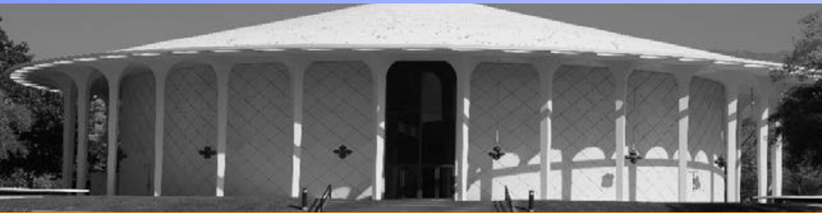
CALIFORNIA INSTITUTE OF TECHNOLOGY



universitätbonn

(Model-Free) Data-Driven Computing

Michael Ortiz
California Institute of Technology and
Rheinische Friedrich-Wilhelms Universität Bonn

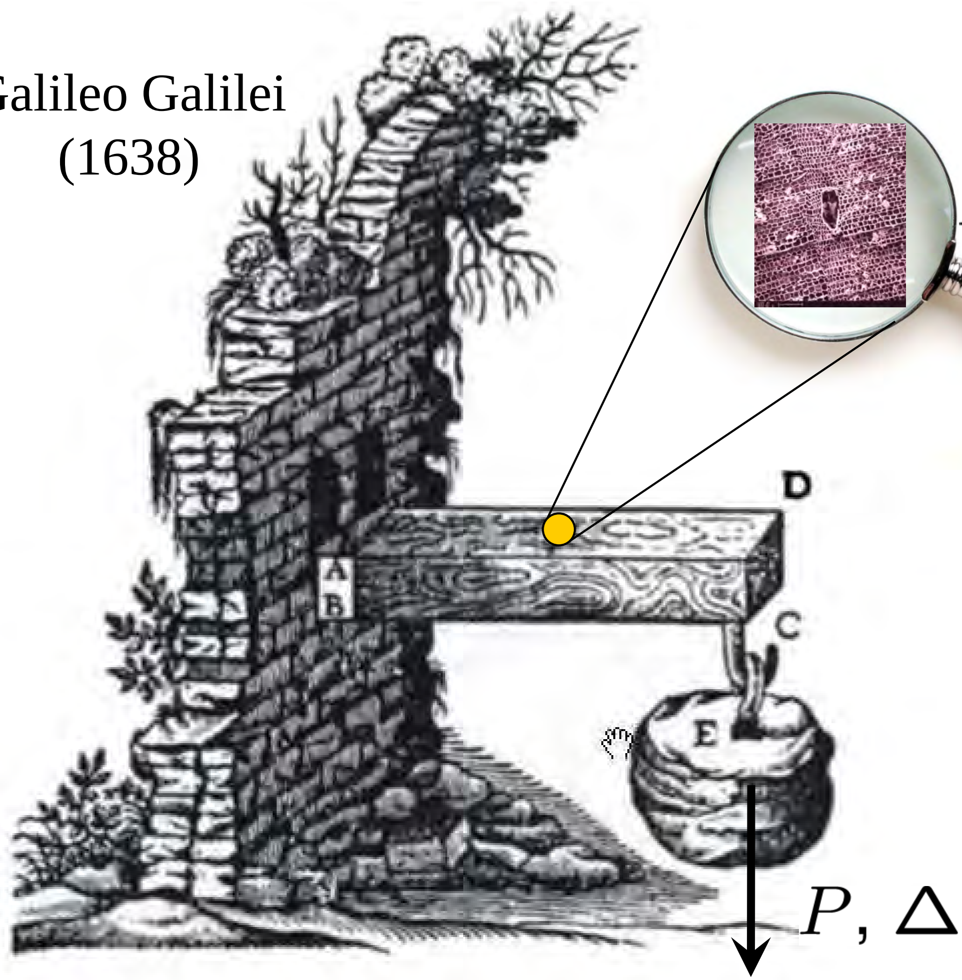
Collaborators: T. Kirchdoerfer (Caltech); L. Stainier (Central Nantes); S. Conti, S. Müller (Bonn); R. Eggersmann, S. Reese (RWTH Aachen)

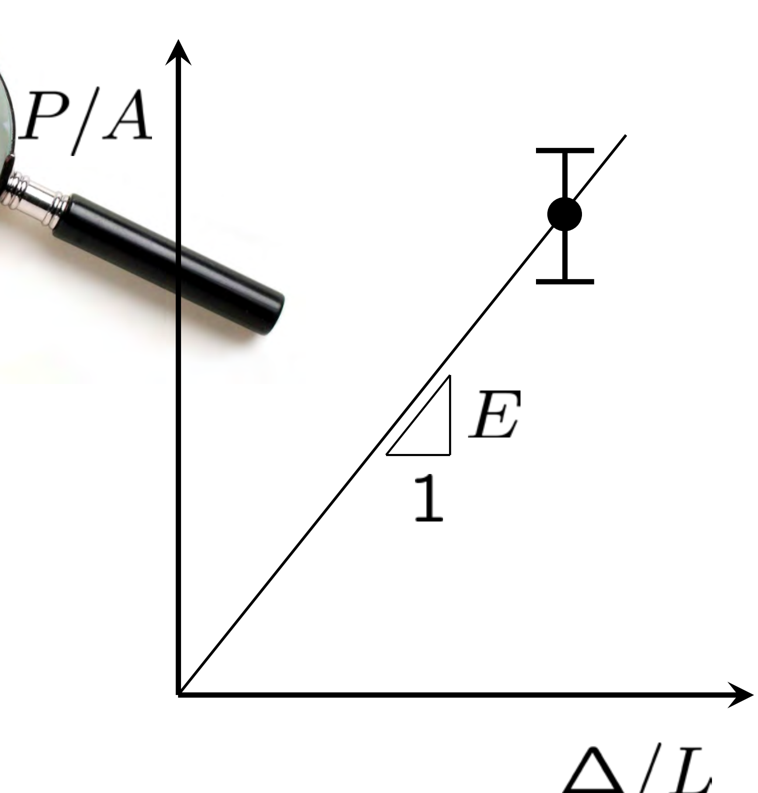
13th World Congress on Computational Mechanics
New York City, July 25, 2018

Materials data through the ages…

Traditionally, mechanics of materials has been *data starved*…

Galileo Galilei
(1638)

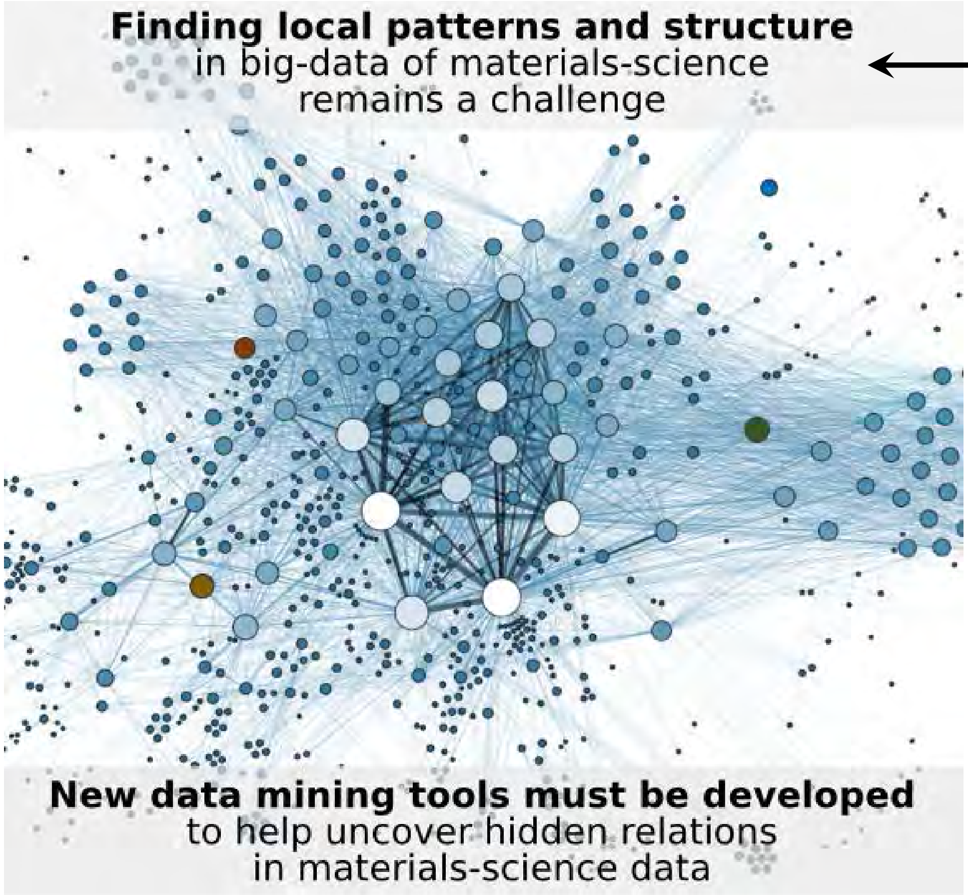




Michael Ortiz
WCCM 2018

Material data through the ages…

At present, mechanics of materials is *data rich!*



(E. Munch, 1893)

NOMAD
https://www.nomad-coe.eu/

Data Science, Big Data…

http://olap.com/forget-big-data-lets-talk-about-all-data/

Michael Ortiz
WCCM 2018

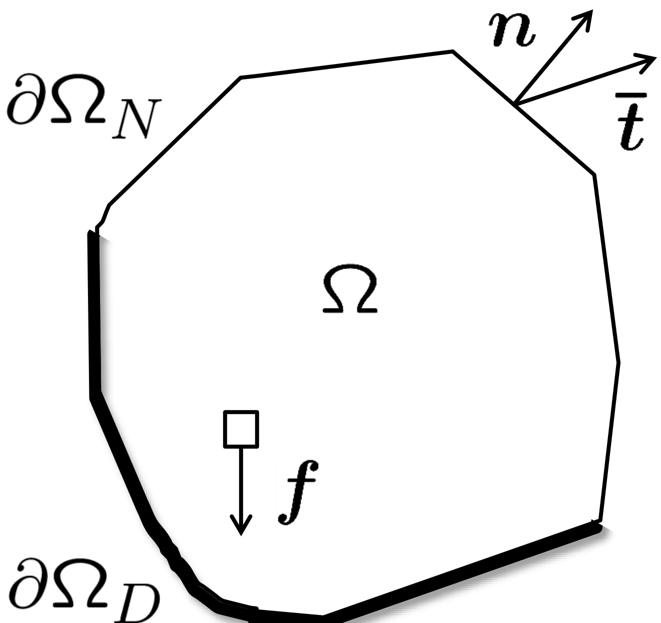
Data Science, Big Data…

- *Data Science* is the extraction of '*knowledge*' from large volumes of unstructured data[1]
- Data science requires sorting through *big-data* sets and extracting '*insights*' from these data
- Data science uses data management, statistics and machine learning to derive *mathematical models* for subsequent use in decision making
- Data science influences (*non-STEM*) fields such as marketing, advertising, finance, social sciences, security, policy, medical informatics…
- *But… What's in it for us? (STEM folk)*

[1]Dhar, V., *Communications of the ACM*, **56**(12) (2013) p. 64.

Where does Data Science intersect with (computational) mechanics?

- Anatomy of a field-theoretical *STEM* problem:



i) Kinematics + Dirichlet:

$$\epsilon(\boldsymbol{u}) = 1/2(\nabla \boldsymbol{u} + \nabla \boldsymbol{u}^T)$$
$$\boldsymbol{u} = \bar{\boldsymbol{u}}, \quad \text{on } \partial\Omega_D$$

ii) Equilibrium + Neumann:

$$\text{div}\,\boldsymbol{\sigma} + \boldsymbol{f} = \boldsymbol{0}$$
$$\boldsymbol{\sigma}\boldsymbol{n} = \bar{\boldsymbol{t}}, \quad \text{on } \partial\Omega_N$$

iii) Material law: $\boldsymbol{\sigma}(\boldsymbol{x}) = \boldsymbol{\sigma}\Big(\epsilon(\boldsymbol{x})\Big)$

Where does Data Science intersect with (computational) mechanics?

- Anatomy of a field-theoretical *STEM* problem:

Universal laws!
(Newton's laws,
Schrodinger's eq.,
Maxwell's eqs.,
Einstein's eqs…)
Exactly known!
Uncertainty-free!
(epistemic)

i) Kinematics + Dirichlet:

$$\epsilon(\boldsymbol{u}) = 1/2(\nabla\boldsymbol{u} + \nabla\boldsymbol{u}^T)$$
$$\boldsymbol{u} = \bar{\boldsymbol{u}}, \quad \text{on } \partial\Omega_D$$

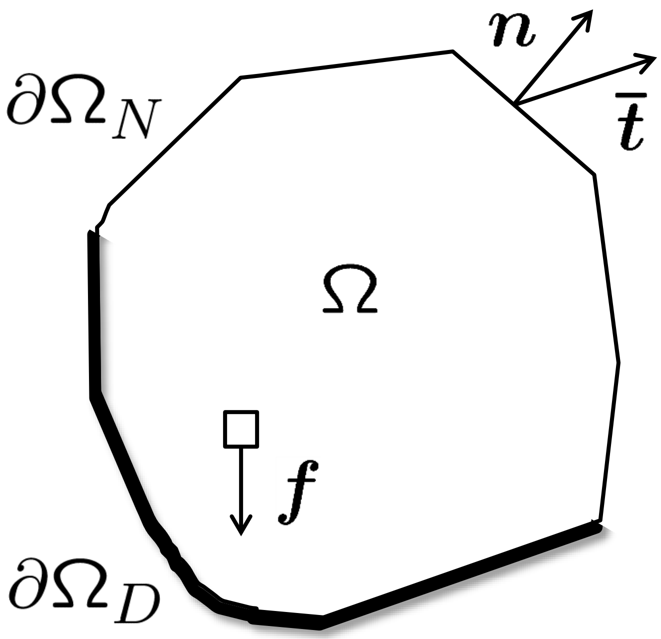
ii) Equilibrium + Neumann:

$$\text{div}\,\boldsymbol{\sigma} + \boldsymbol{f} = \boldsymbol{0}$$
$$\boldsymbol{\sigma}\boldsymbol{n} = \bar{\boldsymbol{t}}, \quad \text{on } \partial\Omega_N$$

iii) Material law: $\boldsymbol{\sigma}(\boldsymbol{x}) = \boldsymbol{\sigma}\Big(\epsilon(\boldsymbol{x})\Big)$

Where does Data Science intersect with (computational) mechanics?

- Anatomy of a field-theoretical *STEM* problem:



i) Kinematics + Dirichlet:

$$\begin{rcases} \epsilon(\boldsymbol{u}) = 1/2(\nabla \boldsymbol{u} + \nabla \boldsymbol{u}^T) \\ \boldsymbol{u} = \bar{\boldsymbol{u}}, \quad \text{on } \partial\Omega_D \end{rcases}$$

ii) Equilibrium + Neumann:

$$\begin{rcases} \text{div}\, \boldsymbol{\sigma} + \boldsymbol{f} = \boldsymbol{0} \\ \boldsymbol{\sigma n} = \bar{\boldsymbol{t}}, \quad \text{on } \partial\Omega_N \end{rcases}$$

iii) Material law: $\underbrace{\boldsymbol{\sigma}(\boldsymbol{x}) = \boldsymbol{\sigma}\Big(\epsilon(\boldsymbol{x})\Big)}$

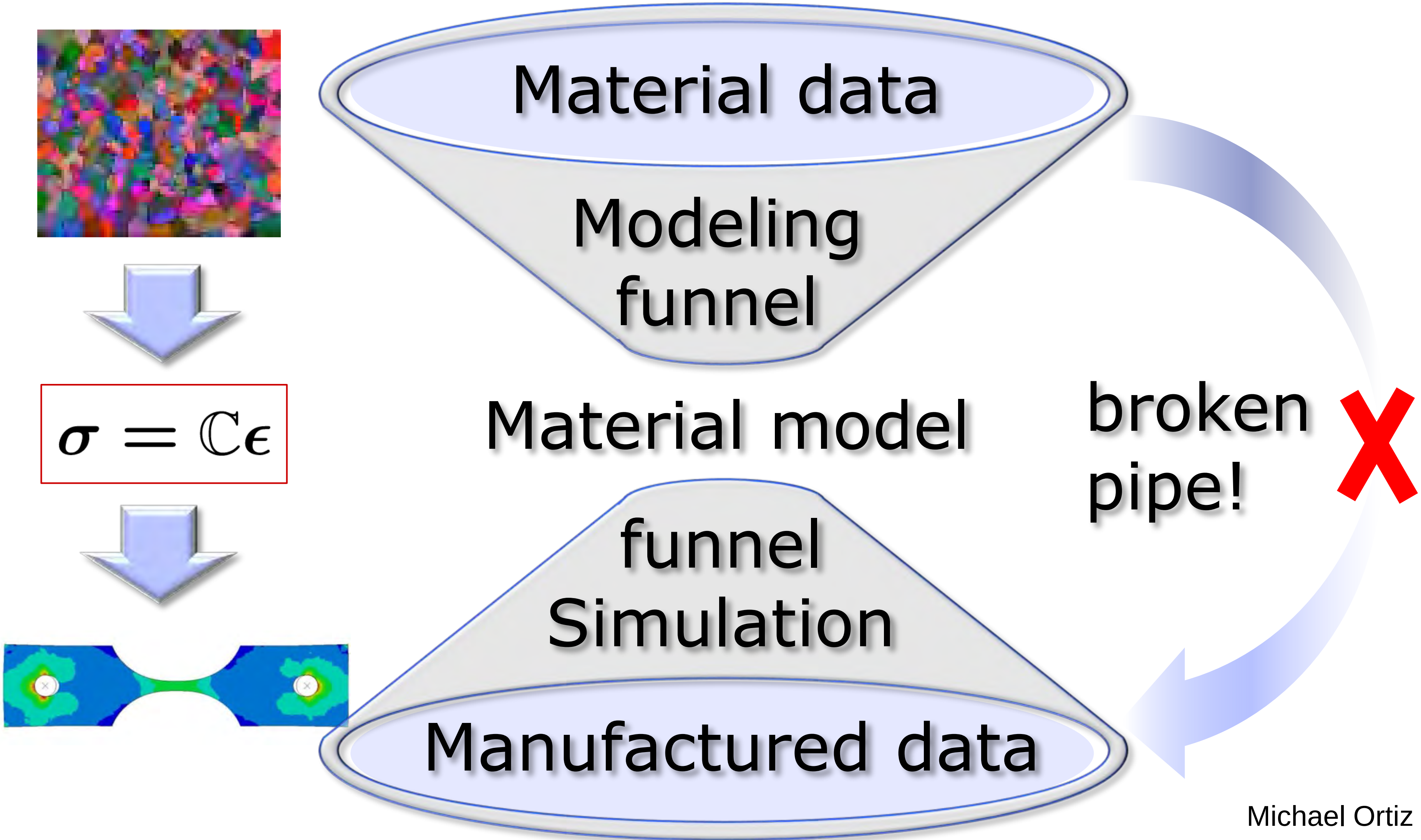
Unknown! Epistemic uncertainty!

Michael Ortiz
WCCM 2018

Classical Modeling & Simulation

- Need to generate (epistemic) ‘*knowledge*’ about material behavior to close BV problems…
- Traditional *modeling paradigm*: Fit data (a.k.a. regression, machine learning, model reduction, central manifolds…), use calibrated empirical models in BV problems

Classical Modeling & Simulation

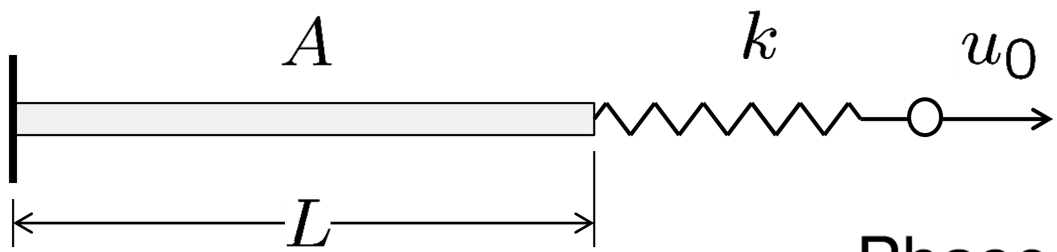


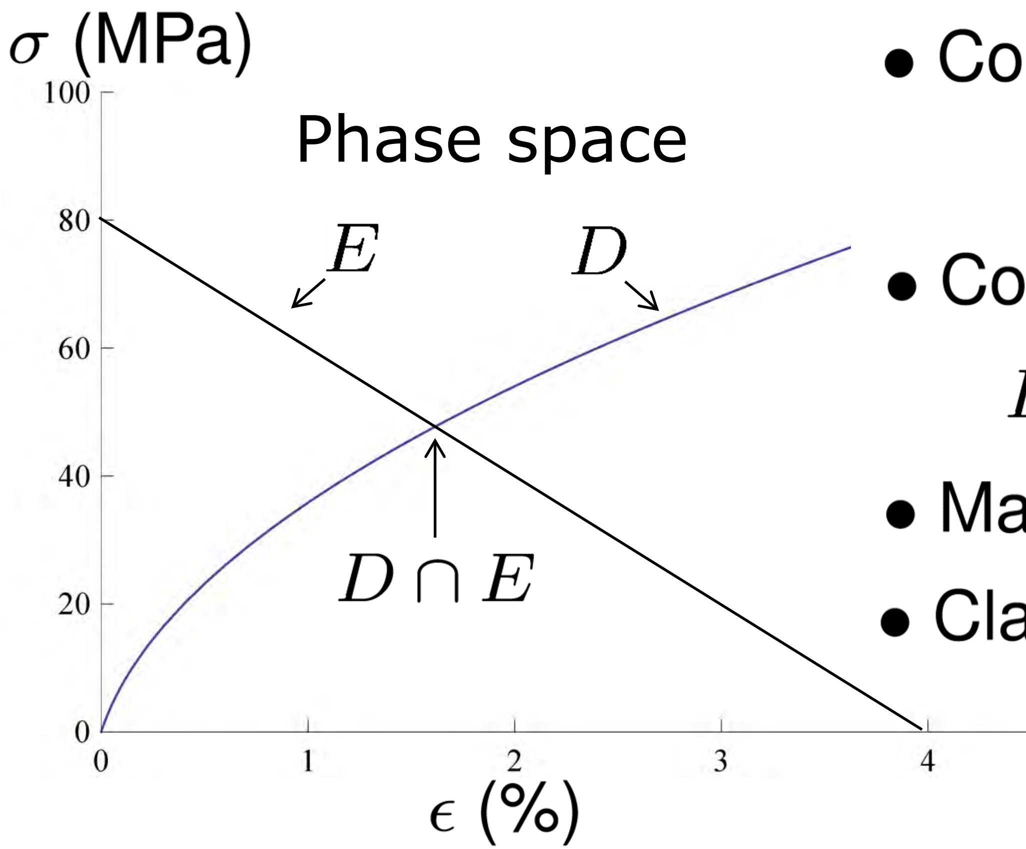
Michael Ortiz
WCCM 2018

Data-Driven (model-free) mechanics

- Need to generate (epistemic) '*knowledge*' about material behavior to close BV problems…
- Traditional *modeling paradigm*: Fit data (a.k.a. regression, machine learning, model reduction, central manifolds…), use calibrated empirical models in BV problems
- *But:* We live in a *data-rich world* (full-field diagnostics, data mining from first principles…)
- *Extreme Data-Driven paradigm* (*model-free!*): Use material data directly in BVP (no fitting by any name, no loss of information, no broken pipe between material and manufactured data)
- *How?*

Elementary example: Bar and spring





- Phase space: $\{(\epsilon, \sigma)\} \equiv Z$
- Compatibility + equilibrium:

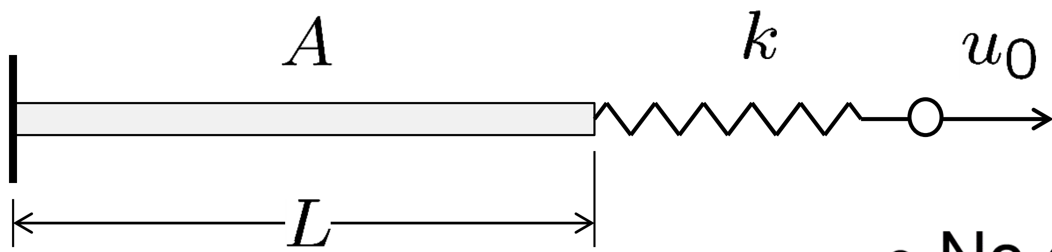
$$\sigma A = k(u_0 - \epsilon L)$$

- Constraint set:

$$E = \{\sigma A = k(u_0 - \epsilon L)\}$$

- Material data set: $D \subset Z$
- Classical solution set: $D \cap E$

Elementary example: Bar and spring

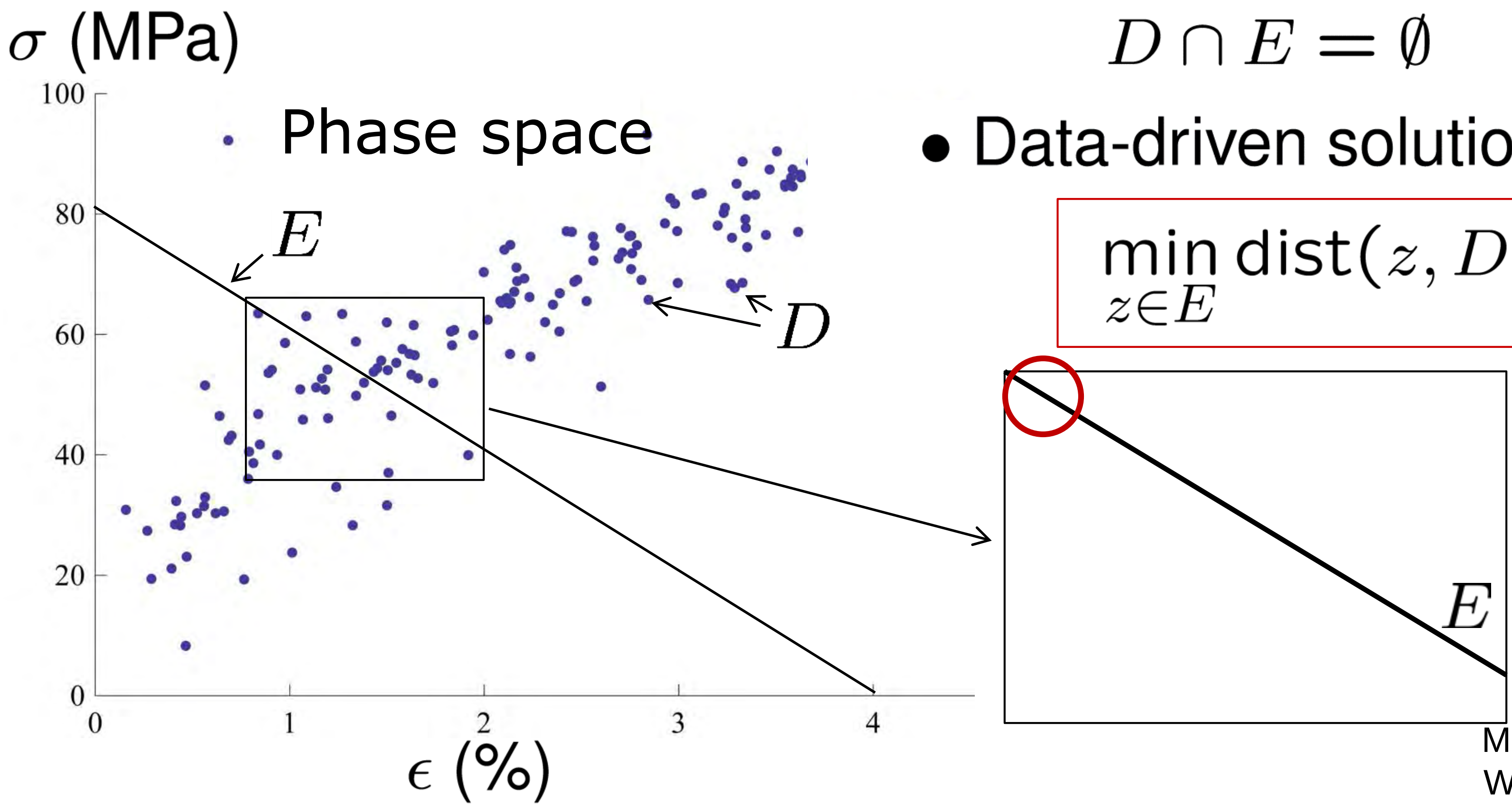


- No classical solutions!

$$D \cap E = \emptyset$$

- Data-driven solution:

$$\min_{z \in E} \operatorname{dist}(z, D)$$



Michael Ortiz
WCCM 2018

The general Data-Driven (DD) problem

- The Data-Driven paradigm[1]: Given,
  - *D = {fundamental material data}*,
  - *E = {compatibility + equilibrium}*,

  Find: argmin$\{d(z, D),\ z \in E\}$
- *The aim of Data-Driven analysis is to find the compatible strain field and the equilibrated stress field closest to the material data set*
- No material modeling, no data fitting, no V&V…
- Raw *fundamental* (stress *vs* strain) material data is used (unprocessed) in calculations
- No assumptions, artifacts, loss of information…

[1]T. Kirchdoerfer and M. Ortiz (2015) arXiv:1510.04232.
[1]T. Kirchdoerfer and M. Ortiz, *CMAME*, **304** (2016) 81–101

Data-Driven Computing: Issues

- *Data-driven (model-free!) computing: Use material data sets directly in calculations!*
- Is the Data-Driven reformulation of classical BVPs (possibly off of noisy data) *well-posed*?
- Implementation of *Data-Driven solvers*?
- *Numerical convergence* (iterative solvers, mesh size, time step…)
- Convergence with respect to *material data set*
- Extension to *time-dependent* problems
- Extension to *history-dependent* materials
- *Phase-space sampling* in high dimension
- *Data management*, repositories, outlook…

Data-Driven elasticity – Well-posedness

Definition (Constraint set)

i) Compatibility,

$$\epsilon = 1/2(\nabla u + \nabla u^T),$$
$$u = g, \quad \text{on } \Gamma_D.$$

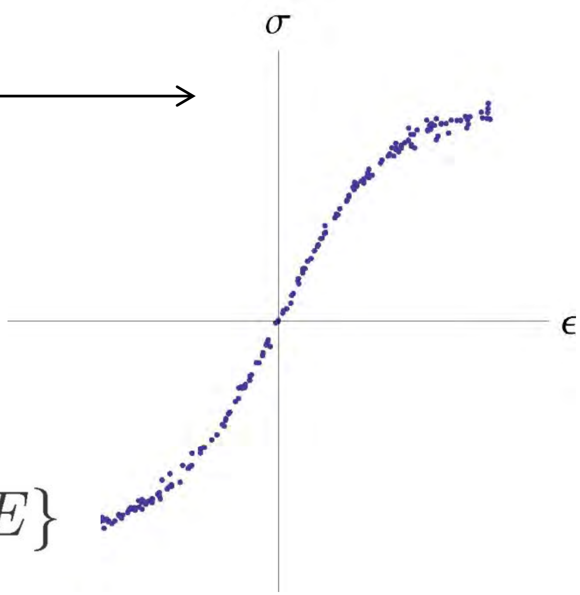
ii) Equilibrium,

$$\text{div}\sigma + f = 0,$$
$$\sigma\nu = h, \quad \text{on } \Gamma_N.$$

$\Big\} \; E$

Definition (Material data set)

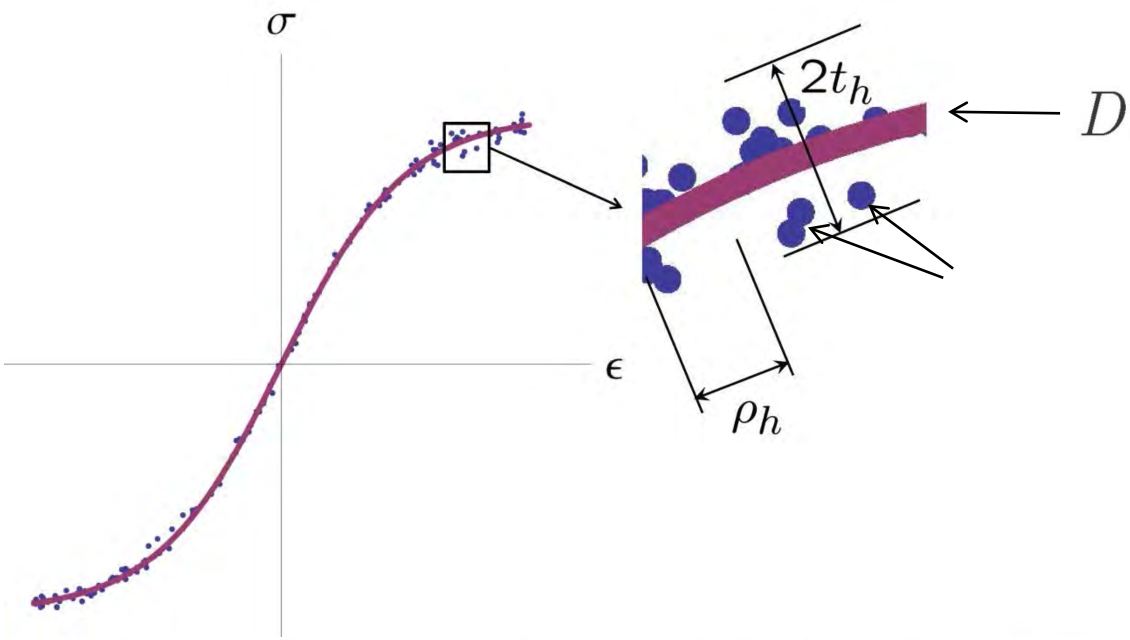
Hooke's law (linear) $D = \{\sigma = \mathbb{C}\epsilon\}$.
Hooke's law (monotone) $D = \{\sigma = \sigma(\epsilon)\}$.



$$\min\{d(z, D),\ z \in E\}$$

Conti, S., Müller, S. & Ortiz, M., *ARMA*, **229** (2018) 79-123.

Data-Driven elasticity – Δ-convergence



Theorem

Suppose D *monotone graph,* $\rho_h \downarrow 0$ *and* $t_h \downarrow 0$ *such that:*

i) *Fine approximation:* $d(\xi, D_h) \leq \rho_h$, $\forall \xi \in D$.

ii) *Uniform approximation:* $d(\xi, D) \leq t_h$, $\forall \xi \in D_h$.

Then, $(\epsilon_h, \sigma_h) \to (\epsilon, \sigma)$.

Conti, S., Müller, S. & Ortiz, M., *ARMA*, **229** (2018) 79-123.

Data-Driven elasticity

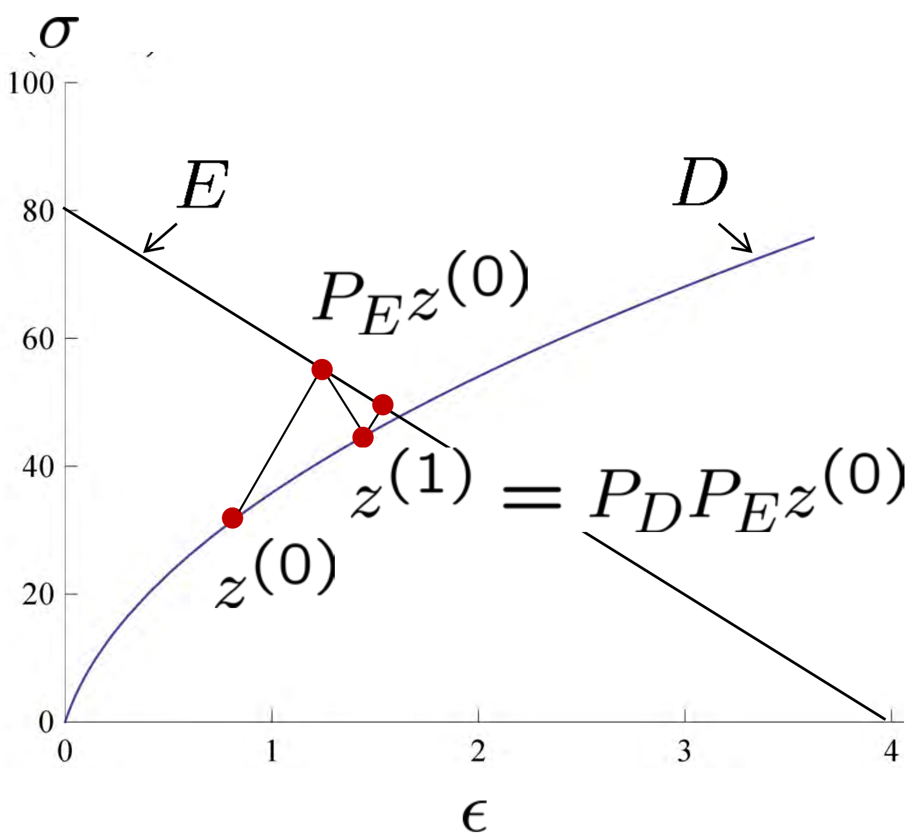
- Data-driven problems represent a *reformulation of the classical problems of mechanics*, in which compatibility and equilibrium are enforced as *differential constraints*, and the aim is to *minimize discrepancy with a material data set*
- Material data sets can be graphs (classical), point sets, fat sets… *Data-Driven problems subsume and extend the classical problems*
- Data-driven problems are well-posed (existence) and *solutions depend continuously on data sets* (in the sense of Δ-convergence)
- *Variational character* provides strong basis for *approximation* (numerical, data reduction…)

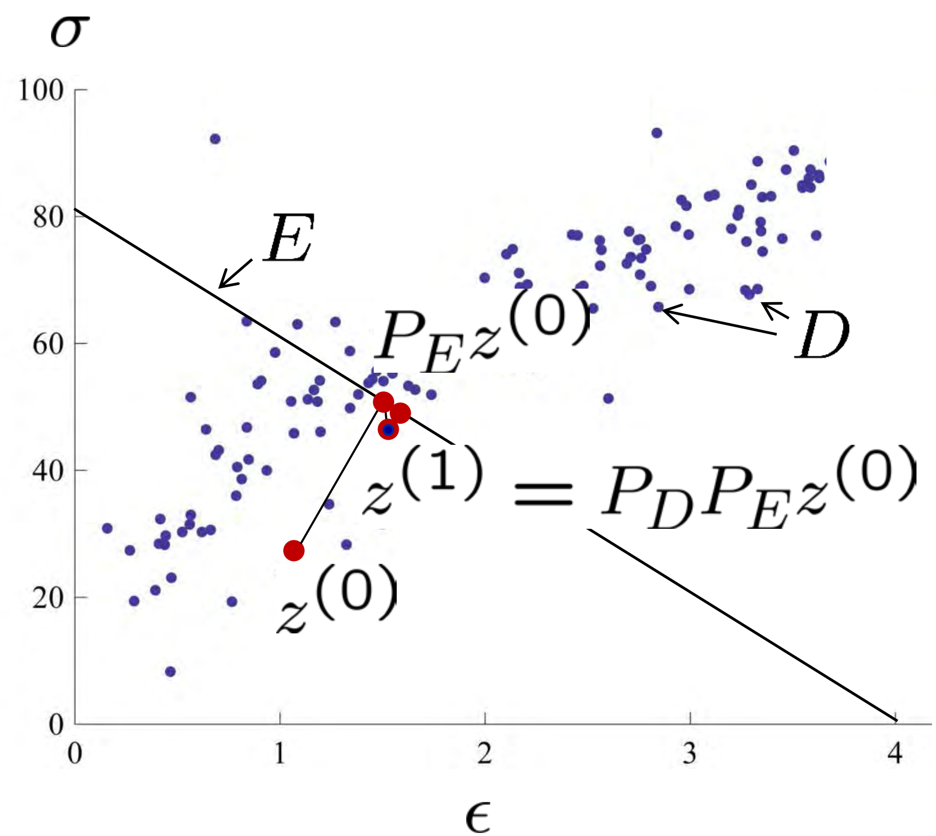
Data-Driven Computing: Issues

- *Data-driven (model-free!) computing: Use material data sets directly in calculations!*
- Is the Data-Driven reformulation of classical BVPs (possibly off of noisy data) *well-posed*?
- Implementation of *Data-Driven solvers*?
- *Numerical convergence* (iterative solvers, mesh size, time step…)
- Convergence with respect to *material data set?*
- Extension to *time-dependent* problems
- Extension to *history-dependent* materials
- *Phase-space sampling* in high dimension
- *Data management*, repositories, outlook…

DD solvers: Fixed-point iteration

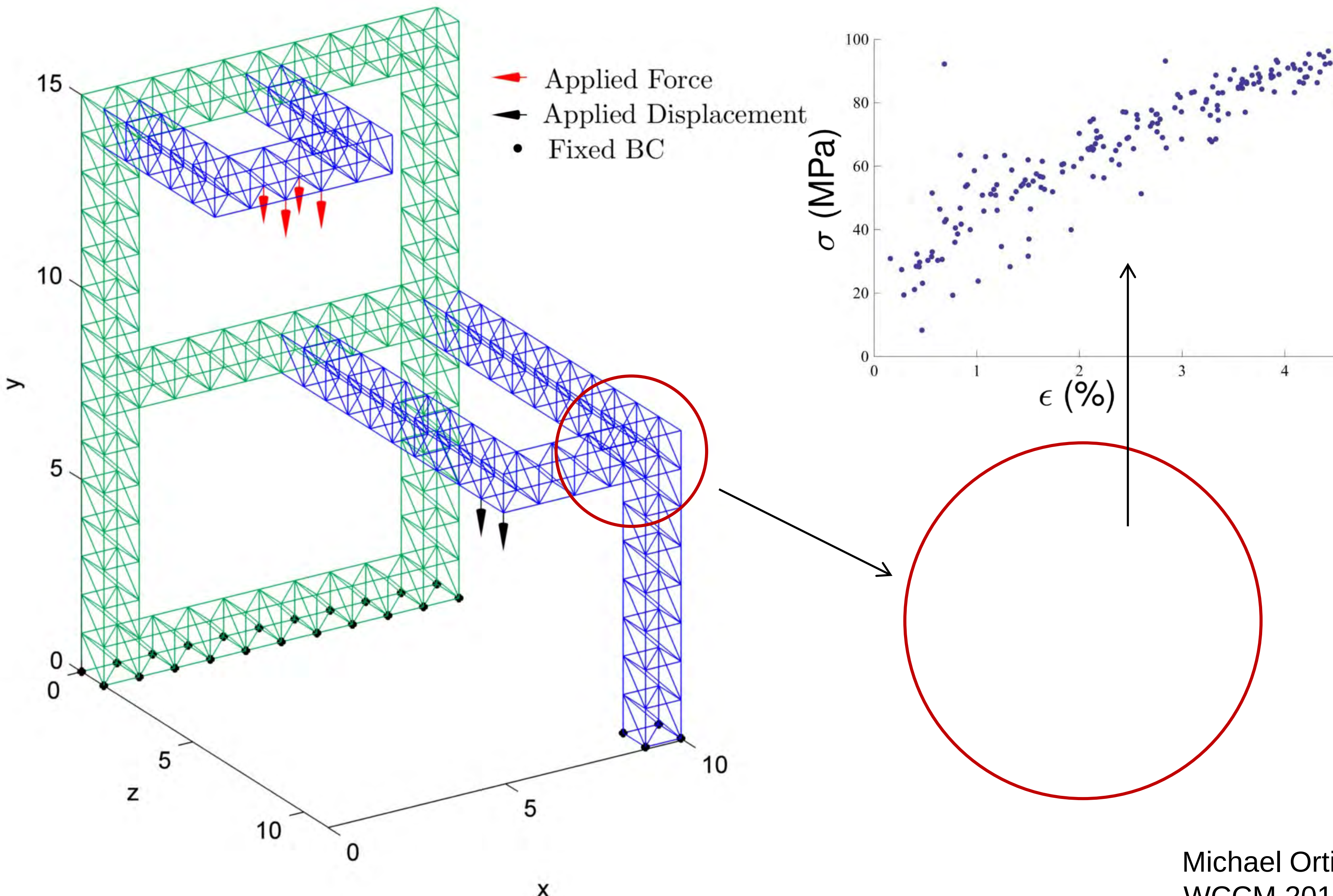
- Find: $\text{argmin}\{d(z, D),\ z \in E\}$
- Fixed-point iteration[1]: $z^{(k+1)} = P_D P_E z^{(k)}$





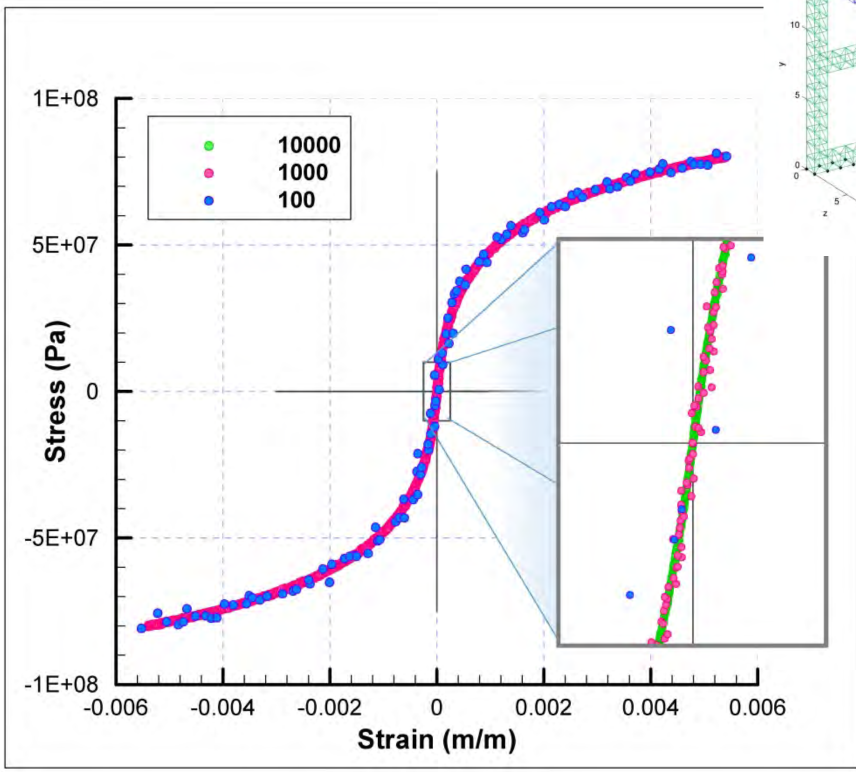
[1]T. Kirchdoerfer and M. Ortiz (2015) arXiv:1510.04232.
[1]T. Kirchdoerfer and M. Ortiz, *CMAME*, **304** (2016) 81–101

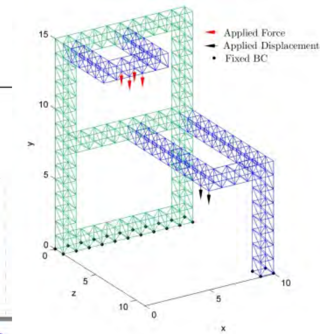
Test case: 3D Truss

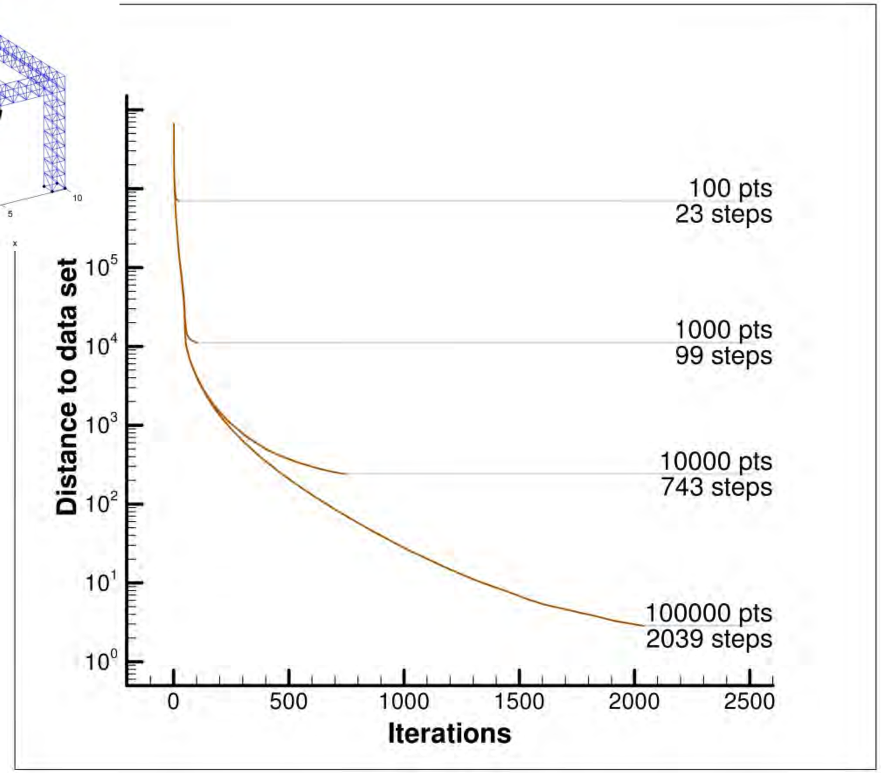


Michael Ortiz
WCCM 2018

Truss test: Convergence of solver





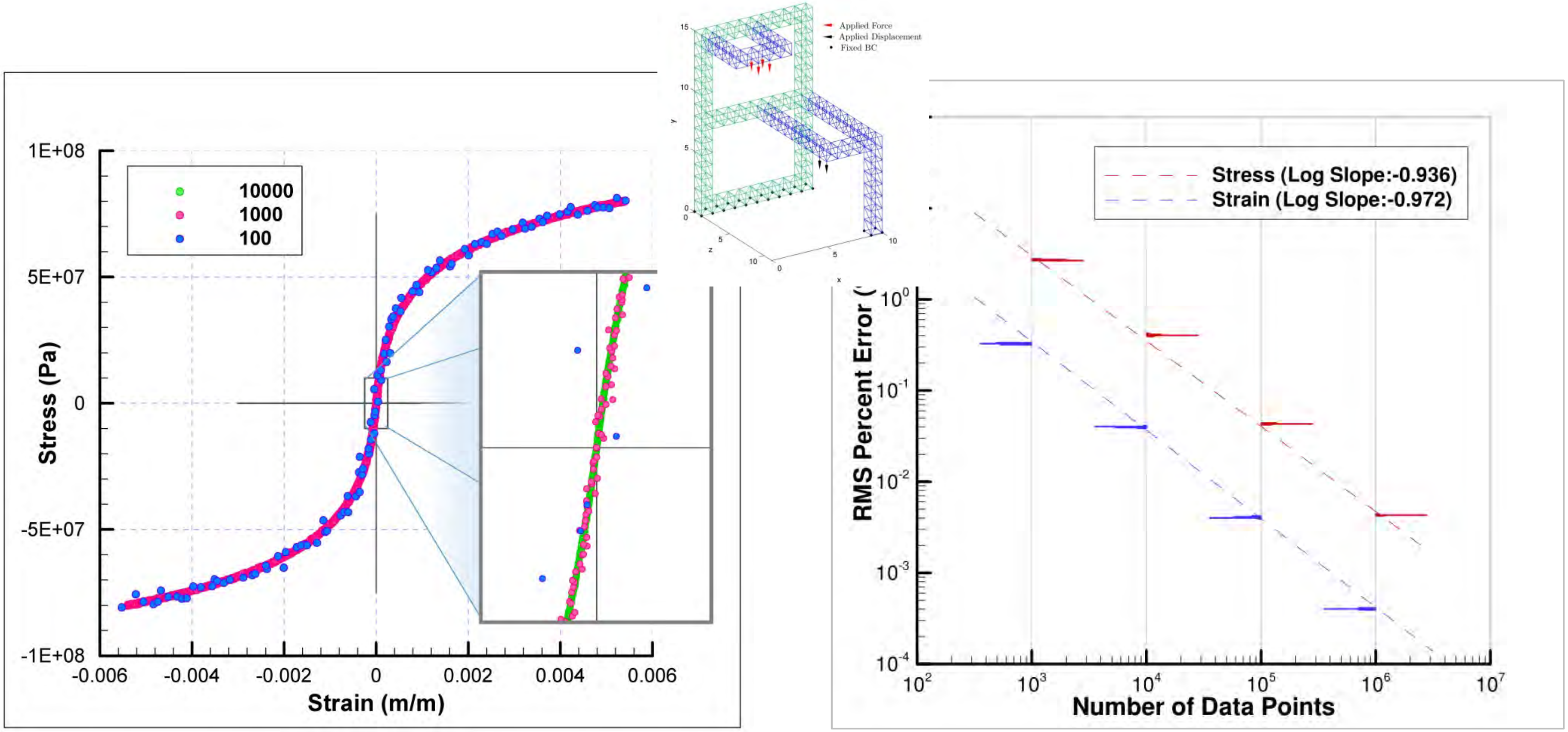


Material-data sets
of increasing size
and decreasing scatter

Convergence,
local data assignment
iteration

T. Kirchdoerfer and M. Ortiz, *CMAME*, **304** (2016) 81–101.

Truss test: Convergence wrt data



Material-data sets
of increasing size
and decreasing scatter

Convergence
with respect to sample size
(with initial Gaussian noise)

Michael Ortiz
WCCM 2018

T. Kirchdoerfer and M. Ortiz, *CMAME*, **304** (2016) 81–101.

Distance-based DD solvers

- Distance-based DD solvers exhibit *good convergence* wrt to material-data association
- Distance-based DD solvers exhibit *good convergence* wrt uniformly converging data
- Data search structures are a form of *machine learning* (supervised classification)
- But distance-based DD solvers can be overly sensitive to *outliers* in the data (non-uniform data convergence)
- If outliers cannot be ruled out, distance-based *DD solvers can be generalized and extended* (cluster analysis, k-means, max-ent inference)

[1]T. Kirchdoerfer and M. Ortiz, *CMAME*, **326** (2017) 622-641.

Data-Driven Computing: Issues

- *Data-driven (model-free!) computing: Use material data sets directly in calculations!*
- Is the Data-Driven reformulation of classical BVPs (possibly off of noisy data) *well-posed*?
- Implementation of *Data-Driven solvers*?
- *Numerical convergence* (iterative solvers, mesh size, time step…)
- Convergence with respect to *material data set?*
- Extension to *time-dependent* problems
- Extension to *history-dependent* materials
- *Phase-space sampling* in high dimension
- *Data management*, repositories, outlook…

Time-dependent problems: Dynamics

- Time discretization: $t_0, \ldots, t_{k+1} = t_k + \Delta t, \ldots$
- Constraint set (time dependent): $E_{k+1} =$

$$\Big\{\underbrace{\epsilon_{e,k+1} = B_e u_{k+1},}_{\text{compatibility}} \quad \underbrace{\sum_{e=1}^{m} w_e B_e^T \sigma_{e,k+1} = f_{k+1}^{\text{ext}} - M a_{k+1}}_{\text{dynamic equilibrium}}\Big\}$$

- Newmark algorithm (3-point multistep scheme):

$$u_{k+1} = u_k + \Delta t\, v_k + \Delta t^2 \Big((1/2 - \beta) a_k + \beta a_{k+1}\Big)$$
$$v_{k+1} = v_k + \Delta t \Big((1 - \gamma) a_k + \gamma a_{k+1}\Big)$$

T. Kirchdoerfer and M. Ortiz, *IJNME*, **113**(11) (2018) 1697-1710.

Time-dependent problems: Dynamics

- Time discretization: $t_0, \ldots, t_{k+1} = t_k + \Delta t, \ldots$
- Constraint set (time dependent): $E_{k+1} =$

$$\left\{ \epsilon_{e,k+1} = B_e u_{k+1}, \quad \sum_{e=1}^{m} w_e B_e^T \sigma_{e,k+1} = f_{k+1}^{\text{ext}} - M a_{k+1} \right\}$$

- Constraint set representation (3-point scheme):

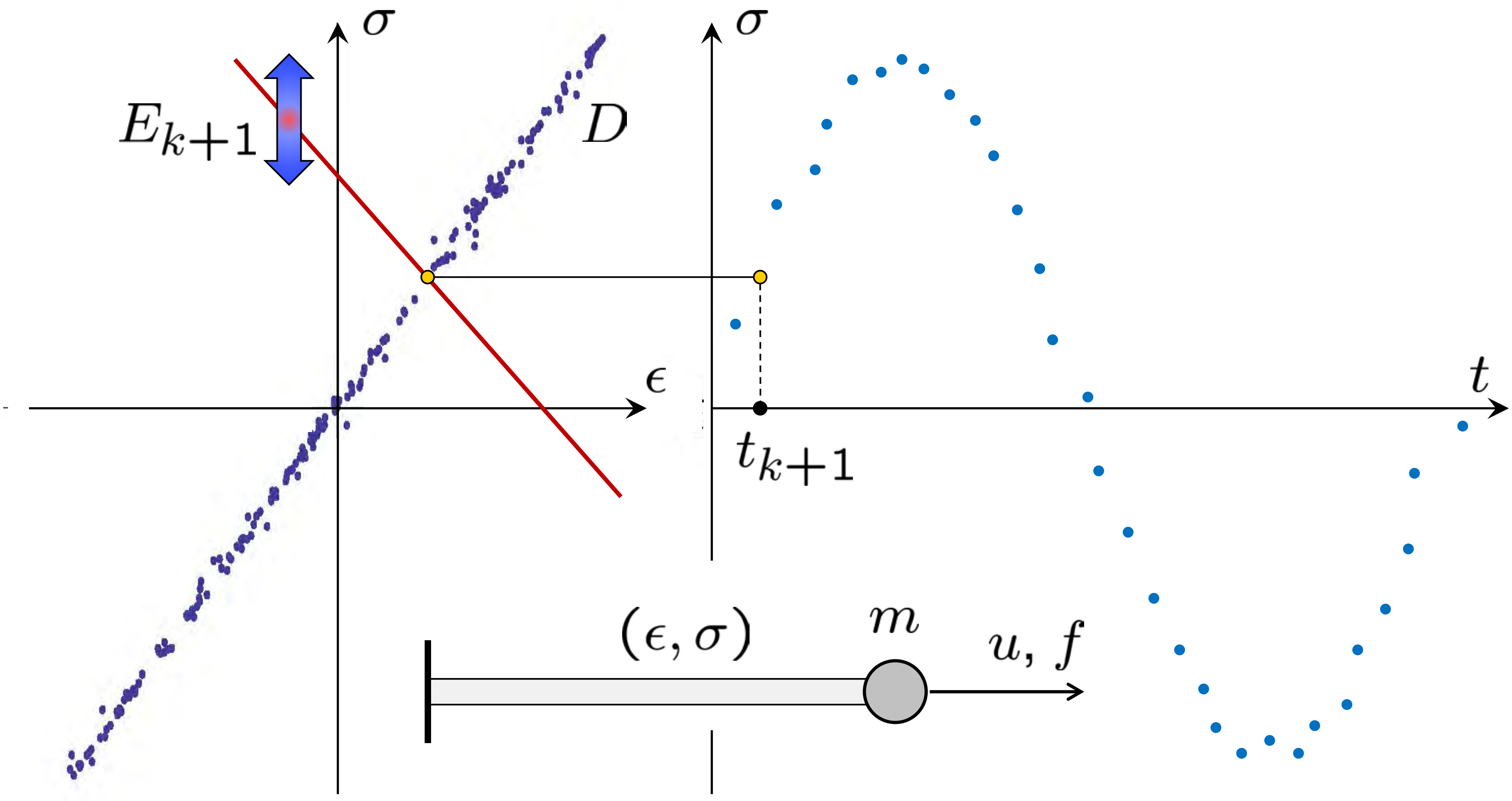
$$E_{k+1} = \{(\epsilon_{k+1}, \sigma_{k+1}) \, : \, \underbrace{(u_k, f_k), \; (u_{k-1}, f_{k-1})}_{\text{causality}}\}$$

- Data-driven problem:

$$\min_{(\epsilon^*,\sigma^*)\in D} \left(\min_{(\epsilon_{k+1},\sigma_{k+1})\in E_{k+1}} |(\epsilon_{k+1}-\epsilon^*, \sigma_{k+1}-\sigma^*)|^2 \right)$$

T. Kirchdoerfer and M. Ortiz, *IJNME*, **113**(11) (2018) 1697-1710.

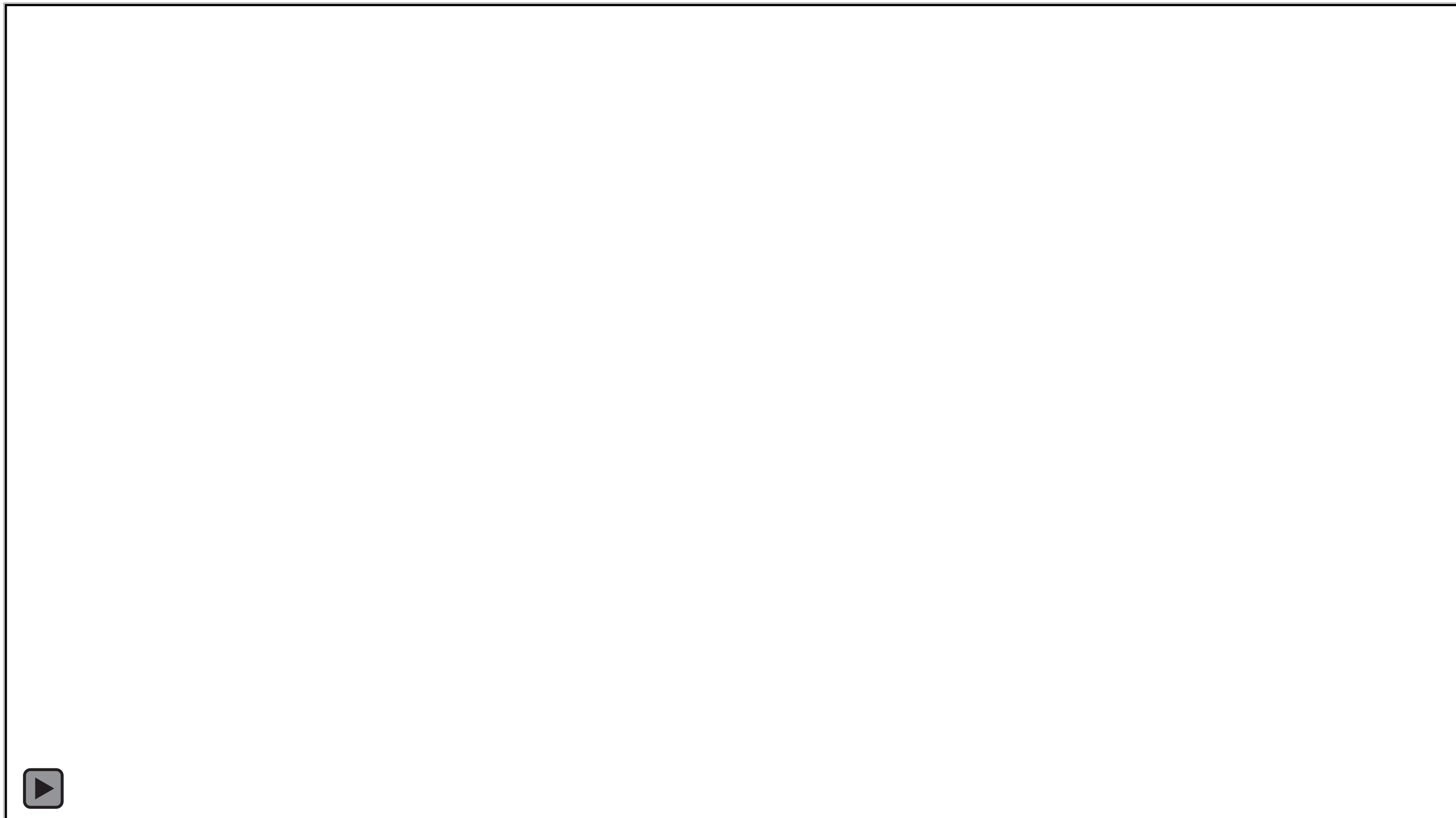
Time-dependent problems: Dynamics



T. Kirchdoerfer and M. Ortiz, *IJNME*, **113**(11) (2018) 1697-1710.

Michael Ortiz
WCCM 2018

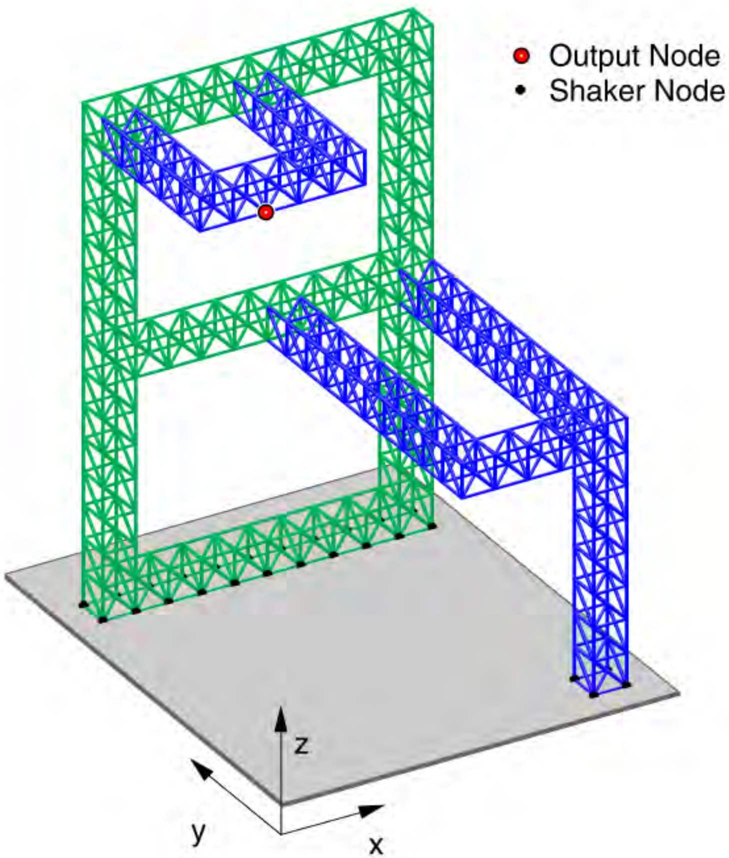
Test case: Truss dynamics

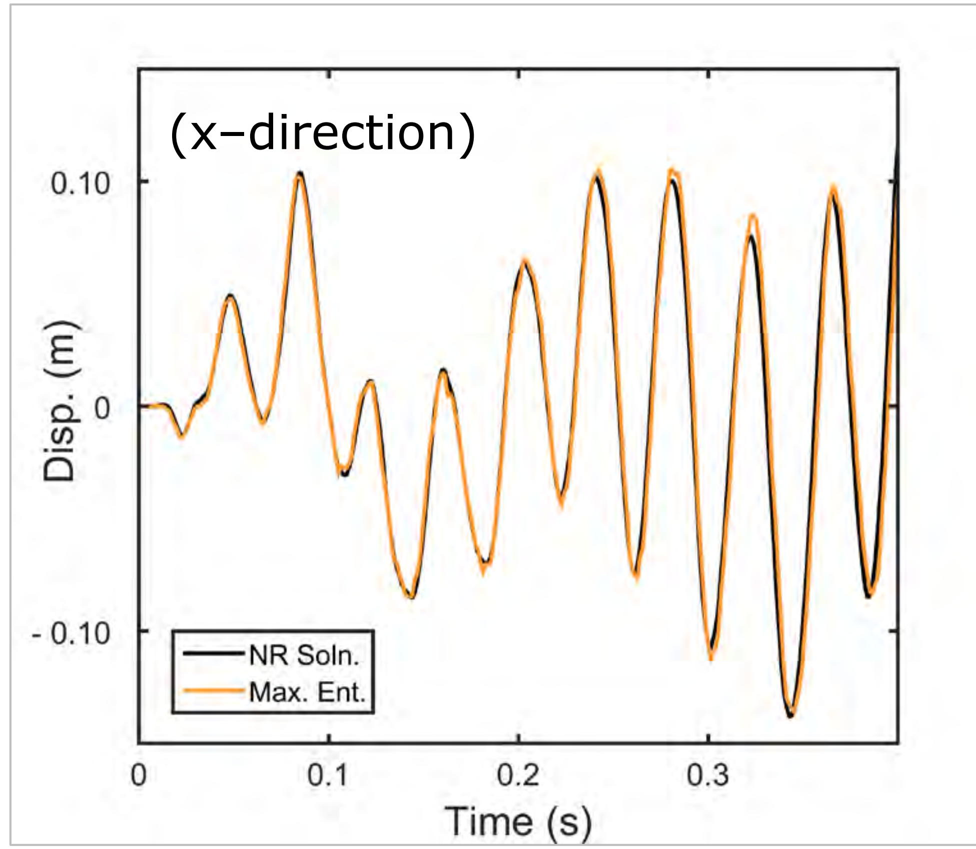


Data-Driven dynamics solution and data coverage

T. Kirchdoerfer and M. Ortiz, *IJNME*, **113**(11) (2018) 1697-1710.

Test case: Truss dynamics





Data-Driven solution vs. direct Newmark solution

T. Kirchdoerfer and M. Ortiz, *IJNME*, **113**(11) (2018) 1697-1710.

Data-Driven Dynamics

- Distance-based DD solvers carry over to *time-dependent problems*, in particular *dynamics*
- *Convergence properties* of solvers (fixed-point iteration, convergence with respect to time-step, convergence with respect to data…) *are identical to the static case*
- *Essential difference*: For time-dependent problems the *constraint set is time-dependent* due to dependence on initial conditions

[1]T. Kirchdoerfer and M. Ortiz, *CMAME*, **326** (2017) 622-641.

Data-Driven Computing: Issues

- *Data-driven (model-free!) computing: Use material data sets directly in calculations!*
- Is the Data-Driven reformulation of classical BVPs (possibly off of noisy data) *well-posed*?
- Implementation of *Data-Driven solvers*?
- *Numerical convergence* (iterative solvers, mesh size, time step…)
- Convergence with respect to *material data set?*
- Extension to *time-dependent* problems
- Extension to *history-dependent* materials
- *Phase-space sampling* in high dimension
- *Data management*, repositories, outlook…

Inelasticity and history dependence

- Inelastic materials have memory[1,2,3]:

  "*The characteristic property of inelastic solids which distinguishes them from elastic solids is the fact that the stress measured at time t depends not only on the instantaneous value of the deformation but also on the entire history of deformation.*"

- Data-Driven interpretation: *Time-dependent, evolving material data set*:

$$D(t) = \{(\epsilon(t), \sigma(t)) \ : \ \text{past material history}\}$$

[1]A.E. Green and R.S. Rivlin, *ARMA*, **1** (1957) 1.
[2]A.E. Green, R.S. Rivlin and A.J.M. Spencer, *ARMA*, **3** (1959) 82.
[3]A.E. Green and R.S. Rivlin, *ARMA*, **4** (1960) 387.

Data-driven inelasticity

- Constraint set (time dependent): $E_{k+1} =$

$$\left\{ \epsilon_{e,k+1} = B_e u_{k+1}, \quad \sum_{e=1}^{m} w_e B_e^T \sigma_{e,k+1} = f_{k+1}^{\text{ext}} \right\}$$

- History-dependent (local) material data sets:

$$D_{e,k+1} = \left\{ (\epsilon_{e,k+1}, \sigma_{e,k+1}) \ : \ \text{past material history} \right\}$$

- Data-driven problem:

$$\min_{(\epsilon^*,\sigma^*)\in D_{k+1}} \left(\min_{(\epsilon_{k+1},\sigma_{k+1})\in E_{k+1}} |(\epsilon_{k+1}-\epsilon^*, \sigma_{k+1}-\sigma^*)|^2 \right)$$

- ***Fundamental question: Data representability!***

Eggersmann, R., Kirchdoerfer, T., Stainier, L.,
Reese, S. and Ortiz, M., arXiv:1808.10859 **[physics.comp-ph]**

Data representation paradigms

- Hereditary representation (fading memory):

$$D_{e,k+1} = \Big\{ (\epsilon_{e,k+1}, \sigma_{e,k+1}) \ : \ \{\epsilon_{e,h}\}_{h \leq k} \Big\}$$

- History-variable representation:

Keep: $q_{e,k} = \widehat{q}_e(\{\epsilon_{e,h}, \sigma_{e,h}\}_{h \leq k})$ *(ad hoc)*

Then: $D_{e,k+1} = \Big\{ (\epsilon_{e,k+1}, \sigma_{e,k+1}) \ : \ q_{e,k} \Big\}$

- Differential representation: $D_{e,k+1} =$

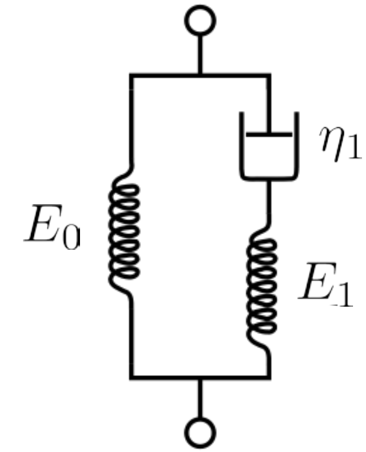
$$\Big\{ f_e(\epsilon_{e,k+1}, \epsilon_{e,k}, \epsilon_{e,k-1} \cdots \sigma_{e,k+1}, \sigma_{e,k}, \sigma_{e,k-1} \cdots \Big\} = 0\}$$

- ***Open question: Data convergence?***

S. Flügge and C. Truesdell, *Handbuch der Physik* (1972).

Data-driven viscoelasticity

- Smooth kinetics (linear or nonlinear)
- Allows for differential representation
- Example: Standard Linear Solid,



$$\sigma + \tau_1 \dot{\sigma} - E_0 \epsilon - (E_0 + E_1) \tau_1 \dot{\epsilon} = 0$$

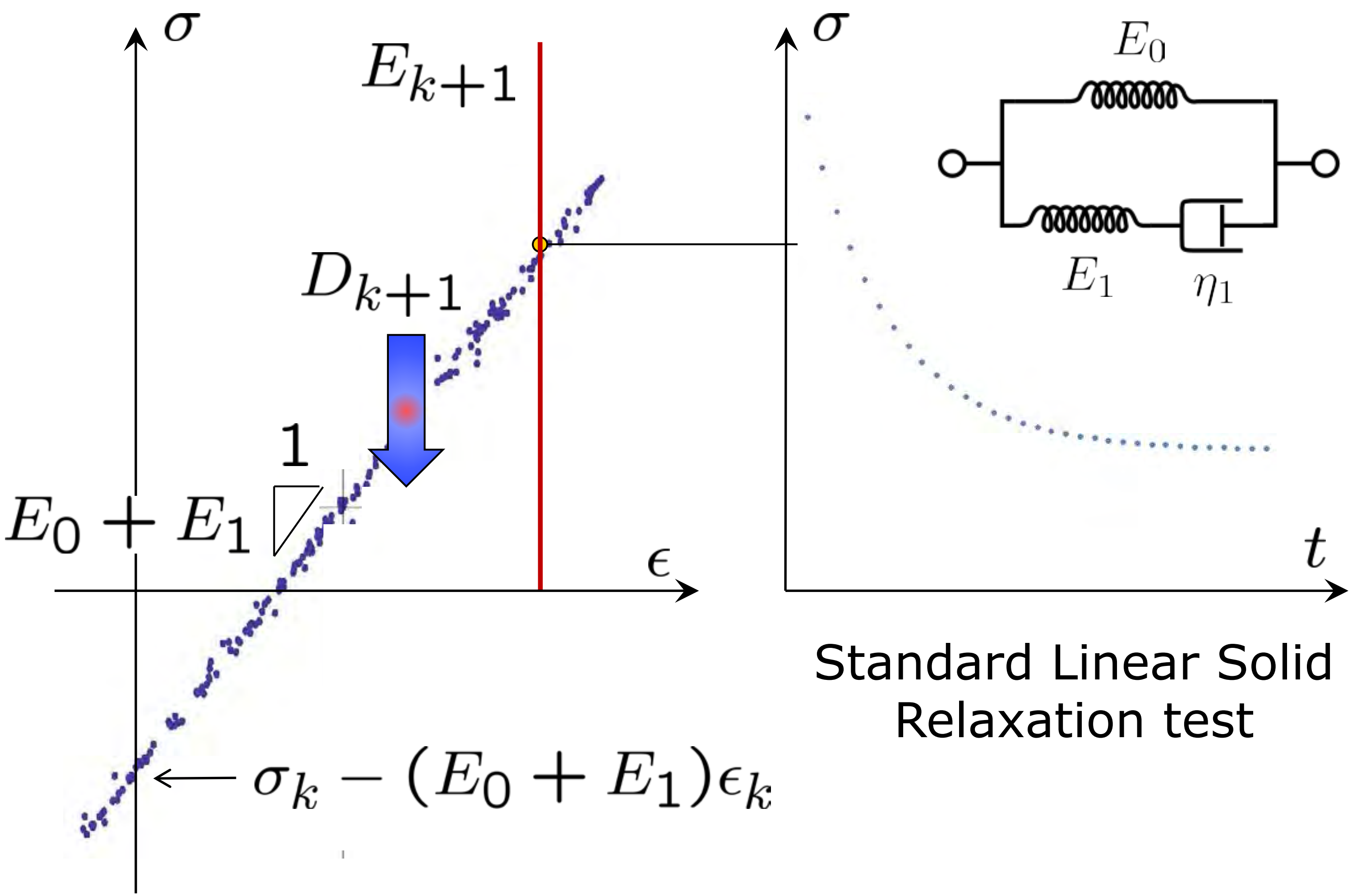
- Time discretization: $D_{k+1} =$

$$\left\{ \sigma_{k+1} + \tau_1 \frac{\sigma_{k+1} - \sigma_k}{t_{k+1} - t_k} - E_0 \epsilon_{k+1} - (E_0 + E_1) \tau_1 \frac{\epsilon_{k+1} - \epsilon_k}{t_{k+1} - t_k} = 0 \right\}$$

- General first-order differential materials:

$$D_{k+1} = \left\{ (\epsilon_{k+1}, \sigma_{k+1}) \, : \, (\epsilon_k, \sigma_k) \right\}$$

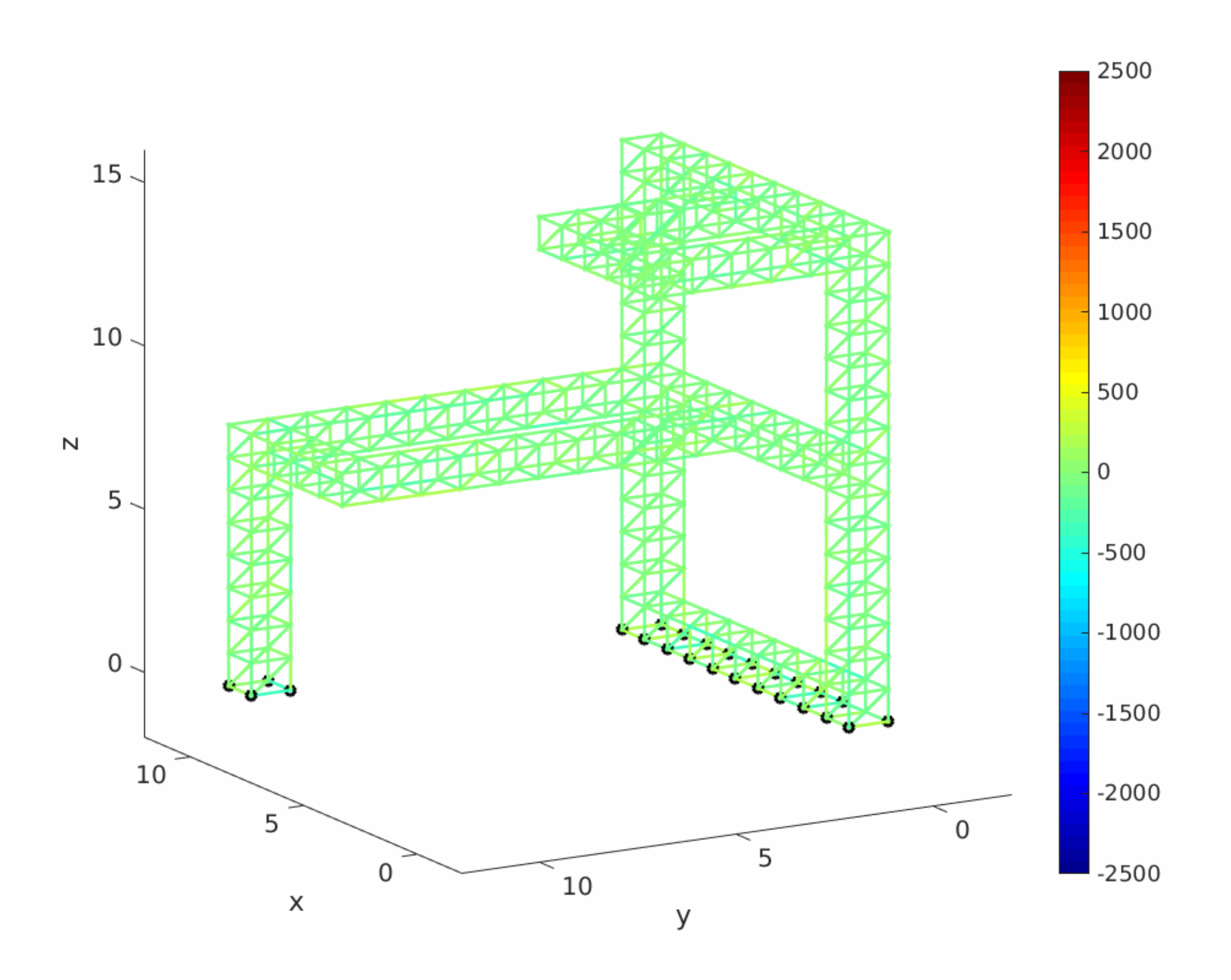
Data-Driven viscoelasticity



Standard Linear Solid
Relaxation test

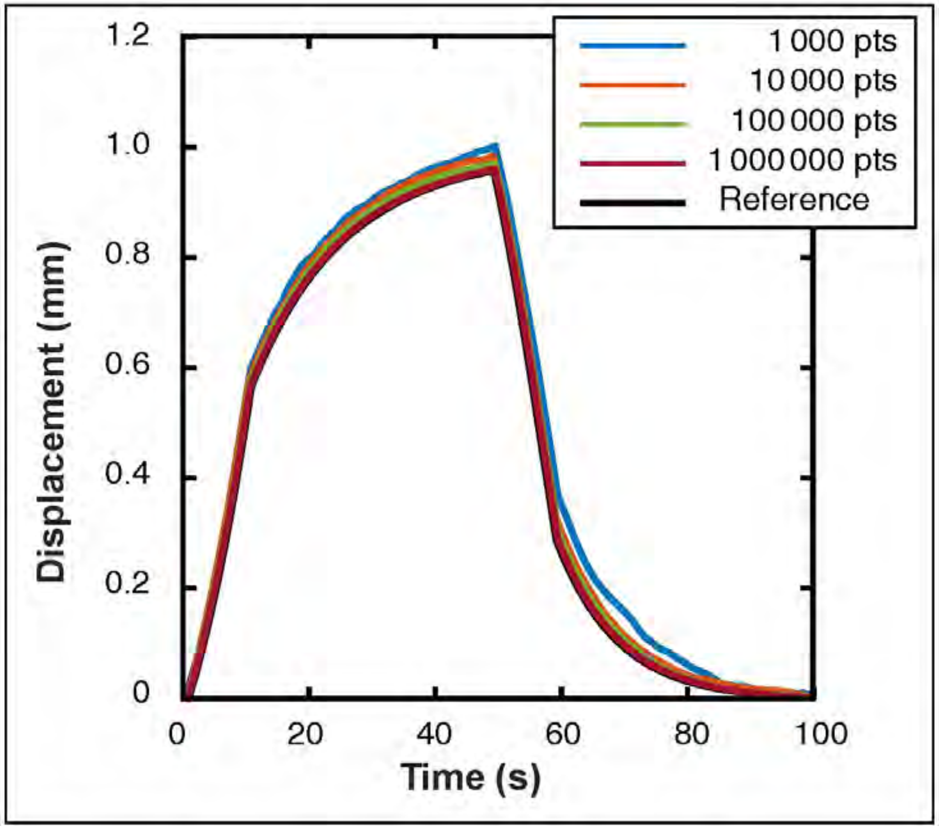
Michael Ortiz
WCCM 2018

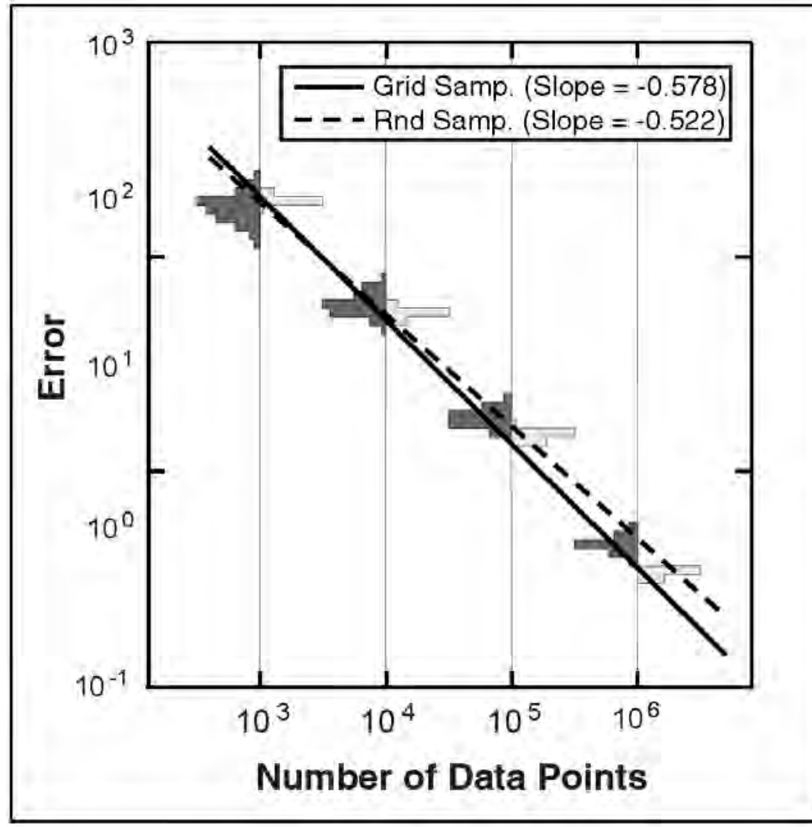
Data-Driven viscoelasticity



Michael Ortiz
WCCM 2018

Data-Driven viscoelasticity

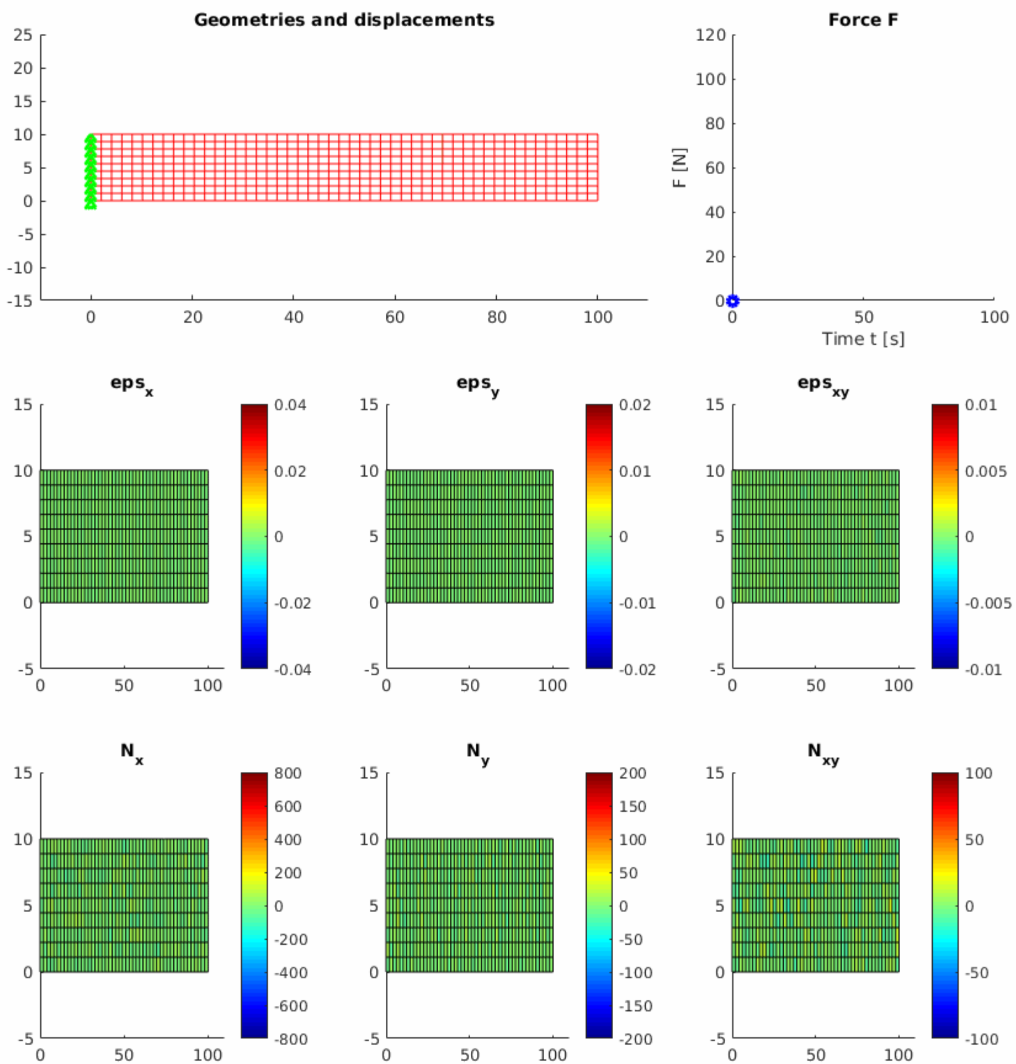




Convergence with respect to the data set

Eggersmann, R., Kirchdoerfer, T., Stainier, L.,
Reese, S. and Ortiz, M., arXiv:1808.10859 **[physics.comp-ph]**

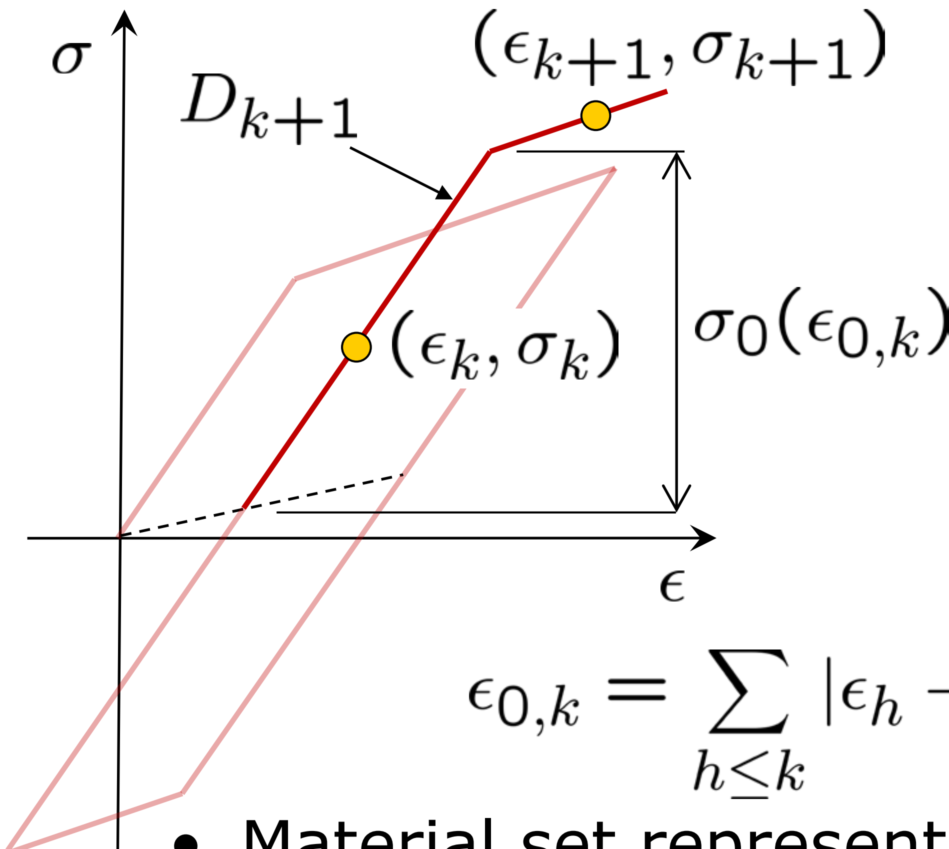
Data-Driven viscoelasticity - FE

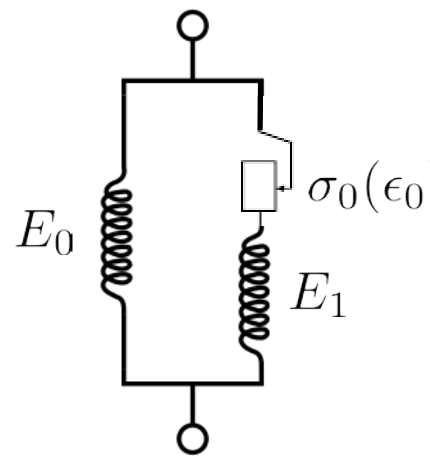


Michael Ortiz
WCCM 2018

Data-Driven plasticity

- Example: Isotropic/kinematic hardening





- History variable:

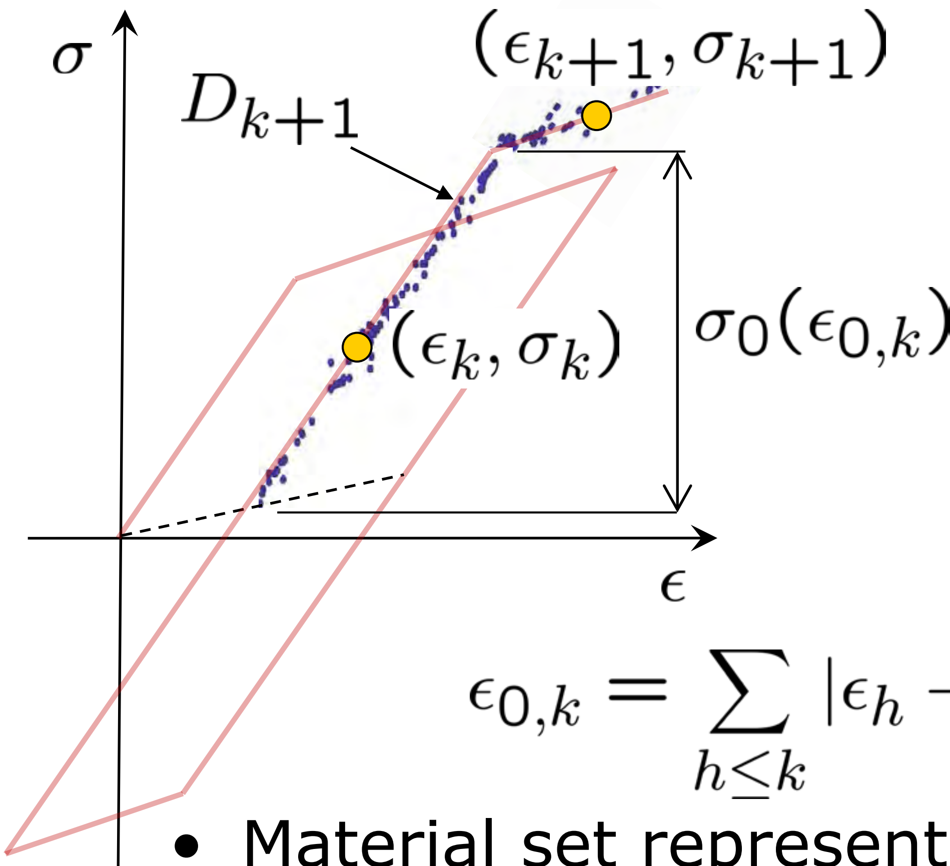
$$\epsilon_{0,k} = \sum_{h \leq k} |\epsilon_h - \epsilon_{h-1} - (\sigma_h - \sigma_{h-1})/E|$$

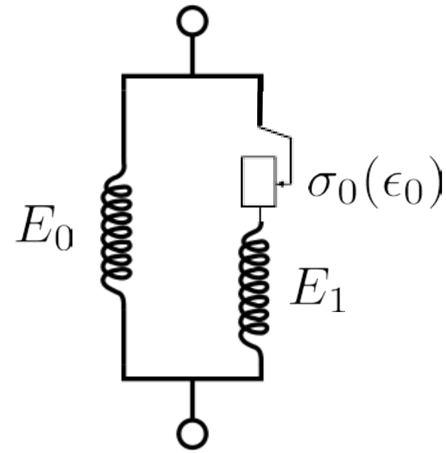
- Material set representation (diff + history):

$$D_{k+1} = \left\{ (\epsilon_{k+1}, \sigma_{k+1}) \, : \, (\epsilon_k, \sigma_k), \epsilon_{0,k} \right\}$$

Data-Driven plasticity

- Example: Isotropic/kinematic hardening





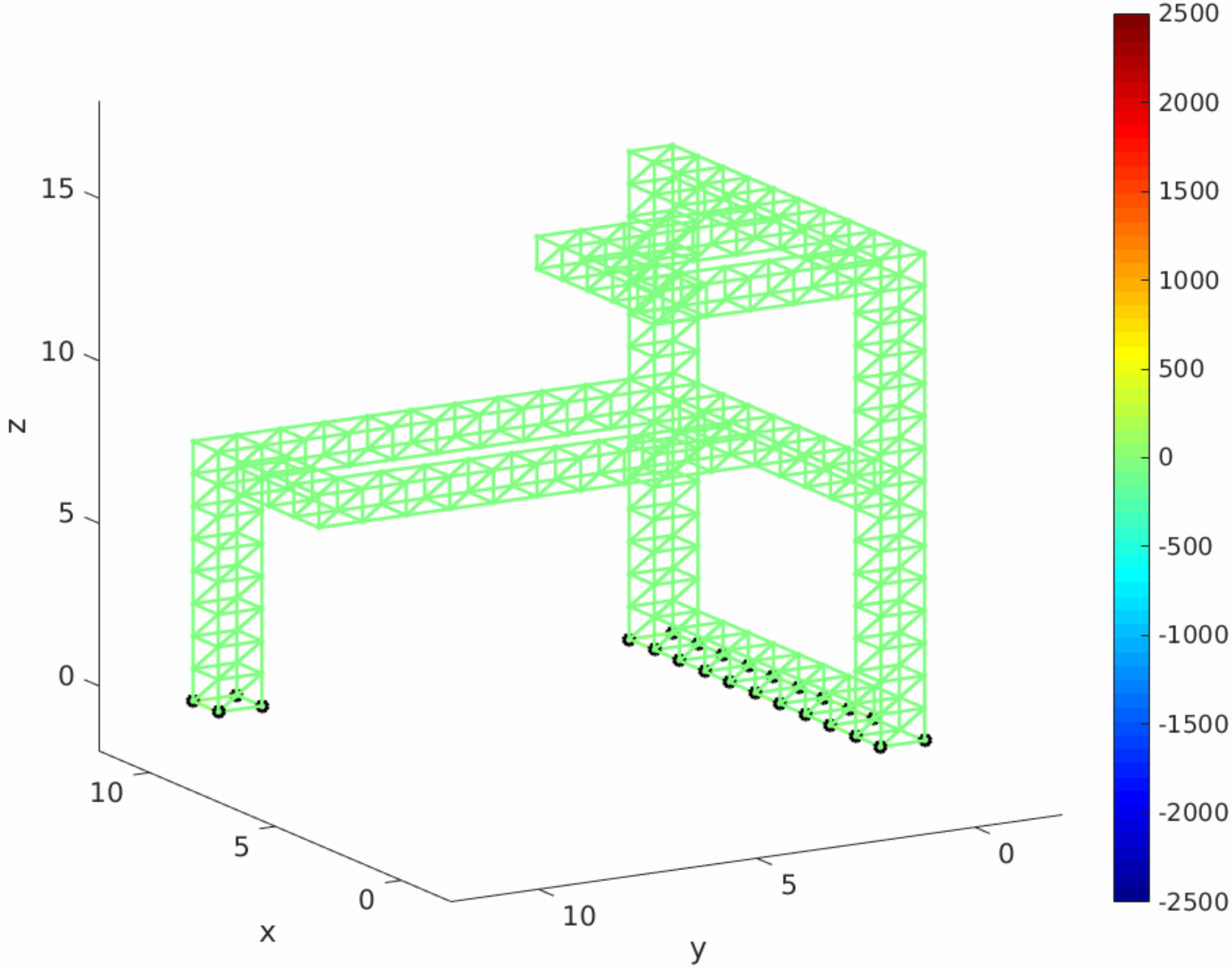
- History variable:

$$\epsilon_{0,k} = \sum_{h \leq k} |\epsilon_h - \epsilon_{h-1} - (\sigma_h - \sigma_{h-1})/E|$$

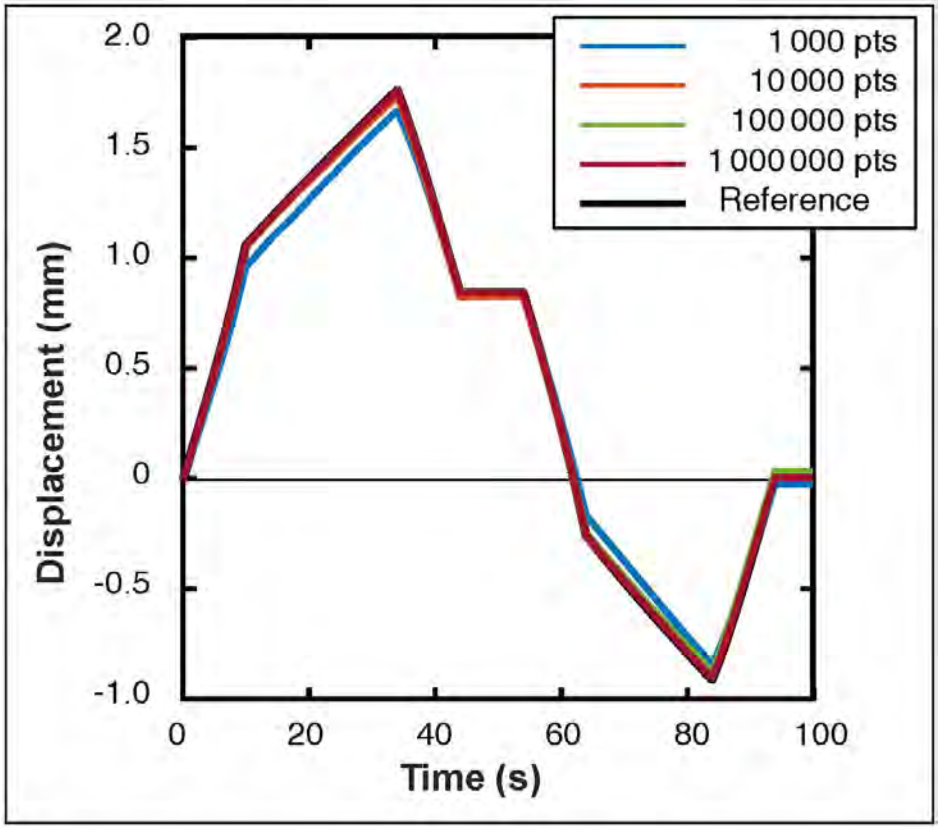
- Material set representation (diff + history):

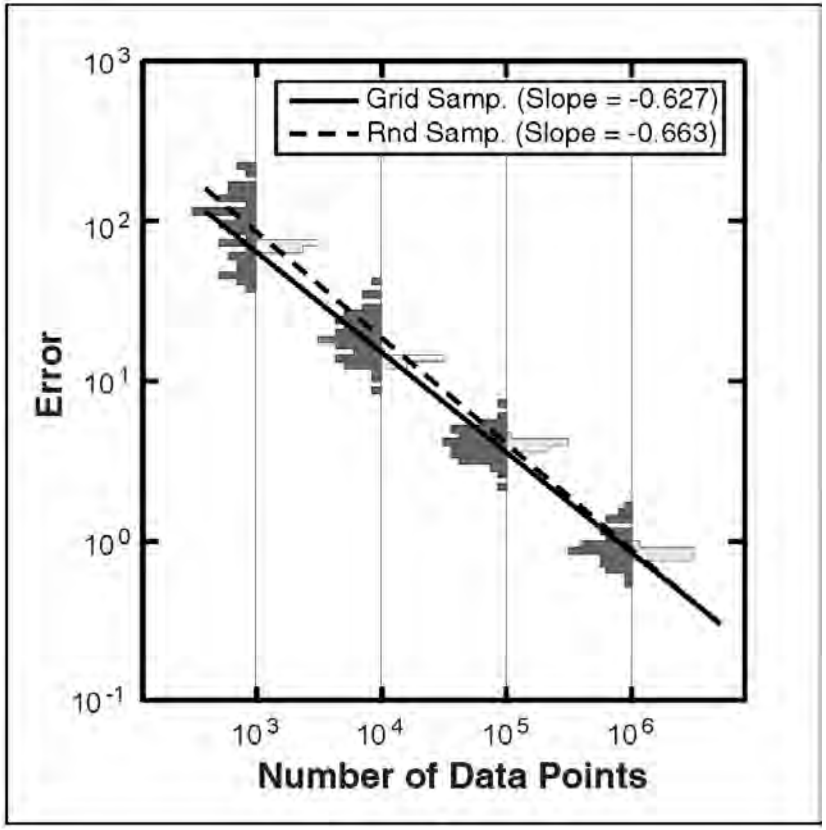
$$D_{k+1} = \left\{ (\epsilon_{k+1}, \sigma_{k+1}) \,:\, (\epsilon_k, \sigma_k), \epsilon_{0,k} \right\}$$

Data-Driven plasticity



Data-Driven plasticity



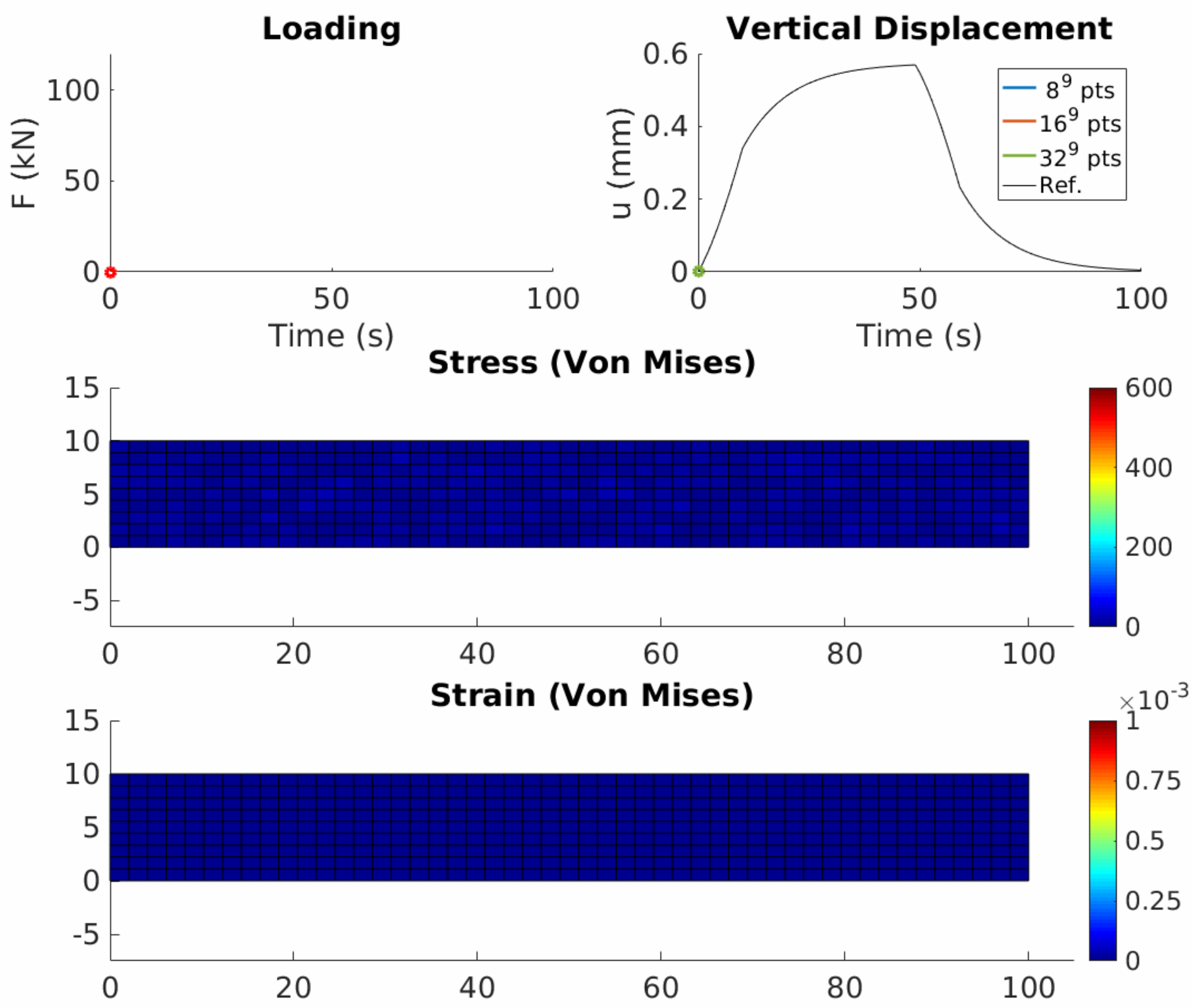


Convergence with respect to the data set

Eggersmann, R., Kirchdoerfer, T., Stainier, L., Reese, S. and Ortiz, M., arXiv:1808.10859 [physics.comp-ph]

Michael Ortiz
WCCM 2018

Data-Driven plasticity - FE



Concluding remarks

- *Data-driven computing* is emerging as an alternative to model-based computing
- Data-Driven mechanics entails a *comprehensive reformulation* of classical problems of mechanics
- Data-driven computing can reliably supply *model-free* solutions from material data sets
- Data can be *mined* from lower-scale calculations, used in upper-scale calculations (*DD upscaling*)
- Data can also be extracted from *full-field experimental data* (TEM, SEM, DIC, EBSD…)
- High-dimensional phase spaces: *Self-consistent importance sampling*[1] (problem specific)

J. Rethore, HAL Id: hal-01454432, Feb. 2017.
J. Rethore and A. Leygue, HAL Id: hal-01452494, Feb. 2017.

Concluding remarks

- Reliance on *fundamental data* (stress and strain only, no model-dependent data) makes *material data fungible*, mergeable, interchangeable…
- *Publicly editable material data repository?* (*Wikimat*, a material data Wikipedia…)
- *Data-driven computing* is likely to be a *growth area* in an increasingly *data-rich world!*

Thank you!

Michael Ortiz
WCCM 2018