CALIFORNIA INSTITUTE OF TECHNOLOGY

universitätbonn

Model-Free Data-Driven Science: Cutting out the Middleman

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

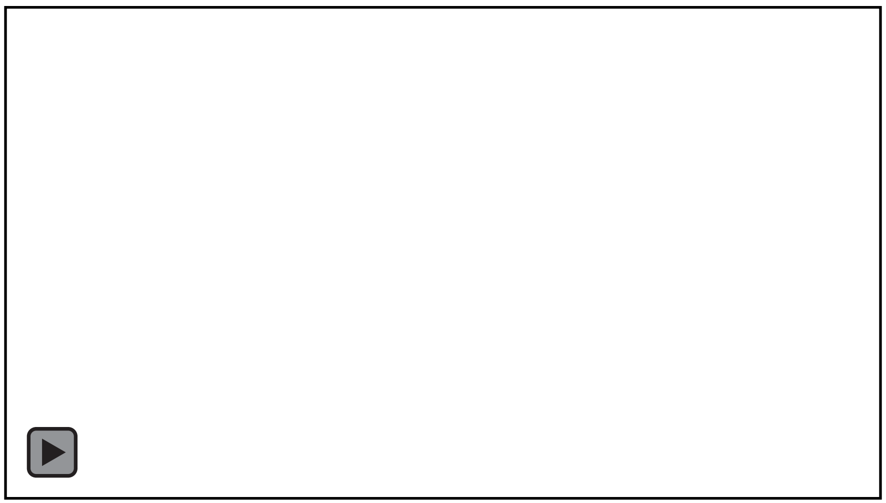
BiGmax (virtual) Workshop 2021
on Big-Data-Driven Materials Science
Max-Planck Research Network, April 15, 2021

Data-Driven Science: Talking points

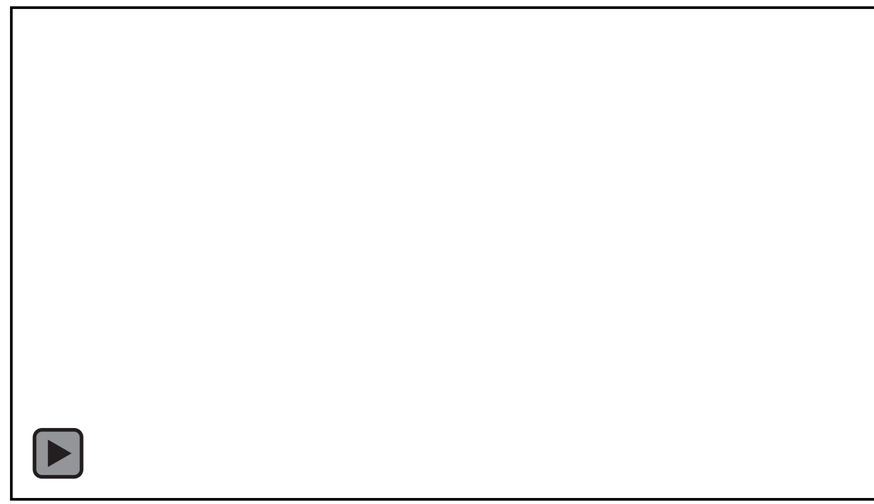
- New emerging paradigm: *Data-Driven Science*
- Why now? *What has changed*?
- How does Data Science intersect with the *physical sciences*? With *experimental science*?
- Is there *intellectual depth* to Data-Driven Science? Are there *opportunities* for fundamental, enduring, contributions?
- Data-Driven Science: *Theory vs. practice*
- Is Data-Driven Science likely to change *engineering practice*? Industry?
- Data-Driven ecosystem, *infrastructure*…

The new data-rich world…

- Material *data is currently plentiful* due to dramatic advances in experimental science (DIC, EBSD, microscopy, tomography…) and multiscale computing (DFT → MD → DDD → SM → Hom)



3D tomographic reconstruction of particles in battery electrode

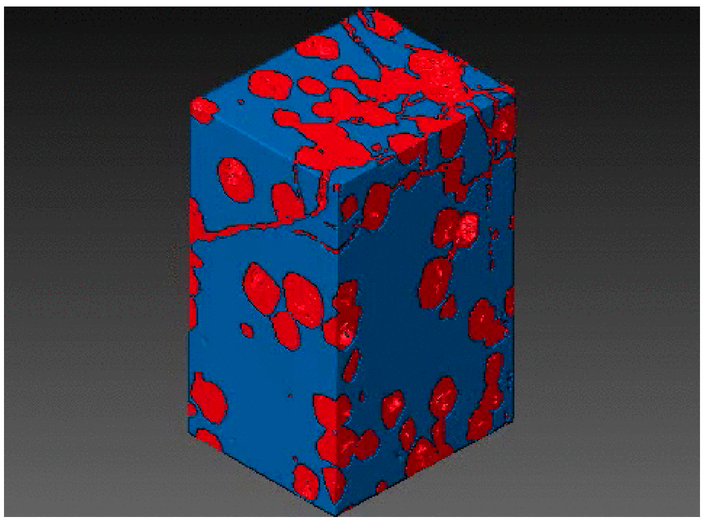


3D DIC-measured internal-strain full-field compressed PDMS sample

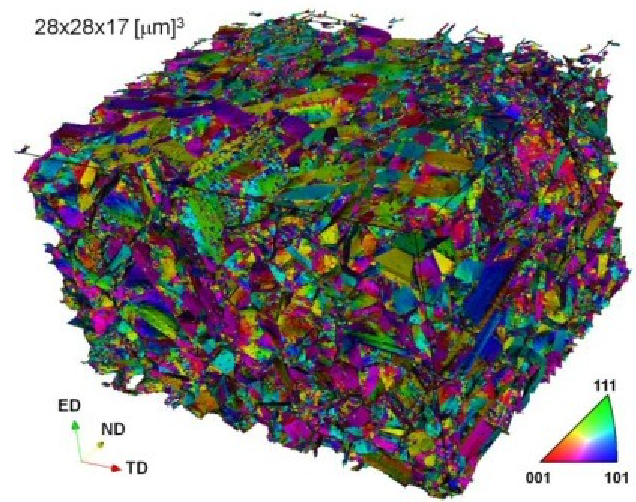
John Lambros, UIUC,
https://lambros.ae.illinois.edu/moviesimages/

The new data-rich world…

- Material *data is currently plentiful* due to dramatic advances in experimental science (DIC, EBSD, microscopy, tomography…) and multiscale computing (DFT → MD → DDD → SM → Hom)



Two-phase μCT analysis of Ti2AlC/Al composite[1]



3D EBSD microstructure in Cu-0.17wt%Zr after ECAP[2]

[1]Hanaor *etal*, *Mater Sci Eng A*, **672** (2019) 247.
[2]Khorashadizadeh, *Adv Eng Mater*, **13** (2011) 237.

Michael Ortiz
BiGmax 2021

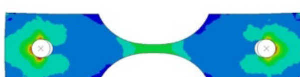
Adapting to a new data-rich world...

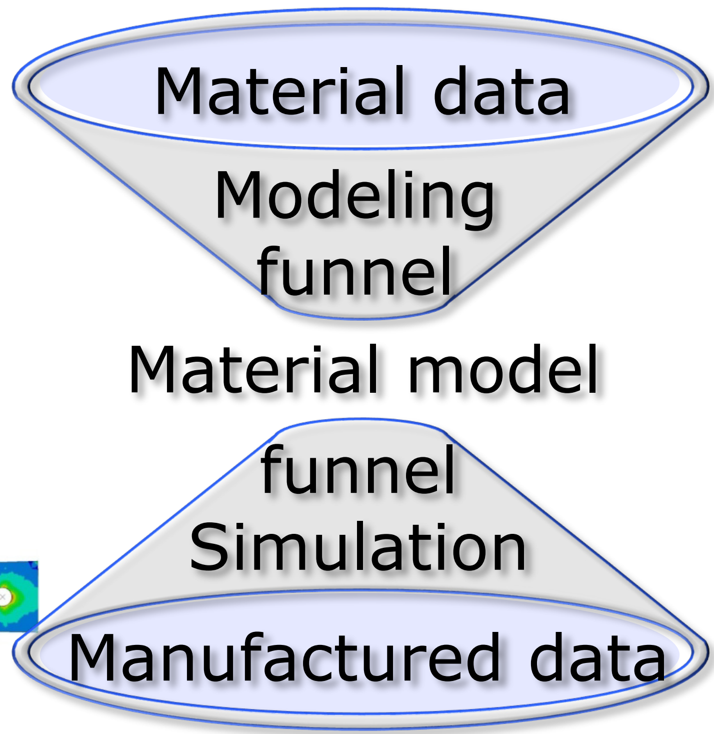
Classical Model-Based Computational Science...

Modern microscopy generates massive data sets

$$\sigma = \mathbb{C}\epsilon$$

Modelling entails massive loss of information!





Adapting to a new data-rich world...

- Modeling entails *massive loss of information* from material data sets, inserts *uncertainty*
- Modeling is *ad hoc, open ended, ill-posed*
- There is *no theory* that determines models from first principles to a desired level of accuracy
- Modeling requires *heuristics and intuition*: Models are *only as good as the modeler*'s physical intuition, empirical knowledge
- *Deep Learning (ANN) – hype = piecewise-linear regression* (cf., e.g., Gilbert Strang, 2019)
- Can a more *direct connection between data and prediction* be forged? (*the data, all the data, nothing but the data!*) How?

Adapting to a new data-rich world...

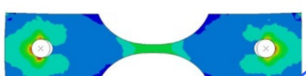
Classical Model-Based
Computational Science...

Alternative:
Model-Free
Data-Driven
Computing!

Modern
microscopy
generates
massive
data sets

$$\sigma = \mathbb{C}\epsilon$$

Modelling
entails
massive
loss of
information!



Material data

Modeling
funnel

Material model

funnel
Simulation

Manufactured data

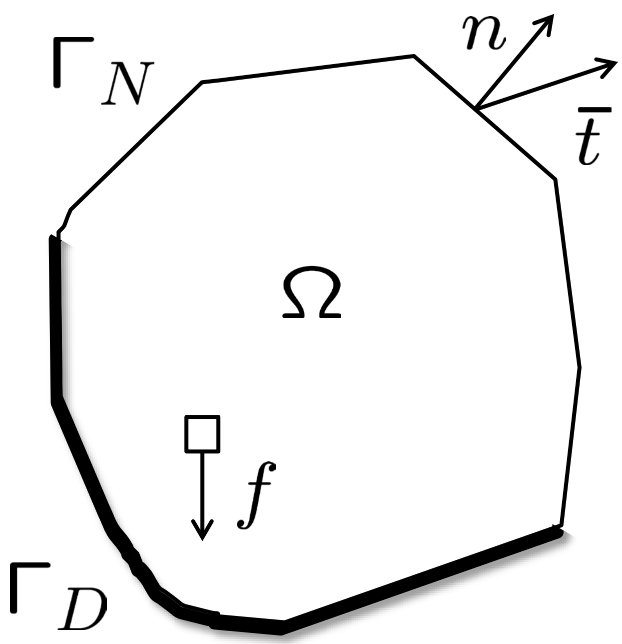
Set/solve
problems
directly
from data!

Eliminate
modeling
bottleneck!

Michael Ortiz
BiGmax 2021

Differential structure of field equations

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



i) Kinematics + Dirichlet BC:

$$\left.\begin{aligned} &\epsilon(u) = 1/2(\nabla u + \nabla u^T) \\ &u = \bar{u}, \quad \text{on } \Gamma_D \end{aligned}\right\}$$

ii) Equilibrium + Neumann BC:

$$\left.\begin{aligned} &\operatorname{div}\sigma + f = 0 \\ &\sigma\, n = \bar{t}, \quad \text{on } \Gamma_N \end{aligned}\right\}$$

iii) *Material law:* $\sigma(x) = \sigma(\epsilon(x))$

Differential structure of field equations

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

Universal laws!
Material independent!
Exactly known!
Uncertainty-free!
Similarly:
Maxwell, Einstein, Schrödinger
Mass transport…

i) Kinematics + Dirichlet BC:

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

δ!

ii) Equilibrium + Neumann BC:

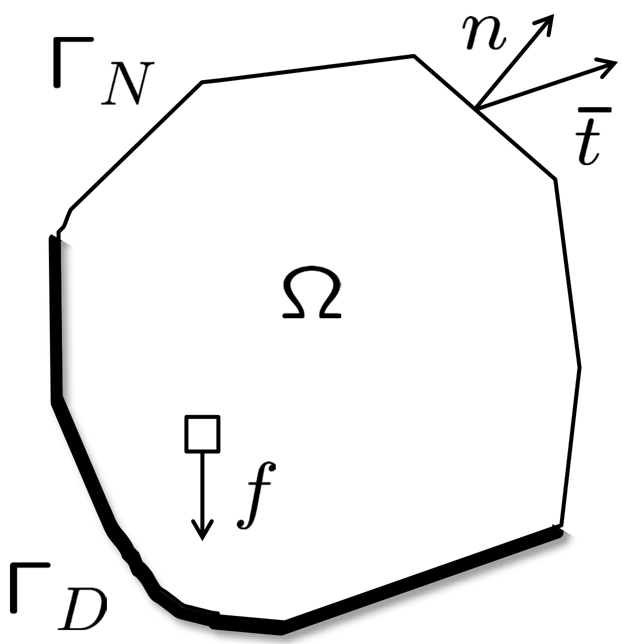
$$\text{div}\,\sigma + f = 0$$
$$\sigma\, n = \bar{t}, \quad \text{on } \Gamma_N$$

∂!

iii) *Material law:* $\sigma(x) = \sigma(\epsilon(x))$

Differential structure of field equations

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



i) Kinematics + Dirichlet BC:

$$\left.\begin{aligned} &\epsilon(u) = 1/2(\nabla u + \nabla u^T) \\ &u = \bar{u}, \quad \text{on } \Gamma_D \end{aligned}\right\}$$

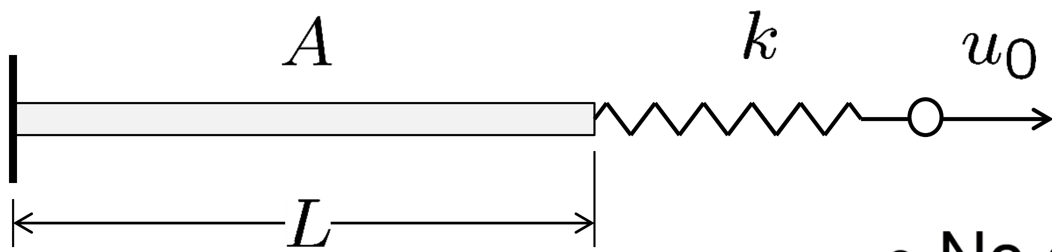
ii) Equilibrium + Neumann BC:

$$\left.\begin{aligned} &\operatorname{div}\sigma + f = 0 \\ &\sigma\, n = \bar{t}, \quad \text{on } \Gamma_N \end{aligned}\right\}$$

iii) *Material law:* $\sigma(x) = \sigma(\epsilon(x))$

Unknown! Epistemic uncertainty!

Elementary example: Bar and spring

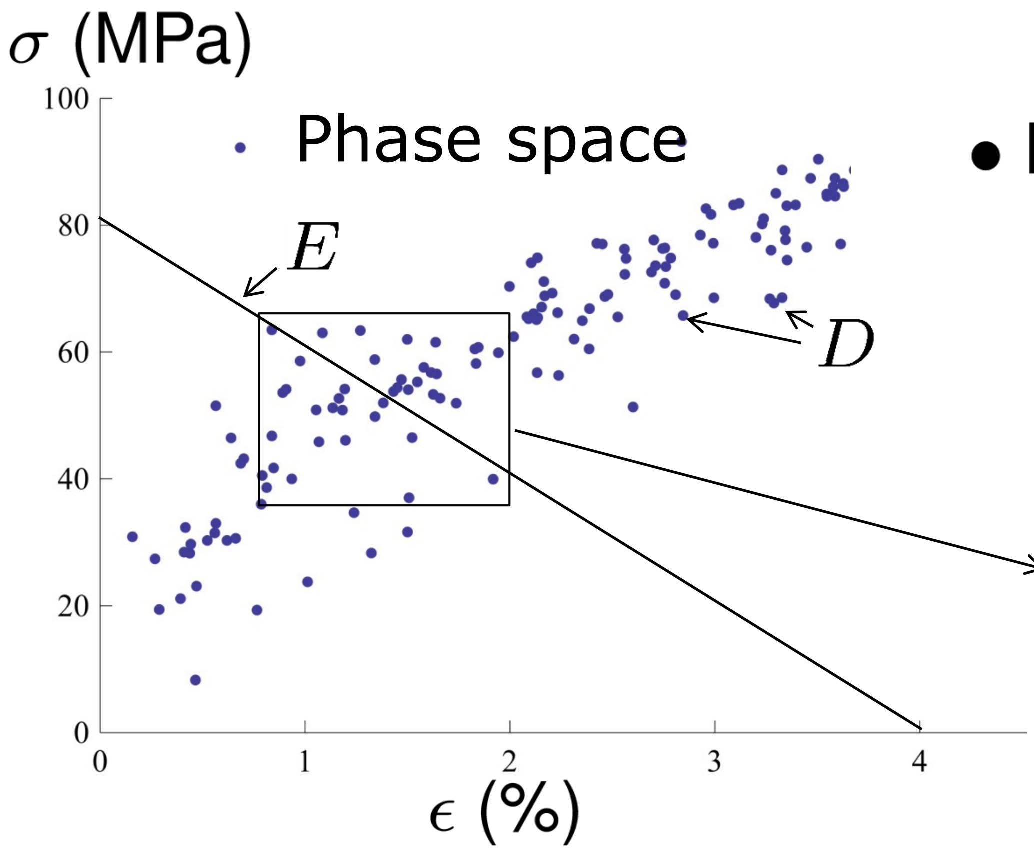


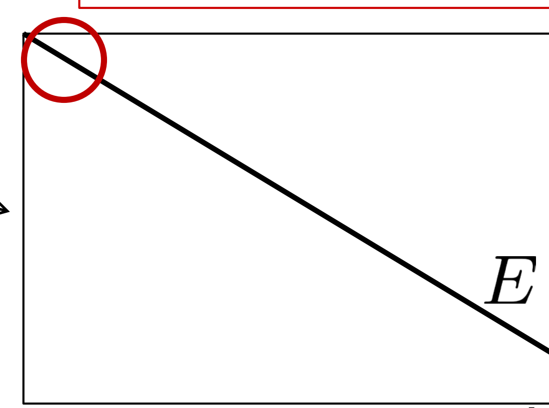
- No classical solutions!

$$D \cap E = \emptyset$$

- Data-driven solution:

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





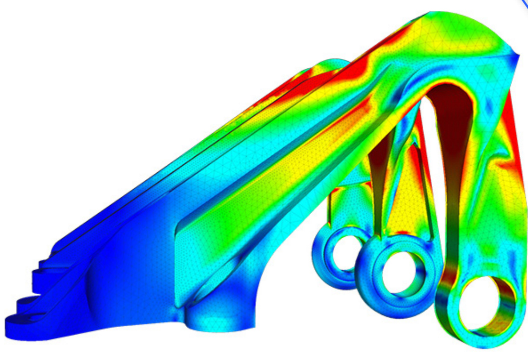
The general Data-Driven (DD) problem

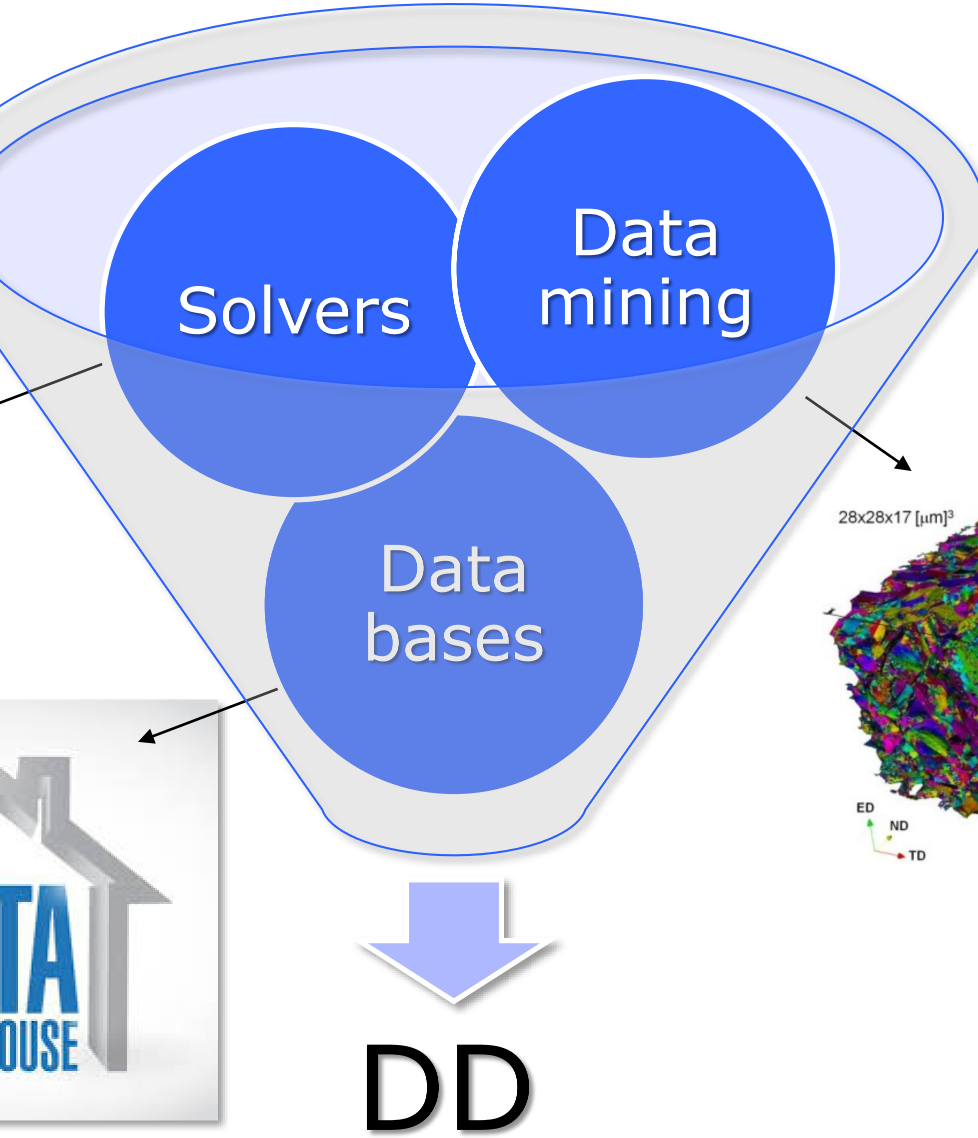
- The *Data-Driven paradigm*[1]: Given,
  - *D = {fundamental material data},*
  - *E = {compatibility + equilibrium},*

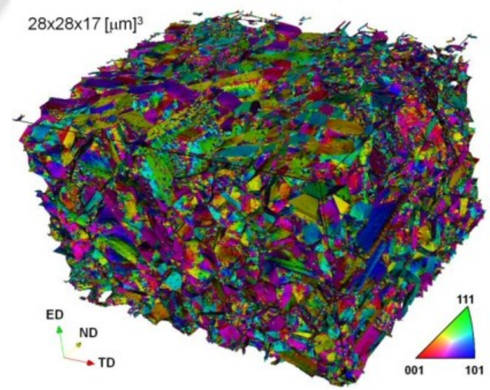
  Find: $\text{argmin}\{d(z, D),\ z \in E\}$
- *The aim of Data-Driven analysis is to find the admissible state (compatibility and equilibrium) closest to the material data set*
- Raw *fundamental material data* (stress *and* strain) is used (unprocessed) in calculations
- *No material modeling, no loss of information (the data, all the data, nothing but the data!)*

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

The (model-free) DD ecosystem…

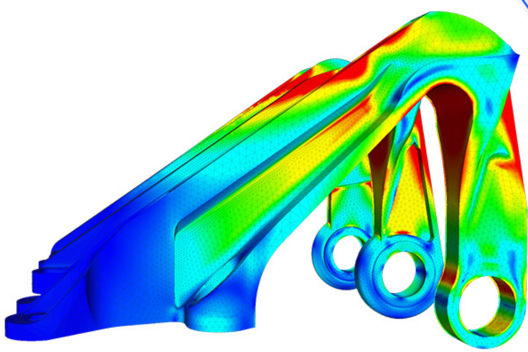


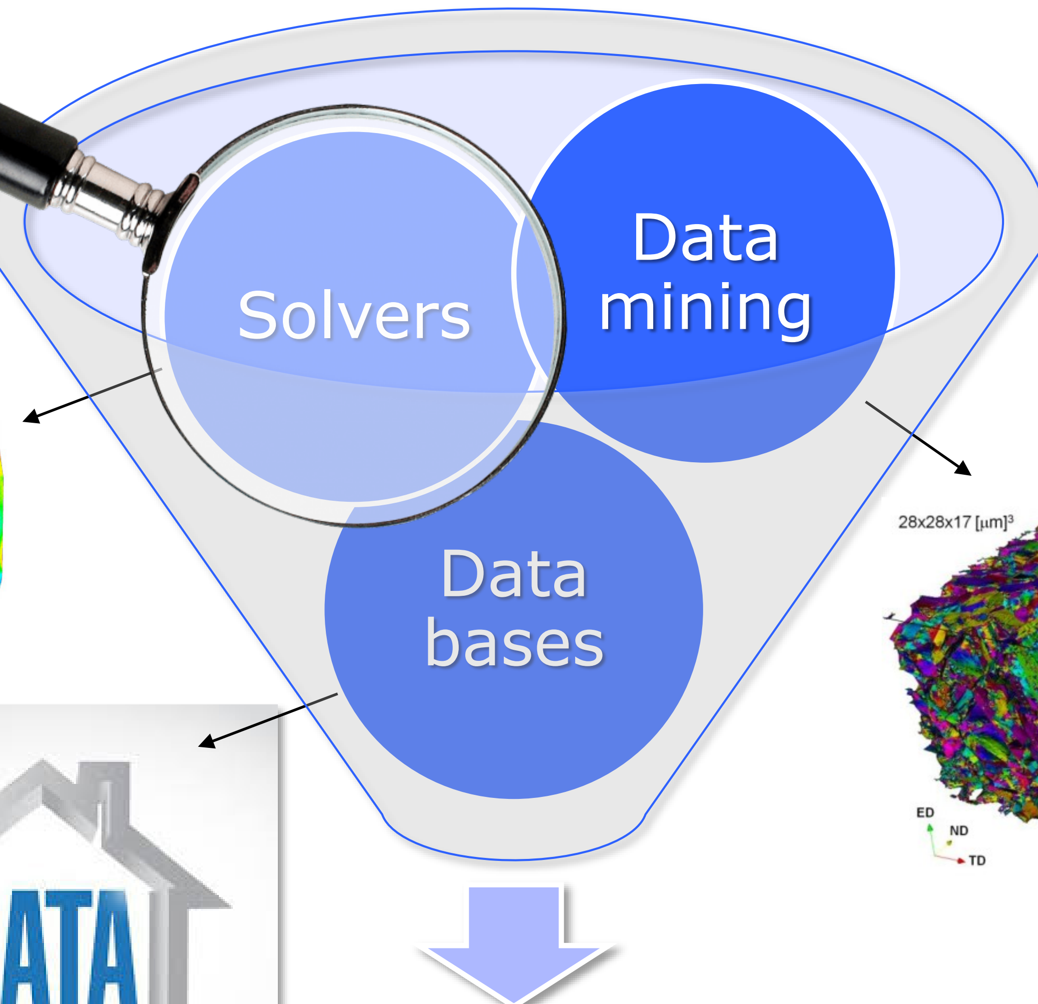


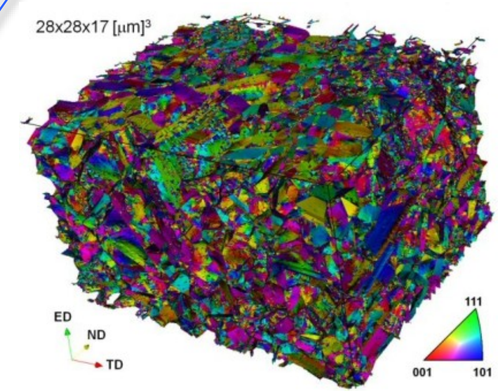


Michael Ortiz
BiGmax 2021

The (model-free) DD ecosystem…



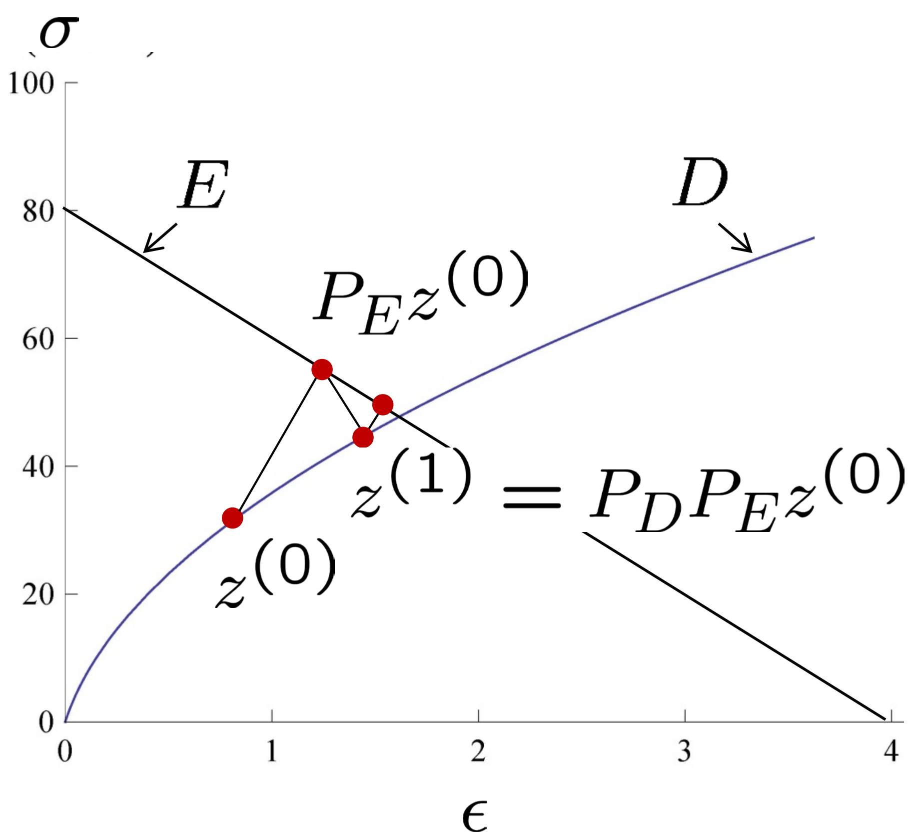


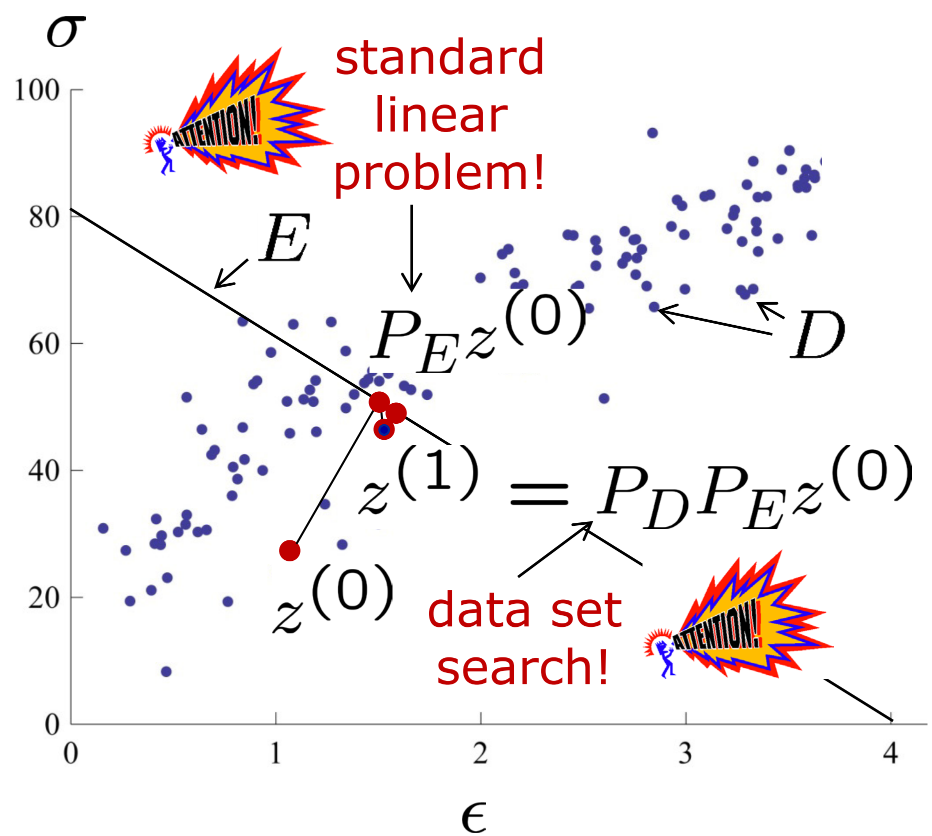


DD

DD solvers: Fixed-point iteration

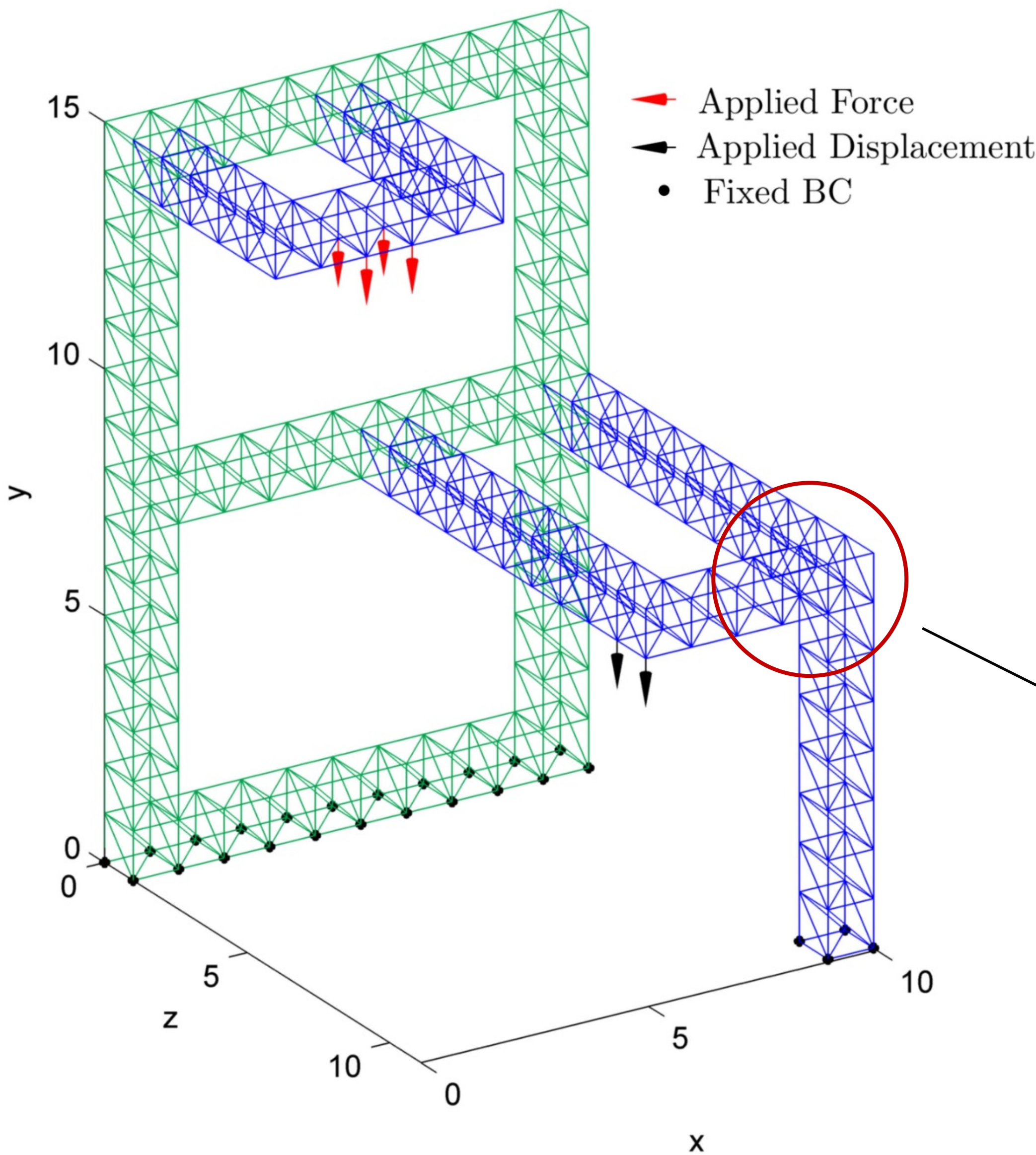
- Find: $\text{argmin}\{d(z, D),\ z \in E\}$
- Fixed-point iteration[1]: $z^{(k+1)} = P_D P_E z^{(k)}$

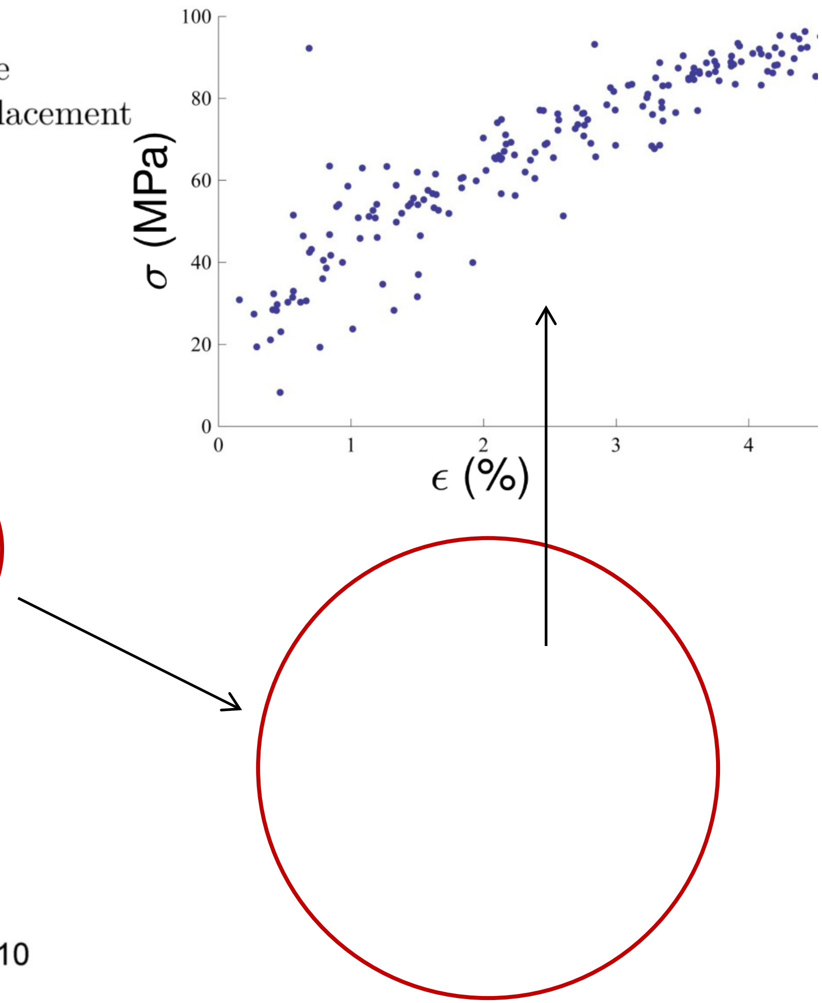




[1]T. Kirchdoerfer and M. Ortiz (2015) arXiv:1510.04232.
[1]T. Kirchdoerfer and M. Ortiz, *CMAME*, **304** (2016) 81–101

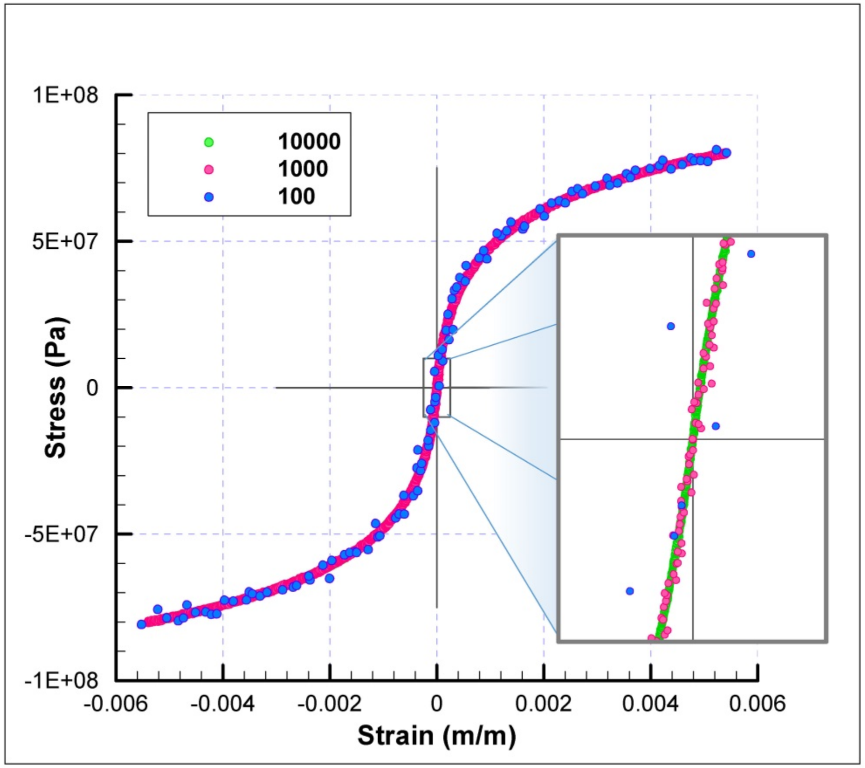
Example: 3D Truss



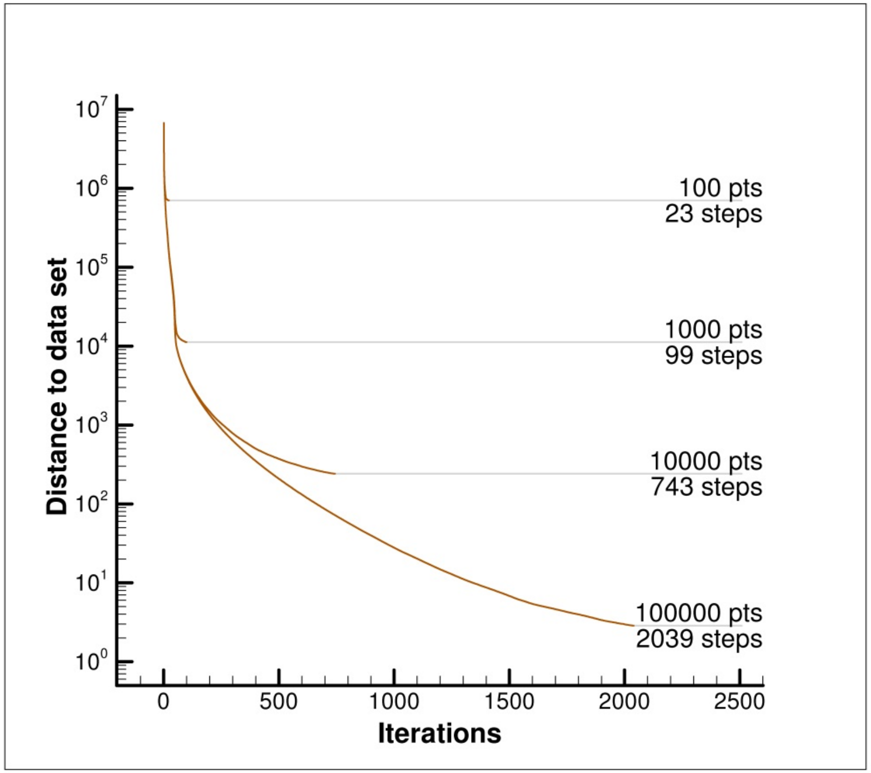


Michael Ortiz
BiGmax 2021

Convergence of fixed-point solver



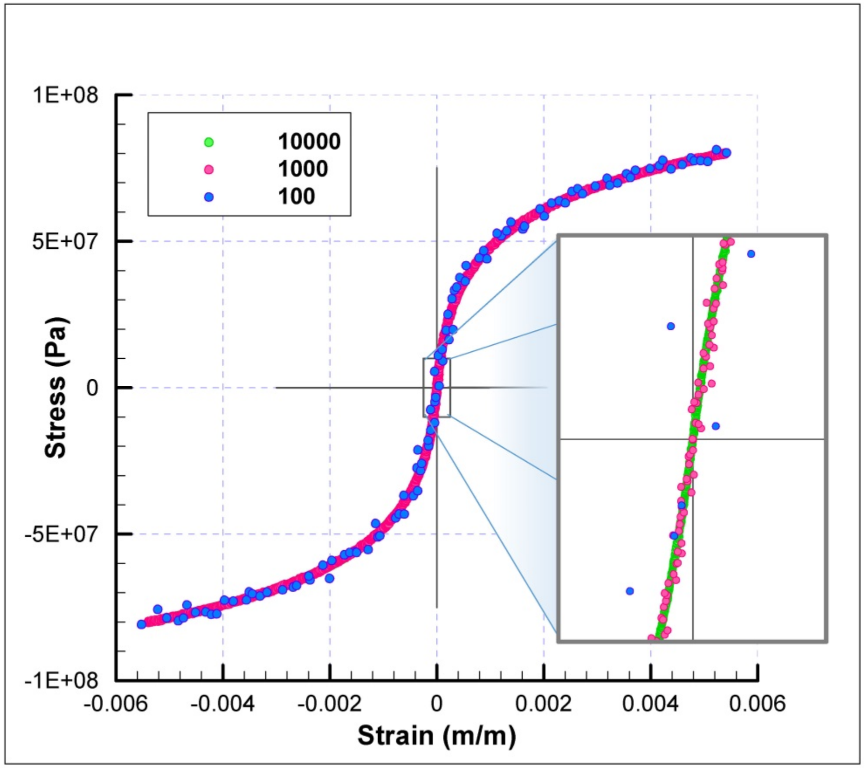
Randomized material-data sets of increasing size



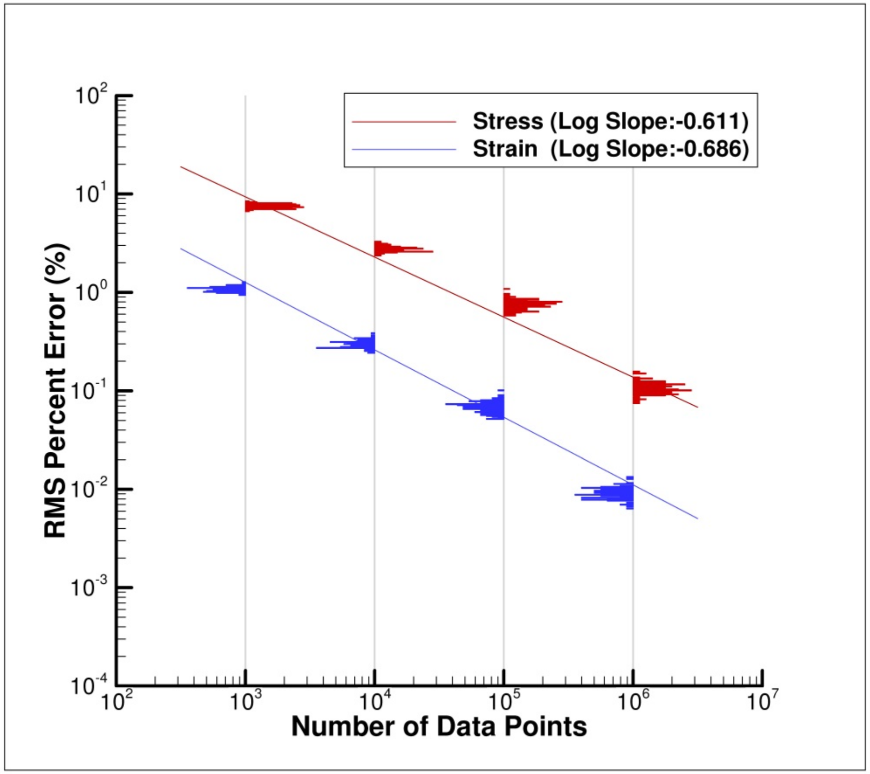
Convergence of fixed-point iteration

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

Convergence of fixed-point solver



Randomized material-data sets of increasing size and decreasing scatter



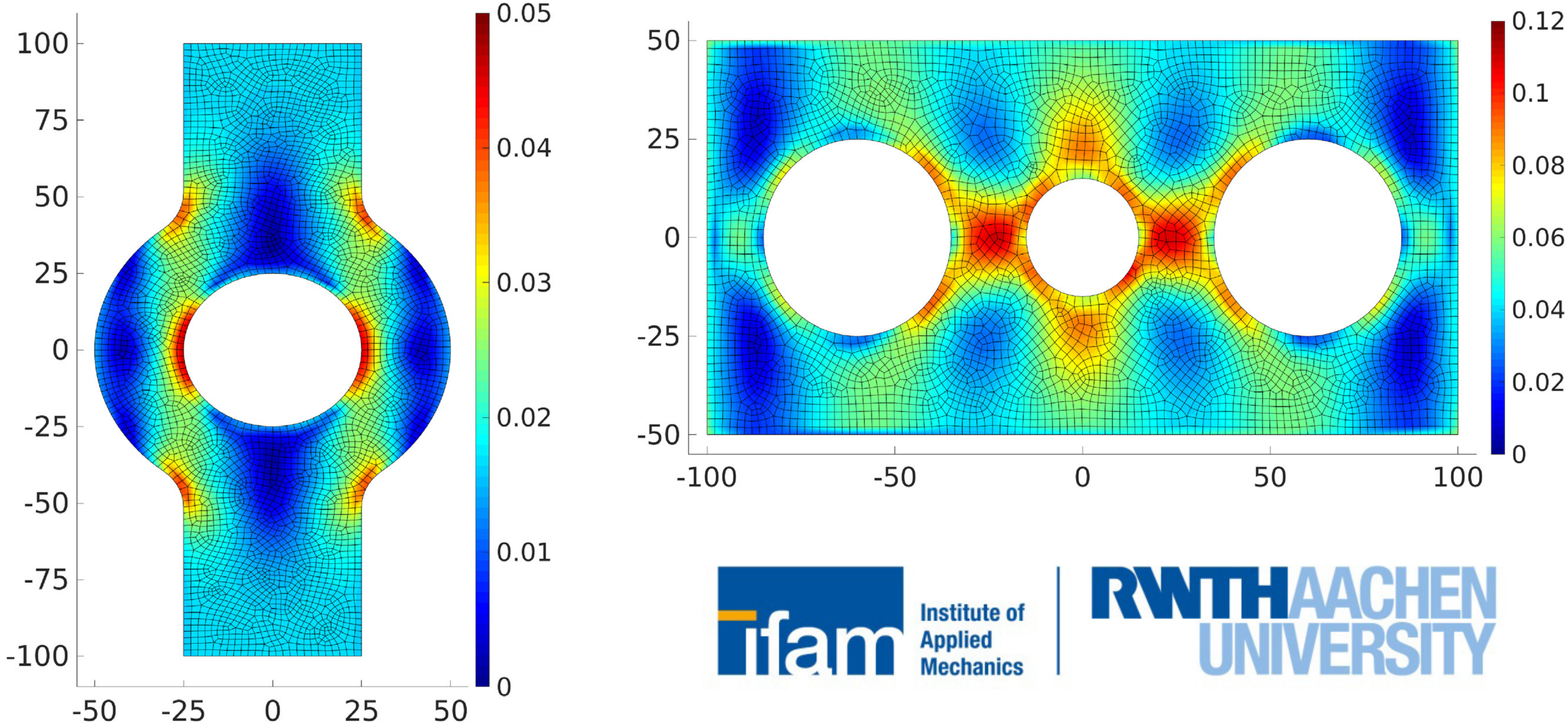
Convergence with respect to data set

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

Michael Ortiz
BiGmax 2021

Data-Driven solvers and FE analysis

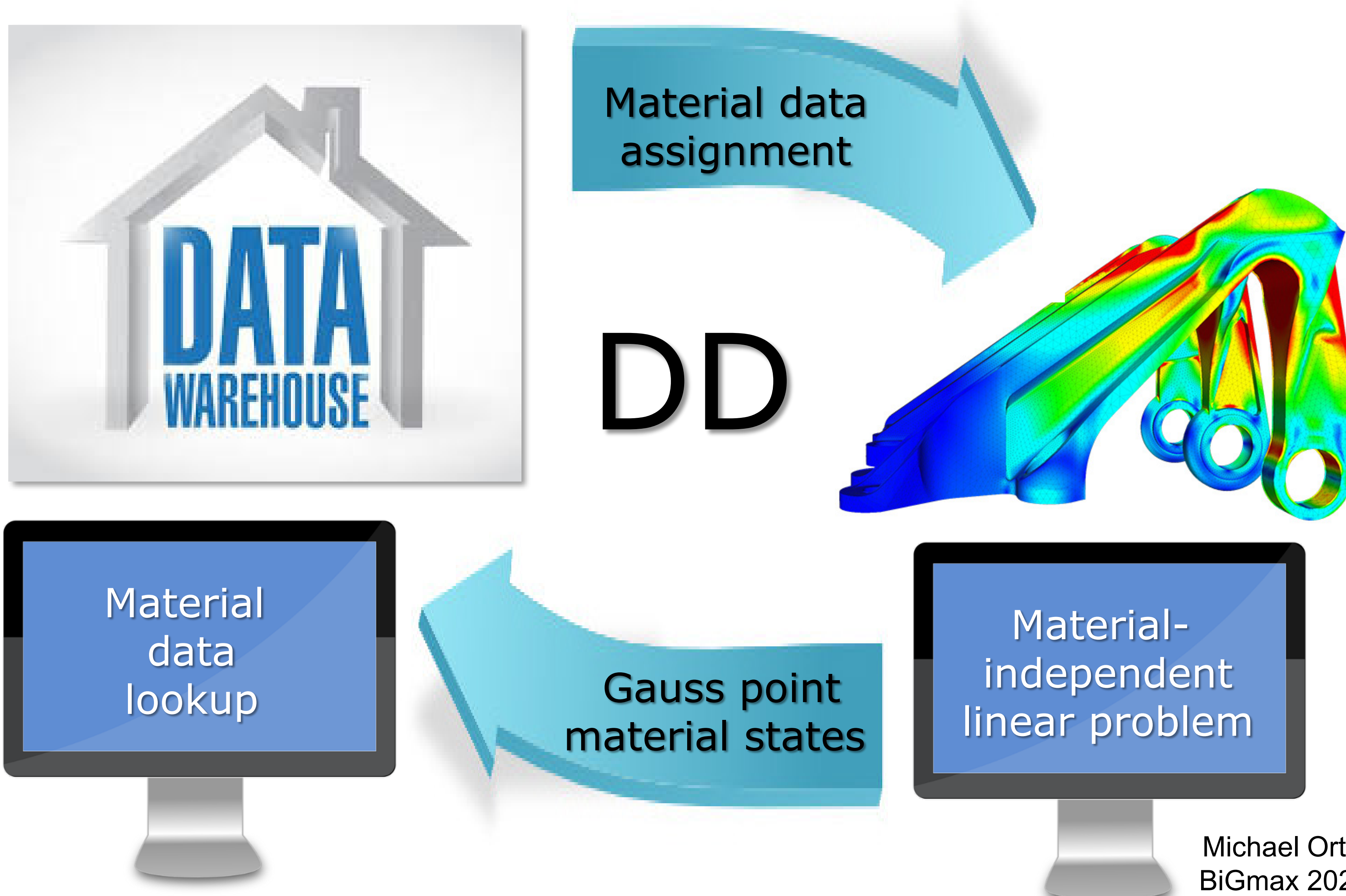
- (Model-free) Data-Driven solvers can be easily implemented within a finite-element framework



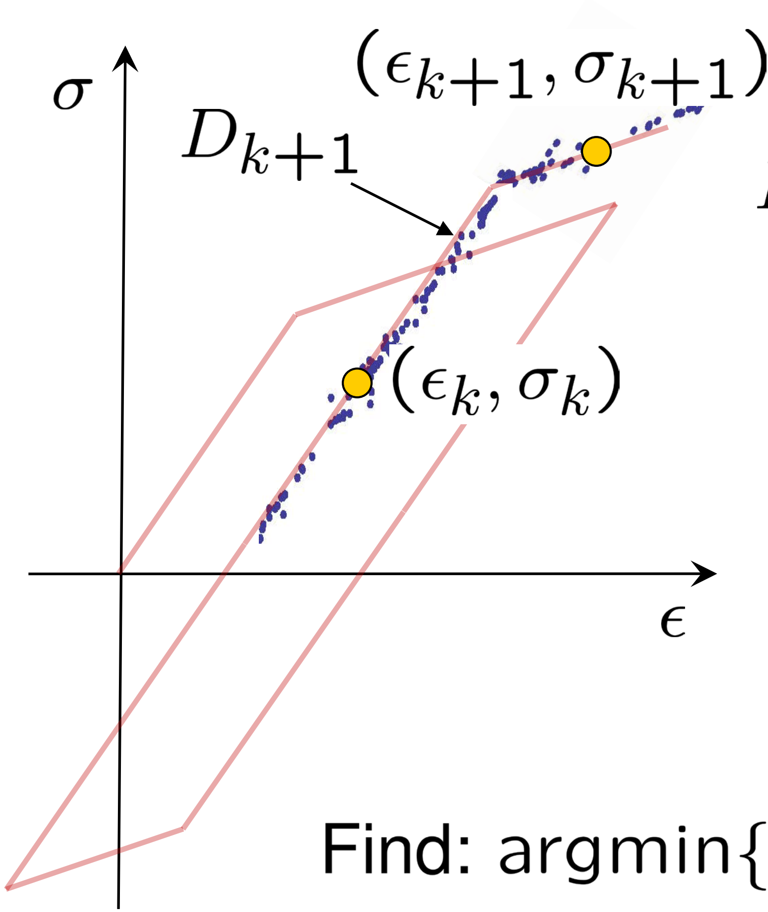
(R. Eggersmann & S. Reese, 2019)

Institute of Applied Mechanics
RWTH AACHEN UNIVERSITY

The DD information flow



Data-Driven inelasticity



- Material set representation:

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

- Need ***material history data***! (from material testing along selected loading paths…)
- History data must provide adequate ***path coverage***…
- Data-driven problem:

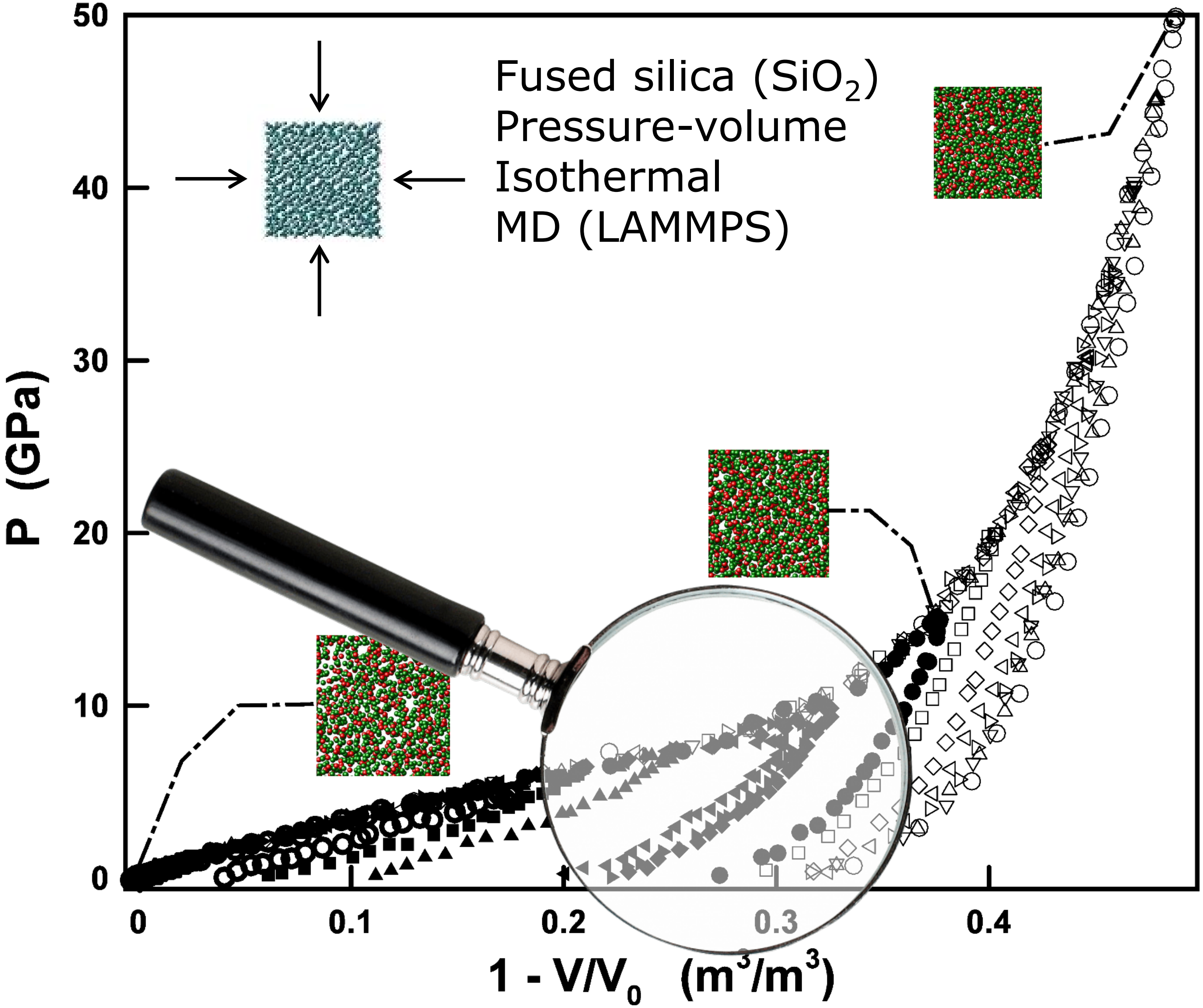
Find: $\text{argmin}\{d(z, D_{k+1}),\ z \in E_{k+1}\}$

time dependent!

R. Eggersmann, T. Kirchdoerfer, L. Stainier, S. Reese and M. Ortiz, *CMAME*, **350** (2019) 81-99.

Michael Ortiz
BiGmax 2021

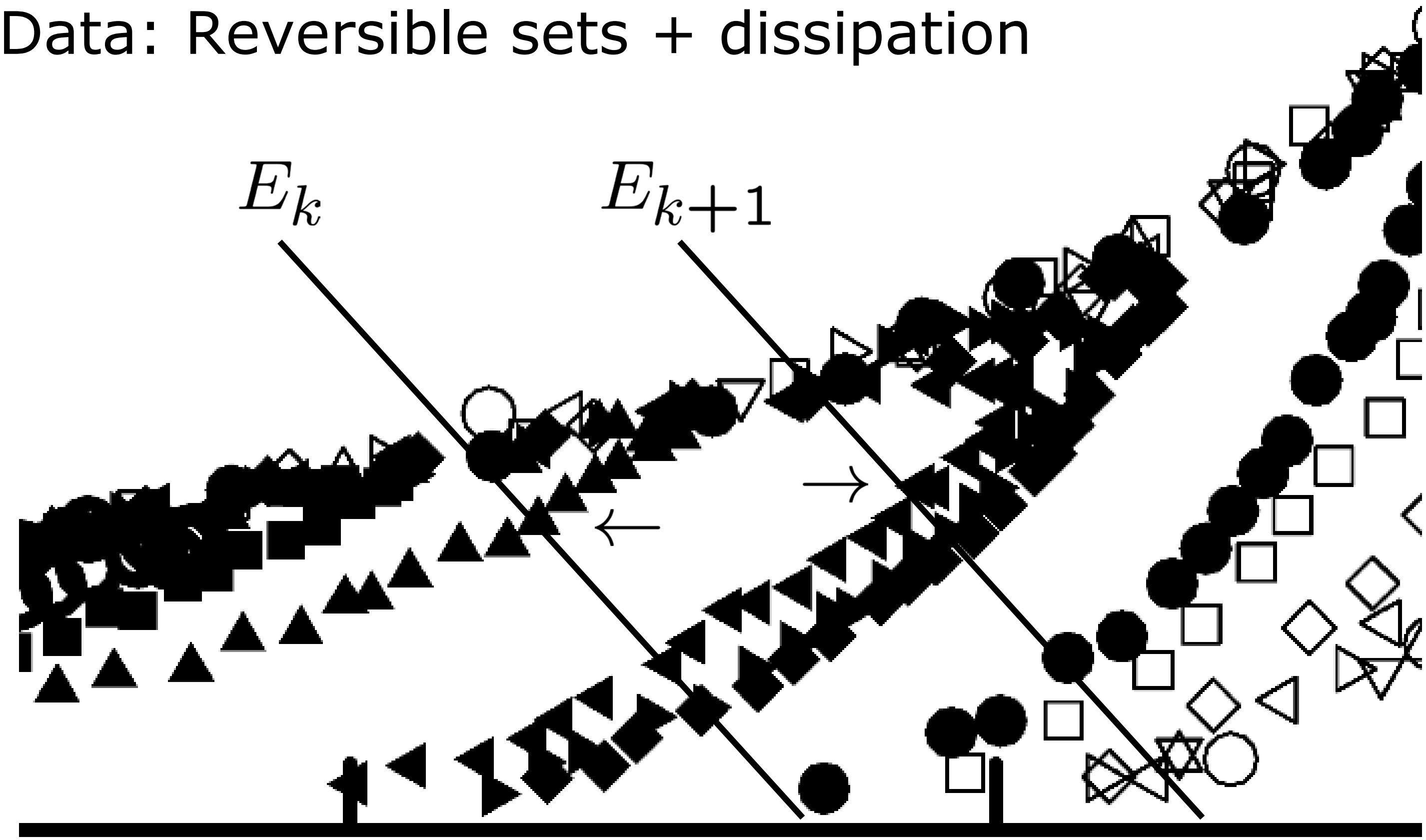
Data-Driven inelasticity



ael Ortiz
ıax 2021

Data-Driven inelasticity

- Phase space: Add temperature and entropy
- Field eqs: Add energy conservation + second law
- Data: Reversible sets + dissipation



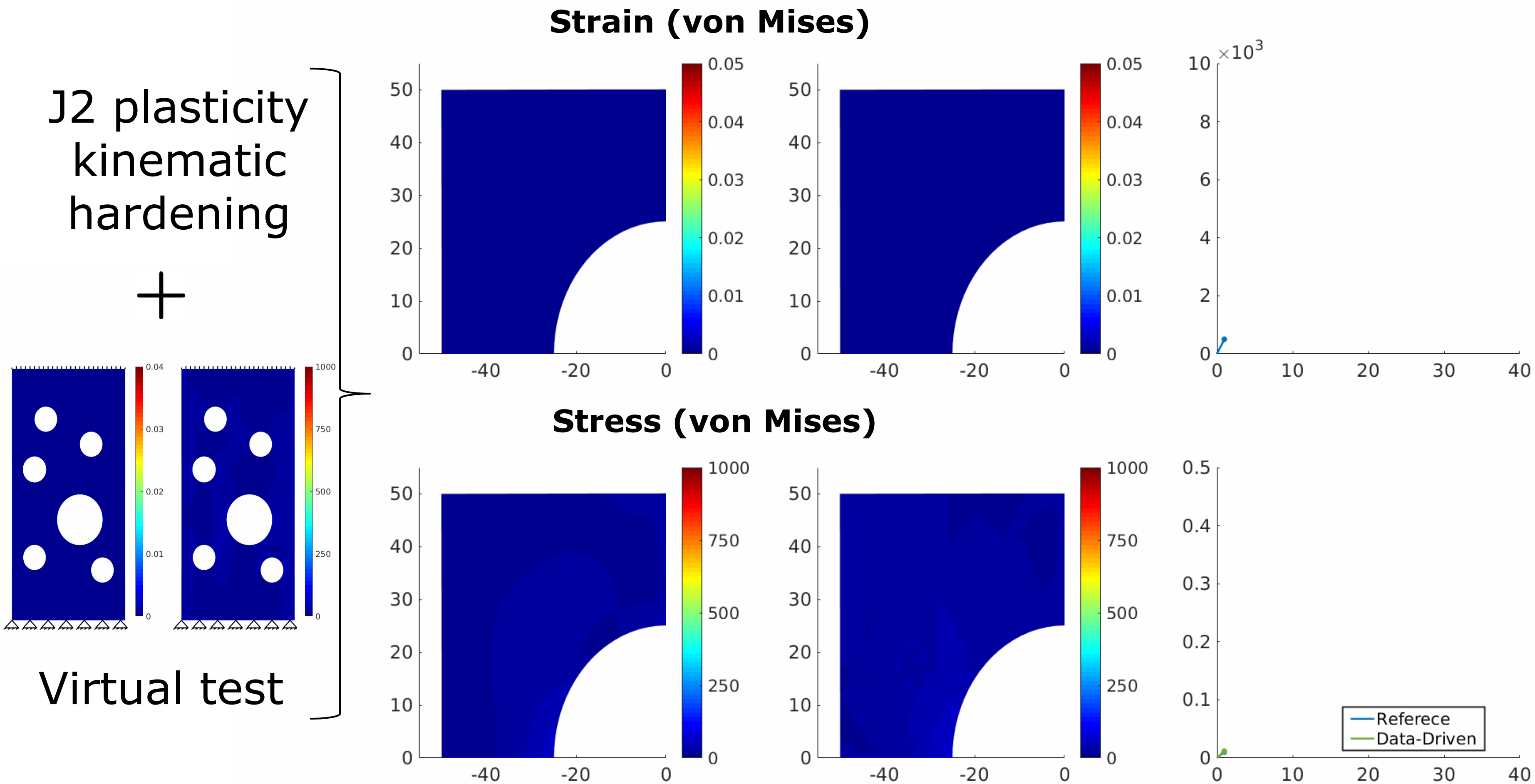
Michael Ortiz
BiGmax 2021

Data-Driven plasticity

J2 plasticity
kinematic
hardening

+

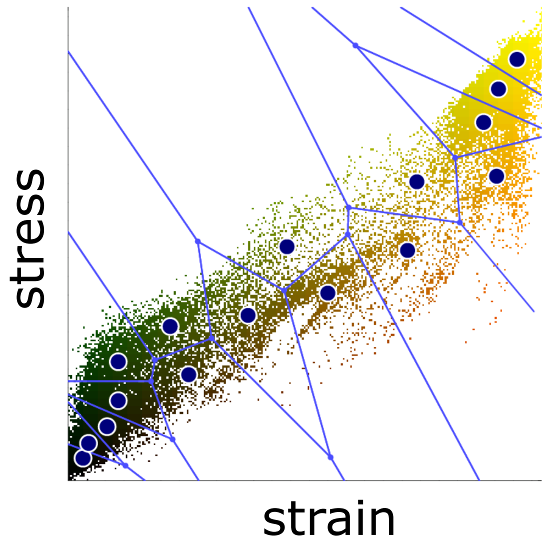
Virtual test

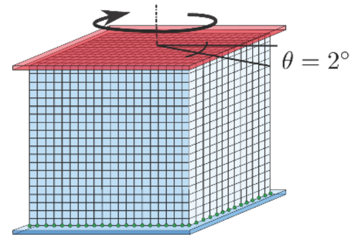


(R. Eggersmann & S. Reese, 2019)

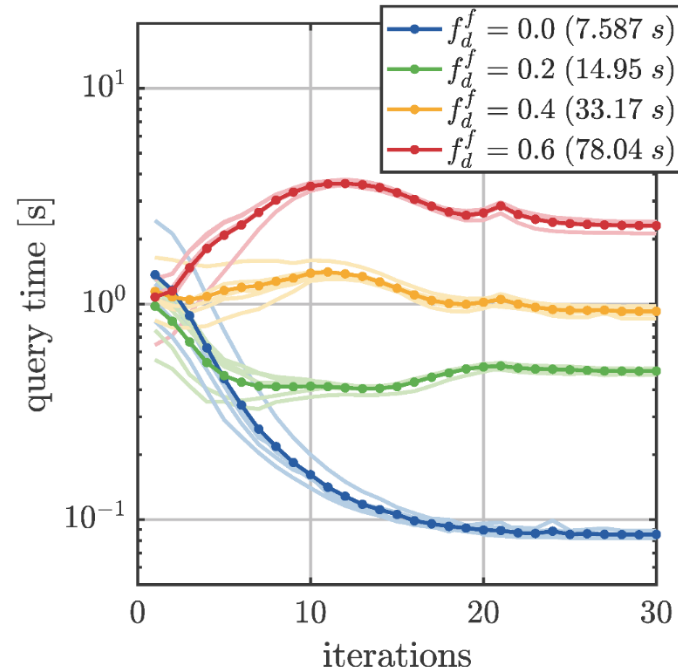
Michael Ortiz
BiGmax 2021

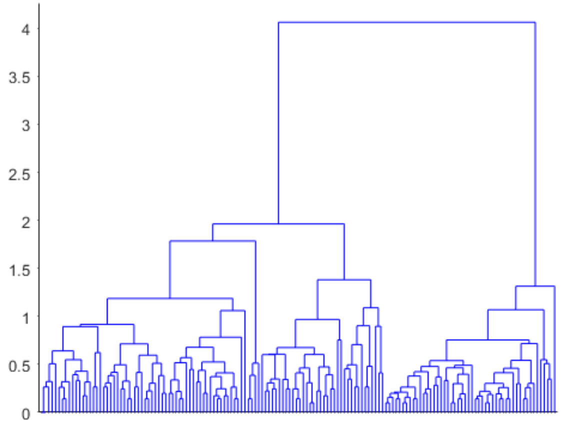
Connection to Machine Learning





Test problem:
Torsion of
20x20x20 cube
64000 mat pts





K-means hierarchical structure

- Material data set: *1 billion points*
- Approx k-means search, 0.1 secs
- *Set-oriented machine learning!*
- We learn the *structure of the data set*
- No regression, *no loss of information!*
- *The data, all the data, nothing but the data!*

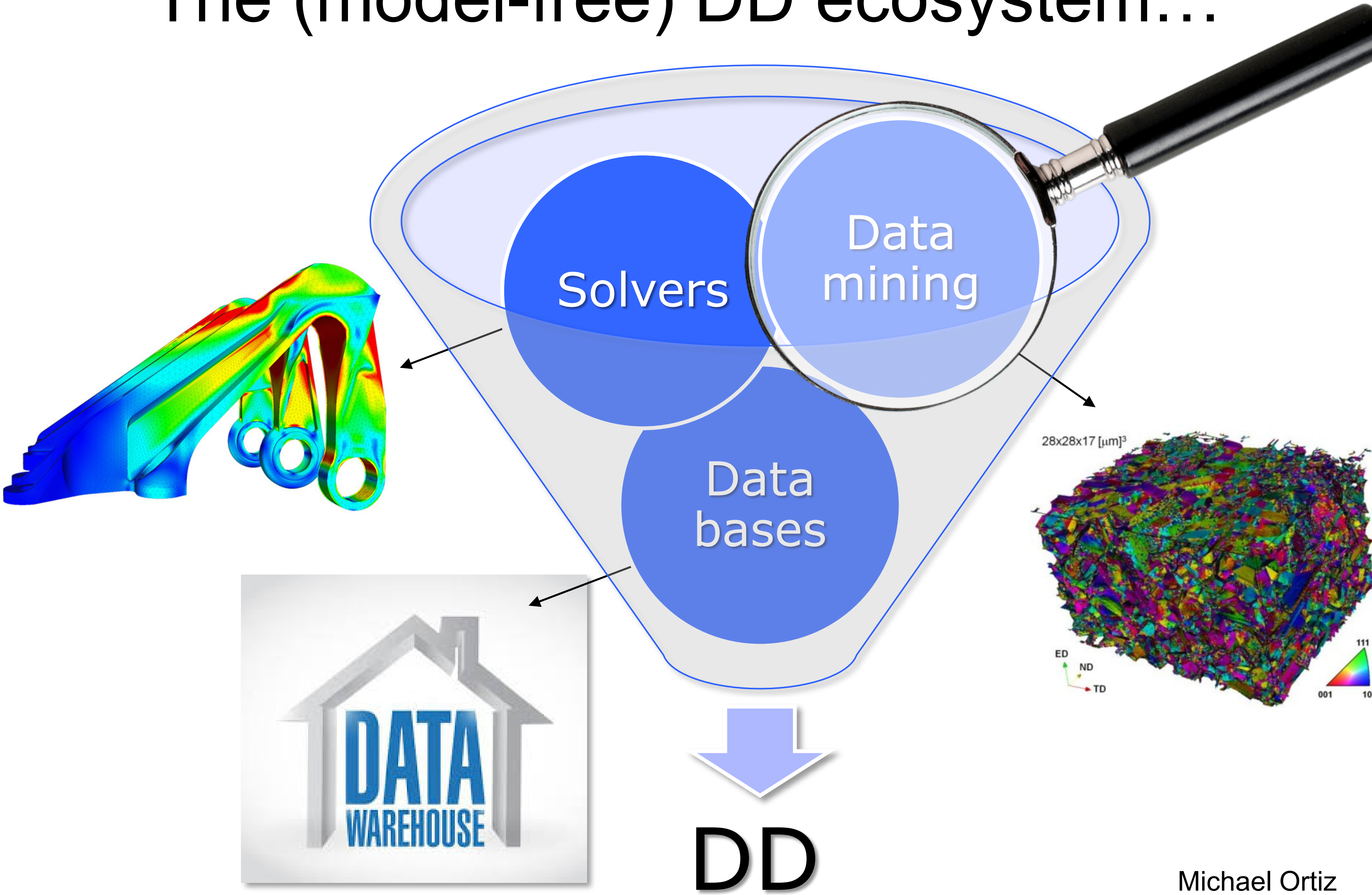
Michael Ortiz
BiGmax 2021

Model-Free Data-Driven (DD) solvers

- Model-Free Data-Driven solvers forge a *direct connection* between material data and prediction
- *No material modeling, no loss of information (the data, all the data, nothing but the data)*
- *Standardization of solvers:* Material-independent linear back-substitutions (geometry+loading FE)
- *Standardization of data:* Fundamental, model-independent data (e.g., stress and strain)
- *Lossless, set-oriented Machine Learning*: graph structure of data sets (e.g., hierarchical clustering, NN trees/graphs…)

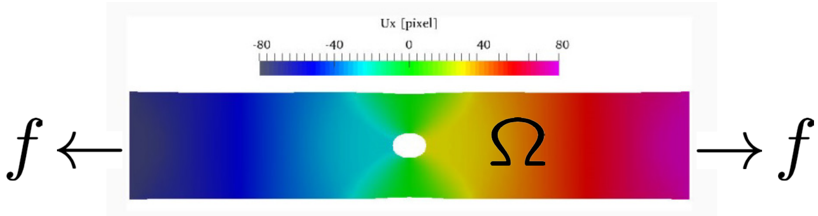
[1]T. Kirchdoerfer and M. Ortiz (2015) arXiv:1510.04232.
[1]T. Kirchdoerfer and M. Ortiz, *CMAME*, **304** (2016) 81–101

The (model-free) DD ecosystem…



Michael Ortiz
BiGmax 2021

DD material identification (DDMI)



- Full-field measurements (DIC): $(\epsilon'(x), f(x))$
- *Stress field cannot be measured directly!*
- *Density of states* in phase space:

$$\rho(\epsilon, \sigma) = \frac{1}{|\Omega|} \int_\Omega \delta(\epsilon - \epsilon'(x), \sigma - \sigma'(x))\, dx$$

- Maximize *information-theoretical entropy*: $S(\rho)$

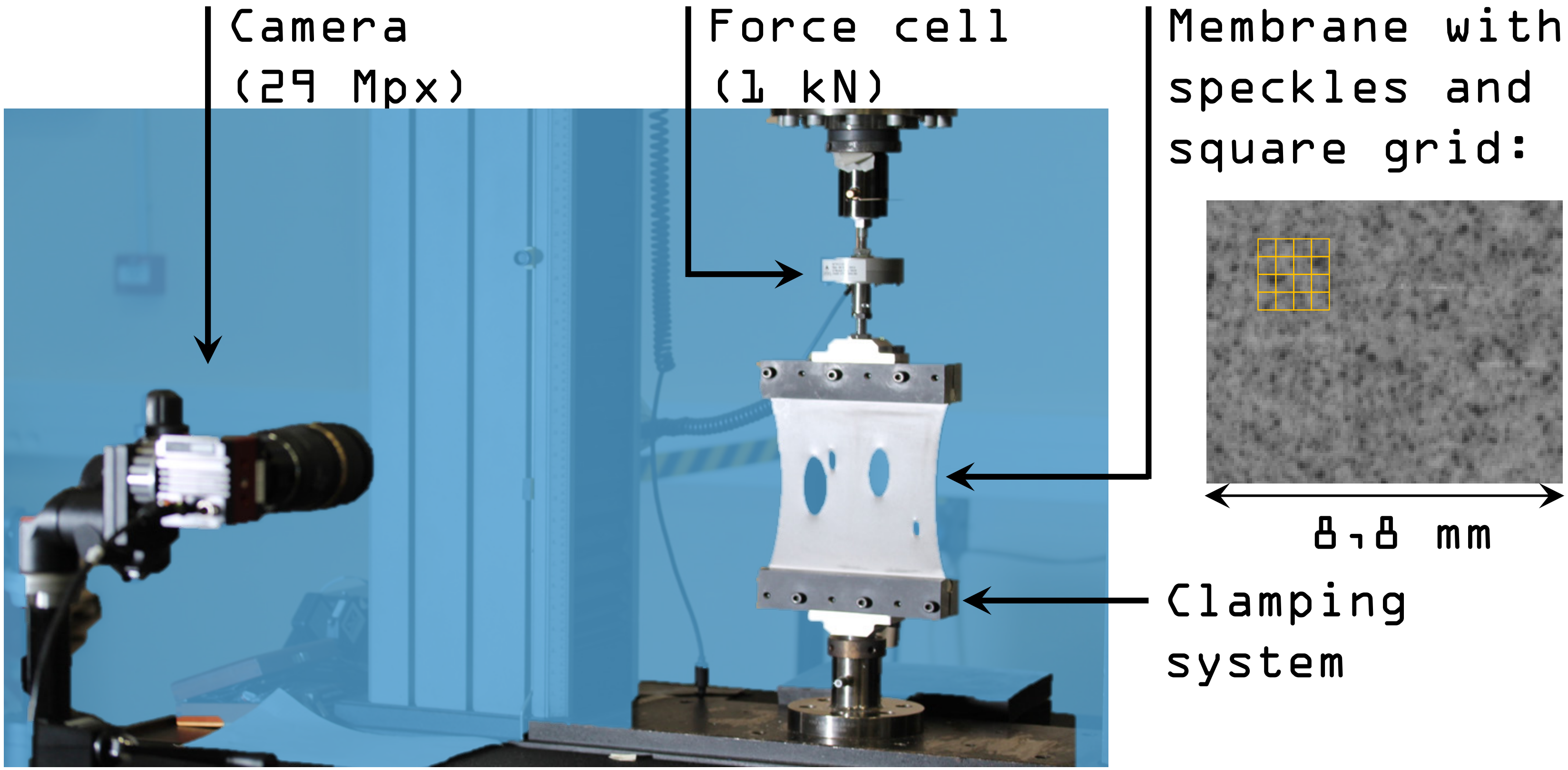
  subject to *equilibrium*: $\text{div}\, \sigma'(x) + f(x) = 0$

- Solve (inverse) problem using *DD k-means*

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

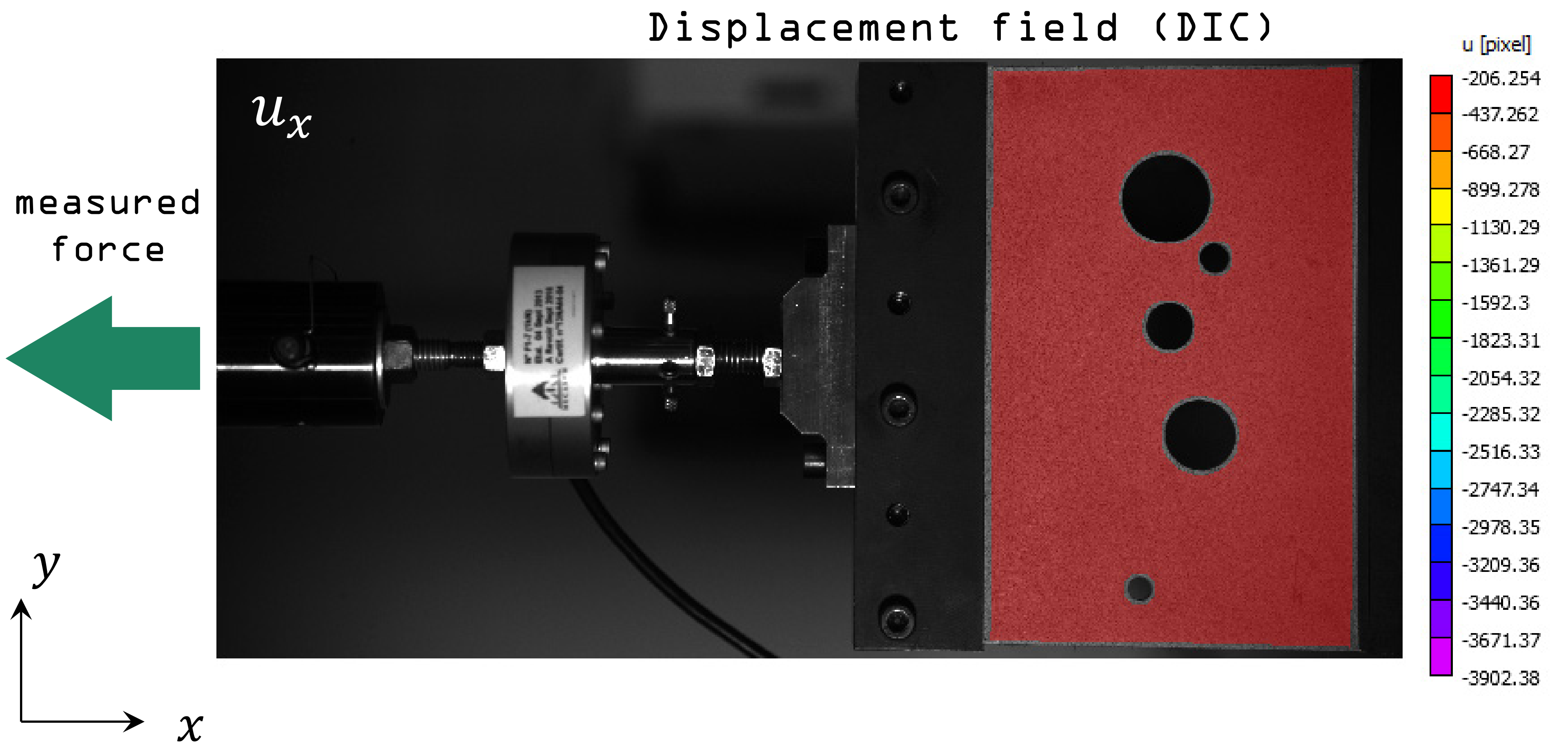
DDMI at Central Nantes

- Tests on perforated silicone sheets



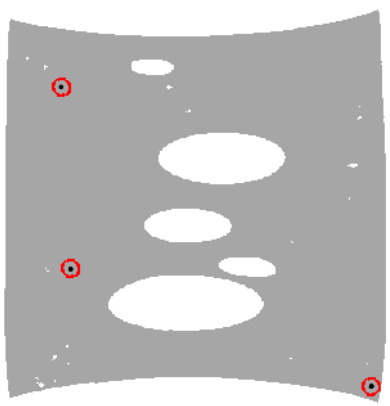
M. Dalémat, M. Coret, A. Leygue and E. Verron, *Mechanics of Materials*, **136** (2019) 103087.

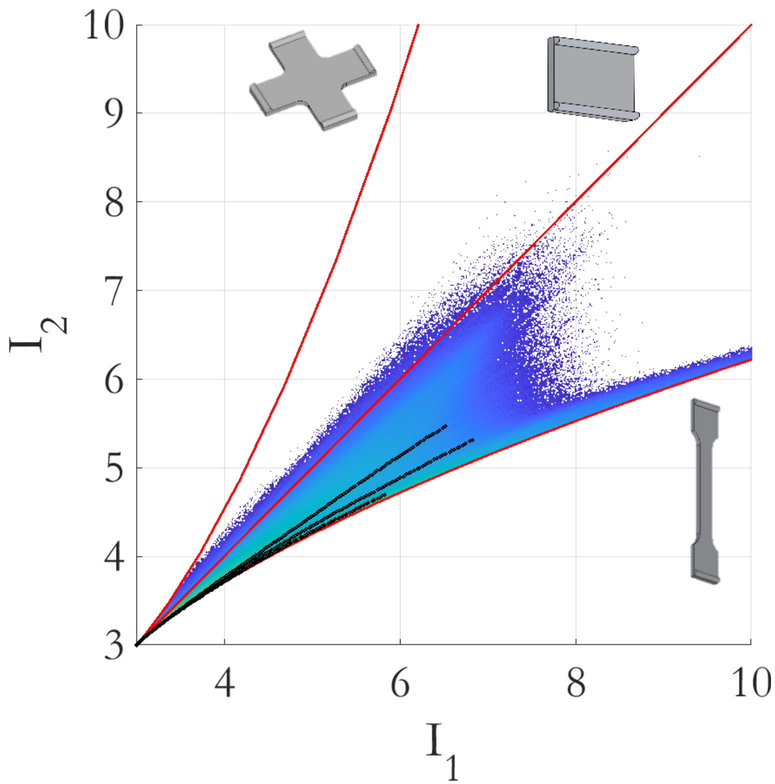
DDMI at Central Nantes



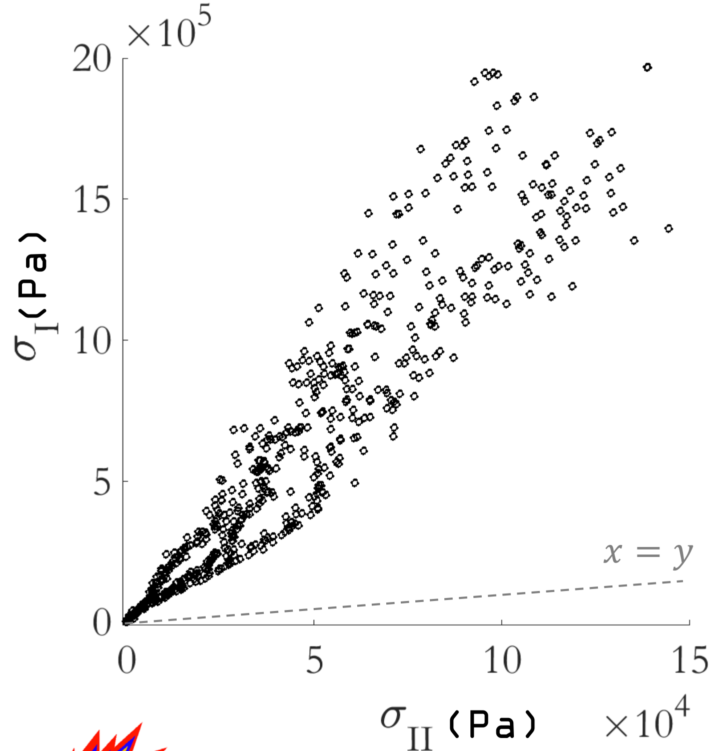
M. Dalémat, M. Coret, A. Leygue and E. Verron, *Mechanics of Materials*, **136** (2019) 103087.

DDMI at Central Nantes





Strain data

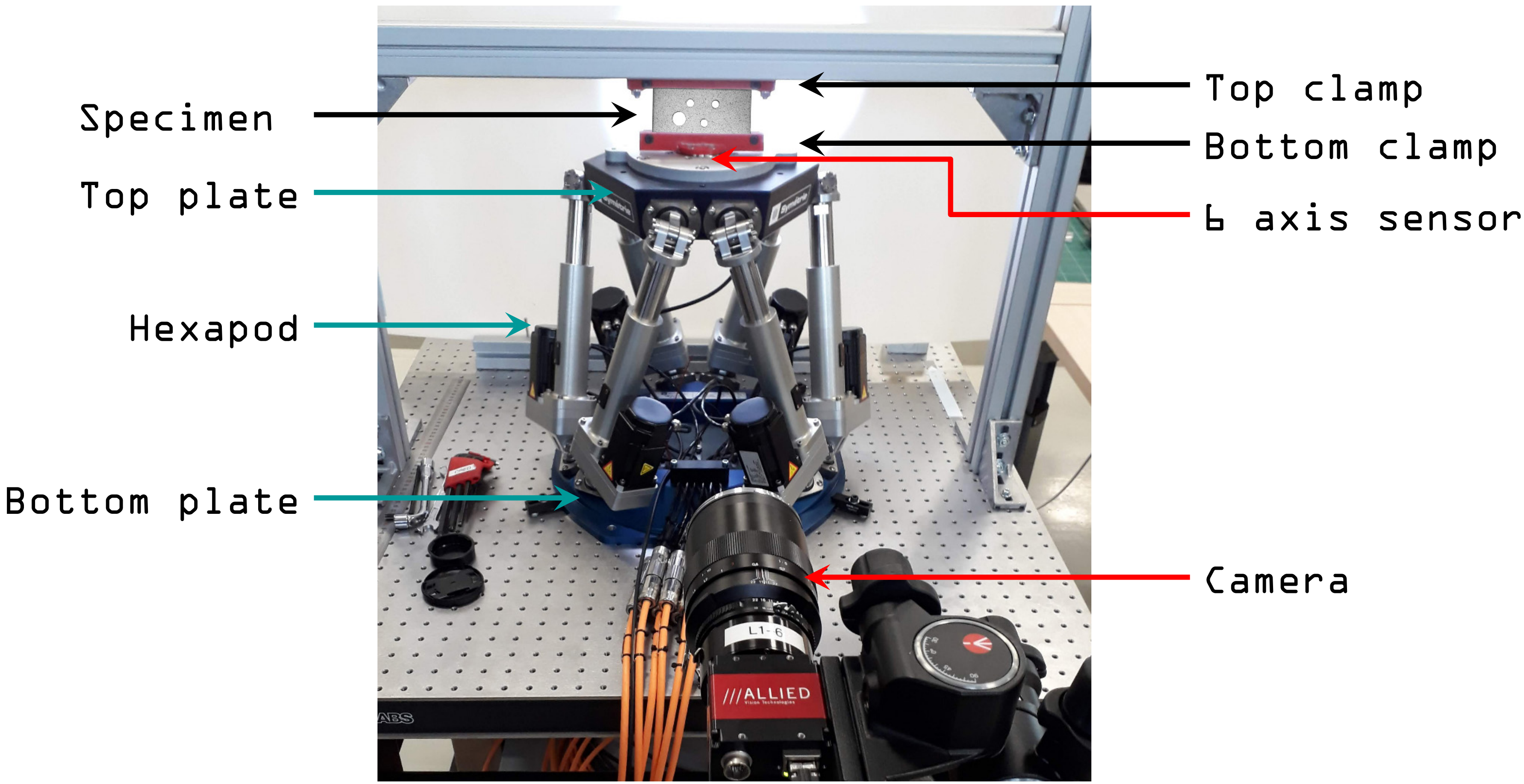


Stress data

M. Dalémat, M. Coret, A. Leygue and E. Verron, *Mechanics of Materials*, **136** (2019) 103087.

DDMI at Central Nantes

- Hexapod testing machine



M. Dalémat, M. Coret, A. Leygue and E. Verron, *Mechanics of Materials*, **136** (2019) 103087.

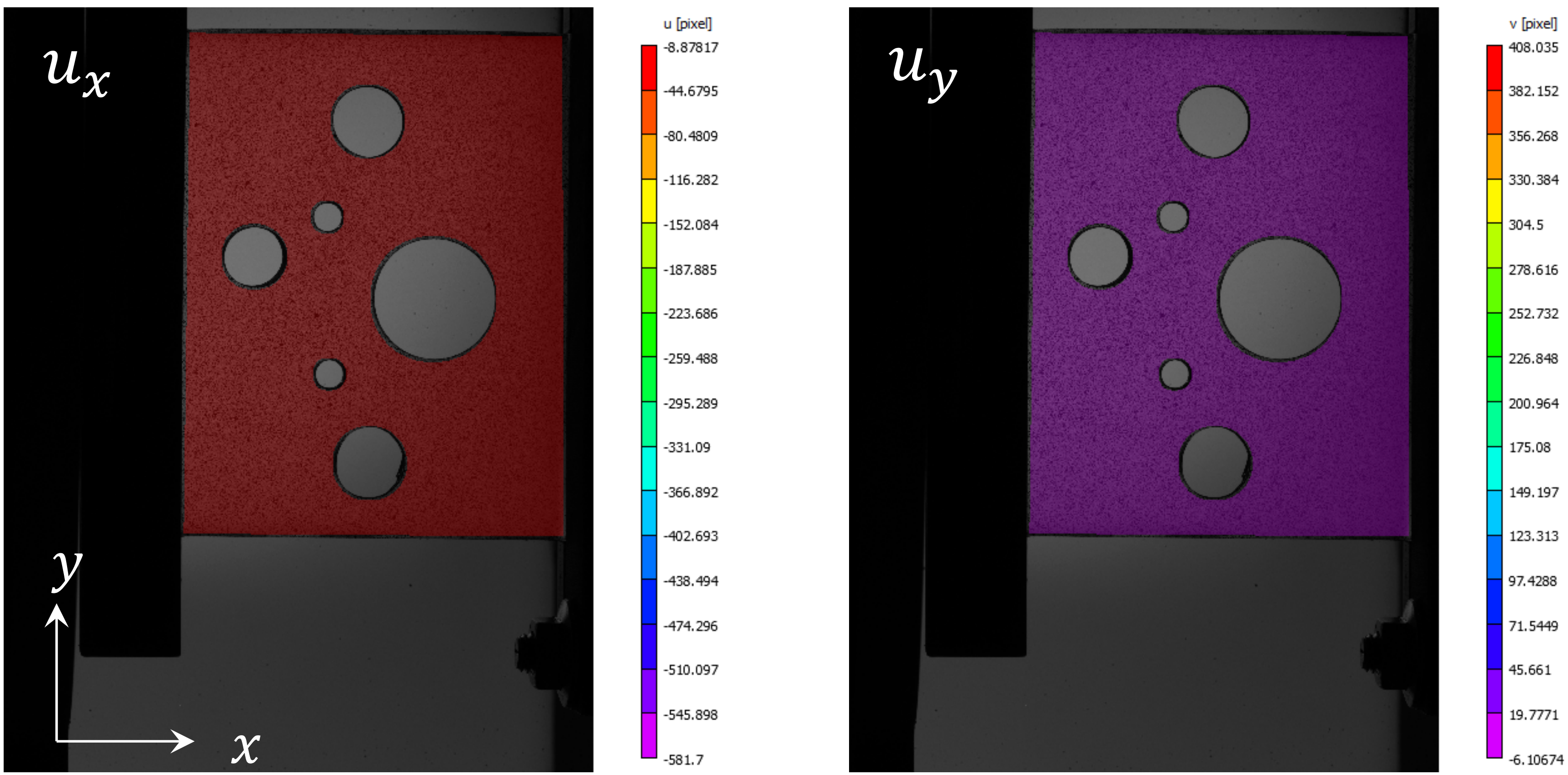
DDMI at Central Nantes

Hexapod in operation

M. Dalémat, M. Coret, A. Leygue and E. Verron, *Mechanics of Materials*, **136** (2019) 103087.

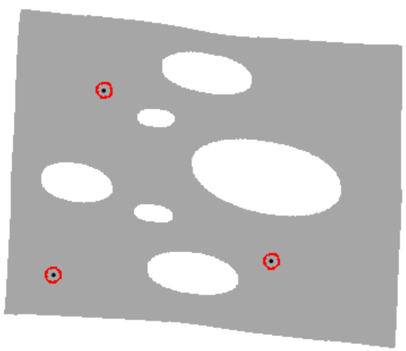
DDMI at Central Nantes

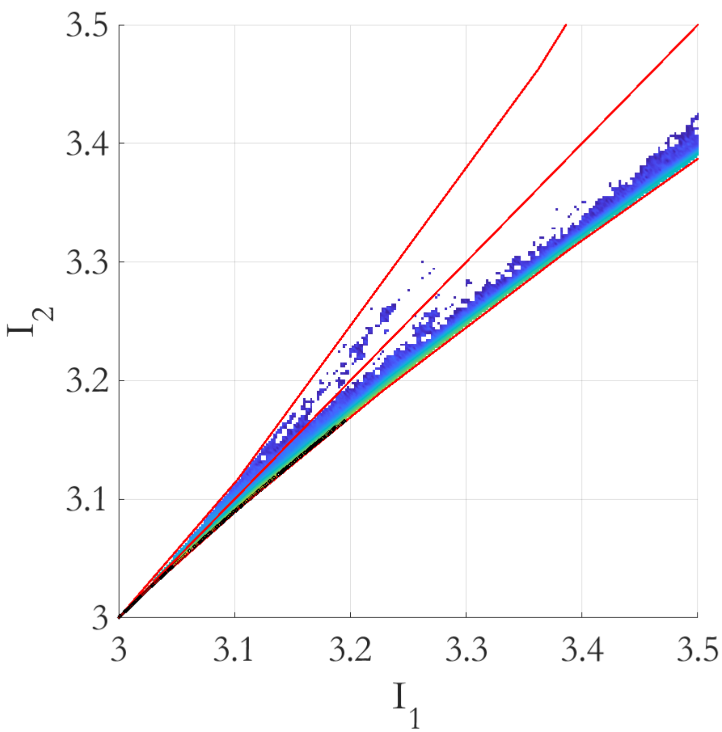
Displacement field (DIC)



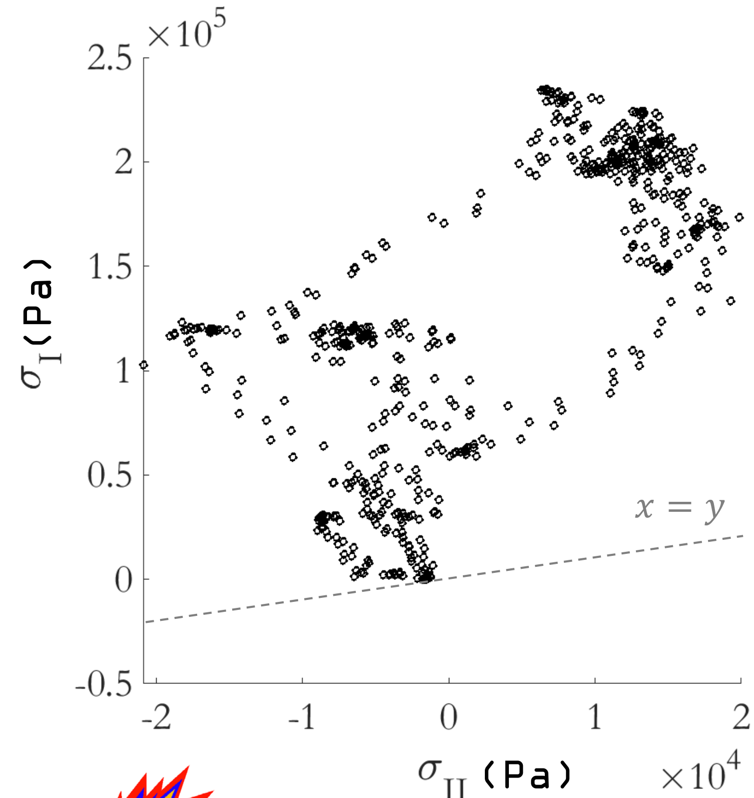
M. Dalémat, M. Coret, A. Leygue and E. Verron, *Mechanics of Materials*, **136** (2019) 103087.

DDMI at Central Nantes





Strain data

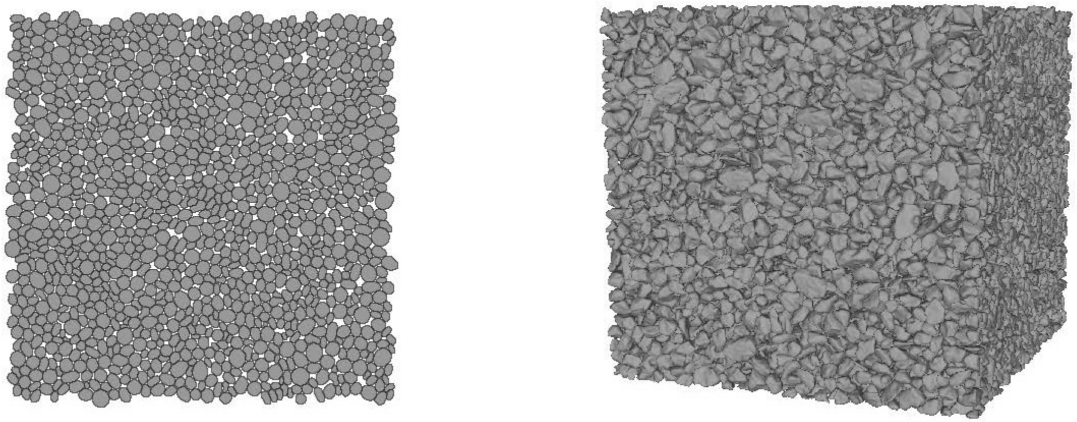


Stress data

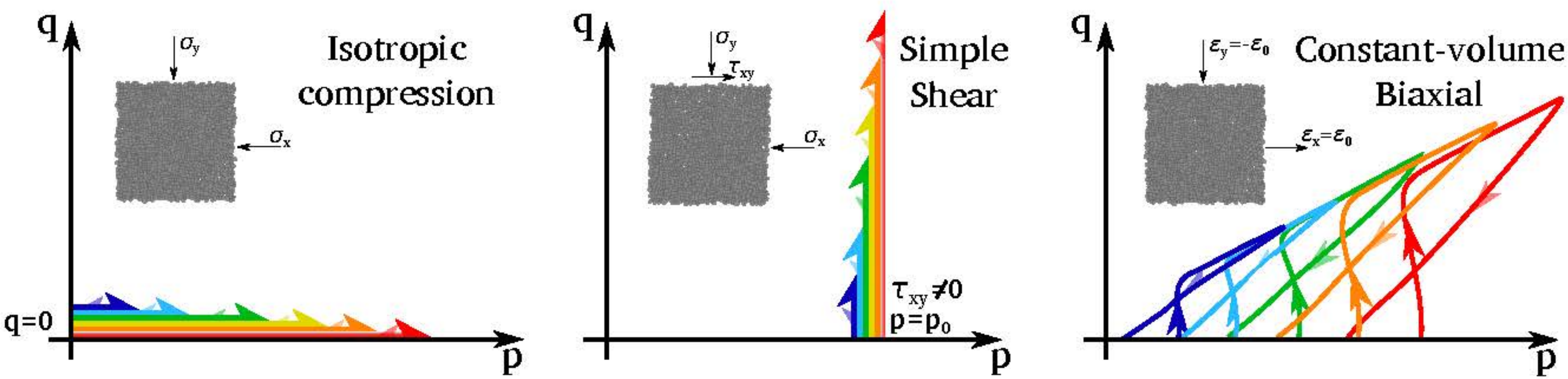
M. Dalémat, M. Coret, A. Leygue and E. Verron,
Mechanics of Materials, **136** (2019) 103087.

Michael Ortiz
BiGmax 2021

Multiscale data mining



2D and 3D RVE and granular assemblies for LS-DEM analysis

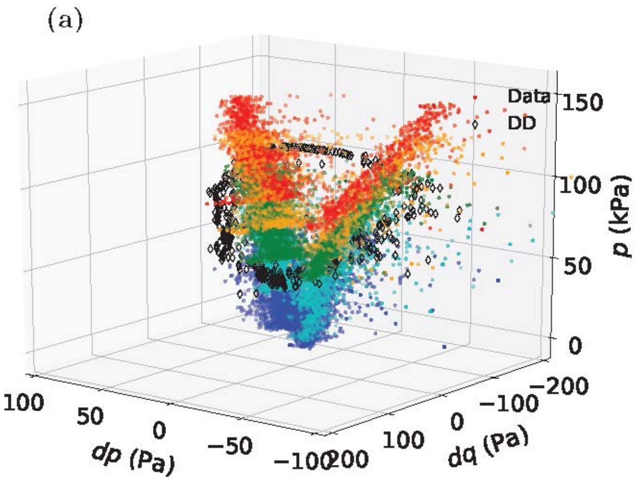


Stress paths for material point simulations and data mining

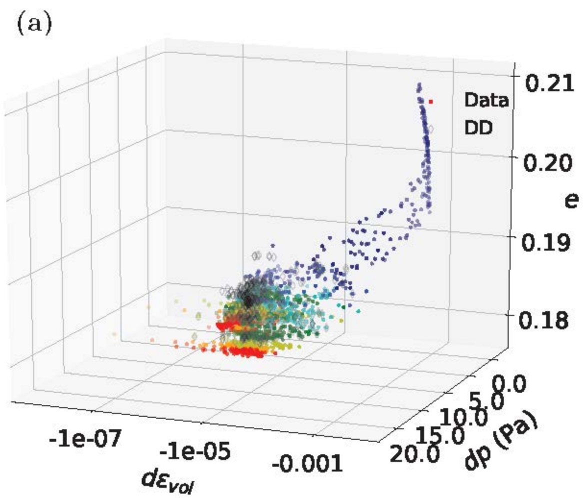
K. Karapiperis, L. Stainier, M. Ortiz, J.E. Andrade, *JMPS*, (2020) 104239.

Michael Ortiz
BiGmax 2021

Multiscale data mining

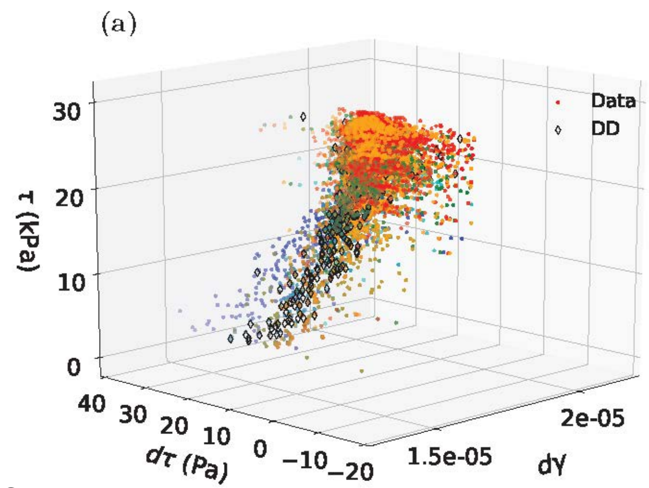


Isotropic compression



Biaxial compression

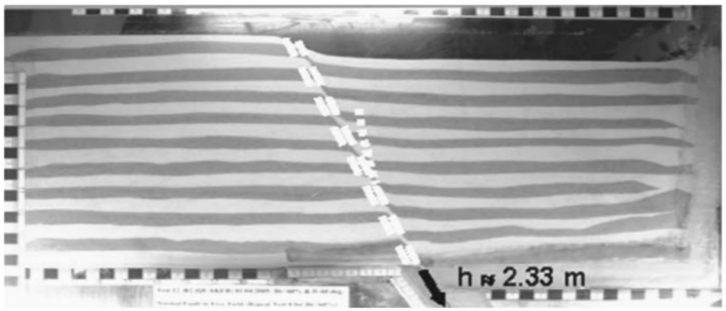
Simple shear



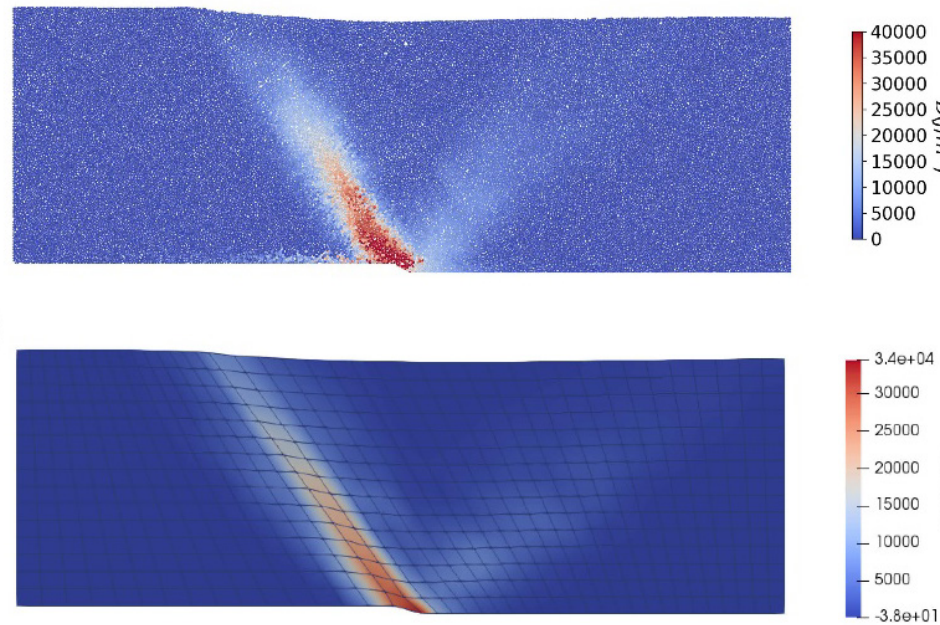
Data computed from RVE LS-DEM calculations

K. Karapiperis, L. Stainier, M. Ortiz, J.E. Andrade, *JMPS*, (2020) 104239.

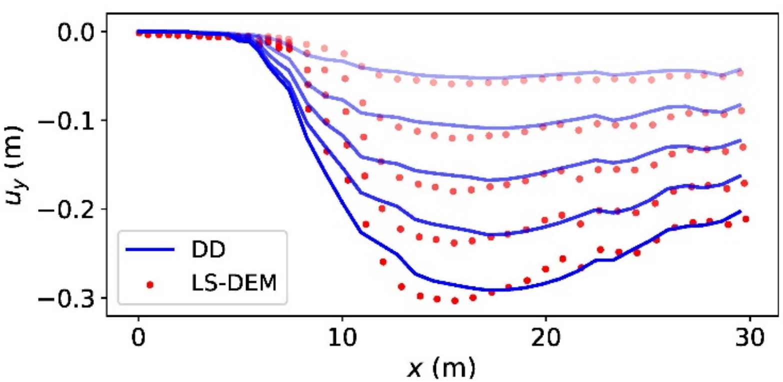
(Model-free) Data-Driven – Sand



Experimental fault rupture experiment (Anastapoulos etal.. 2007)



Top: LS-DEM simulation
Bottom: DD simulation



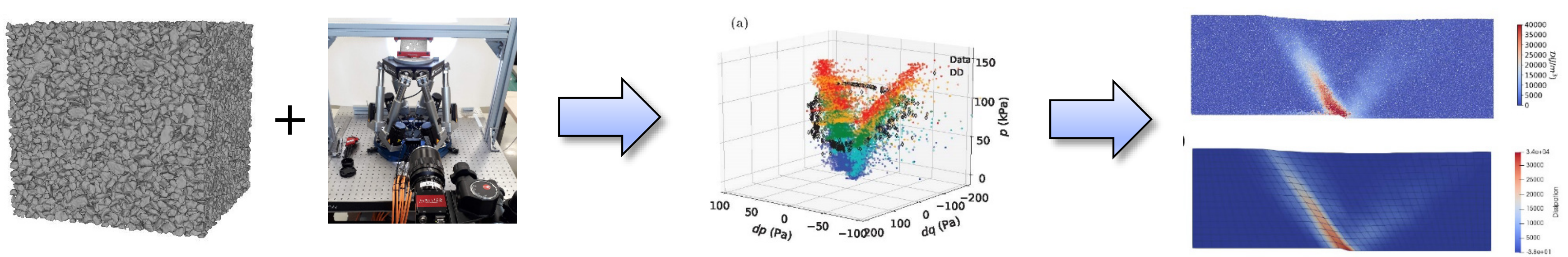
Evolution of surface settlement, LS-DEM vs. DD simulations

K. Karapiperis, L. Stainier, M. Ortiz, J.E. Andrade, *JMPS*, (2020) 104239.

Michael Ortiz
BiGmax 2021

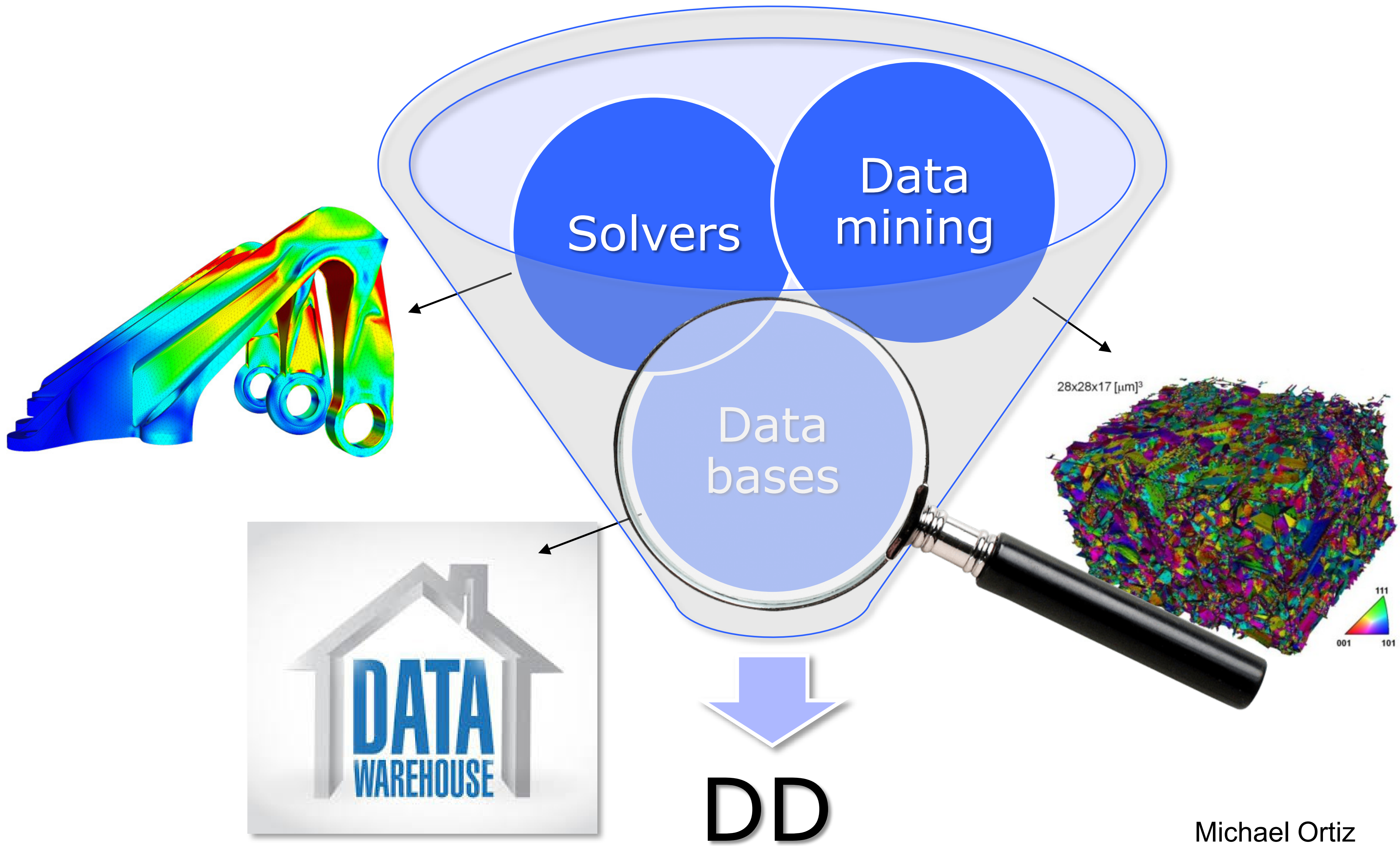
Data mining, generation, upscaling

- DDMI can extract data from *full-field experimental microscopy* (TEM, SEM, DIC, EBSD…)
- DDMI generates *fundamental data without prior assumptions* on the form of constitutive response
- Data can be *mined* from lower-scale calculations, used in upper-scale calculations (*DD upscaling*)
- *DD sets forth new opportunities for synergism between experimental science and scientific computing, new multiscale analysis paradigms*



J. Rethore and A. Leygue, HAL Id: hal-01452494, Feb. 2017.
K. Karapiperis, L. Stainier, M. Ortiz, J.E. Andrade, *JMPS*, (2020) 104239.

The (model-free) DD ecosystem…



Michael Ortiz
BiGmax 2021

Publically-editable data repositories

- Reliance on *fundamental data* (stress and strain only, no model-dependent data) makes *material data fungible*, mergeable, interchangeable…
- *Publically editable repositories* have proven exceptional capacity for *organic growth*…
- *Publicly-editable material data repository! (Wikimat)*
  – *Fundamental, model-independent data:*
    - *Stress-strain*
    - *Temperature-entropy*
    - *Gradients, rates*
  – *Tools/Scripts for interfacing with commercial FE packages*

WIKIPEDIA
The Free Encyclopedia

Concluding remarks

- *Model-Free Data-Driven computing: The data, all the data, nothing but the data!*
- *The Model-Free Data-Driven ecosystem*:
  - Solver standardization (material independent)
  - Data standardization (model-independent data)
  - Lossless, set-oriented Machine Learning
  - Data identification from full-field microscopy
  - Data identification from multiscale analysis
  - Data repositories (fungible, publically editable)
- *Data-driven computing is likely to be a growth area in an increasingly data-rich world and to change the way in which data is mined, stored, exchanged, disseminated and utilized in science and in industry!*

Concluding remarks

Thank you!