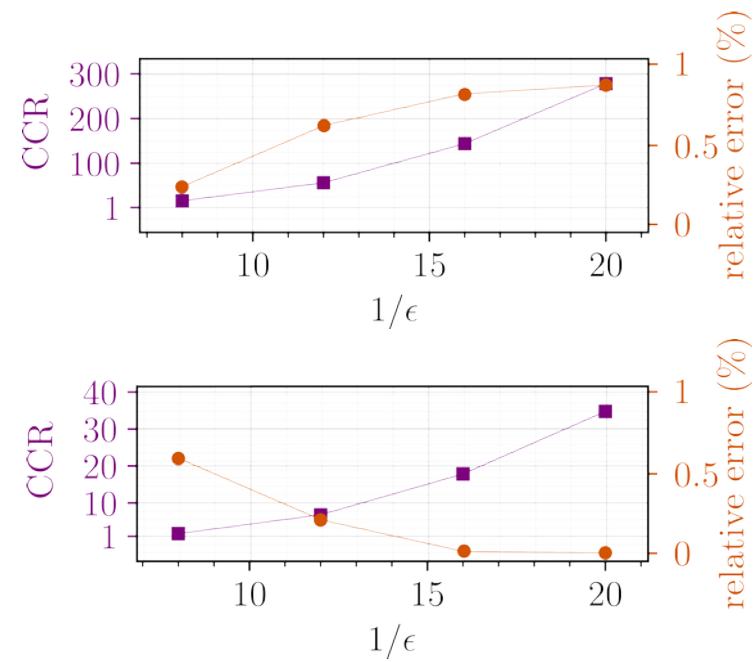
Masi and Stefanou. *Comput Methods Appl Mech Eng.* 2022



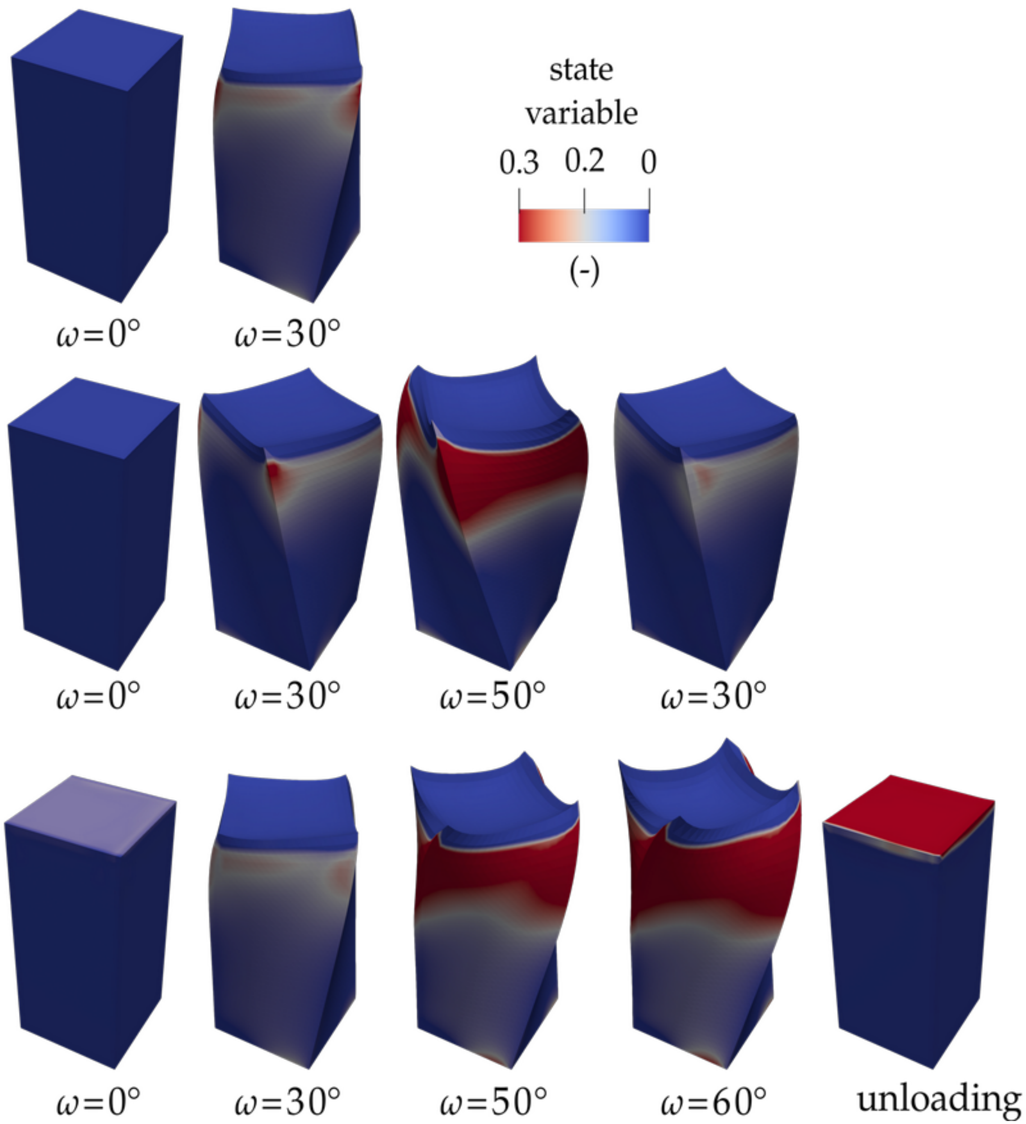
Micromechanical vs FEM×TANN

- High accuracy (despite the first-order homogenization scheme)

- Computational Cost Ratio

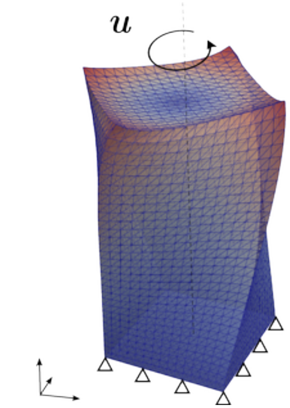


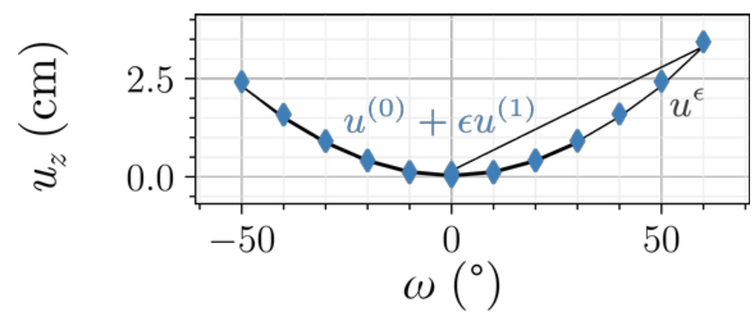
FEM×TANN – Cyclic loading

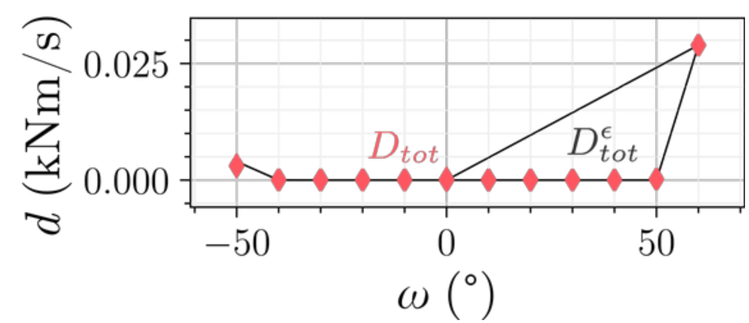


Macro

Boundary Value Problem (BVP)



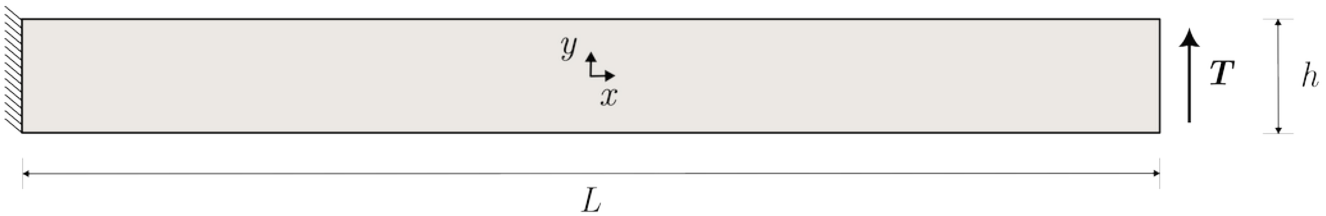


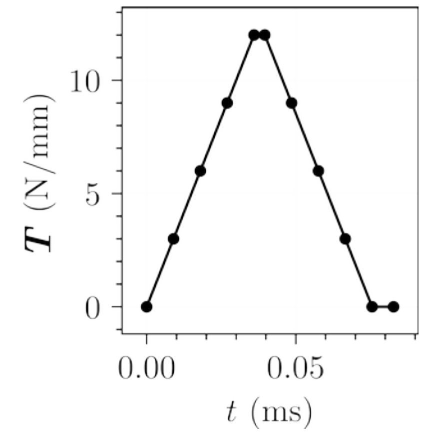


FEM×TANN – Transient phenomena

Masi, Stefanou. *J Mech Phys Solids*, 2023

Panel subjected to a shearing load

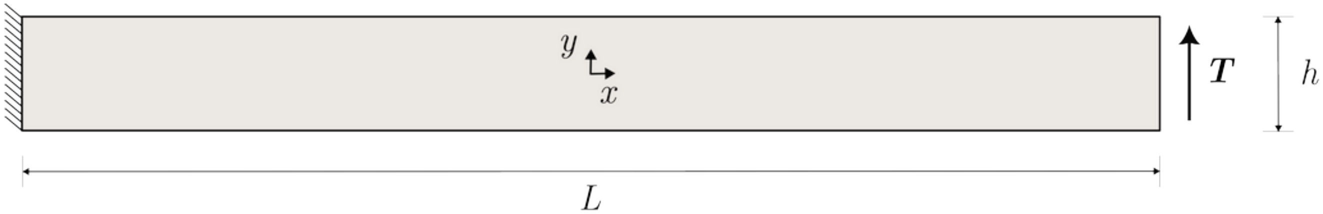


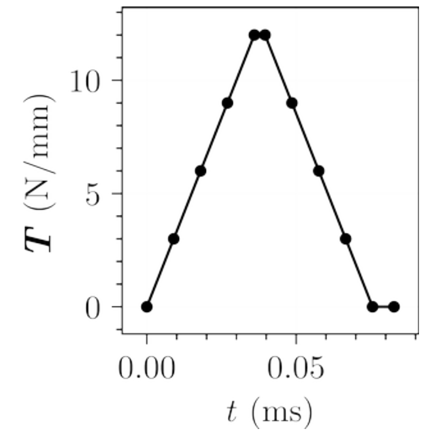


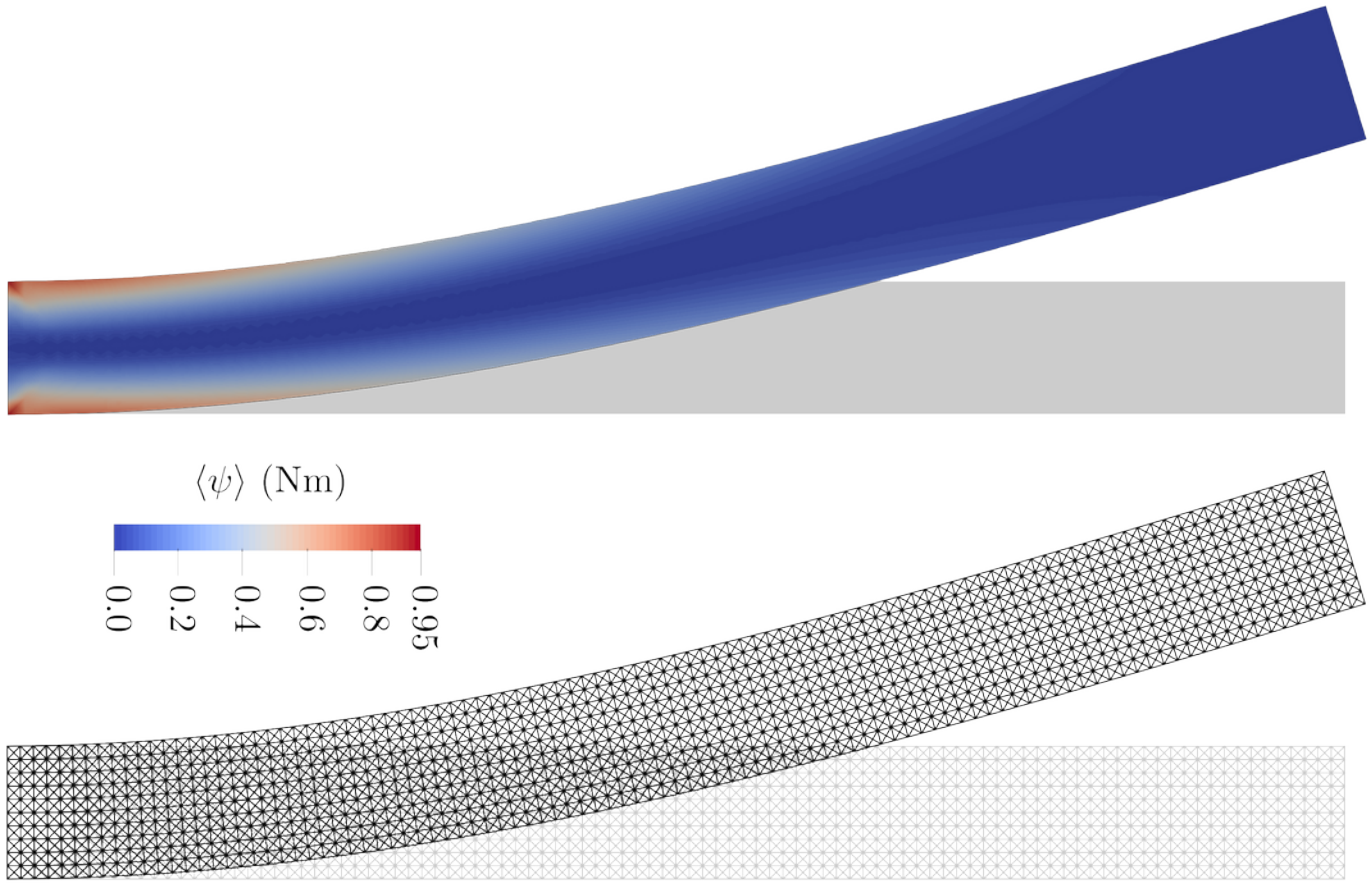
FEM×TANN – Transient phenomena

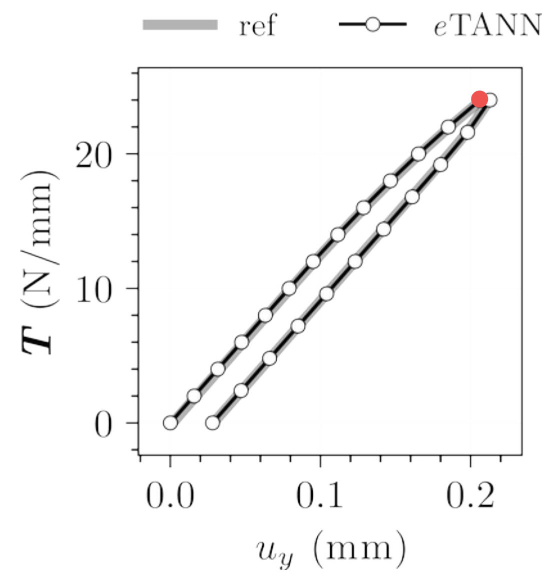
Masi, Stefanou. *J Mech Phys Solids*, 2023

Panel subjected to a shearing load





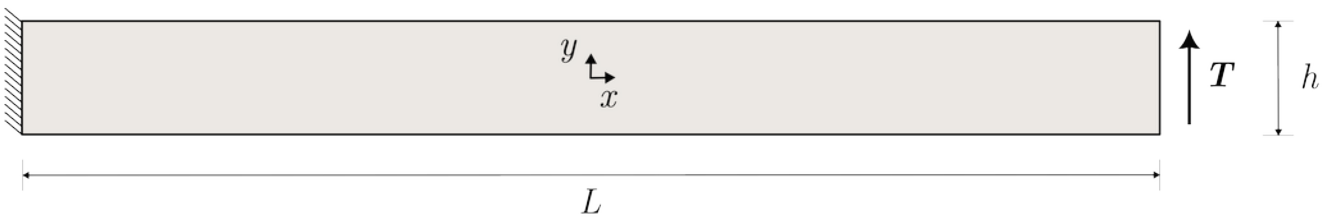


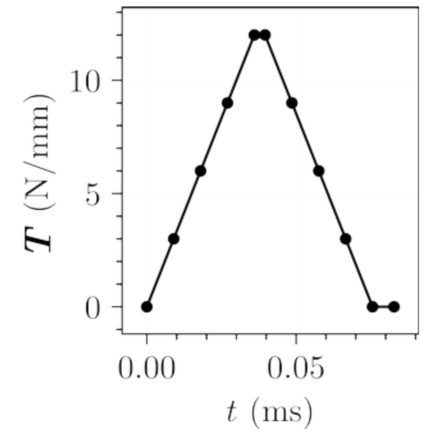


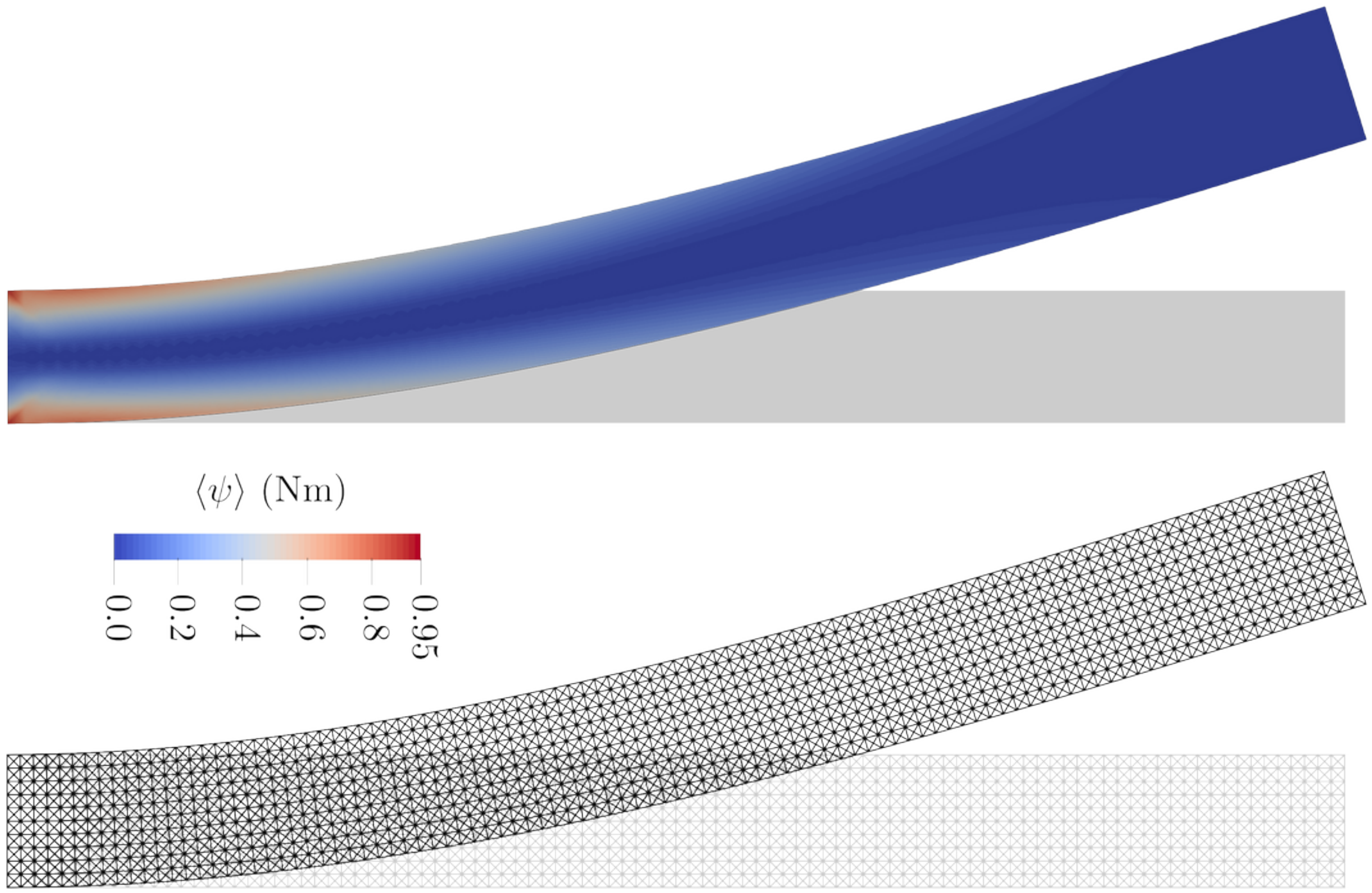
FEM×TANN – Transient phenomena

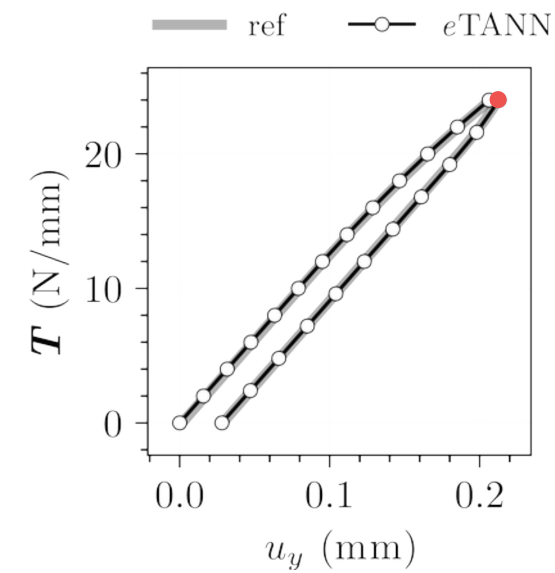
Masi, Stefanou. *J Mech Phys Solids*, 2023

Panel subjected to a shearing load





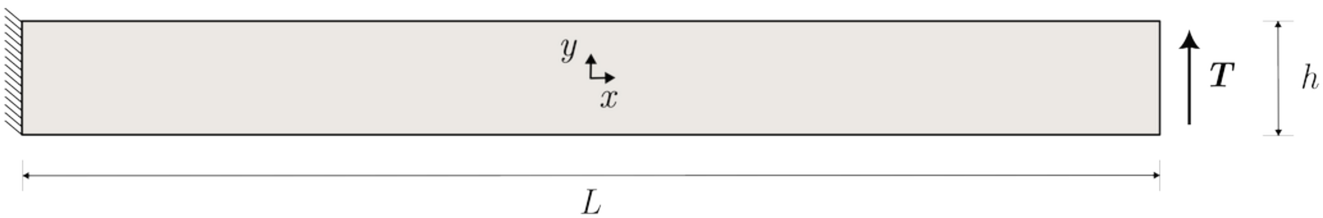


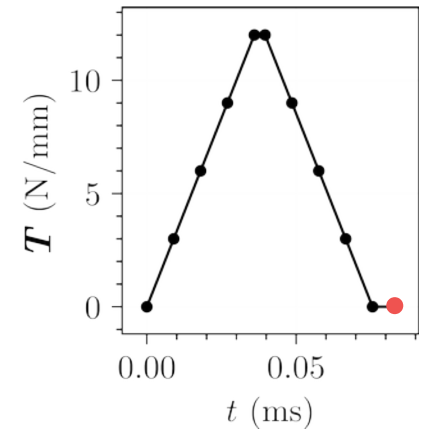


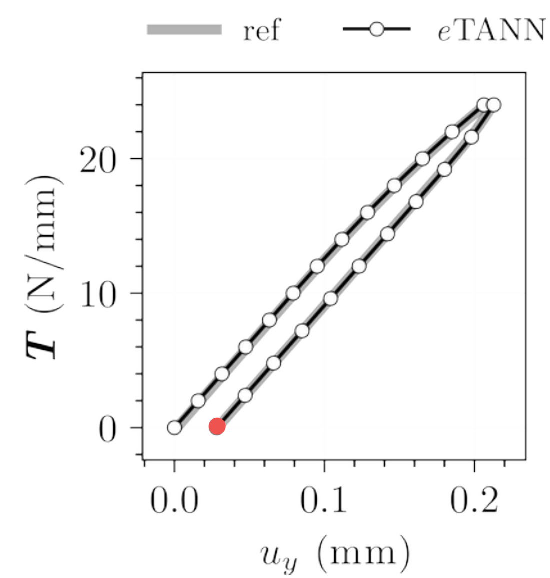
FEM×TANN – Transient phenomena

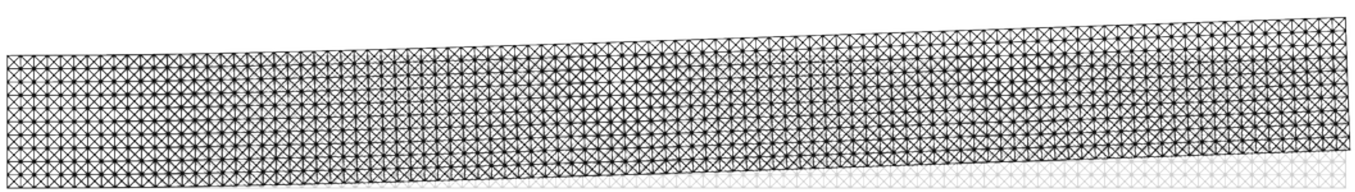
Masi, Stefanou. *J Mech Phys Solids*, 2023

Panel subjected to a shearing load

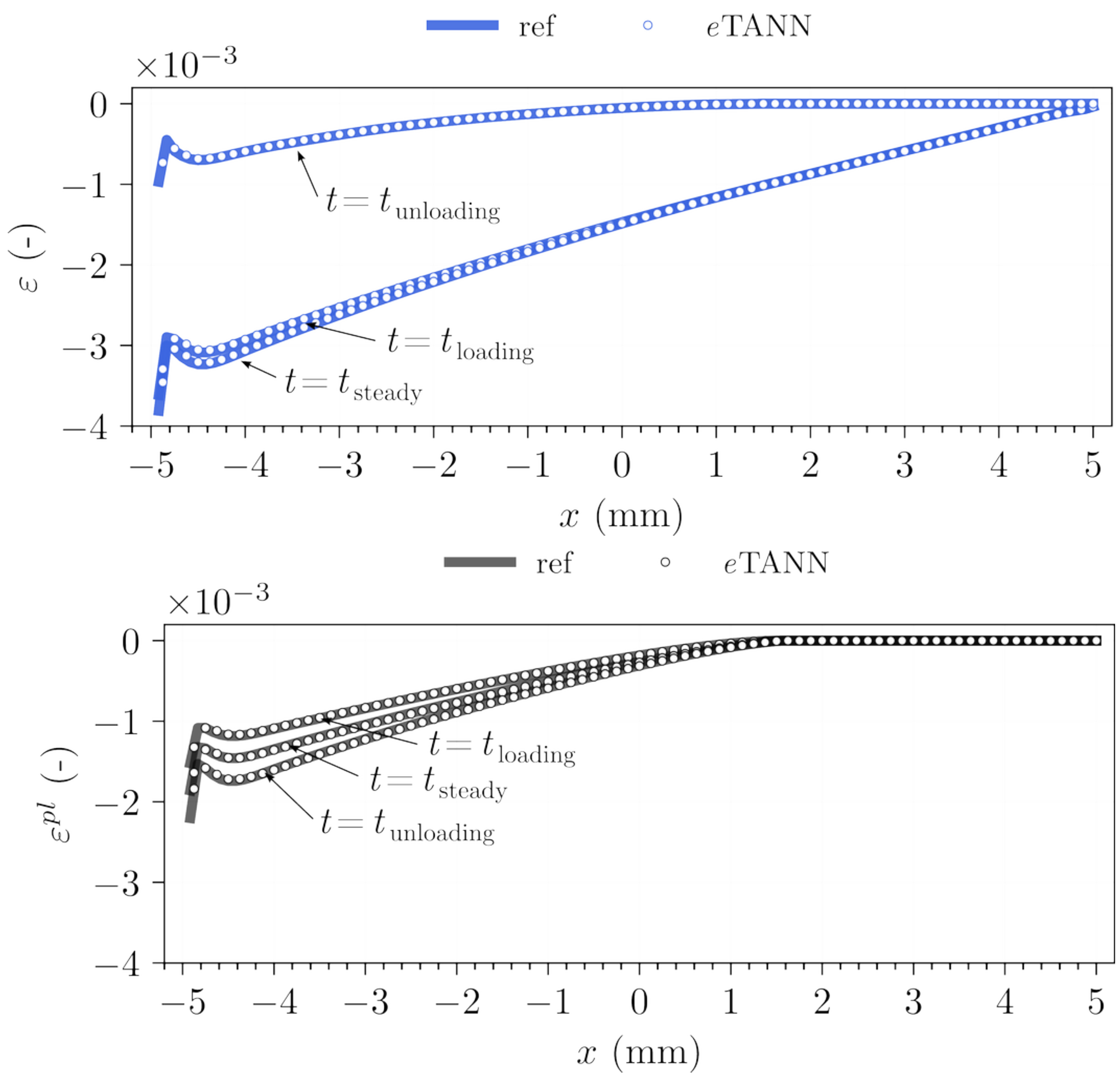








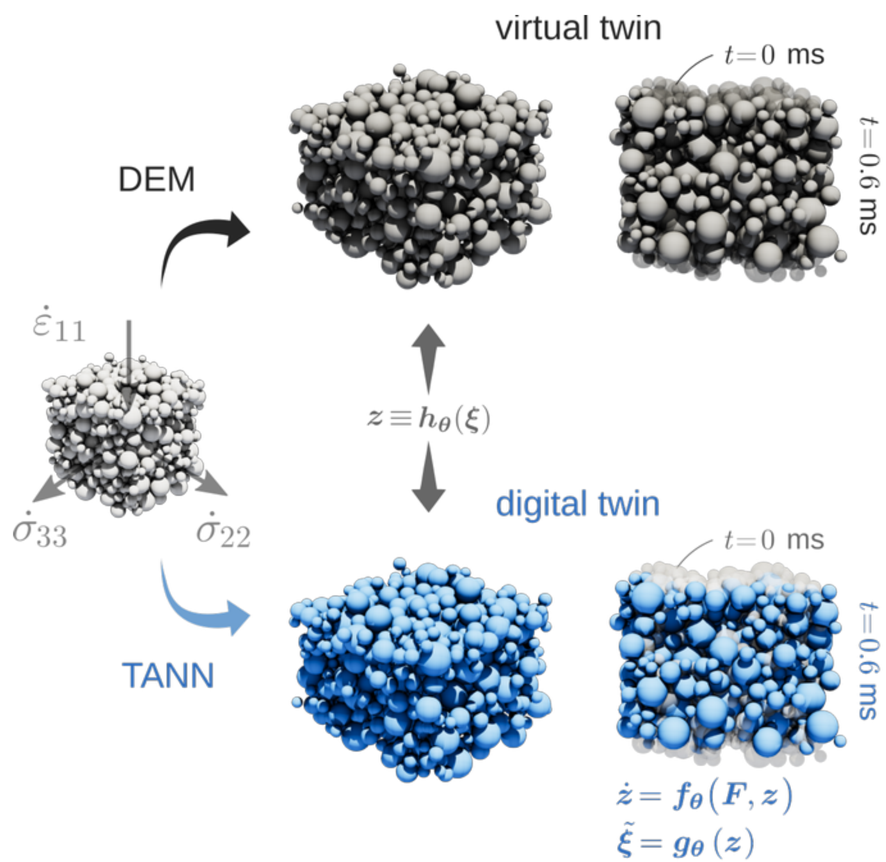
FEM×TANN – Transient phenomena



Conclusions

TANN and PINN are reliable, accurate and scalable DL methods

- TANN predict path-dependency and rate-effects in complex materials, by enforcing the universal laws of thermodynamics
- FEM$\times$TANN: speed-up state-of-the-art multiscale simulations
- Automatic discovery of internal variables and evolution equations



References

- F Masi, I Stefanou (2023). Evolution TANN and the identification of the internal variables and evolution equations in solid mechanics. *J Mech Phys Solids* 174, 105245. doi: 10.48550/arXiv.2209.13269.
- F Masi, I Stefanou (2022). Multiscale modeling of inelastic materials with Thermodynamics-based Artificial Neural Networks (TANN). *Comput Methods Appl Mech Eng* 398, 115190. doi: 10.1016/j.cma.2022.115190.
- F Masi, I Stefanou, P Vannucci, V Maffi-Berthier (2021). Thermodynamics-based Artificial Neural Networks for constitutive modeling. *J Mech Phys Solids* 147, 104277. doi: 10.1016/j.jmps.2020.104277.

Codes

- TANN – *Thermodynamics-based Artificial Neural Networks*

  repository (TensorFlow, PyTorch)
  url: filippo-masi/Thermodynamics-Neural-Networks
- TANN-multiscale – *Multiscale modeling of inelastic materials with TANN*

  repository (TensorFlow, FeniCS, Numerical Geolab)
  url: filippo-masi/TANN-multiscale

Questions?

European Research Council
Established by the European Commission

CoQuake.com

The authors acknowledge the support of the European Research Council (ERC) under the European Union Horizon 2020 research and innovation program (Grant agreement ID 757848 CoQuake).