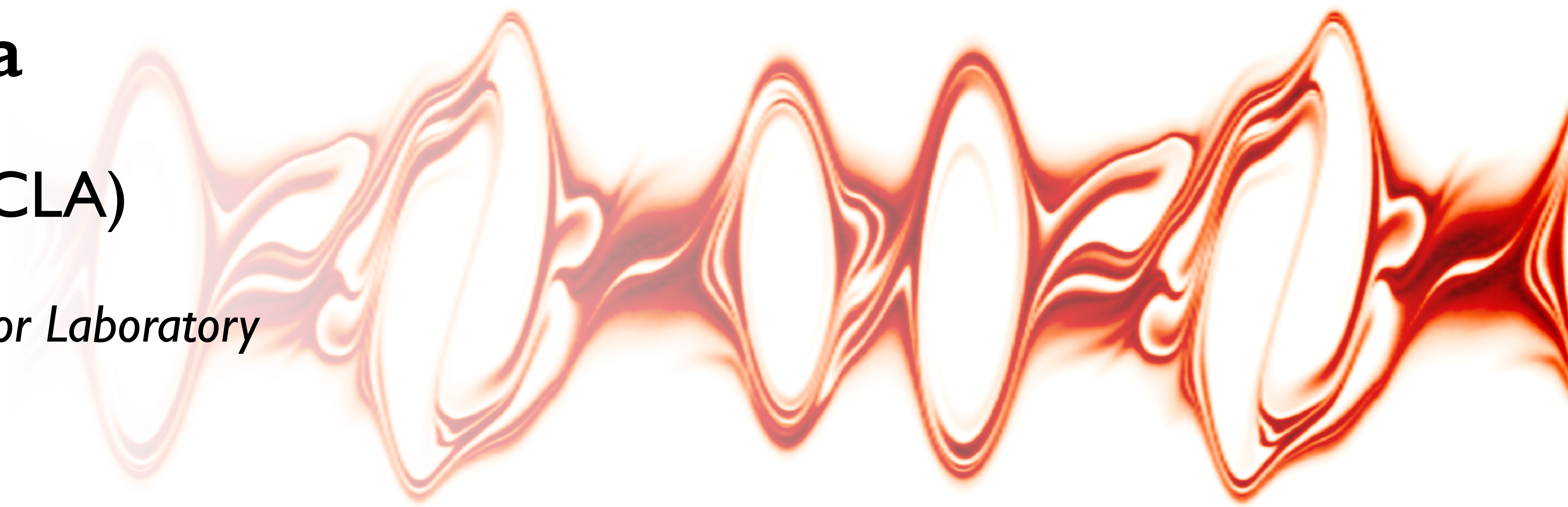
Accelerating the understanding of nonlinear dynamical systems using machine learning

Frederico Fiuza

with E. P. Alves (UCLA)

SLAC National Accelerator Laboratory
Stanford University
fiuza@slac.stanford.edu



Outline

- **Data-driven models of nonlinear dynamical systems**
  the importance of interpretability and generalizability

- **Discovering reduced plasma models from first-principles particle simulations**
  noisy data and the importance of an integral formulation of sparse regression

- **Recovering the hierarchy of plasma equations (kinetic to fluid) and fluid closures**
  prospects for improved fluid-kinetic closures and more efficient multi-scale algorithms

Outline

Data-driven models of nonlinear dynamical systems

the importance of interpretability and generalizability

Discovering reduced plasma models from first-principles particle simulations

noisy data and the importance of an integral formulation of sparse regression

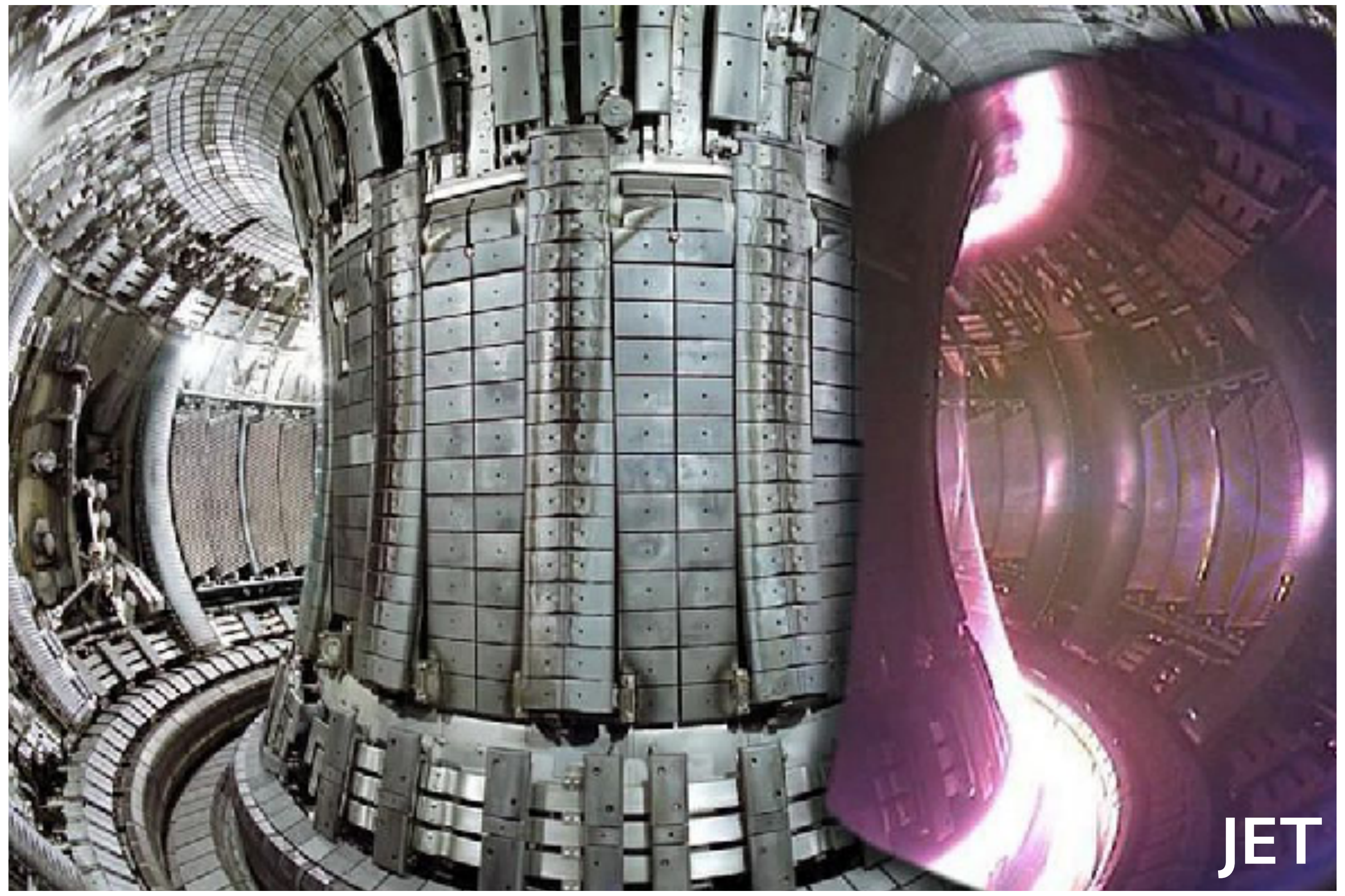
Recovering the hierarchy of plasma equations (kinetic to fluid) and fluid closures

prospects for improved fluid-kinetic closures and more efficient multi-scale algorithms

Machine learning is offering powerful new ways of harnessing wealth of scientific data

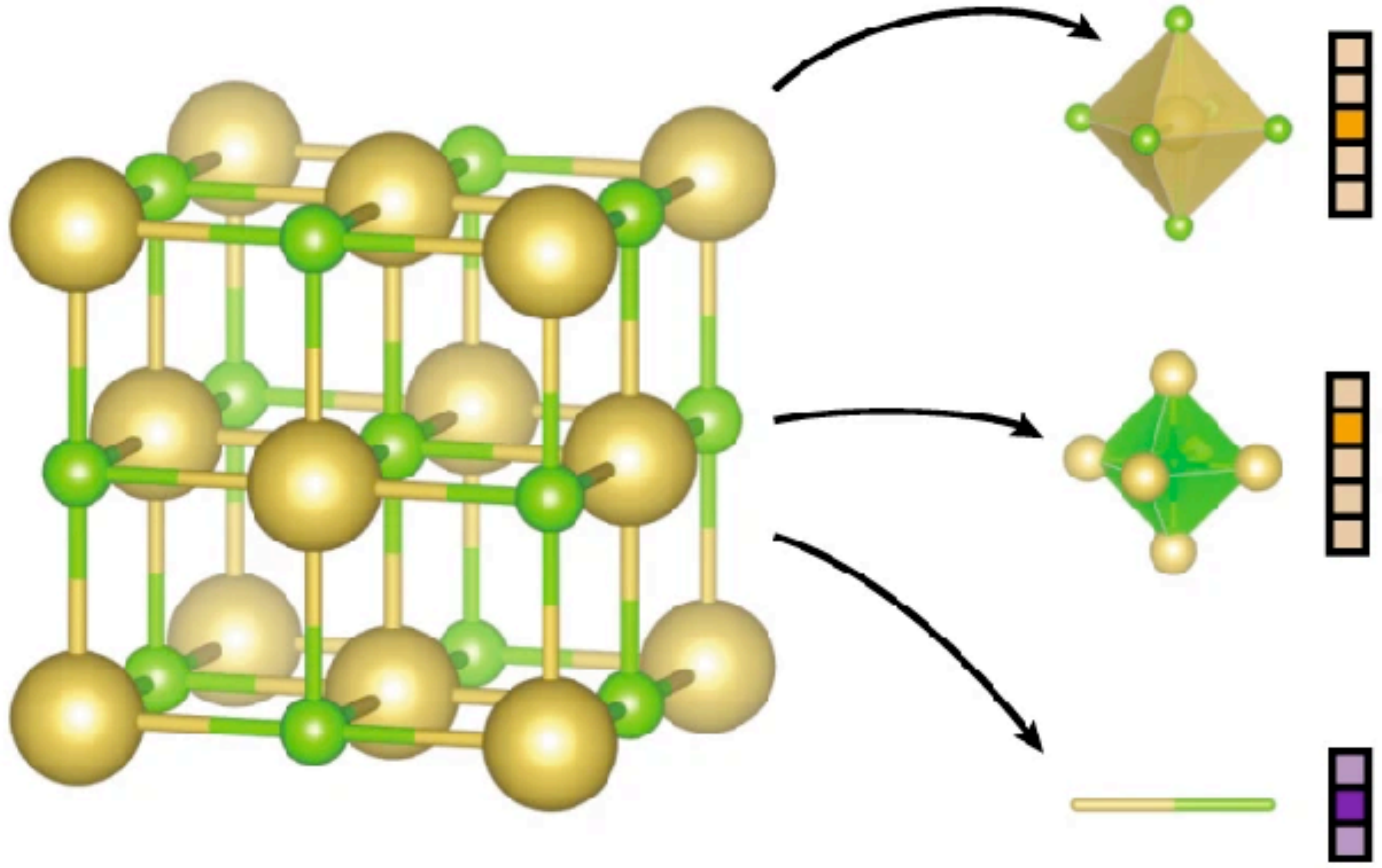
Deep learning is enabling the development of highly-predictive models from data

Disruption prediction of fusion plasmas



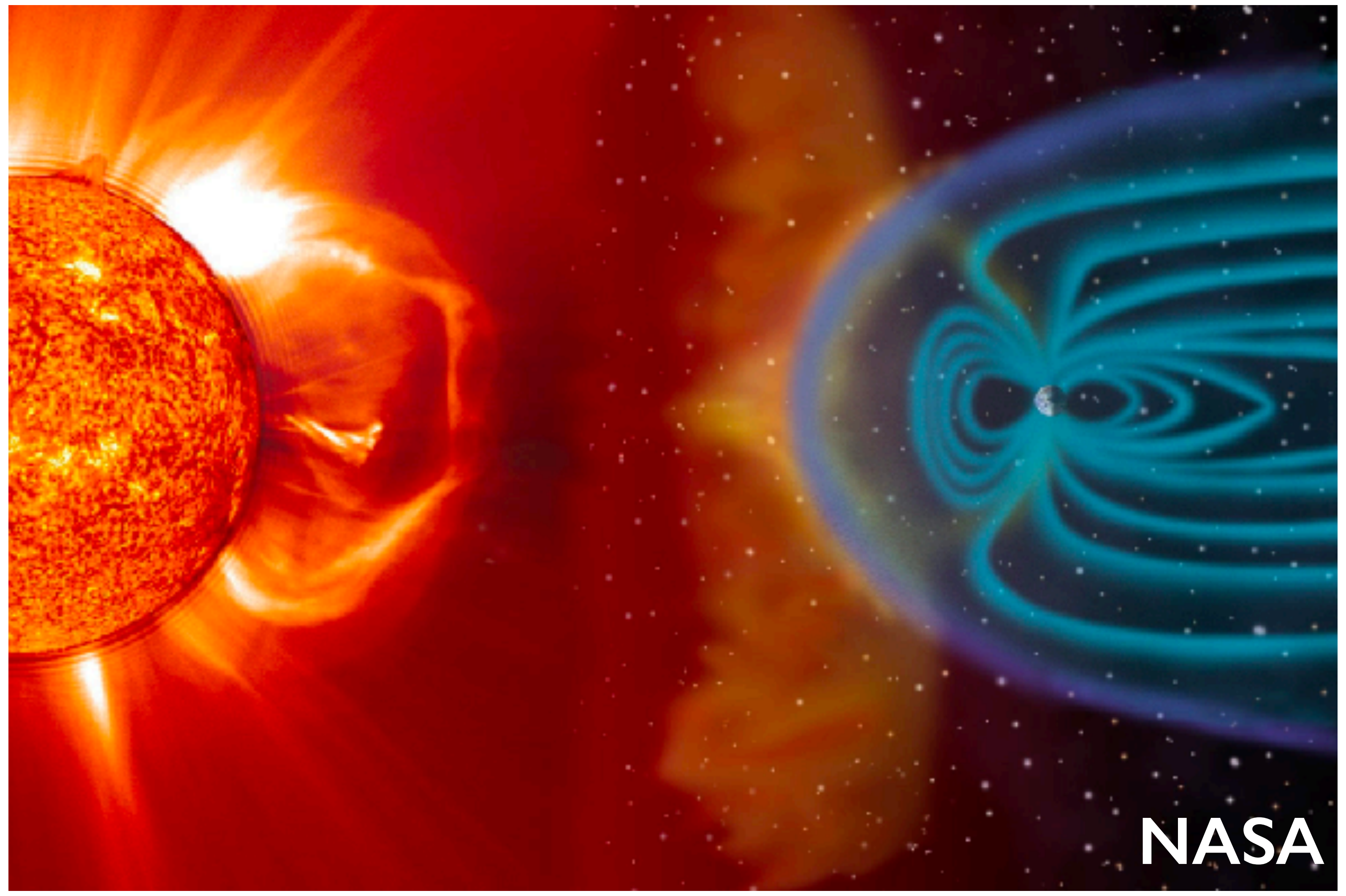
Kates-Harbeck et al. 2019
Vega et al. 2022

Discovery of stable materials



Schmidt et al. 2019

Space weather prediction



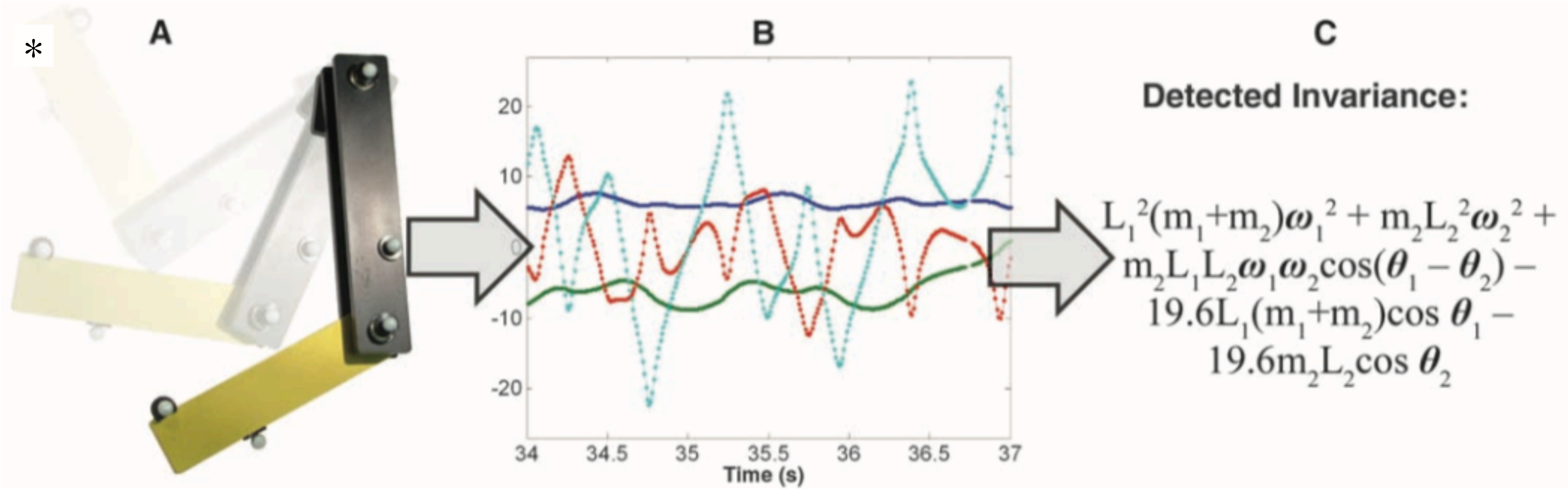
Camporeale et al. 2018

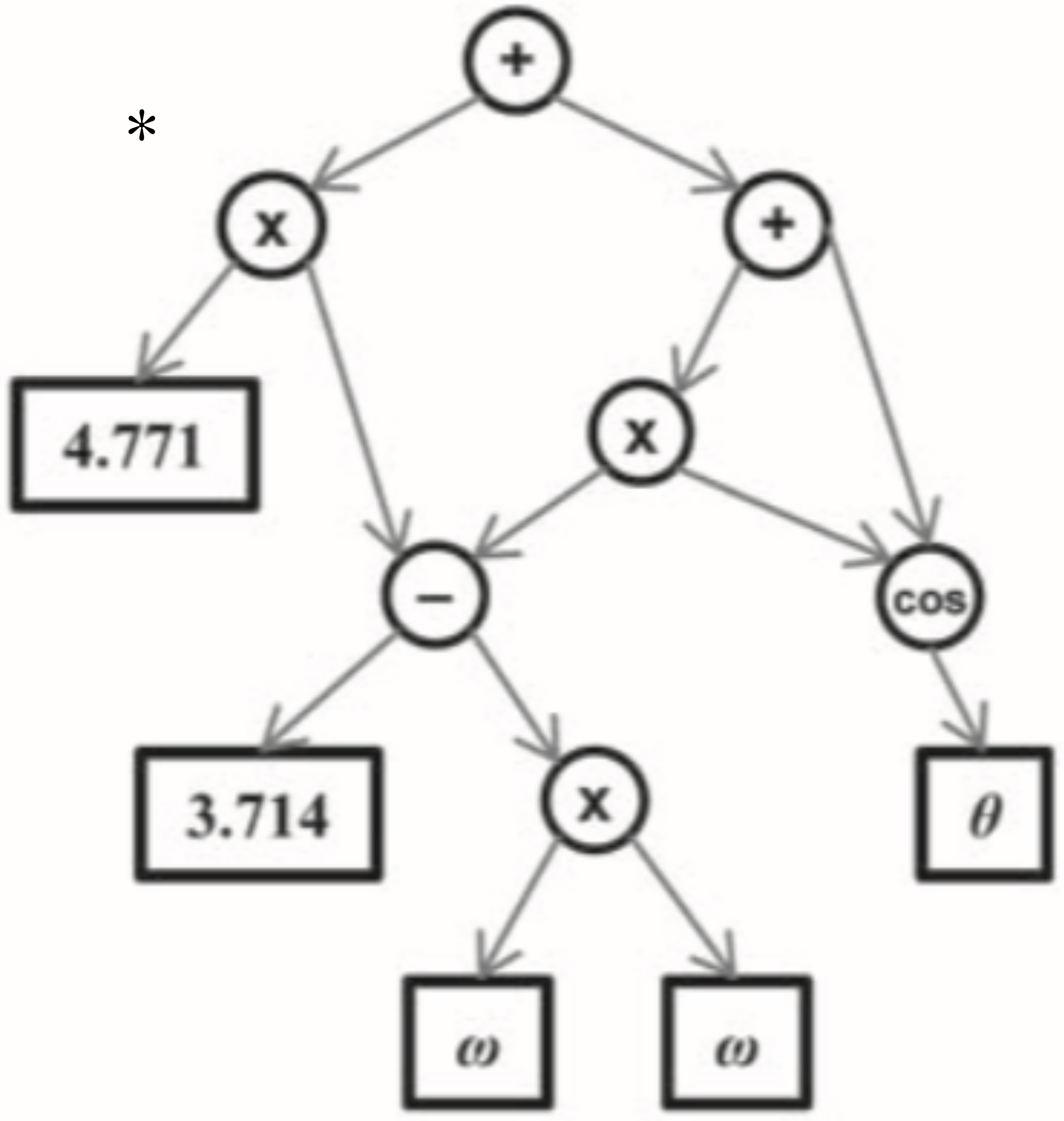
Prediction accuracy of deep learning often comes at expense of poor interpretability and generalizability

Can we discover governing equations directly from data?

Symbolic regression methodology

- Candidate symbolic expressions generated via genetic programming
- Expressions are evaluated on the data
- Evolutionary algorithms to optimize expression accuracy while balancing complexity





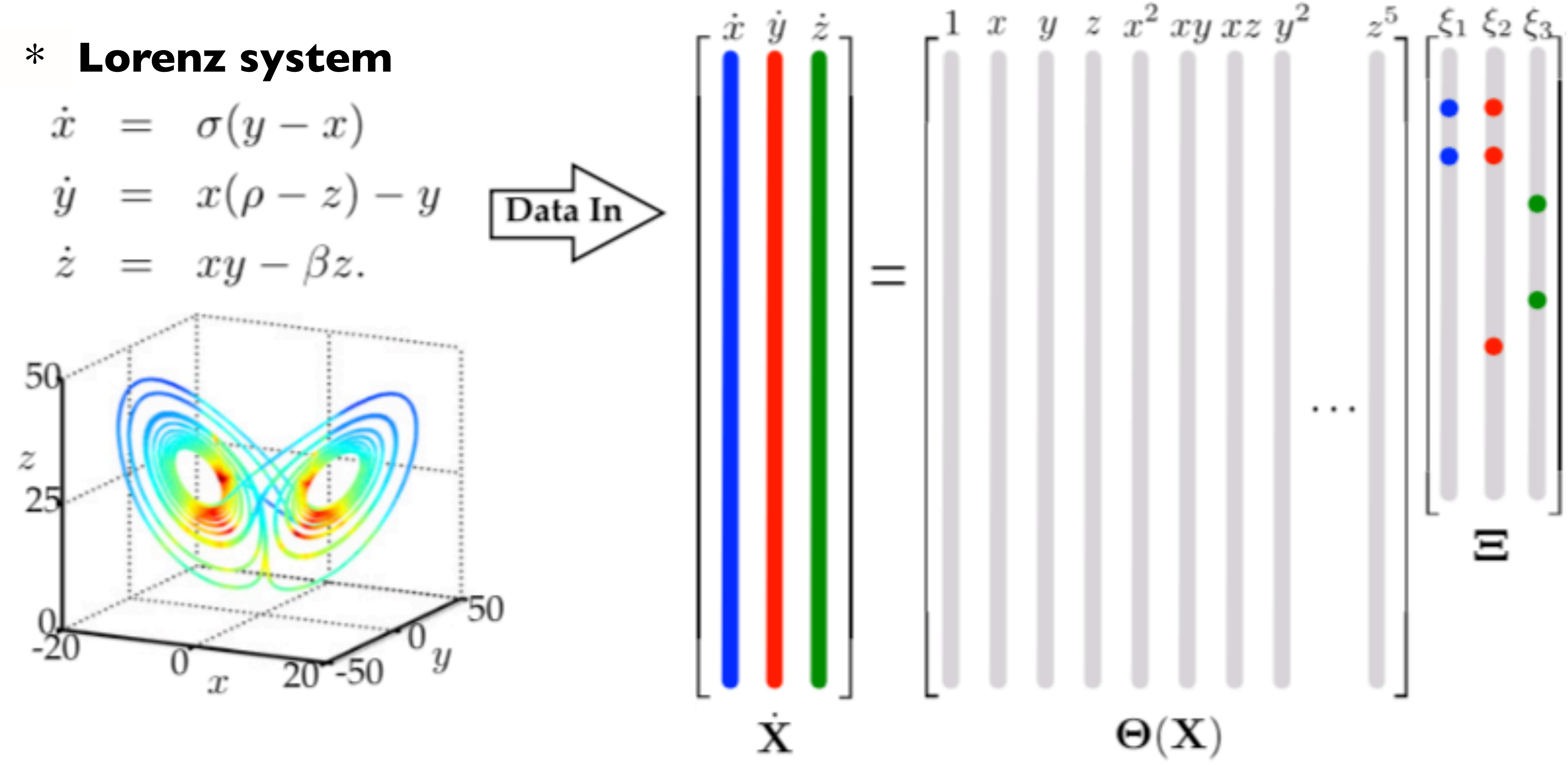
Bongard & Lipson, PNAS (2007)
*Schmidt & Lipson, Science (2009)

Combinatorially large search space and does not scale well to multi-variate high-dimensional systems

Sparse regression offers efficient approach for nonlinear dynamical systems

Sparse regression methodology (SINDy)

- ODE/PDE identification by selecting from library of candidate terms
- Use sparsity-promoting regularizers to select important terms



* **Lorenz system**

$$\begin{aligned}\dot{x} &= \sigma(y-x)\\ \dot{y} &= x(\rho-z)-y\\ \dot{z} &= xy-\beta z.\end{aligned}$$

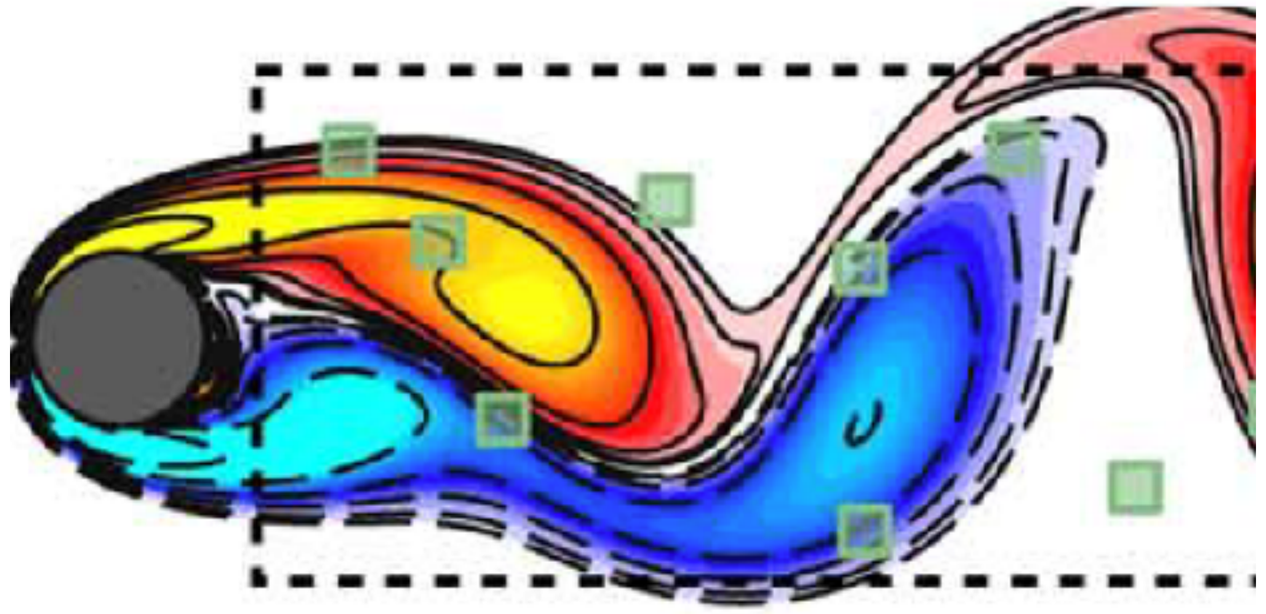
Wang et al., PRL (2011); *Brunton et al., PNAS (2016); Schaeffer, PRS A (2017); Rudy et al. Science Adv. (2017)

$$\text{argmin}_{\xi}||\dot{\mathbf{X}}-\Theta\xi||_2^2+\lambda||\xi||_0$$

Ability to efficiently handle multi-dimensional, multi-variate data makes sparse regression a potentially suitable approach for complex nonlinear dynamics, such as plasma physics problems

SINDy has been successfully applied to different nonlinear dynamical systems

Nonlinear fluid dynamics

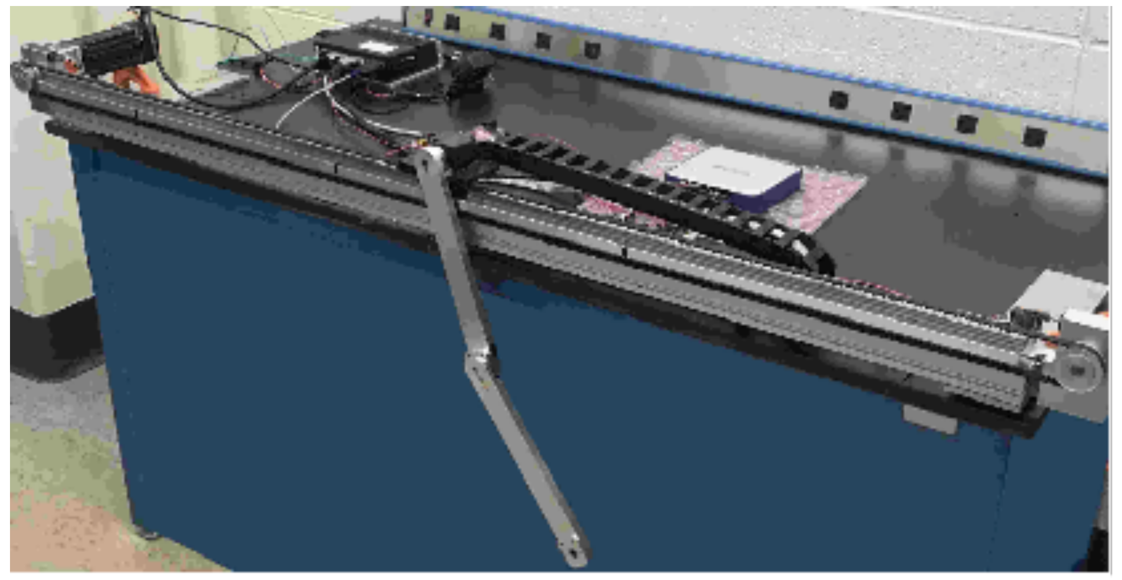


Brunton et al., PNAS (2016)
Rudy et al. Science Adv. (2017)

Chemical-kinetics in human body

Manga et al., IEEE Trans. (2016)

Mechanical systems



Kaheman et al., CDC (2019)

- Low-dimensional representation of nonlinear fluid models
- Model selection in biological networks
- Control of dynamical systems

In many physical systems of interest we know the fundamental equations, but these are too expensive to solve for global simulations; Can we find more efficient reduced models?

Outline

Data-driven models of nonlinear dynamical systems

the importance of interpretability and generalizability

Discovering reduced plasma models from first-principles particle simulations

noisy data and the importance of an integral formulation of sparse regression

Recovering the hierarchy of plasma equations (kinetic to fluid) and fluid closures

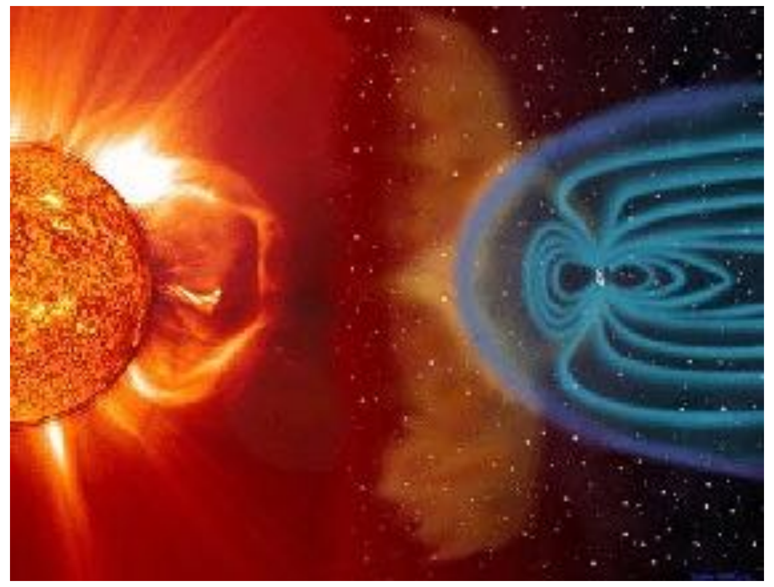
prospects for improved fluid-kinetic closures and more efficient multi-scale algorithms

Multi-scale modeling of nonlinear plasma dynamics is a long-standing challenge

Galaxy cluster

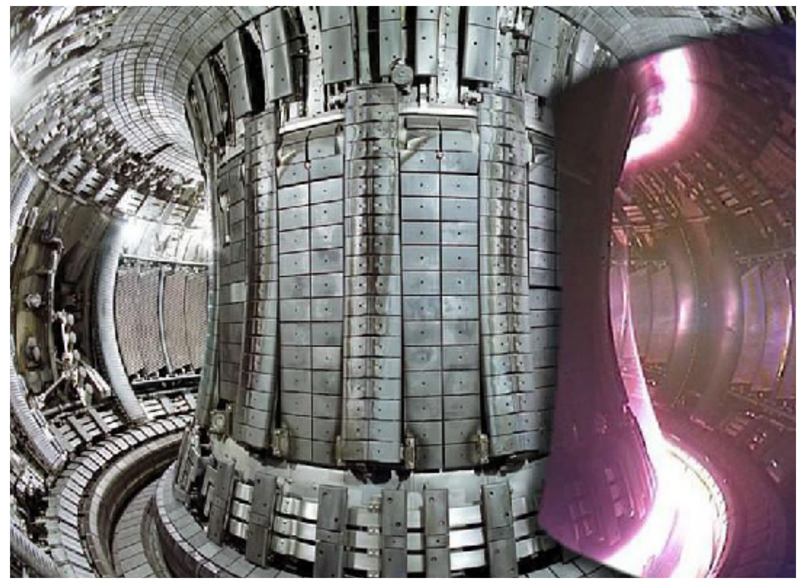
L/d_e = 10^{16}

Solar wind



L/d_e = 10^8

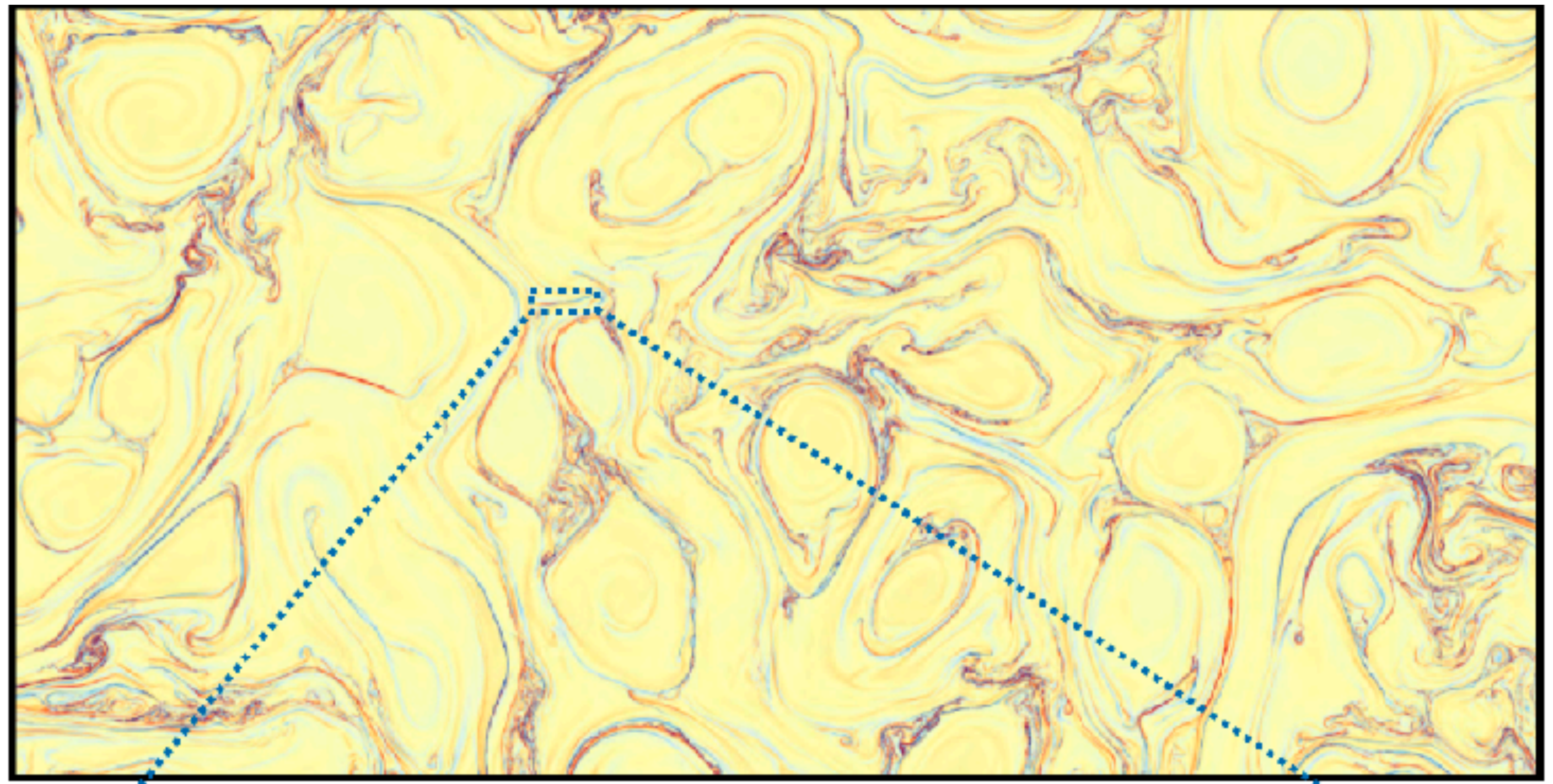
Fusion experiments



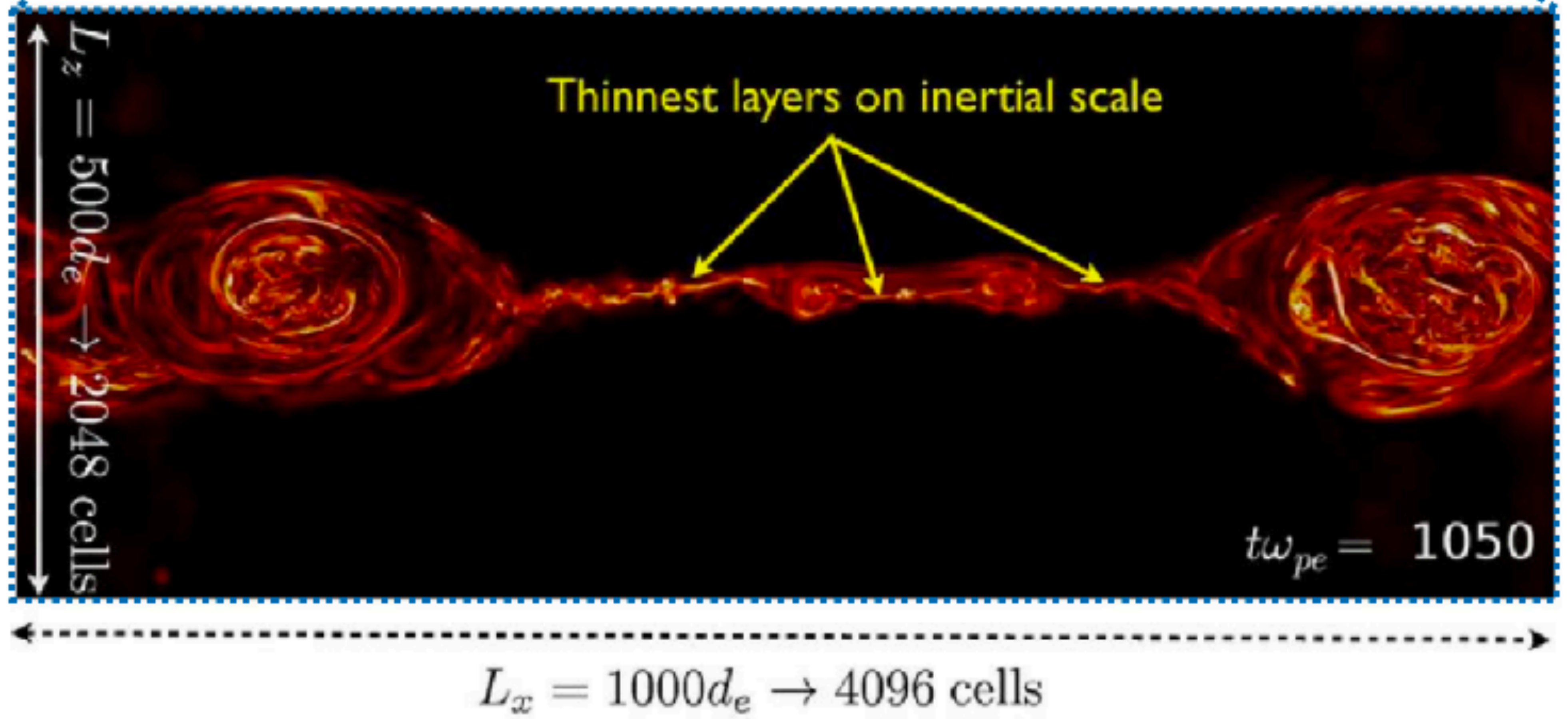
L/d_e = 10^4

- Multi-scale due to long-range nature of electromagnetic force
- Nonlinear interplay between large-scale (fluid) and microscopic N-body phenomena
- Studies are typically compartmentalized between kinetic and fluid approaches

MHD simulation with Rm = 10^6 (Dong et al. 2018)



PIC simulation (Daughton 2019)



Multi-scale modeling of nonlinear plasma dynamics is a long-standing challenge

Kinetic model (Vlasov equation):

$$\frac{\partial f_s(\mathbf{x},\mathbf{v},t)}{\partial t} + \mathbf{v}\cdot\nabla_{\mathbf{x}} f_s + \frac{q_s}{m_s}\left(\mathbf{E}+\frac{\mathbf{v}\times\mathbf{B}}{c}\right)\cdot\nabla_{\mathbf{v}} f_s = 0$$

Fluid model (built from velocity moments of the Vlasov equation):

Continuity eq.

$$\frac{\partial n_s}{\partial t} = -\nabla\cdot(n_s\langle\mathbf{v}\rangle_s)$$

Momentum eq.

$$\frac{\partial(n_s\langle\mathbf{v}\rangle_s)}{\partial t} = -\nabla\cdot(n_s\langle\mathbf{v}\mathbf{v}\rangle_s) + \frac{q_s n_s}{m_s}(\mathbf{E}+\langle\mathbf{v}\rangle_s\times\mathbf{B})$$

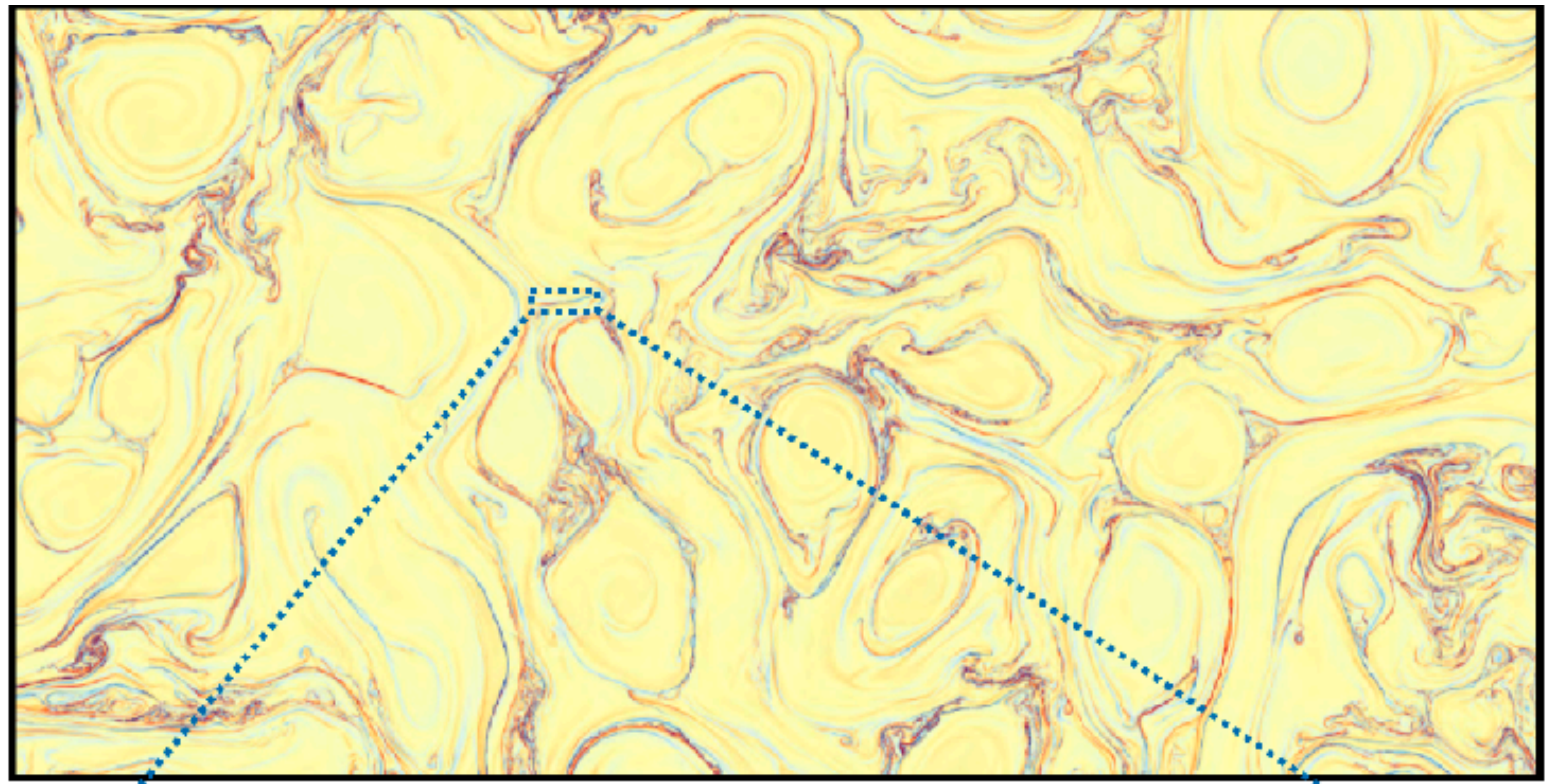
Energy eq.

$$\frac{\partial p_s}{\partial t} = -\nabla\cdot(p_s\langle\mathbf{v}\rangle_s) - (\mathcal{P}_s\cdot\nabla)\cdot\langle\mathbf{v}\rangle_s - \nabla\cdot\mathbf{q}_s$$

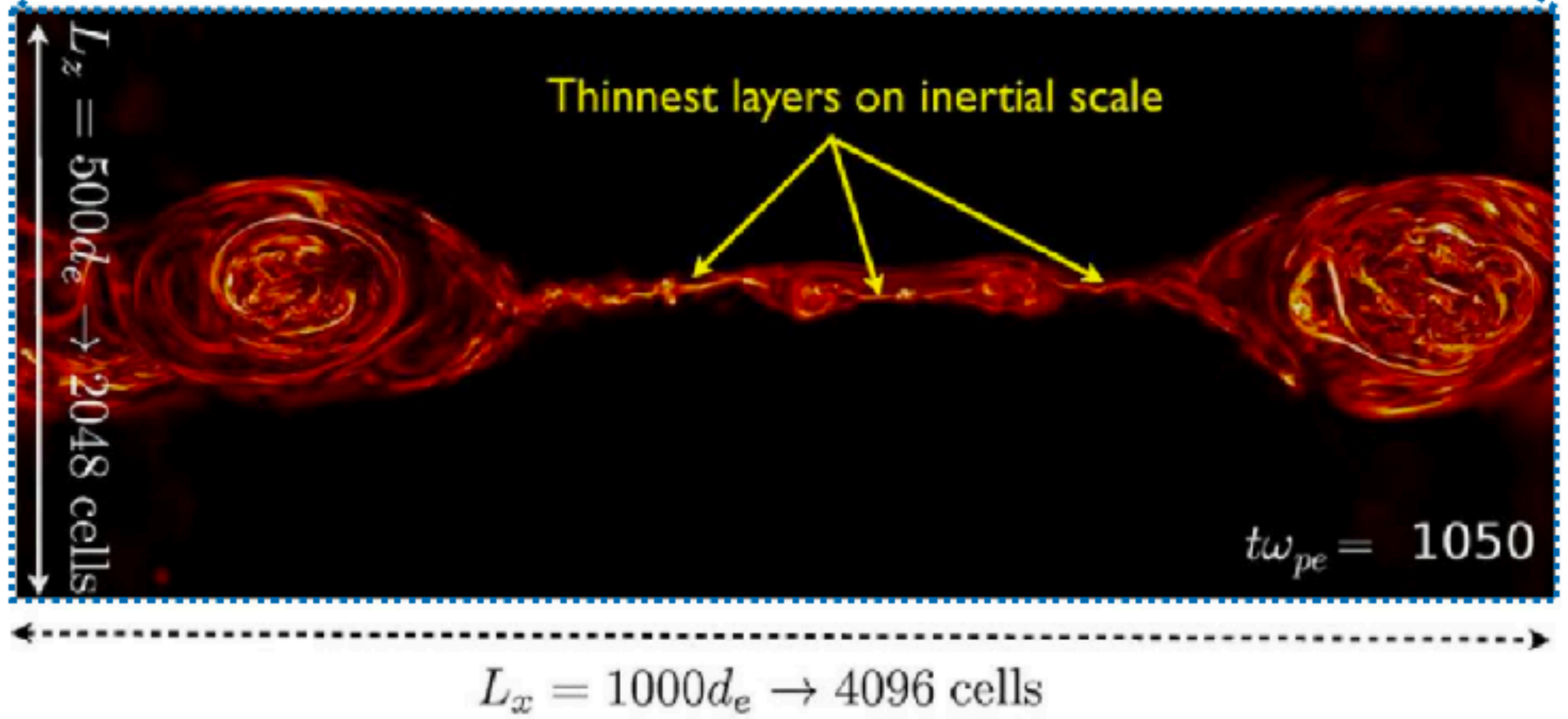
…

Truncation/closure of fluid hierarchy is needed, typically at the first three moments

MHD simulation with Rm = 10^6 (Dong et al. 2018)



PIC simulation (Daughton 2019)



Can we use data-driven techniques to develop better multi-scale plasma models?

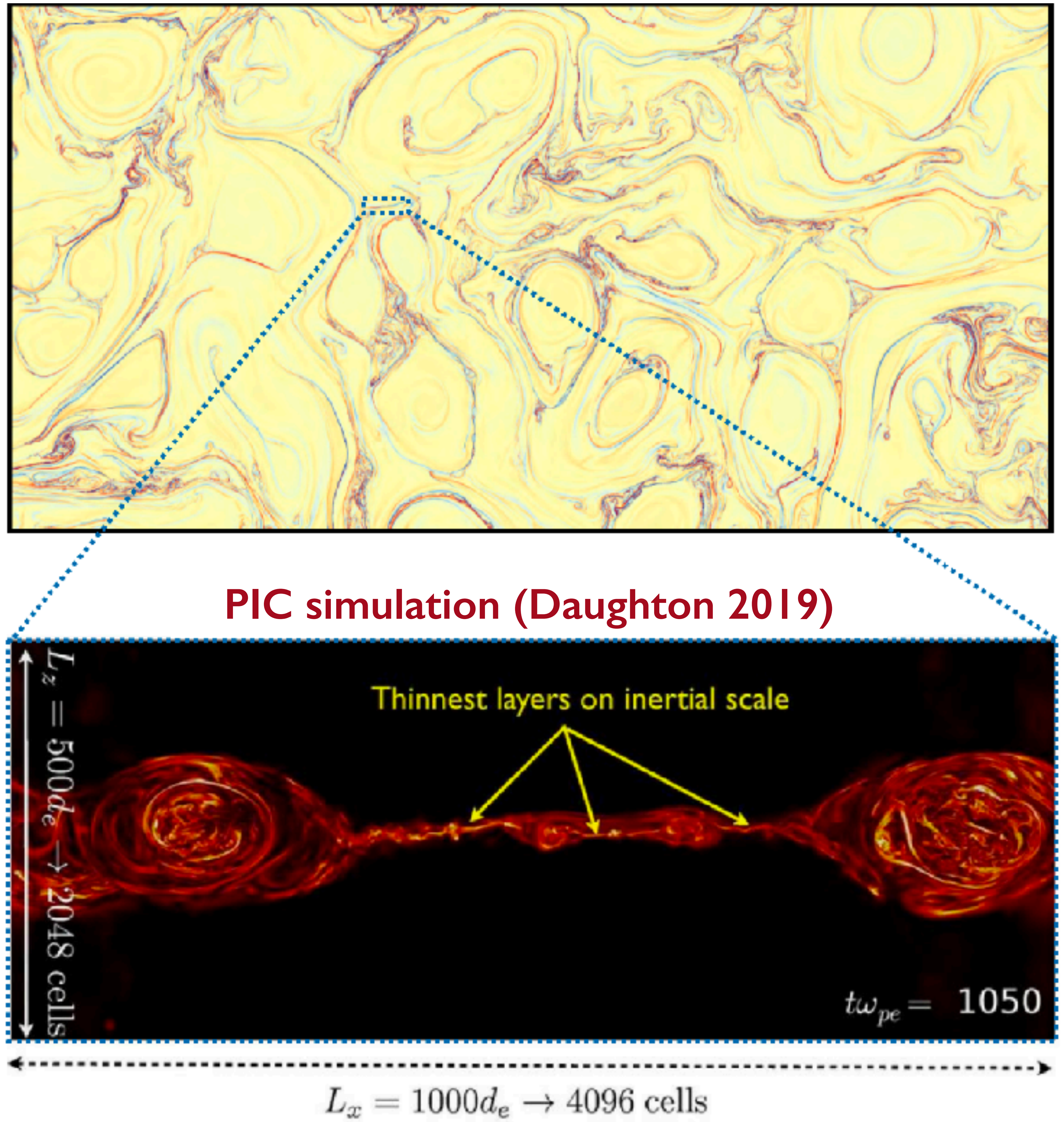
Can we use SR applied to data from first-principles kinetic simulations to

- recover hierarchy of progressively more reduced models (from kinetic to fluid)?
- inform the development of better closures?

To develop more efficient models of

- the impact of microscopic kinetic physics on large-scale (fluid) dynamics
- coupling between kinetic and fluid descriptions
- development of nonthermal distributions due to micro-turbulence
- nonlinear evolution of plasma instabilities
- …

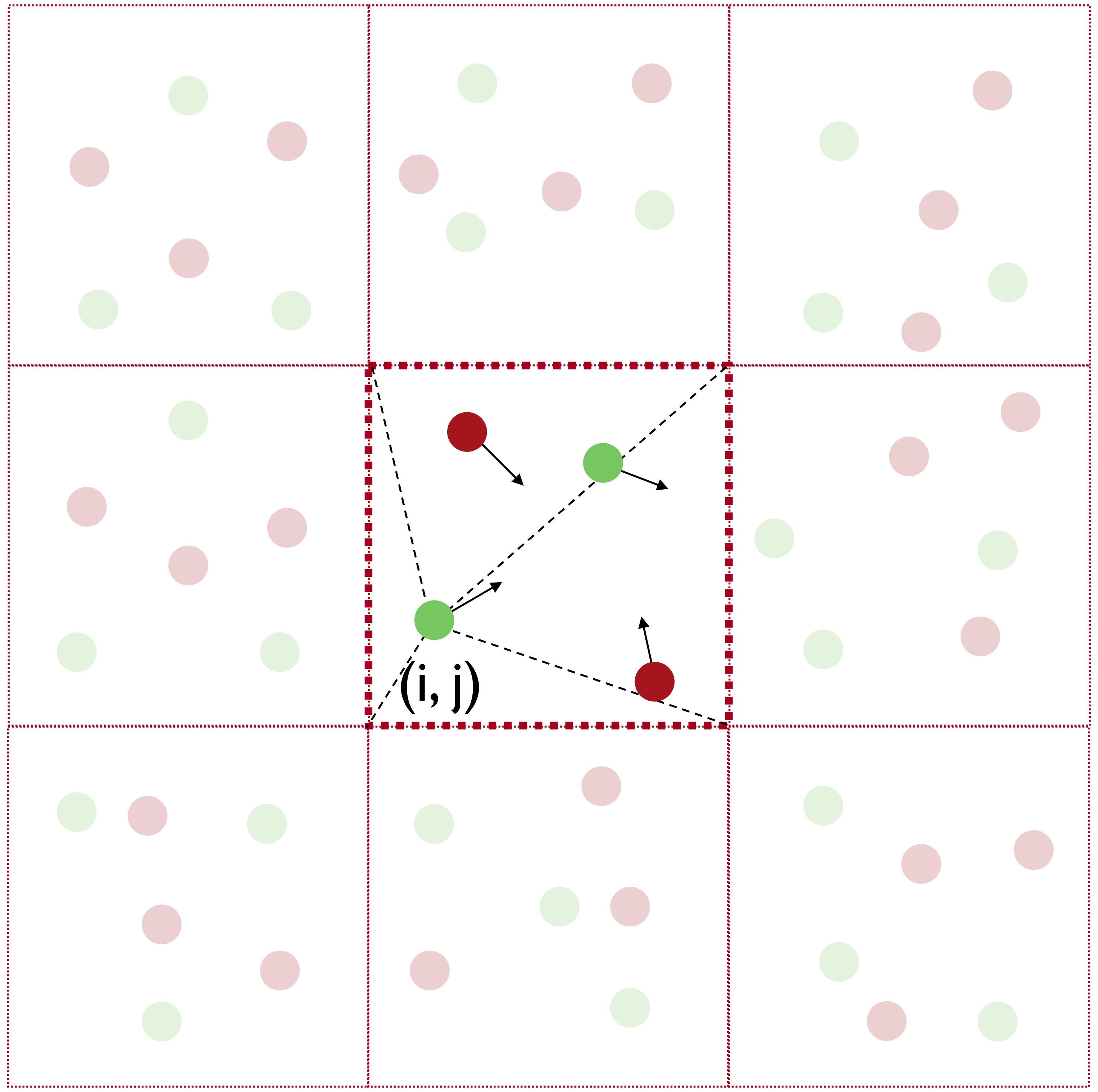
MHD simulation with Rm = 10^6 (Dong et al. 2018)



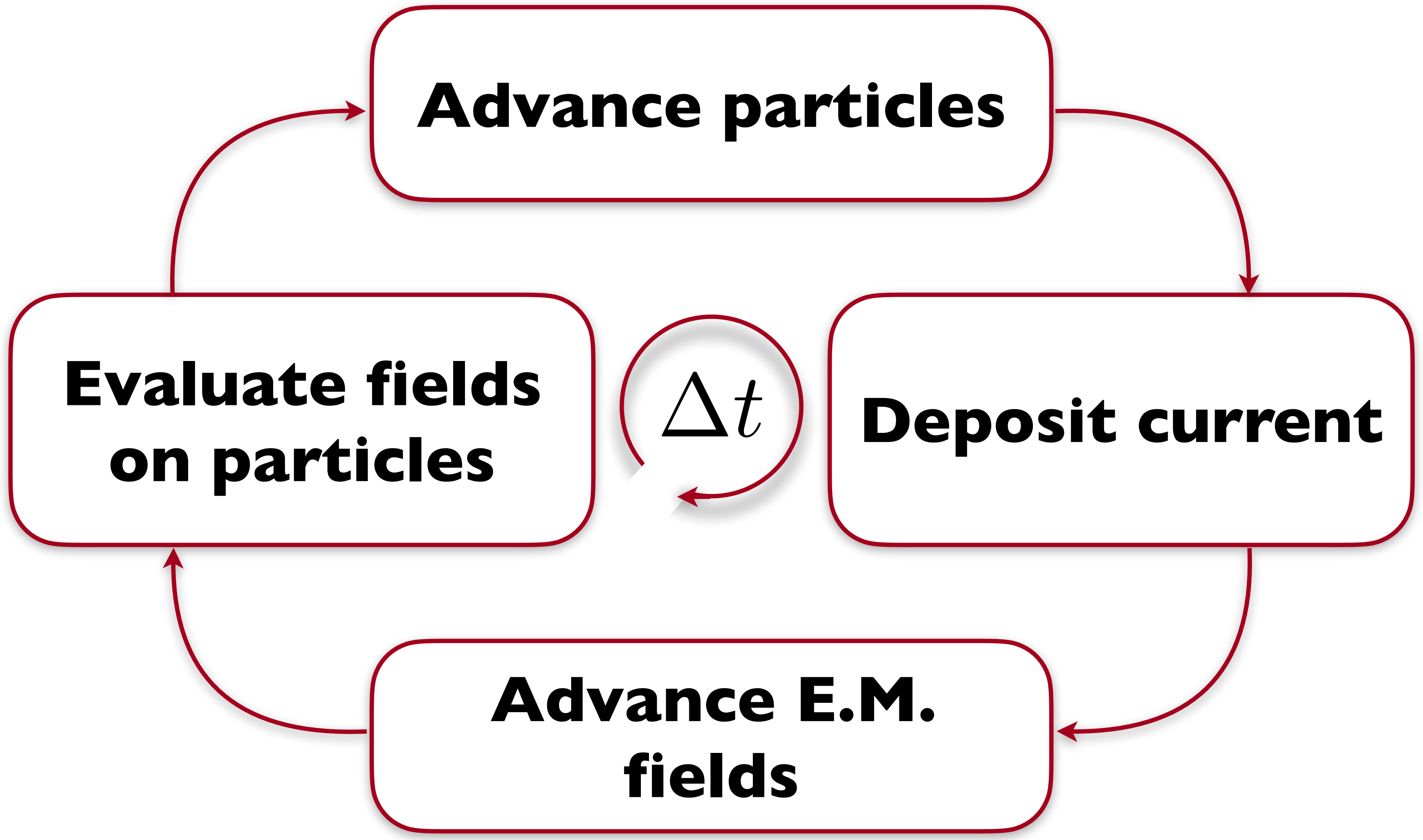
PIC simulation (Daughton 2019)

Data-driven discovery of plasma physics from first-principles simulations

Particle-in-cell (PIC) method provides first-principles description of a plasma



$$\frac{\mathrm{d}\mathbf{p}}{\mathrm{d}t} = q\left(\mathbf{E} + \frac{\mathbf{v}}{c} \times \mathbf{B}\right)$$



Advance particles

Deposit current

Advance E.M. fields

Evaluate fields on particles

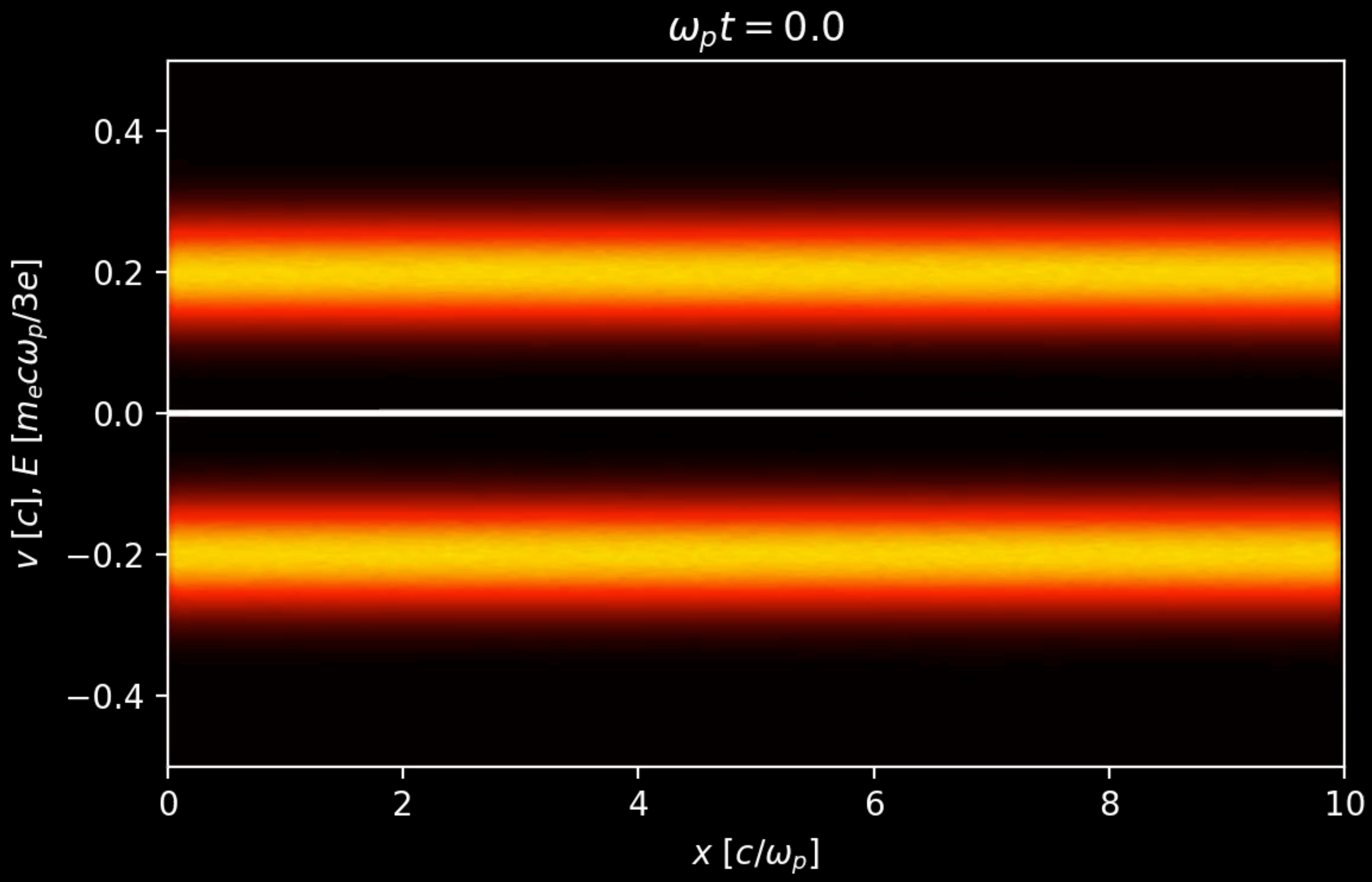
Δt

$$\frac{\partial \mathbf{E}}{\partial t} = c\nabla \times \mathbf{B} - 4\pi \mathbf{j}$$

$$\frac{\partial \mathbf{B}}{\partial t} = -c\nabla \times \mathbf{E}$$

Can we use sparse regression to discover reduced interpretable models from particle data?

Nonlinear dynamics of counter-streaming electron beams (2-stream instability)



- Evolution of the plasma distribution function in phase-space *f*(x, v, t)
- Evolution of the plasma-generated electric field E(x, t)

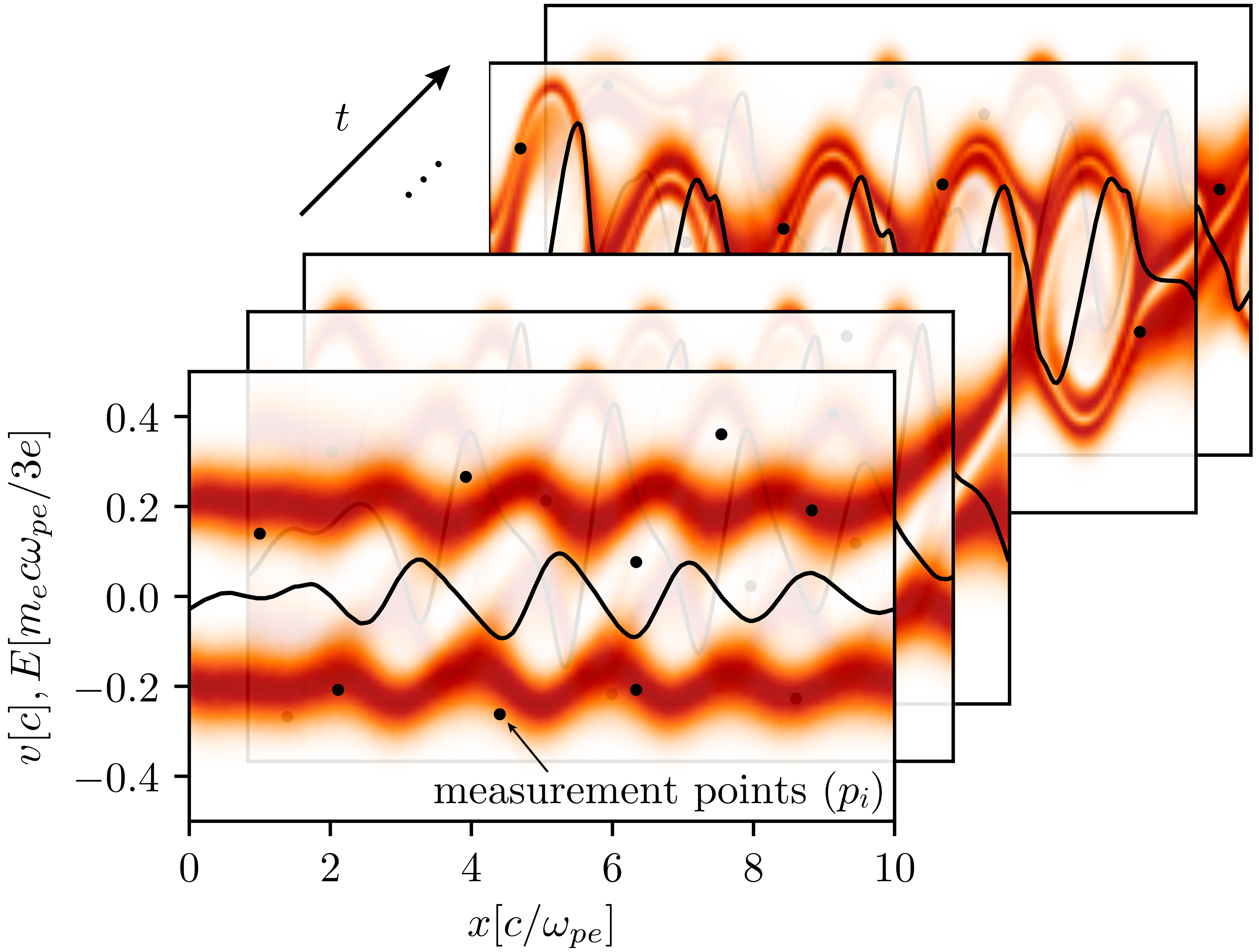
Can we determine the PDE that governs the evolution of the plasma distribution function?

$$\frac{\partial f}{\partial t} = ?$$

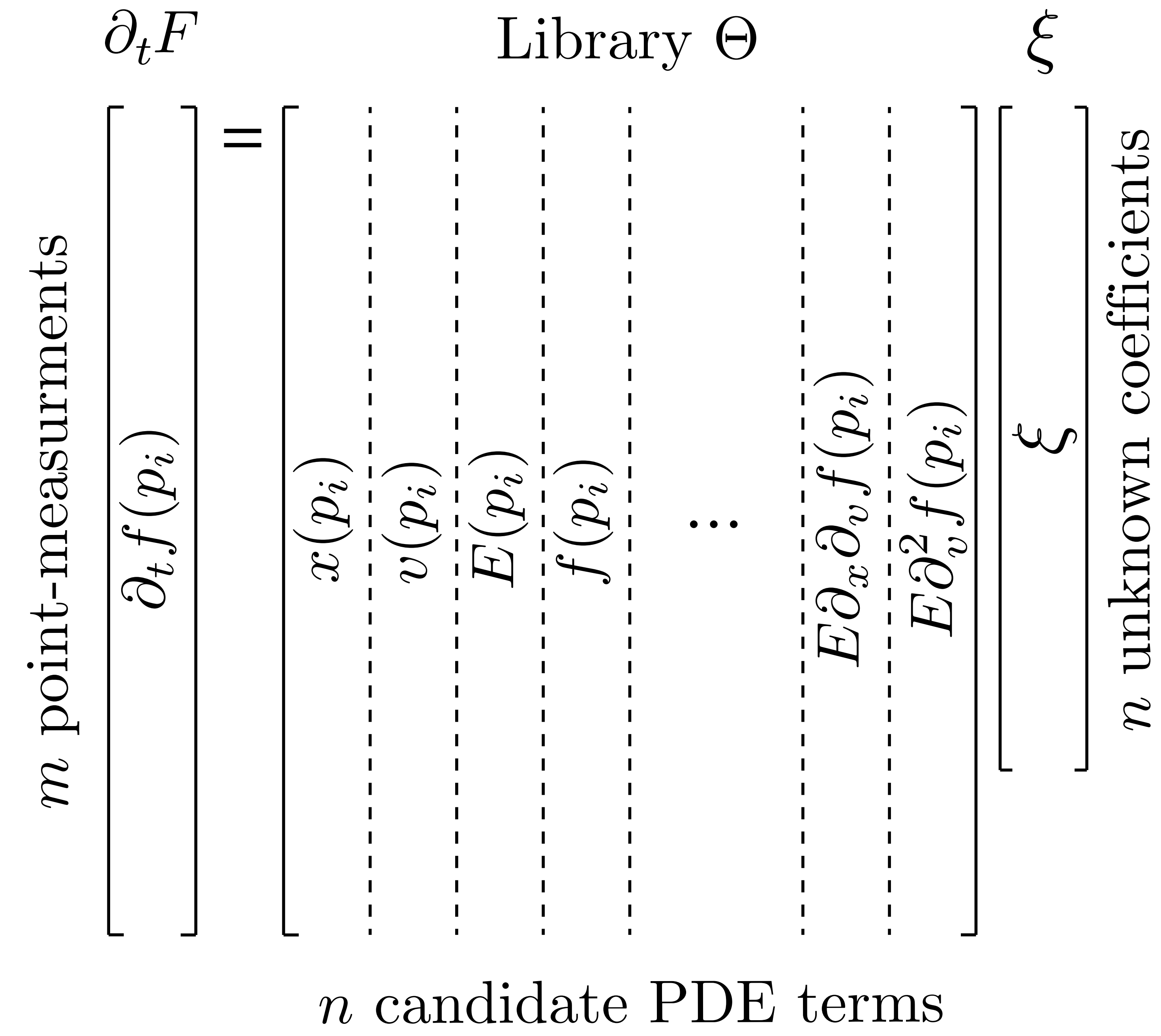
Note that PIC does not explicitly solve for *f*(x, v, t)

Inferring the kinetic Vlasov equation via sparse regression (SINDy)

1. Sampling measurement points from data



2. PDE discovery as a regression problem



3. Solve sparse optimization problem*

$$\text{argmin}_\xi ||\partial_t F - \Theta\xi||_2^2 + \lambda||\xi||_0$$

4. Recovery of correct form of the Vlasov equation

Inferred PDE: $\partial_t f = -0.803 v\partial_x f + 0.673 E\partial_v f$

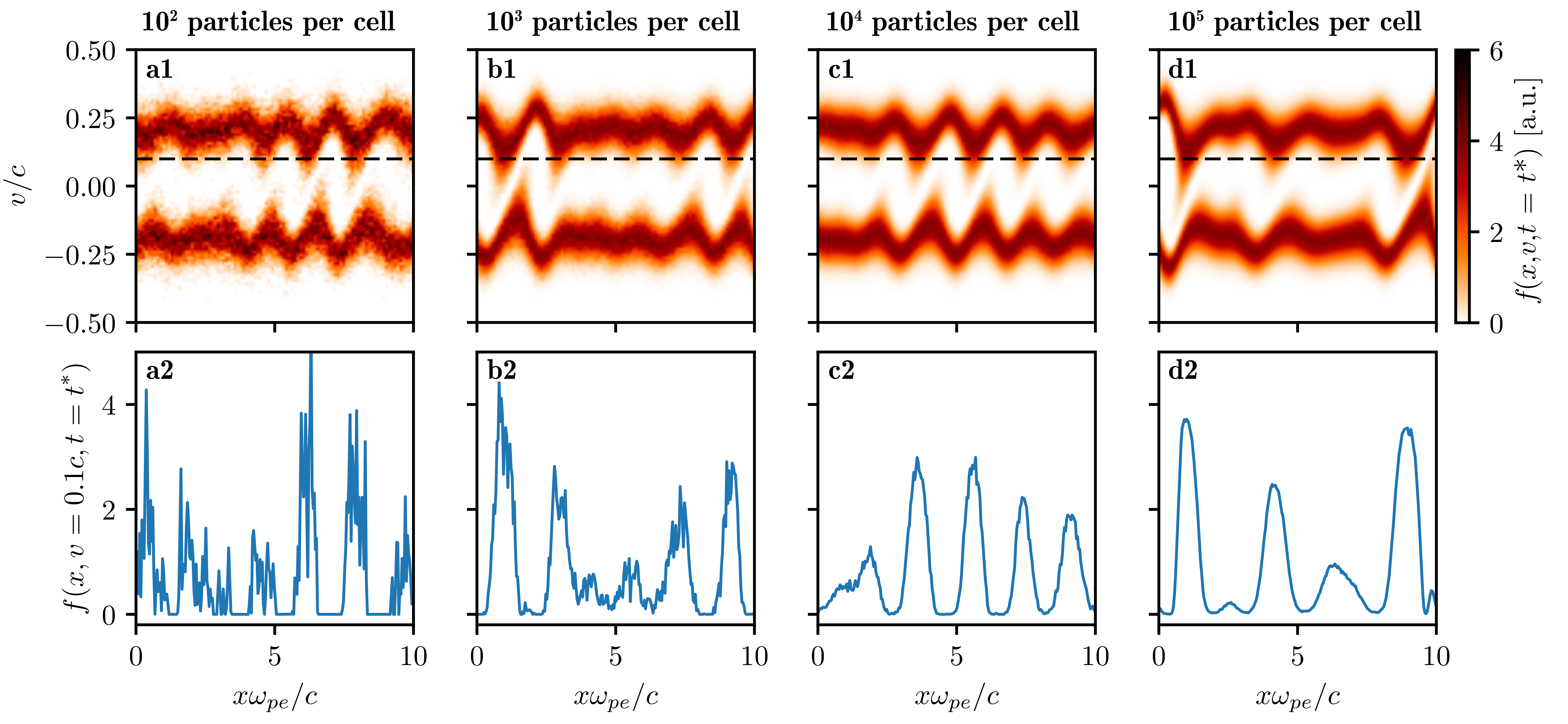
True Vlasov eq.: $\partial_t f = -v\partial_x f - \dfrac{q}{m}E\partial_v f$

Large error (~20%) in the inferred PDE coefficients

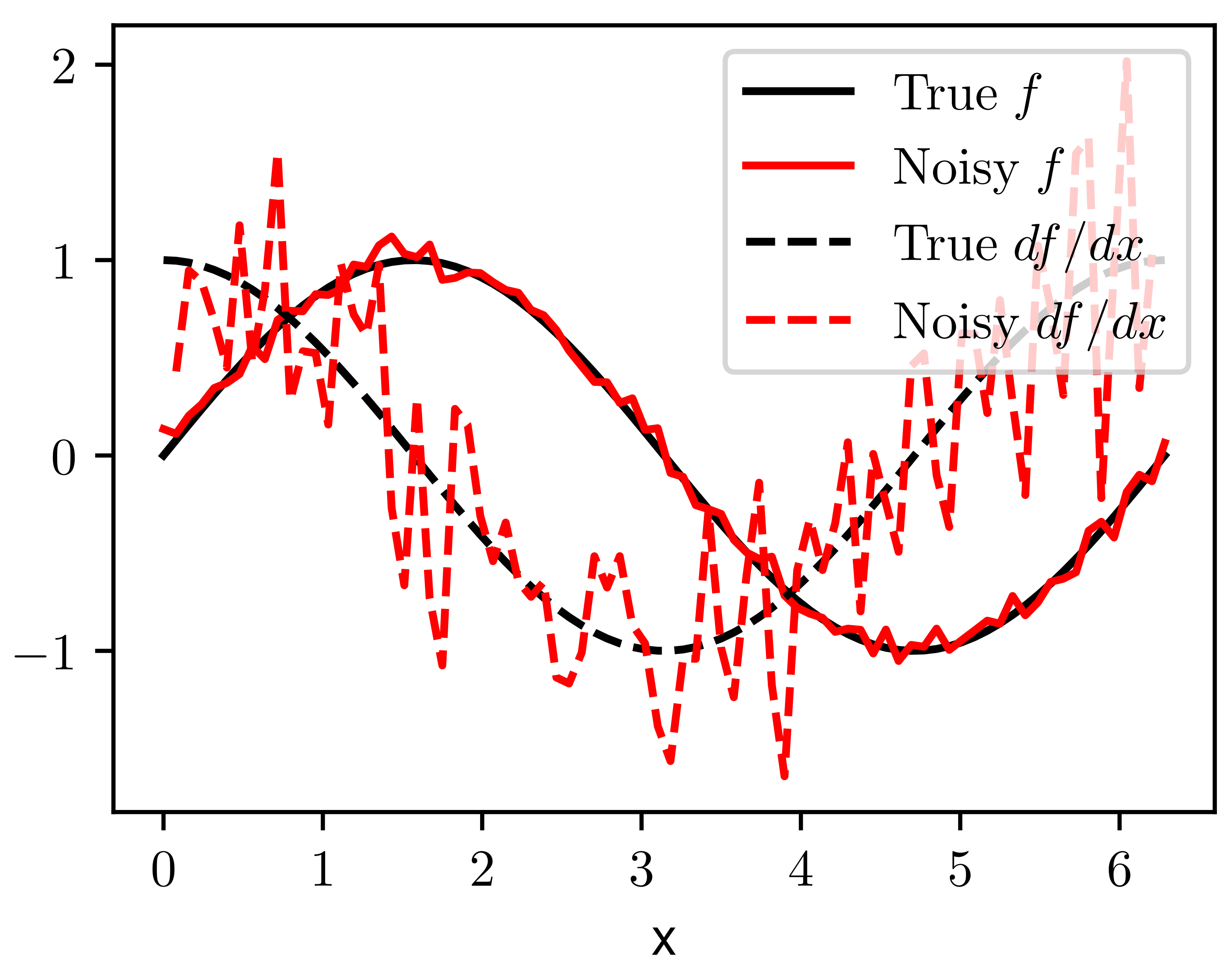
* we use sequential threshold least squares

Particle-based data is inherently noisy due to finite number of particles

Noise dependence on number of particles per cell used



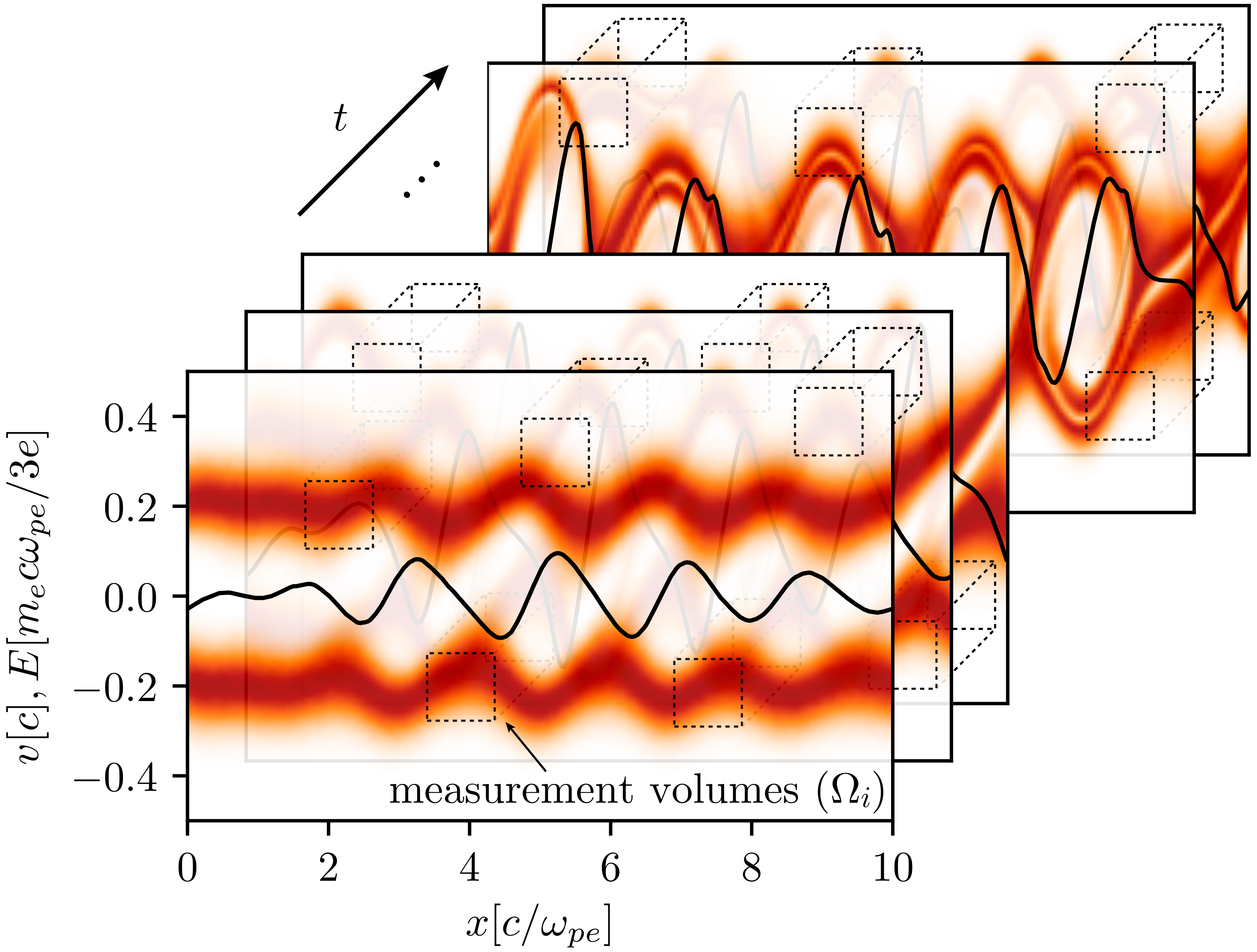
Straightforward numerical differentiation amplifies data noise



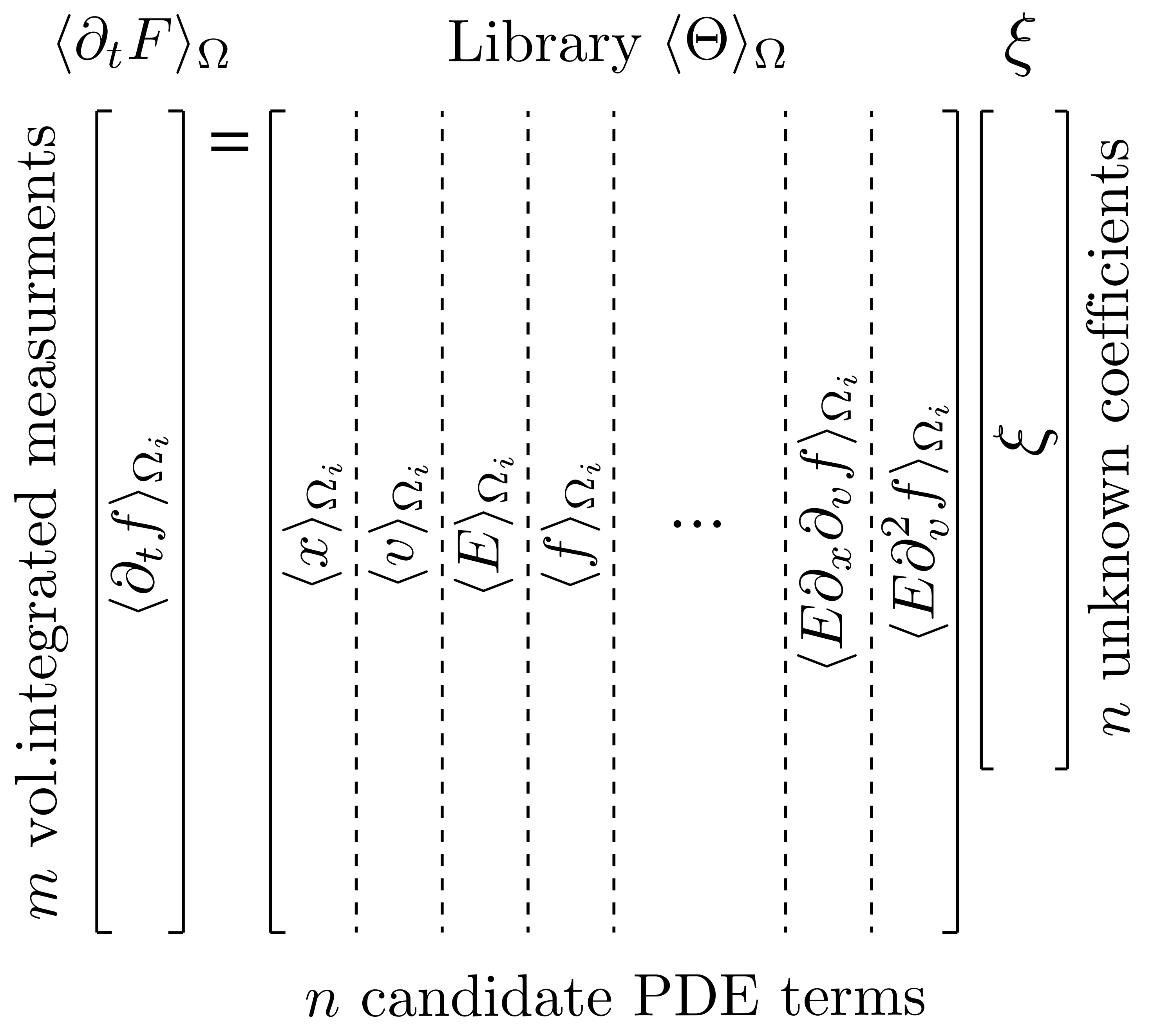
Regularized differentiation techniques can mitigate noise but can also introduce bias in the data

Integral formulation of PDE identification via sparse regression

1. Sampling measurement volumes from data



2. PDE discovery as a regression problem



3. Solve sparse optimization problem

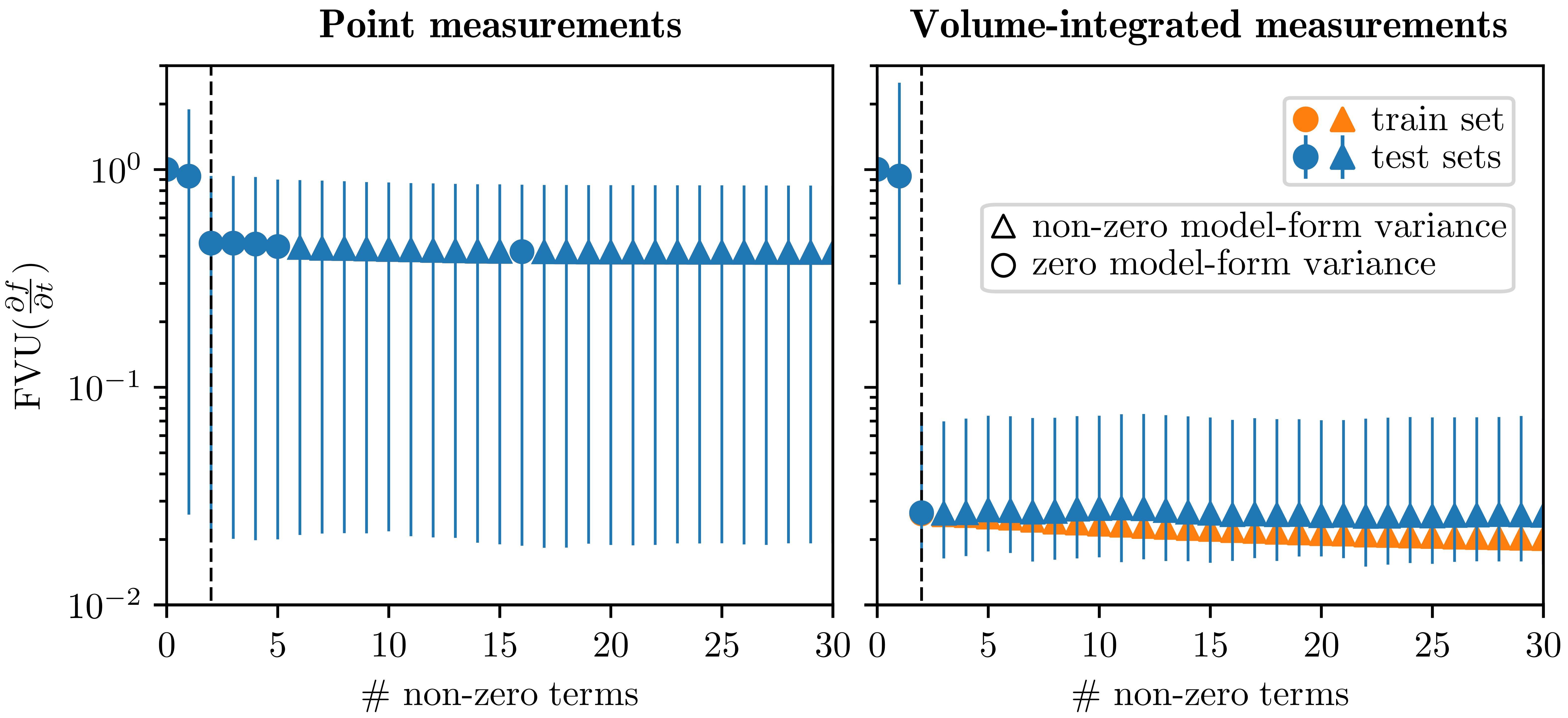
$$\text{argmin}_{\xi}||\langle\partial_t F\rangle_\Omega - \langle\Theta\rangle_\Omega\xi||_2^2 + \lambda||\xi||_0$$

Crutchfield et al. (1987); Schaeffer et al. (2017); Reinbold et al. (2020); Messenger et al. (2021)

E. P. Alves and F. Fiuza, Physical Review Research (2022); arXiv:2011.01927

Improved accuracy and robustness using the integral formulation

Pareto analysis of the integral formulation reveals much steeper Pareto-front



Accurate recovery of the underlying PDE

Inferred PDE:

$$\partial_t f = -1.006 v\partial_x f + 0.997 E\partial_v f$$

True Vlasov eq.:

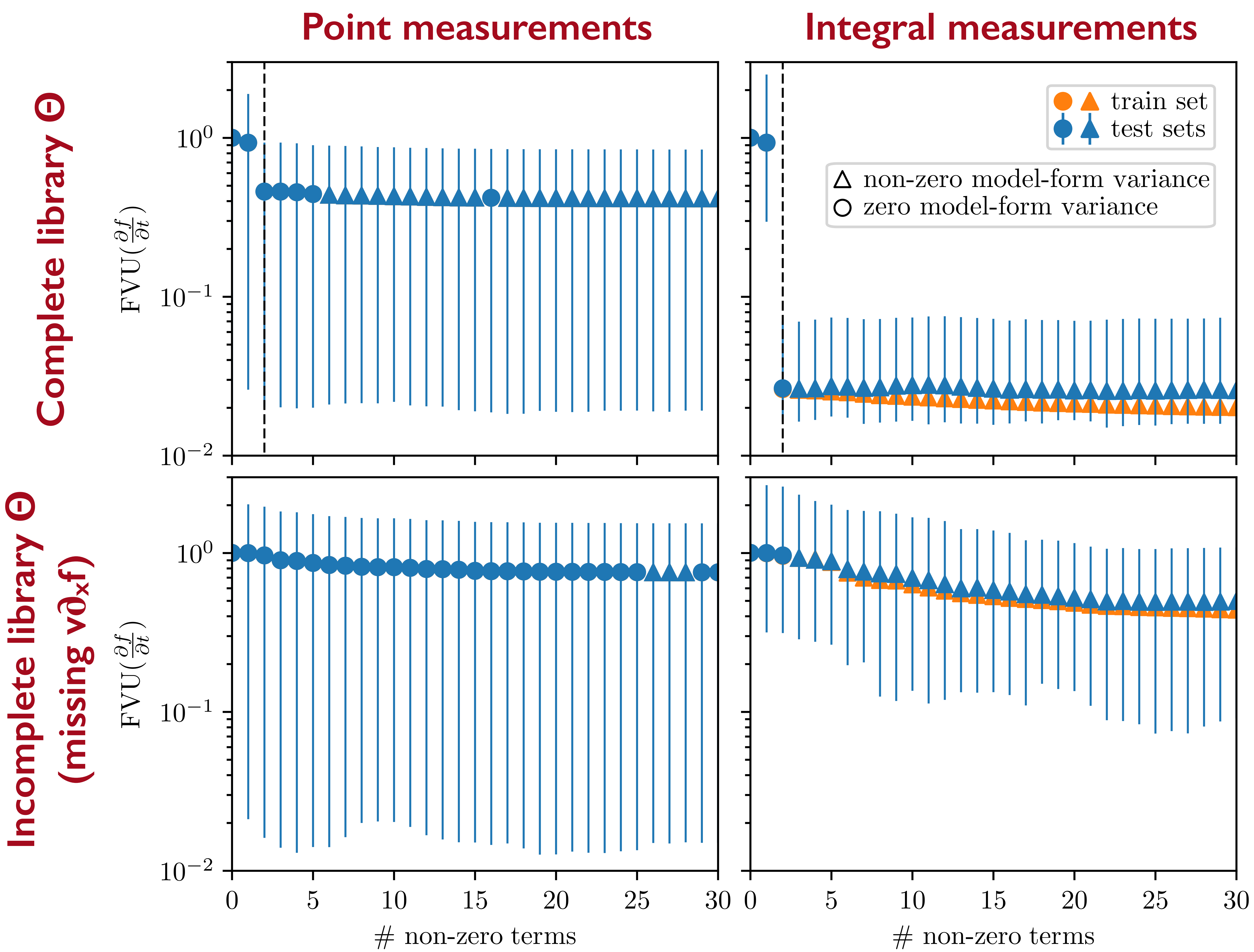
$$\partial_t f = -v\partial_x f - \frac{q}{m}E\partial_v f$$

- <1% error in inferred coefficients
- **Vlasov equation was not directly solved in simulation**

The integral formulation strategy mitigates differentiation errors in noisy data without introducing unwanted bias in the data

E. P. Alves and F. Fiuza, Physical Review Research (2022); arXiv:2011.01927

Improved accuracy and robustness using the integral formulation



Indicators of unreliable PDE identification

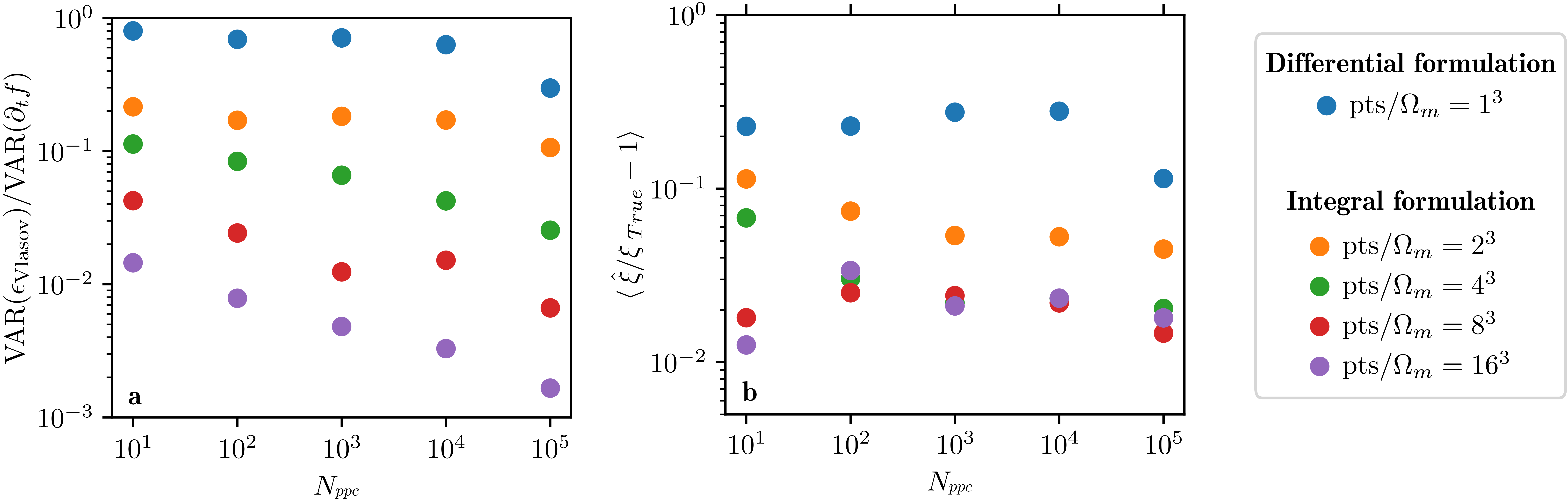
- No pronounced inflection in Pareto front
- High variance in model error
- High variance in model form

- Inadequate choice of basis terms
- Missing dynamical terms (missing physics!)

Redesign/expand library of candidate terms

E. P. Alves and F. Fiuza, Physical Review Research (2022); arXiv:2011.01927

Improved accuracy and robustness using the integral formulation



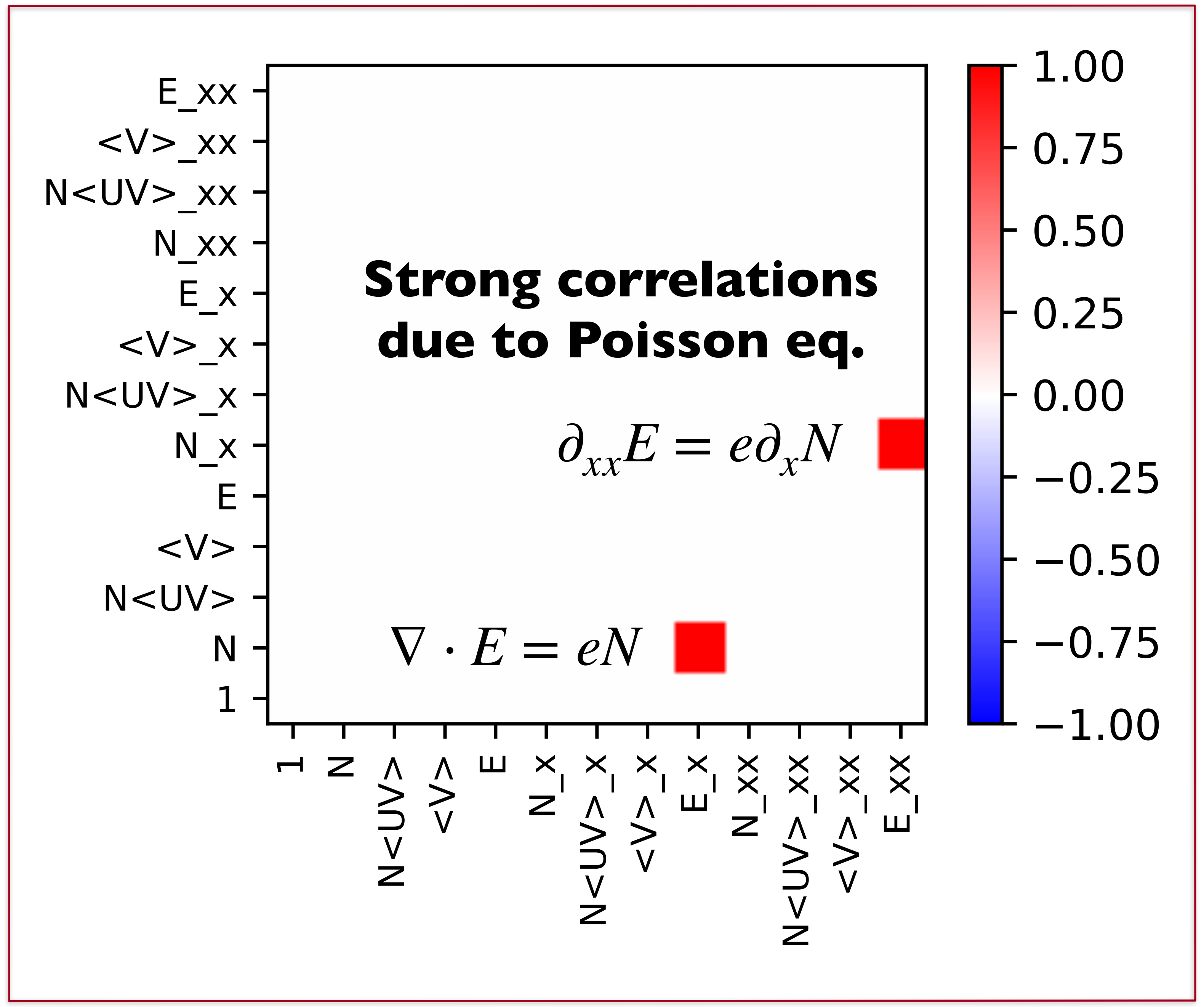
Deviation error from true Vlasov eq.

Average coefficient error

E. P. Alves and F. Fiuza, Physical Review Research (2022); arXiv:2011.01927

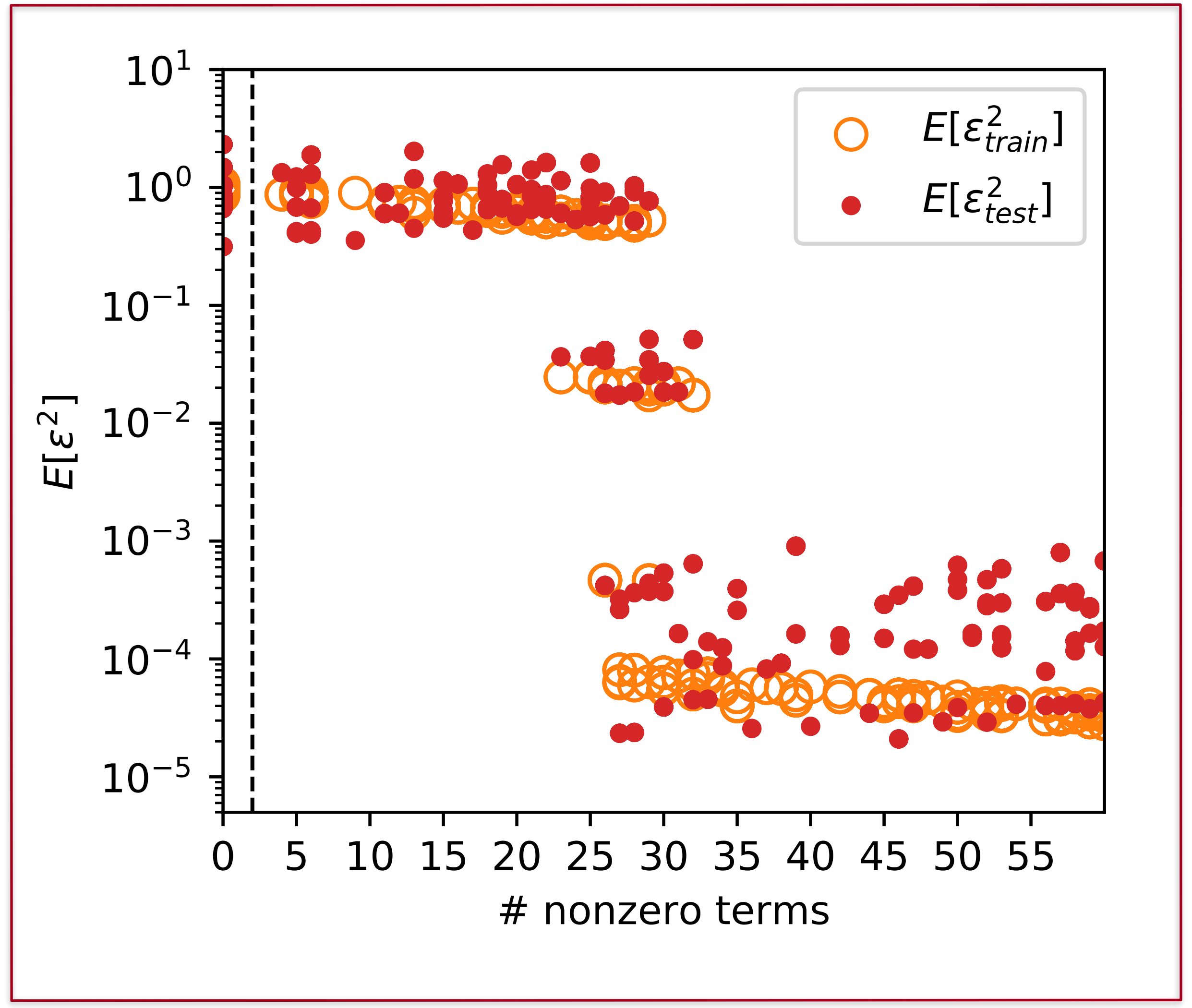
Strongly correlated variables (features) can corrupt sparse model identification

Correlation matrix of primary features used for model discovery

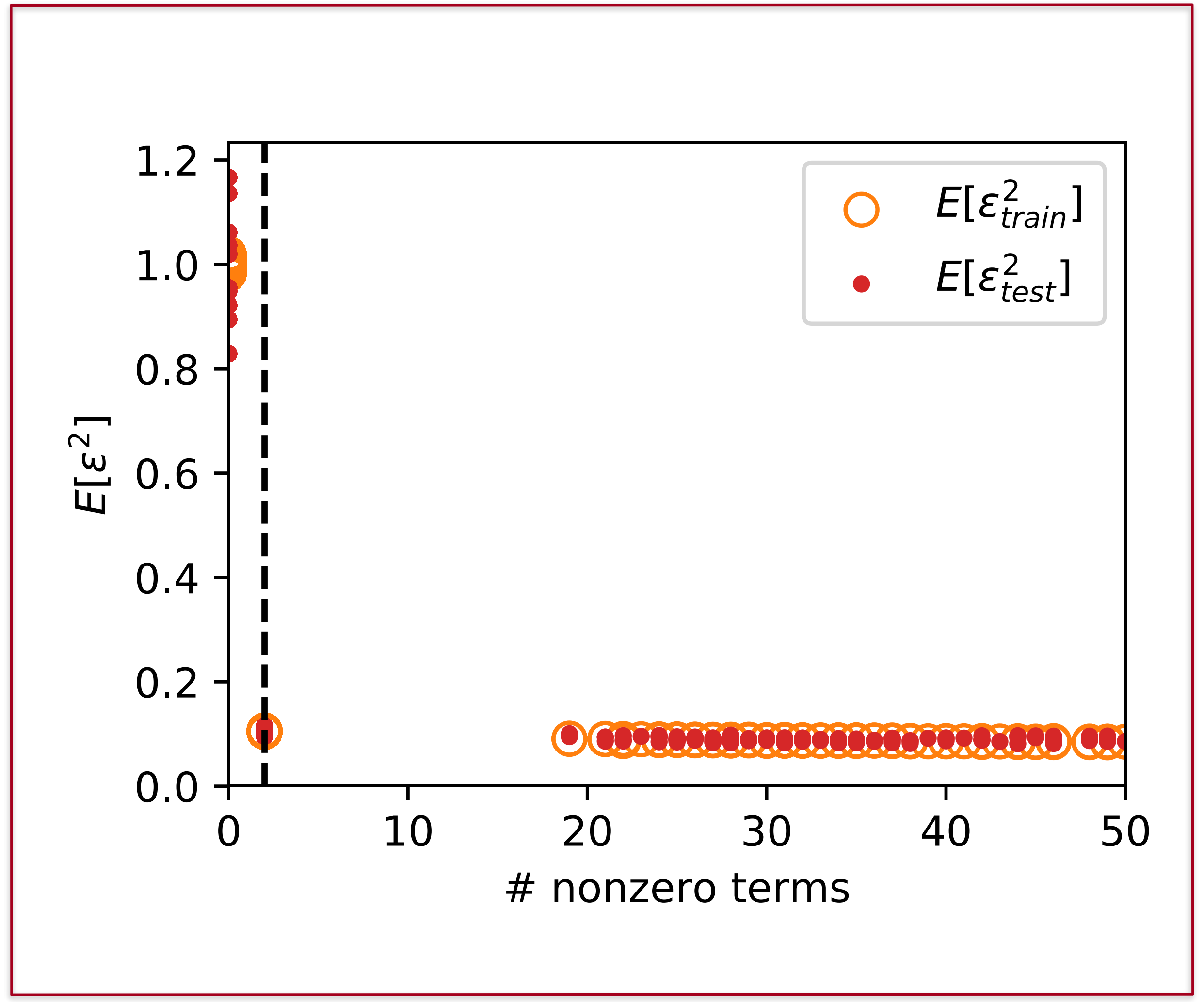


Correlations can arise due to fundamental physical laws/symmetries (e.g. ∇·E = eN, ∇·B = 0)

Sparse model identification with strongly correlated features



Successful model identification after removing E_x & E_{xx} features



E. P. Alves and F. Fiuza, Physical Review Research (2022); arXiv:2011.01927

Outline

Data-driven models of nonlinear dynamical systems

the importance of interpretability and generalizability

Discovering reduced plasma models from first-principles particle simulations

noisy data and the importance of an integral formulation of sparse regression

Recovering the hierarchy of plasma equations (kinetic to fluid) and fluid closures

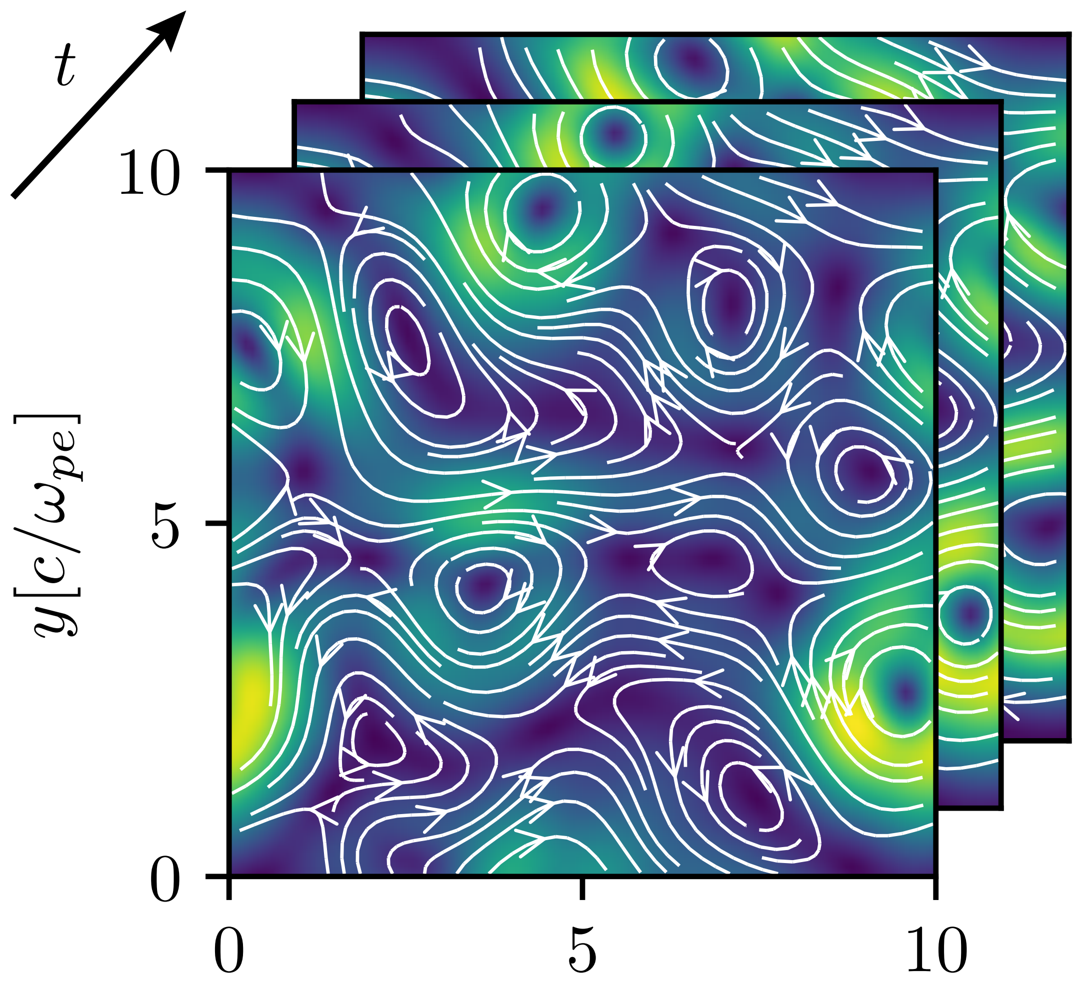
prospects for improved fluid-kinetic closures and more efficient multi-scale algorithms

Recovery of the multi-fluid equations from nonlinear plasma instability dynamics

Multi-fluid equations

Data

Weibel instability dynamics

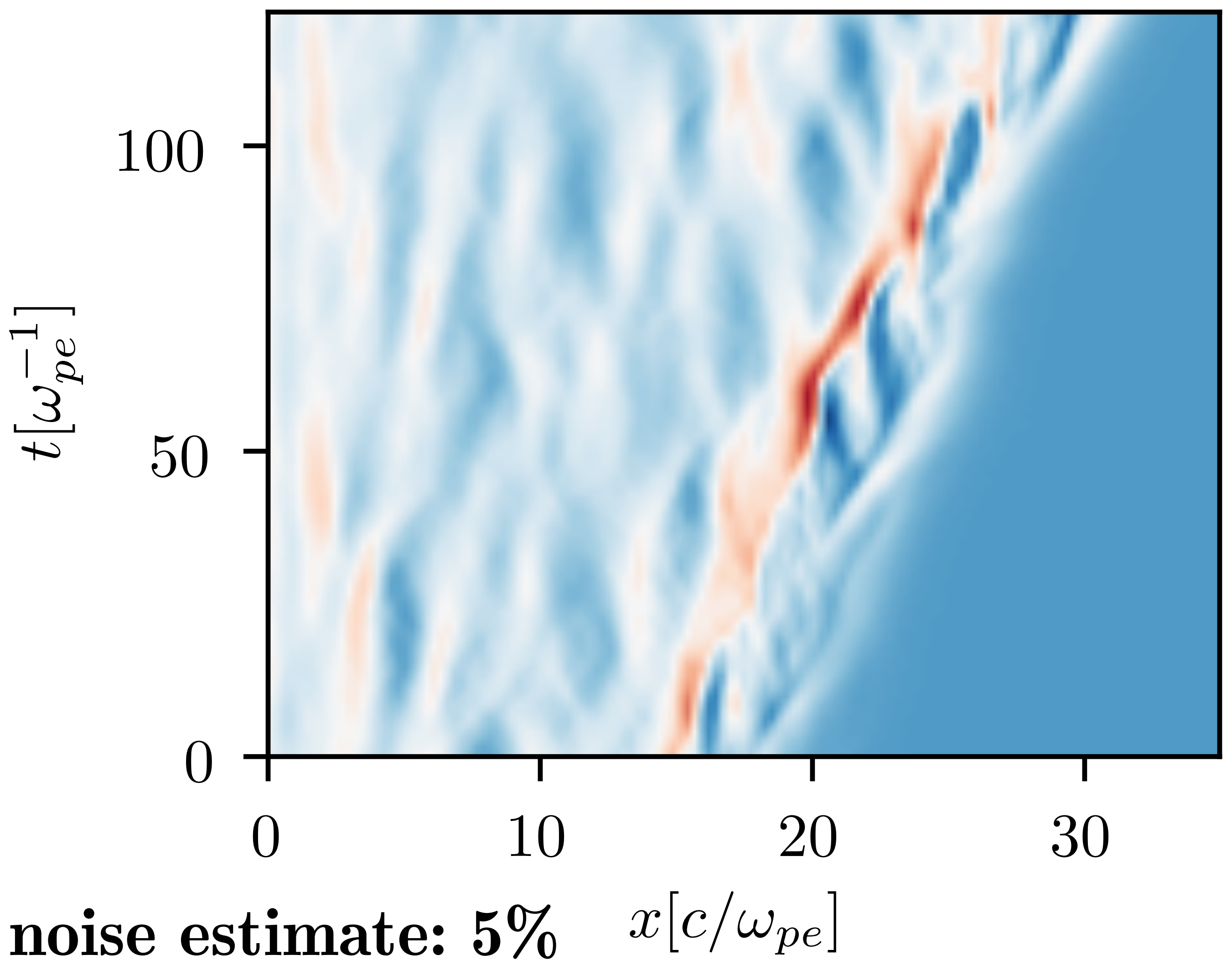


noise estimate: 18%

PDEs	PDE identification and inference accuracy	
	a) Point strategy	b) Integral strategy
Continuity eq. $\partial_t n_s = -\nabla \cdot (n_s \langle \mathbf{v} \rangle_s)$	Sucessful identification Mean coeff. error $\approx 20\%$	**Sucessful identification** **Mean coeff. error $\approx$ 1%**
Momentum eq. $\partial_t (n_s \langle \mathbf{v} \rangle_s) = -\nabla \cdot (n_s \langle \mathbf{vv} \rangle_s) + \frac{q_s n_s}{m_s}(\mathbf{E} + \langle \mathbf{v} \rangle_s \times \mathbf{B})$	Unsucessful identification	**Sucessful identification** **Mean coeff. error $\approx$ 2%**
Energy eq. $\frac{f}{2}\partial_t p_s = -\frac{f}{2}\nabla \cdot (p_s \langle \mathbf{v} \rangle_s) - (\mathcal{P}_s \cdot \nabla) \cdot \langle \mathbf{v} \rangle_s - \nabla \cdot \mathbf{q}_s$	Unsucessful identification	**Sucessful identification** **Mean coeff. error $\approx$ 1%**

E. P. Alves and F. Fiuza, Physical Review Research (2022); arXiv:2011.01927

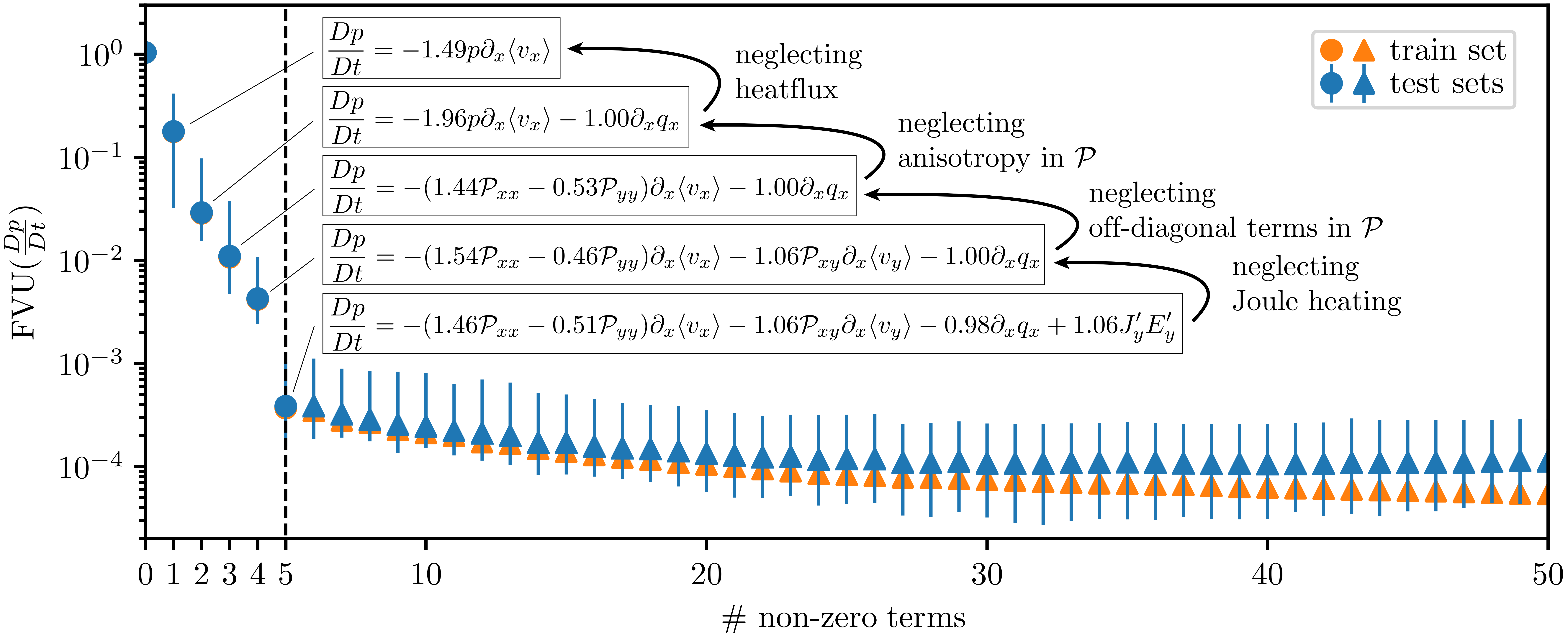
Recovery of the single-fluid (MHD) equations from dynamics of magnetized shocks

	Data	PDEs	PDE identification and inference accuracy	
			a) Point strategy	b) Integral strategy
Single-fluid equations (Magnetohydrodynamics)	**Collisionless magnetized shock dynamics**  **noise estimate: 5%**	**Continuity eq.** $\partial_t \rho_m = -\nabla \cdot (\rho_m \langle \mathbf{v} \rangle)$	Unsucessful identification	**Sucessful identification** **Mean coeff. error $\approx$ 1%**
		Momentum eq. $\partial_t (\rho_m \langle \mathbf{v} \rangle) = -\nabla \cdot (\rho_m \langle \mathbf{v}\mathbf{v} \rangle) + \rho_c \mathbf{E} + \mathbf{J} \times \mathbf{B}$	Unsucessful identification	**Sucessful identification** **Mean coeff. error $\approx$ 1%**
		Energy eq. $\frac{f}{2} \partial_t p = -\frac{f}{2} \nabla \cdot (p \langle \mathbf{v} \rangle) - (\mathcal{P} \cdot \nabla) \cdot \langle \mathbf{v} \rangle - \nabla \cdot \mathbf{q} + \mathbf{J}' \cdot \mathbf{E}'$	Unsucessful identification	**Sucessful identification** **Mean coeff. error $\approx$ 4%**

E. P. Alves and F. Fiuza, Physical Review Research (2022); arXiv:2011.01927

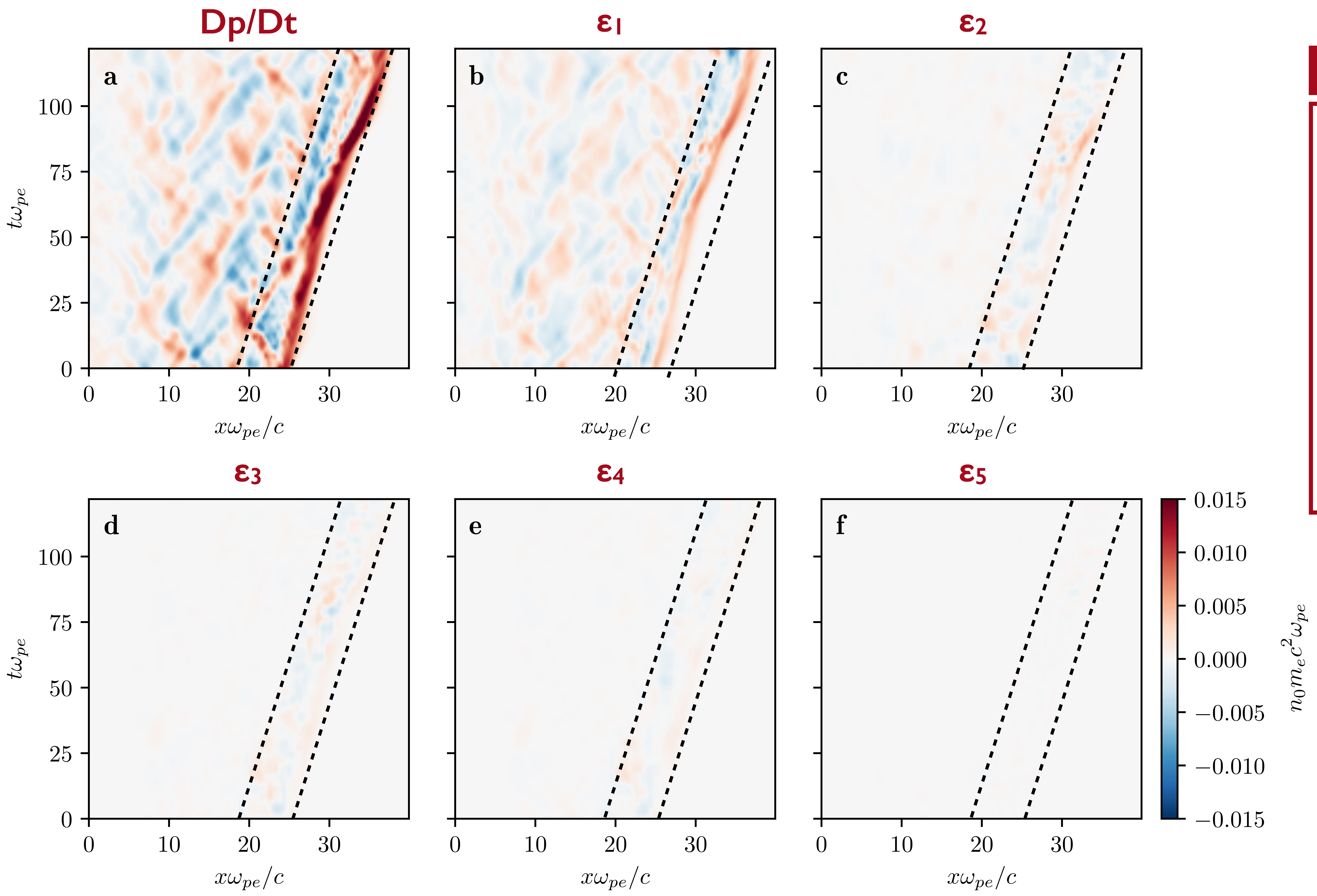
Hierarchy of Pareto-optimal models can be used to identify suitable approximations

Pareto analysis of the MHD energy equation from magnetized shock dynamics



E. P. Alves and F. Fiuza, Physical Review Research (2022); arXiv:2011.01927

Spatiotemporal model-error distribution can help diagnose important missing physics



Discrepancy errors of inferred models

$$\epsilon_1 = \frac{Dp}{Dt} - \left[-p\nabla\cdot\mathbf{v}\right]$$

$$\epsilon_2 = \frac{Dp}{Dt} - \left[-p\nabla\cdot\mathbf{v} - \nabla\cdot\mathbf{q}\right]$$

$$\epsilon_3 = \frac{Dp}{Dt} - \left[-p\nabla\cdot\mathbf{v} - (\mathscr{P}_D\cdot\nabla)\cdot\mathbf{v} - \nabla\cdot\mathbf{q}\right]$$

$$\epsilon_4 = \frac{Dp}{Dt} - \left[-p\nabla\cdot\mathbf{v} - (\mathscr{P}\cdot\nabla)\cdot\mathbf{v} - \nabla\cdot\mathbf{q}\right]$$

$$\epsilon_5 = \frac{Dp}{Dt} - \left[-p\nabla\cdot\mathbf{v} - (\mathscr{P}\cdot\nabla)\cdot\mathbf{v} - \nabla\cdot\mathbf{q} + \mathbf{J}'\cdot\mathbf{E}'\right]$$

E. P. Alves and F. Fiuza, Physical Review Research (2022); arXiv:2011.01927

Towards the discovery of new reduced models and improved fluid-closures

Further extensions to data-driven model discovery methodology

- Expanding library of basis terms for more expressive models
  - e.g. nonlocal (convolutional) terms, rational function nonlinearities
  - coupling with symbolic regression (Schmidt and Lipson, Science 2009; Udrescu and Tegmark, Sci. Adv. 2020)

- Identifying "good" observables to represent the dynamics of interest
  - using NNs to discover coordinates and SINDy to discover dynamics (K. Champion et al. PNAS, 2019)

- Embedding fundamental symmetries and plasma physics constraints
  - e.g. energy conservation (A. Kaptanoglu et al., PRE 2021)
  - embedding Lorentz-covariance (M. McGrae-Menge, J. Pierce and E. P. Alves, in preparation)

Theoretical analysis of discovered equations

- Stability properties; what physics is captured/lost?

- "Reverse-engineer" models it terms of more fundamental principles

Improve integration between simulation software and model discovery techniques

- Efficient ways of learning from large volumes of simulation data

- Verification and validation of data-driven models

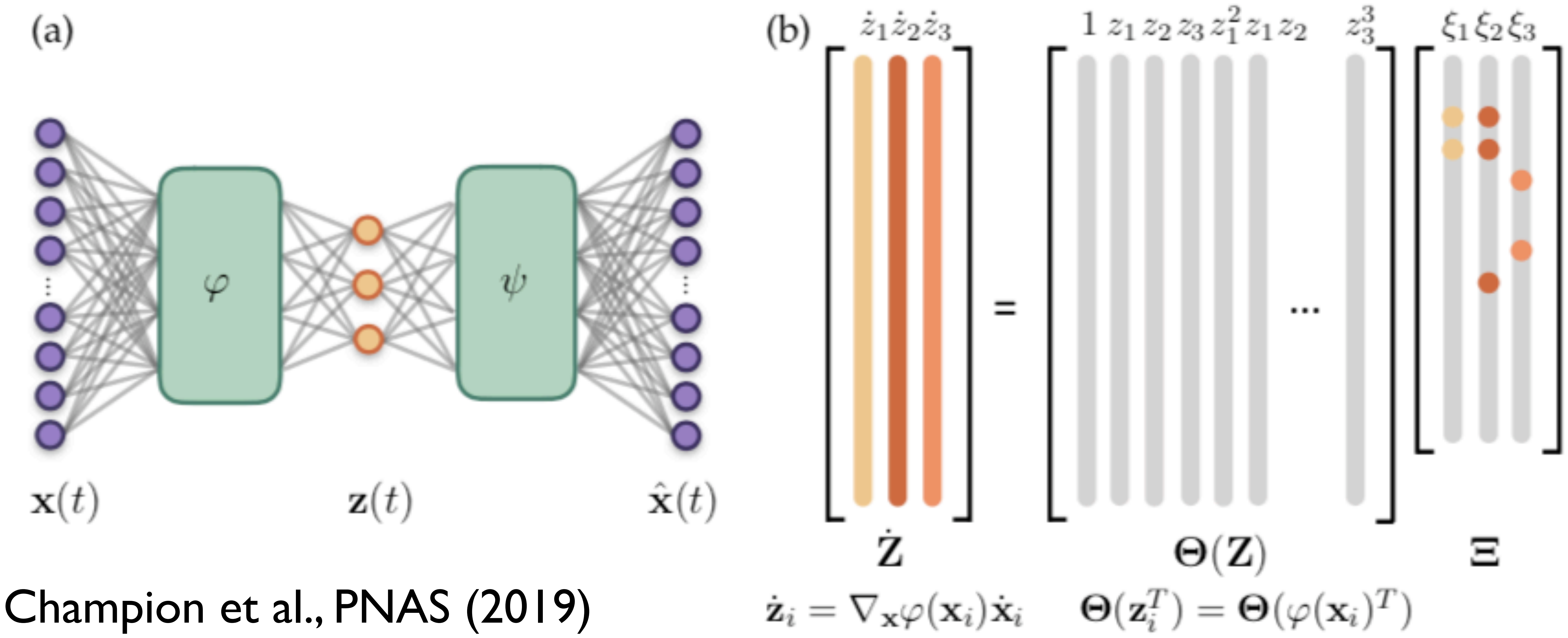
Towards the discovery of new reduced models and improved fluid-closures

Discovery of new "simple" reduced models requires good choice of basis terms

- Domain knowledge can often guide a suitable choice of basis terms
  - e.g. convolutional terms, and rational function nonlinearities, trigonometric nonlinearities

- Limited domain knowledge:
  - symbolic regression
  - deep neural networks

Using deep AEs to discover good coordinates/basis functions for sparse model identification



Champion et al., PNAS (2019)

This data-driven framework can be used to design computationally efficient algorithms for multi-scale modeling of nonlinear dynamical systems

High-fidelity numerical simulations

- Particle-in-cell
- Vlasov-Fokker-Planck
- Gyrokinetics
- Molecular dynamics
- …

Data-driven model discovery framework

Distilling improved reduced plasma models

- Kinetic-fluid closures
- Anomalous resistivity and transport
- Sub-grid models
- …

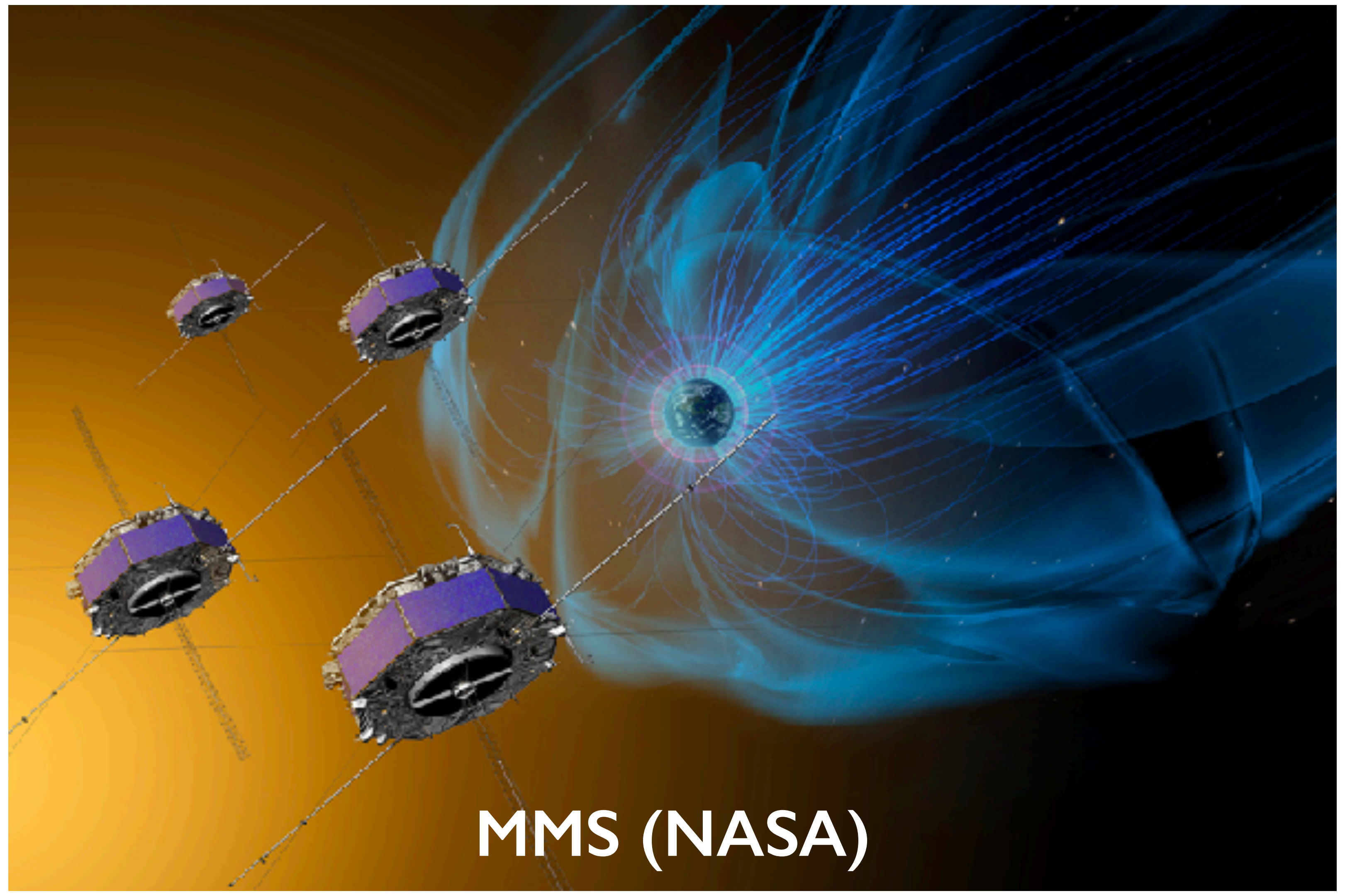
Integration into macroscopic models

Improved multi-scale modeling

- Whole-device modeling for fusion
- Global models of space/astrophysical plasmas

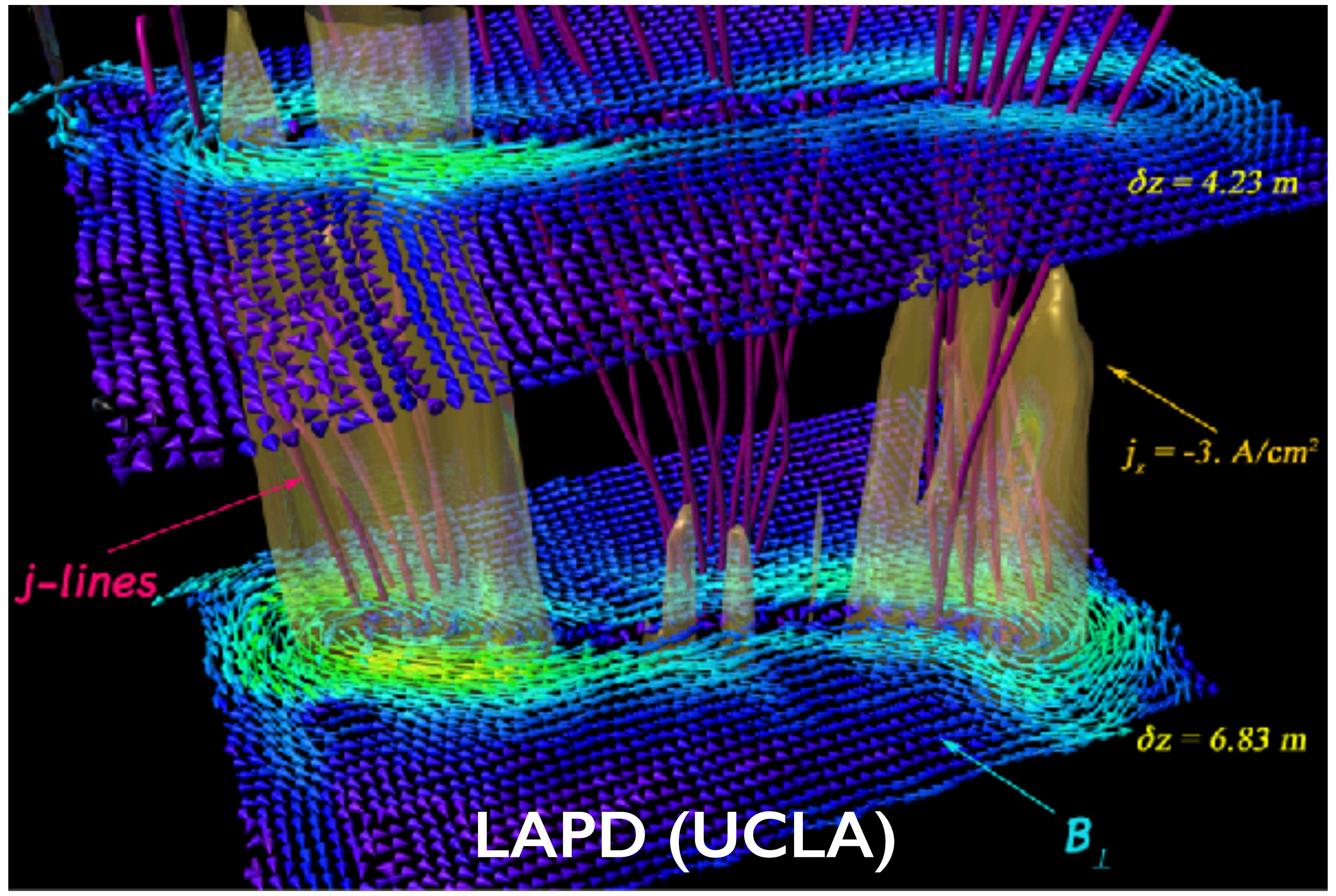
This methodology could also be applied to infer reduced models from noisy experimental data

Spacecraft observations



Burch et al. (2016)

Advanced experimental facilities



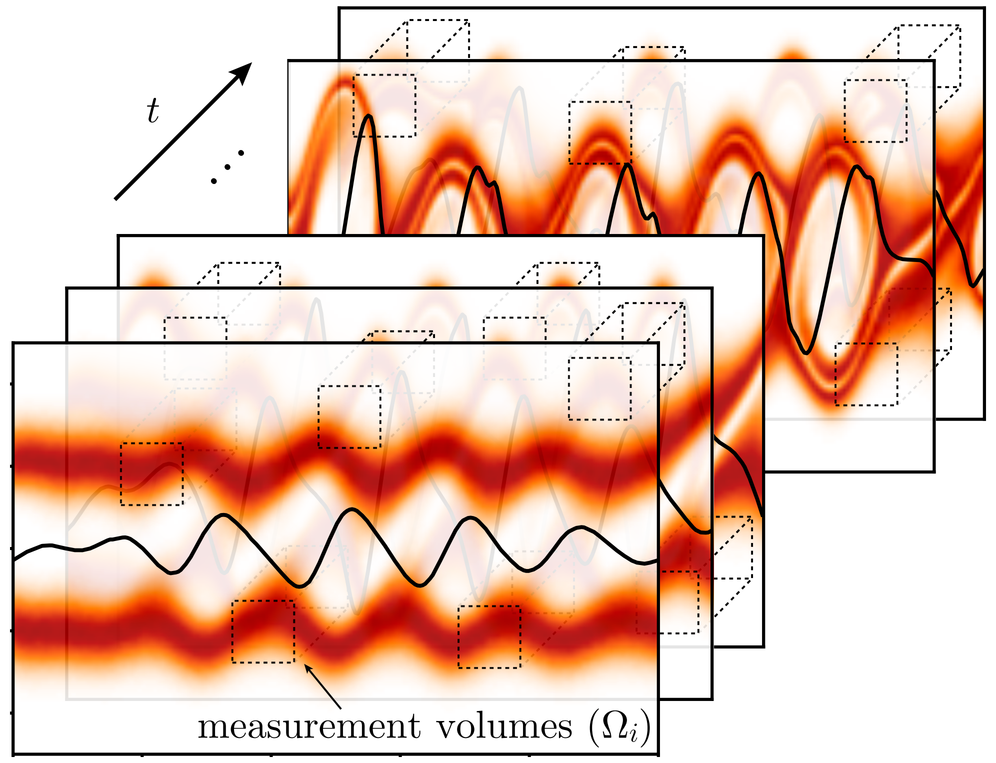
Gekelman et al. (2016)

Fletcher et al. (2015)

Conclusions

- Sparse regression is a viable approach for extracting interpretable and generalizable reduced models from the data of first-principles simulations

- The integral formulation of the PDE discovery procedure (SINDy) is crucial to robustly infer governing equations from noisy particle-based data



- The interpretable form of the inferred data-driven models promotes scientific insight into complex nonlinear dynamics, and can stimulate theoretical effort to "reverse engineer" these from lower-level frameworks

- This data-driven approach can be used to develop more computationally efficient algorithms for different multi-scale scientific applications