The Surprising Robustness and Computational Efficiency of Weak Form System Identification

David M. Bortz

Department of Applied Mathematics
University of Colorado

http://github.com/MathBioCU

Outline

Researchers

Dan Messenger

April Tran

Prof. Vanja Dukic

Nora Heitzman-Breen

Rainey Lyons

Modeling (post 1950)

Problem: Consider $u(x,t)$ which solves:

$$\partial_t u = \mathcal{A}(u)$$

with operator $\mathcal{A}$.

- Data $\mathbf{U} = u + \epsilon$, $\epsilon \sim \mathcal{N}(0, \sigma^2)$

Modeling (post 1950)

Problem: Consider $u(x,t)$ which solves:

$$\partial_t u = \mathcal{A}(u)$$

with operator $\mathcal{A}$.

- Data $\mathbf{U} = u + \epsilon$, $\epsilon \sim \mathcal{N}(0, \sigma^2)$
- Propose $\mathcal{A} = \mathcal{A}(\boldsymbol{w})$, $\boldsymbol{w} \in \mathbb{R}^n$, solve $\boldsymbol{w} \to u(\boldsymbol{w})$
- Numerical discretization $u^h(\boldsymbol{w}) \approx u(\boldsymbol{w})$

Modeling (post 1950)

Problem: Consider $u(x,t)$ which solves:

$$\partial_t u = \mathcal{A}(u)$$

with operator $\mathcal{A}$.

- Data $\mathbf{U} = u + \epsilon$, $\epsilon \sim \mathcal{N}(0, \sigma^2)$
- Propose $\mathcal{A} = \mathcal{A}(\boldsymbol{w})$, $\boldsymbol{w} \in \mathbb{R}^n$, solve $\boldsymbol{w} \to u(\boldsymbol{w})$
- Numerical discretization $u^h(\boldsymbol{w}) \approx u(\boldsymbol{w})$
- Assume $\epsilon = \mathbf{U} - u(\boldsymbol{w}^*) \sim \mathcal{N}(0, \sigma^2)$
- **Parameter Estimation:** Solve NLS problem

$$\min_{\mathbf{w}} \|\mathbf{U} - u^h(\mathbf{w})\|^2$$

$\leftrightarrow$ maximum likelihood estimation

Modeling (post 1950)

Problem: Consider $u(x,t)$ which solves:

$$\partial_t u = \mathcal{A}(u)$$

with operator $\mathcal{A}$.

- Data $\mathbf{U} = u + \epsilon$, $\epsilon \sim \mathcal{N}(0, \sigma^2)$
- Propose $\mathcal{A} = \mathcal{A}(\boldsymbol{w})$, $\boldsymbol{w} \in \mathbb{R}^n$, solve $\boldsymbol{w} \to u(\boldsymbol{w})$
- Numerical discretization $u^h(\boldsymbol{w}) \approx u(\boldsymbol{w})$
- Assume $\epsilon = \mathbf{U} - u(\boldsymbol{w}^*) \sim \mathcal{N}(0, \sigma^2)$
- **Parameter Estimation:** Solve NLS problem

  $$\min_{\mathbf{w}} \|\mathbf{U} - u^h(\mathbf{w})\|^2$$

  $\leftrightarrow$ maximum likelihood estimation
- **Model Selection:** How to select a best model? (t-test, F-test, KL, AIC, BIC, MDL, MML, ...)

Modeling

Problem: Consider $u(x,t)$ which solves:

$$\partial_t u = \mathcal{A}(u)$$

with operator $\mathcal{A}$.

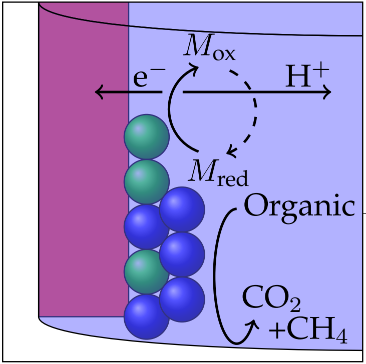
- Data $\mathbf{U} = u + \epsilon$, $\epsilon \sim \mathcal{N}(0, \sigma^2)$
- Propose $\mathcal{A} = \mathcal{A}(\boldsymbol{w})$ $\boldsymbol{w} \in \mathbb{R}^n$, solve $\boldsymbol{w} \to u(\boldsymbol{w})$
- Numerical discretization $u^h(\boldsymbol{w}) \approx u(\boldsymbol{w})$
- Assume $\epsilon = \mathbf{U} - u(\boldsymbol{w}^*) \sim \mathcal{N}(0, \sigma^2)$
- Solve NLS problem

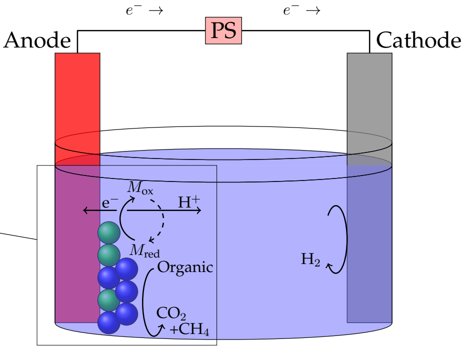
  $$\min_{\mathbf{w}} \|\mathbf{U} - u^h(\mathbf{w})\|^2$$

  $\leftrightarrow$ maximum likelihood estimation
- How to select a best model? (t-test, F-test, KL, AIC, BIC, MDL, MML, ...)

Microbial Fuel Cells (wrong-ish model)

Electroactive bacteria, Methanogenic Microorganims
M = Extracellular Mediator





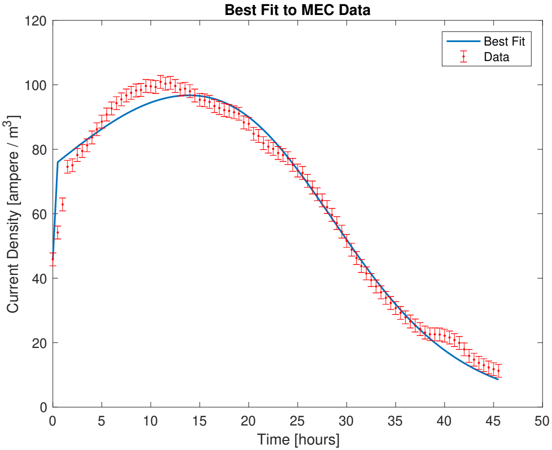
Substrate flows in, External Voltage applied, blue competes with green, decreasing hydrogen production

Microbial Fuel Cells (wrong-ish model)

$$
\begin{aligned}
\frac{dS}{dt} =& D[S_0 - S(t)] - q_e(t)X_e(t) - q_m(t)[X_{m,1}(t) + X_{m,2}(t)],\\
\frac{dX_{m,1}}{dt} =& [\mu_m(t) - K_{d,m} - D\alpha_1(t)]X_{m,1}(t),\\
\frac{dX_e}{dt} =& [\mu_e(t) - K_{d,e} - D\alpha_2(t)]X_e(t),\\
\frac{dX_{m,2}}{dt} =& [\mu_m(t) - K_{d,m} - D\alpha_2(t)]X_{m,2}(t),\\
\frac{dM_{\text{ox}}}{dt} =& -Y_M q_e(t)X_e(t) + \frac{\gamma}{VmF_2} I_{\text{MEC}}(t),\\
I_{\text{MEC}}(t)R_{\text{int}}(t) =& E_{\text{applied}} + E_{\text{CEMF}} - \frac{RT}{mF}\ln\left(\frac{M_{\text{total}}}{M_{\text{total}} - M_{\text{ox}}(t)}\right)\\
& - \frac{RT}{\beta mF}\text{arcsinh}\left(\frac{I_{\text{MEC}}(t)}{A_{\text{sur,A}} i_0}\right),
\end{aligned}
$$

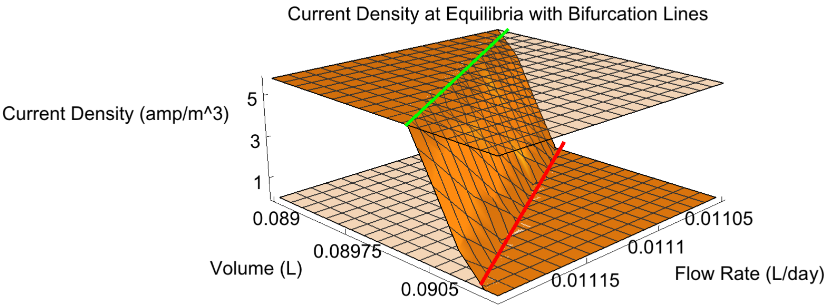
$$
I_{\text{density}}(t) = \frac{1000}{V} I_{\text{MEC}}(t), \qquad R_{\text{int}}(t) = R_{\text{min}} + (R_{\text{max}} - R_{\text{min}})e^{-K_R X_e(t)},
$$

Microbial Fuel Cells (wrong-ish model)



- Dudley, Lu, Ren, & DMB. Sensitivity and Bifurcation Analysis of a Differential-Algebraic Equation Model for a Microbial Electrolysis Cell. SIADS 18(2):709-728, 2019.
- Dudley, Ren, & DMB. Competitive Exclusion in a DAE Model for Microbial Electrolysis Cells. Math. Biosci. & Eng.17(5):6217-6239, 2020.

Microbial Fuel Cells (wrong-ish model)



- Dudley, Lu, Ren, & DMB. Sensitivity and Bifurcation Analysis of a Differential-Algebraic Equation Model for a Microbial Electrolysis Cell. SIADS 18(2):709-728, 2019.
- Dudley, Ren, & DMB. Competitive Exclusion in a DAE Model for Microbial Electrolysis Cells. Math. Biosci. & Eng.17(5):6217-6239, 2020.

Modeling

Problem: Consider $u(x,t)$ which solves:

$$\partial_t u = \mathcal{A}(u)$$

with operator $\mathcal{A}$.

- Data $\mathbf{U} = u + \epsilon$, $\epsilon \sim \mathcal{N}(0, \sigma^2)$
- Propose $\mathcal{A} = \mathcal{A}(\boldsymbol{w})$, $\boldsymbol{w} \in \mathbb{R}^n$, solve $\boldsymbol{w} \to u(\boldsymbol{w})$
- Numerical discretization $u^h(\boldsymbol{w}) \approx u(\boldsymbol{w})$
- Assume $\epsilon = \mathbf{U} - u(\boldsymbol{w}^*) \sim \mathcal{N}(0, \sigma^2)$
- Solve NLS problem

  $$\min_{\mathbf{w}} \|\mathbf{U} - u^h(\mathbf{w})\|^2$$

  $\leftrightarrow$ maximum likelihood estimation
- How to select a best model? (t-test, F-test, KL, AIC, BIC, MDL, MML, ...)

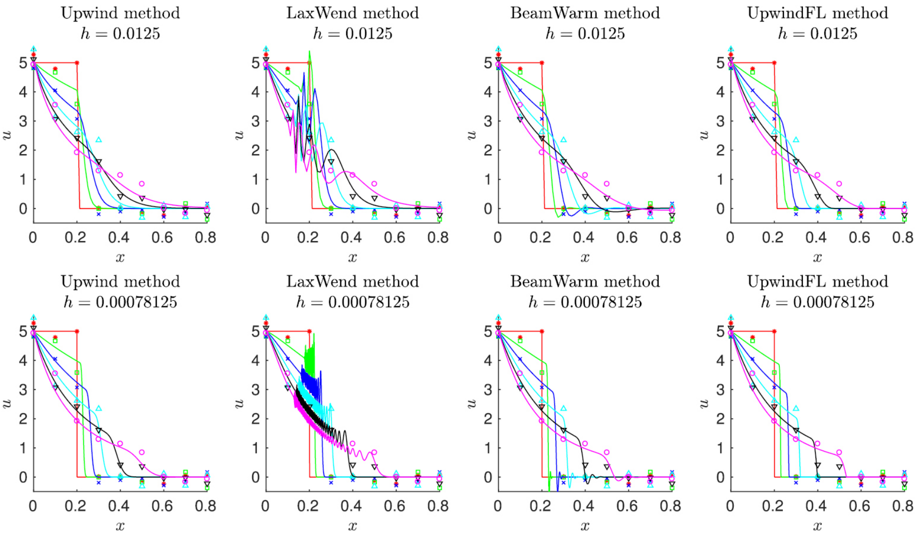
Numerical Error vs. Parameter Estimation

Consider a chemotaxis-like advection equation with parameters α and β

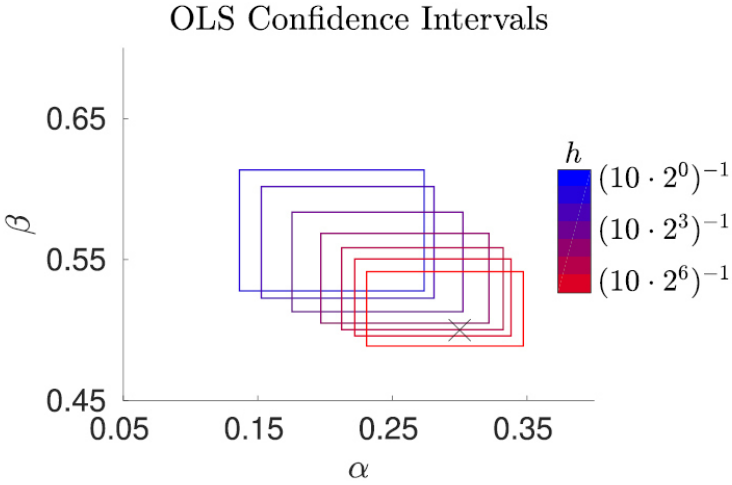
$$\begin{aligned} u_t(t,x) + \nabla \bullet \left(\left(\alpha \sqrt[\beta]{x} \right) u(t,x) \right) &= 0 \\ u(0,x) &= \phi(x) \end{aligned}$$

- Consider creating artificial data with i.i.d. Gaussian noise.
- Fit α and β using different numerical schemes

Numerical Error vs. Parameter Estimation



Numerical Error vs. Parameter Estimation



- Nardini & DMB. The Influence of Numerical Error on Parameter Estimation and Uncertainty Quantification for Advective PDE Models. Inverse Problems 35(6):065003, 2019

Model Selection and Parameter Inference

Problem: Consider $u(x,t)$ which solves:

$$\partial_t u = \mathcal{A}(u)$$

with operator $\mathcal{A}$.

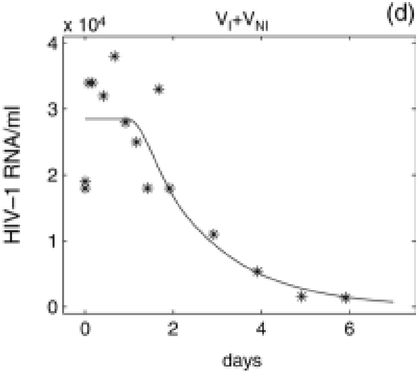
- Data $\mathbf{U} = u + \epsilon$, $\epsilon \sim \mathcal{N}(0, \sigma^2)$
- Propose $\mathcal{A} = \mathcal{A}(\boldsymbol{w})$, $\boldsymbol{w} \in \mathbb{R}^n$, solve $\boldsymbol{w} \to u(\boldsymbol{w})$
- Numerical discretization $u^h(\boldsymbol{w}) \approx u(\boldsymbol{w})$
- Assume $\epsilon = \mathbf{U} - u(\boldsymbol{w}^*) \sim \mathcal{N}(0, \sigma^2)$
- Solve NLS problem

  $$\min_{\mathbf{w}} \|\mathbf{U} - u^h(\mathbf{w})\|^2$$

  $\leftrightarrow$ maximum likelihood estimation
- How to select a best model?
  - t-test, F-test, KL, AIC, BIC, MDL, MML, ...

Model Selection

DMB & Nelson. Model Selection and Mixed-Effects Modeling of HIV Infection Dynamics. Bulletin of Mathematical Biology 68(8):2005-2025, 2006.



Model Selection

$$\text{AIC} = -2\ell(\widehat{\Theta}|\boldsymbol{V}) + 2Q$$

$$\text{TIC} = -2\ell(\widehat{\Theta}|\boldsymbol{V}) + 2\text{Tr}\left(J(\widehat{\Theta})I(\widehat{\Theta})^{-1}\right)$$

$$\text{ICOMP} = -2\ell(\widehat{\Theta}|\boldsymbol{V}) + Q\ln\left(\frac{\lambda_a}{\lambda_g}\right)$$

$$\text{Information Criteria} = \text{goodness of fit} + \text{statistical complexity}$$

Model Selection

Table 2 Sample mean values for the maximum likelihood estimates of the parameters, ICOMP(IFIM), TIC, AIC_C, and AIC. The bold type represents that method's prediction for "best" model.

Model	$\hat{c}$	$\hat{\delta}$	$\hat{\delta}_L$	$\hat{\tau}$	$\ell(\widehat{\Theta})$	ICOMP(IFIM)	TIC	AIC
(10)	13.7	0.35	–	–	−147.5	315.6	338.4	307.0
(11)	31.2	0.35	–	–	−137.3	**271.4**	20939	286.5
(12)	25.3	0.36	0.11	–	−131.2	354.7	971.2	282.4
(13)	24.7	3.6	0.38	–	−132.1	313.5	402.5	284.2
(14)	20.7	0.73	0.67	–	−104.6	329.4	**303.5**	**229.2**
(15)	16.7	0.36	–	0.7	−108.9	303.6	17305	304.5

Model Selection

$$T^* = \text{infectious T-cells}$$
$$V = \text{Virus Particles}$$
$$T = \text{Uninfected T-cells}$$

$$\begin{aligned}
\dot{T}^*(t) &= (1 - n_{rt})kV(t)T(t) - \delta T^*(t), \\
\dot{V}(t) &= N\delta T^*(t) - cV(t), \\
\dot{T}(t) &= S + pT(t)\left(1 - \frac{T(t) + T^*(t)}{T_{\max}}\right) - d_T T(t) - (1 - n_{rt})kV(t)T(t),
\end{aligned} \tag{11}$$

Data-Driven System Identification

Equation Errors: Shinbrot, Greenberg (NACA, 1950's)

- Consider a Differential Equation

$$\partial_t x = w_1 f_1(x) + w_2 f_2(x) + \cdots$$

Equation Errors: Shinbrot, Greenberg (NACA, 1950's)

- Consider a Differential Equation

$$\partial_t x = w_1 f_1(x) + w_2 f_2(x) + \cdots$$

- Plug in data $\boldsymbol{U}$!

$$\boldsymbol{b} := \left[\begin{array}{cccc} \widetilde{\partial}_t U_1 & \widetilde{\partial}_t U_2 & \cdots & \widetilde{\partial}_t U_M \end{array}\right]^T$$

$$\boldsymbol{G} := [f_k(U_i)]_{i,k} = \left[\begin{array}{cccc} f_1(U_1) & f_2(U_1) & \cdots & f_K(U_1) \\ f_1(U_2) & f_2(U_2) & \ddots & \\ \vdots & \ddots & \ddots & \\ f_1(U_M) & \cdots & & f_K(U_M) \end{array}\right]$$

Equation Errors: Shinbrot, Greenberg (NACA, 1950's)

- Consider a Differential Equation

$$\partial_t x = w_1 f_1(x) + w_2 f_2(x) + \cdots$$

- Plug in data $\boldsymbol{U}$!

$$\boldsymbol{b} := \begin{bmatrix} \widetilde{\partial}_t U_1 & \widetilde{\partial}_t U_2 & \cdots & \widetilde{\partial}_t U_M \end{bmatrix}^T$$

$$\boldsymbol{G} := \left[f_k(U_i)\right]_{i,k} = \begin{bmatrix} f_1(U_1) & f_2(U_1) & \cdots & f_K(U_1) \\ f_1(U_2) & f_2(U_2) & \ddots & \\ \vdots & \ddots & \ddots & \\ f_1(U_M) & \cdots & & f_K(U_M) \end{bmatrix}$$

- Plug in data $\boldsymbol{U}$ and solve Equation-Error linear least squares problem

$$\min_{\boldsymbol{w}} \|\boldsymbol{G}\boldsymbol{w} - \boldsymbol{b}\|_2^2$$

Equation Errors

- Varah, J.M., SIAM J. Sci. & Stat. Comput. 3(1), 28–46 (1982).
- Poyton, A., Varziri, M., McAuley, K., McLellan, P., Ramsay, J., Comp. & Chem. Eng. 30(4), 698–708 (2006).
- Ramsay, J.O., Hooker, G., Campbell, D., Cao, J., JRSSB Stat. Method. 69(5), 741–796 (2007).
- Brunel, N.J.B., Electron. J. Stat. 2(0), 1242–1267 (2008)
- Liang, H., Wu, H., JASA 103(484), 1570–1583 (2008)
- Ding, A.A., Wu, H., Stat. Sin. 24(4), 1613–1631 (2014).
- Wang, H., Zhou, X., Int. J. UQ 11(4), 41–57 (2021).
- Weak/integral form
  - Brunel, N.J.B., Clairon, Q., d'Alché-Buc, F., JASA 109(505), 173–185 (2014).
  - Liu, Z., Xu, J., Int. J. Sys. Sci. 49(5), 908–919 (2018).
  - Dattner, I., WIREs Comp. Stat. 13(6) (2021).

SINDy: Brunton, Proctor, Kutz, PNAS (2016)

Consider a Differential Equation

$$\partial_t u = \Theta(u)\mathbf{w}^\star \tag{1}$$

with an uber-model matrix

$$\Theta(u) = \begin{bmatrix} | & | & | & | & | & | \\ u & \partial_x u & \partial_{xx} u & \dots & u\partial_x u & u\partial_{xx} u \dots \\ | & | & | & | & | & | \end{bmatrix},$$

1. **evaluate (1) at U**

SINDy: Brunton, Proctor, Kutz, PNAS (2016)

Consider a Differential Equation

$$\partial_t u = \Theta(u)\mathbf{w}^\star \tag{1}$$

with an uber-model matrix

$$\Theta(u) = \begin{bmatrix} | & | & | & | & | & | \\ u & \partial_x u & \partial_{xx} u & \dots & u\partial_x u & u\partial_{xx} u \dots \\ | & | & | & | & | & | \end{bmatrix},$$

1. **evaluate (1) at U**
2. **solve a sparse regression problem for weights $\mathbf{w}^\star$:**

$$\widehat{\mathbf{w}} = \mathrm{argmin}_{\mathbf{w}} \left\|\Theta(\mathbf{U})\mathbf{w} - \partial_t \mathbf{U}\right\|_2^2 + \lambda \left\|\mathbf{w}\right\|_0$$

SINDy: Brunton, Proctor, Kutz, PNAS (2016)

Consider a Differential Equation

$$\partial_t u = \Theta(u)\mathbf{w}^\star \tag{1}$$

with an uber-model matrix

$$\Theta(u) = \begin{bmatrix} | & | & | & | & | & | \\ u & \partial_x u & \partial_{xx} u & \dots & u\partial_x u & u\partial_{xx} u \dots \\ | & | & | & | & | & | \end{bmatrix},$$

1. **evaluate (1) at U**
2. **solve a sparse regression problem for weights $\mathbf{w}^\star$:**

$$\widehat{\mathbf{w}} = \text{argmin}_{\mathbf{w}} \left\|\Theta(\mathbf{U})\mathbf{w} - \partial_t \mathbf{U}\right\|_2^2 + \lambda \left\|\mathbf{w}\right\|_0$$

PROBLEM!

- Approximating derivatives is unstable:

$$\mathbb{E}\left[\left|\partial_x^k u - \Delta_x^k (u+\epsilon)\right|^2\right] = \mathcal{O}\left(\frac{\sigma^2}{h^{2k}}\right)$$

- True solution $u(x,t)$ may not be smooth

Weak form SINDy (WSINDy)

Consider

$$\partial_t u = \Theta(u)\mathbf{w}^\star \rightarrow \boxed{D^{\alpha^0} u = \sum_{s,j=1}^{S,J} \mathbf{w}^\star_{s,j} D^{\alpha^s} f_j(u)} \quad (2)$$

- Interpret PDE in a weak sense: $\forall \psi \in C_c^{|\boldsymbol{\alpha}|}$,

$$\left\langle (-1)^{|\alpha^0|} D^{\alpha^0} \psi, u \right\rangle = \sum_{s,j=1}^{S,J} \mathbf{w}^\star_{s,j} \left\langle (-1)^{|\alpha^s|} D^{\alpha^s} \psi, f_j(u) \right\rangle \quad (3)$$

Weak form SINDy (WSINDy)

Consider

$$\partial_t u = \Theta(u)\mathbf{w}^\star \rightarrow \boxed{D^{\alpha^0} u = \sum_{s,j=1}^{S,J} \mathbf{w}^\star_{s,j} D^{\alpha^s} f_j(u)} \quad (2)$$

- Interpret PDE in a weak sense: $\forall \psi \in C_c^{|\boldsymbol{\alpha}|}$,

$$\left\langle (-1)^{|\alpha^0|} D^{\alpha^0}\psi, u \right\rangle = \sum_{s,j=1}^{S,J} \mathbf{w}^\star_{s,j} \left\langle (-1)^{|\alpha^s|} D^{\alpha^s}\psi, f_j(u) \right\rangle \quad (3)$$

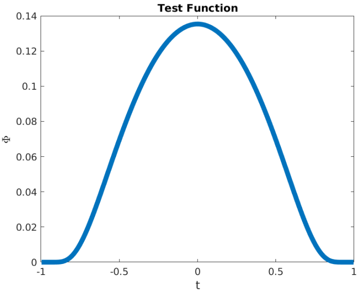
- For test functions $(\psi_k)_{1\leq k \leq K}$, **evaluate** (3) **at $\mathbf{U}$**

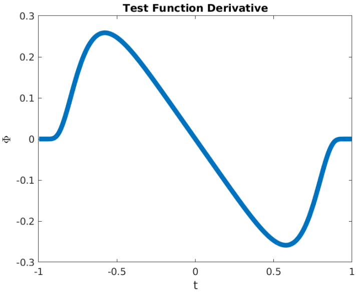
$$\mathbf{b}_k := \left\langle (-1)^{|\alpha^0|} D^{\alpha^0}\psi_k, \mathbf{U} \right\rangle, \quad \mathbf{G}_{k,(j-1)S+s} := \left\langle (-1)^{|\alpha^s|} D^{\alpha^s}\psi_k, f_j(\mathbf{U}) \right\rangle$$

- Solve sparse regression problem for weights $\mathbf{w}^\star$:

$$\widehat{\mathbf{w}} = \text{argmin}_{\mathbf{w}} \left\| \mathbf{G}\mathbf{w} - \mathbf{b} \right\|_2^2 + \lambda^2 \left\| \mathbf{w} \right\|_0$$

Weak form Parameter Estimation





- Equations of Motion method: Shinbrot, NACA TN 3288, (1954).
- Modulating Function method: Loeb & Cahen, IEEE TAC, (1965).
- Preisig & Rippin. Theory & App. of the Modulating Function Method-{I,II,III}, Comp. & Chem. Eng., (1993).

Weak form SINDy (WSINDy)

- Pantazis & Tsamardinos. A Unified Approach for Sparse Dynamical System Inference from Temporal Measurements. Bioinformatics 35(18):3387–96, 2019.
- Wang, Huan, & Garikipati. Variational System Identification of the Partial Differential Equations Governing the Physics of Pattern-Formation: Inference under Varying Fidelity and Noise. CMAME 356:44-74, 2019.
- Gurevich, Reinbold, & Grigoriev. Robust and Optimal Sparse Regression for Nonlinear PDE Models. Chaos 29(10):103113, 2019.

Weak form system identificaiton

ODE DAM & DMB. WSINDy for ODEs. MMS 2021

Weak form system identificaiton

ODE DAM & DMB. WSINDy for ODEs. MMS 2021

PDE DAM & DMB. WSINDy for PDEs,. JCP, 2021.

Weak form system identificaiton

ODE DAM & DMB. WSINDy for ODEs. MMS 2021

PDE DAM & DMB. WSINDy for PDEs,. JCP, 2021.

IPS (order 1) DAM & DMB. WSINDy for SDEs, Physica D, 2022.

Weak form system identificaiton

ODE DAM & DMB. WSINDy for ODEs. MMS 2021

PDE DAM & DMB. WSINDy for PDEs,. JCP, 2021.

IPS (order 1) DAM & DMB. WSINDy for SDEs, Physica D, 2022.

Online DAM, Dall'Anese, & DMB. Online WSINDy, MSML, 2022.

Weak form system identificaiton

ODE DAM & DMB. WSINDy for ODEs. MMS 2021

PDE DAM & DMB. WSINDy for PDEs,. JCP, 2021.

IPS (order 1) DAM & DMB. WSINDy for SDEs, Physica D, 2022.

Online DAM, Dall'Anese, & DMB. Online WSINDy, MSML, 2022.

IPS (order 2) DAM, Wheeler, Liu, & DMB. WSINDy for migration, JRS Interface, 2022.

Weak form system identificaiton

ODE	DAM & DMB. WSINDy for ODEs. MMS 2021
PDE	DAM & DMB. WSINDy for PDEs,. JCP, 2021.
IPS (order 1)	DAM & DMB. WSINDy for SDEs, Physica D, 2022.
Online	DAM, Dall'Anese, & DMB. Online WSINDy, MSML, 2022.
IPS (order 2)	DAM, Wheeler, Liu, & DMB. WSINDy for migration, JRS Interface, 2022.
Theory	DAM & DMB, Asymptotics..., arXiv:2211.16000

Weak form system identificaiton

ODE DAM & DMB. WSINDy for ODEs. MMS 2021

PDE DAM & DMB. WSINDy for PDEs,. JCP, 2021.

IPS (order 1) DAM & DMB. WSINDy for SDEs, Physica D, 2022.

Online DAM, Dall'Anese, & DMB. Online WSINDy, MSML, 2022.

IPS (order 2) DAM, Wheeler, Liu, & DMB. WSINDy for migration, JRS Interface, 2022.

Theory DAM & DMB, Asymptotics..., arXiv:2211.16000

ParEst DMB, DAM, VD, WENDy, Bull. Math. Biol., Oct., 2023

Weak form system identificaiton

- ODE — DAM & DMB. WSINDy for ODEs. MMS 2021
- PDE — DAM & DMB. WSINDy for PDEs,. JCP, 2021.
- IPS (order 1) — DAM & DMB. WSINDy for SDEs, Physica D, 2022.
- Online — DAM, Dall'Anese, & DMB. Online WSINDy, MSML, 2022.
- IPS (order 2) — DAM, Wheeler, Liu, & DMB. WSINDy for migration, JRS Interface, 2022.
- Theory — DAM & DMB, Asymptotics..., arXiv:2211.16000
- ParEst — DMB, DAM, VD, WENDy, Bull. Math. Biol., Oct., 2023
- ROM#1 — AT, He, DAM, Choi, & DMB, WLaSDI, CMAME, 2024

Weak form system identificaiton

ODE	DAM & DMB. WSINDy for ODEs. MMS 2021
PDE	DAM & DMB. WSINDy for PDEs,. JCP, 2021.
IPS (order 1)	DAM & DMB. WSINDy for SDEs, Physica D, 2022.
Online	DAM, Dall'Anese, & DMB. Online WSINDy, MSML, 2022.
IPS (order 2)	DAM, Wheeler, Liu, & DMB. WSINDy for migration, JRS Interface, 2022.
Theory	DAM & DMB, Asymptotics..., arXiv:2211.16000
ParEst	DMB, DAM, VD, WENDy, Bull. Math. Biol., Oct., 2023
ROM#1	AT, He, DAM, Choi, & DMB, WLaSDI, CMAME, 2024
ROM#2	DAM, Burby, DMB, Coarse-graining with WSINDy, SciRep 2024 (revised)

Weak form system identificaiton

ODE	DAM & DMB. WSINDy for ODEs. MMS 2021
PDE	DAM & DMB. WSINDy for PDEs,. JCP, 2021.
IPS (order 1)	DAM & DMB. WSINDy for SDEs, Physica D, 2022.
Online	DAM, Dall'Anese, & DMB. Online WSINDy, MSML, 2022.
IPS (order 2)	DAM, Wheeler, Liu, & DMB. WSINDy for migration, JRS Interface, 2022.
Theory	DAM & DMB, Asymptotics..., arXiv:2211.16000
ParEst	DMB, DAM, VD, WENDy, Bull. Math. Biol., Oct., 2023
ROM#1	AT, He, DAM, Choi, & DMB, WLaSDI, CMAME, 2024
ROM#2	DAM, Burby, DMB, Coarse-graining with WSINDy, SciRep 2024 (revised)
RKHS	Russo, DAM, DMB, & Rosenfeld, submitted, 2024

Weak form system identificaiton

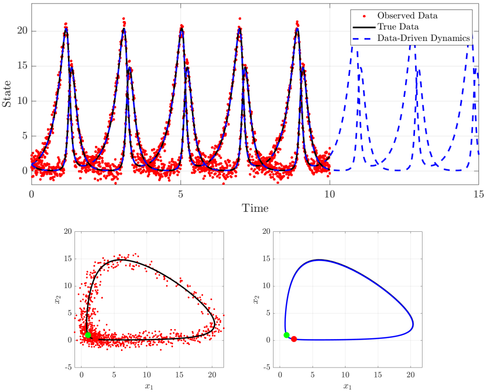
ODE	DAM & DMB. WSINDy for ODEs. MMS 2021
PDE	DAM & DMB. WSINDy for PDEs,. JCP, 2021.
IPS (order 1)	DAM & DMB. WSINDy for SDEs, Physica D, 2022.
Online	DAM, Dall'Anese, & DMB. Online WSINDy, MSML, 2022.
IPS (order 2)	DAM, Wheeler, Liu, & DMB. WSINDy for migration, JRS Interface, 2022.
Theory	DAM & DMB, Asymptotics..., arXiv:2211.16000
ParEst	DMB, DAM, VD, WENDy, Bull. Math. Biol., Oct., 2023
ROM#1	AT, He, DAM, Choi, & DMB, WLaSDI, CMAME, 2024
ROM#2	DAM, Burby, DMB, Coarse-graining with WSINDy, SciRep 2024 (revised)
RKHS	Russo, DAM, DMB, & Rosenfeld, submitted, 2024
MHD	Vasey et al., 2023, arXiv:2312.05339

Weak form system identificaiton

ODE	DAM & DMB. WSINDy for ODEs. MMS 2021
PDE	DAM & DMB. WSINDy for PDEs,. JCP, 2021.
IPS (order 1)	DAM & DMB. WSINDy for SDEs, Physica D, 2022.
Online	DAM, Dall'Anese, & DMB. Online WSINDy, MSML, 2022.
IPS (order 2)	DAM, Wheeler, Liu, & DMB. WSINDy for migration, JRS Interface, 2022.
Theory	DAM & DMB, Asymptotics..., arXiv:2211.16000
ParEst	DMB, DAM, VD, WENDy, Bull. Math. Biol., Oct., 2023
ROM#1	AT, He, DAM, Choi, & DMB, WLaSDI, CMAME, 2024
ROM#2	DAM, Burby, DMB, Coarse-graining with WSINDy, SciRep 2024 (revised)
RKHS	Russo, DAM, DMB, & Rosenfeld, submitted, 2024
MHD	Vasey et al., 2023, arXiv:2312.05339
Missing variables	DMB & DAM, submitted, 2024

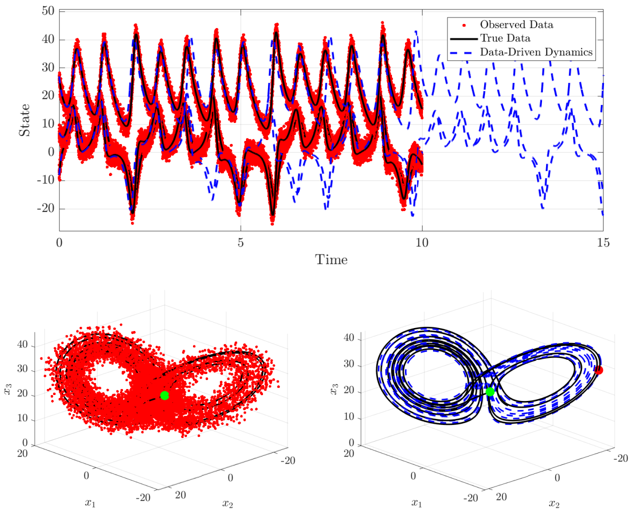
Examples

Lotka-Volterra 10% noise $\|\widehat{\mathbf{w}} - \mathbf{w}^{\star}\|_2 / \|\mathbf{w}^{\star}\|_2 = 0.0013$

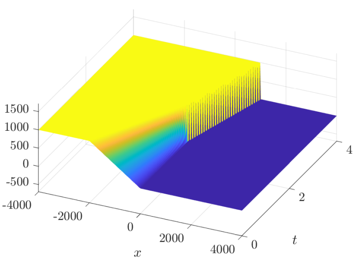


D. A. Messenger and D. M. Bortz. Weak SINDy: Galerkin-Based Data-Driven Model Selection. Multiscale Model. Simul., 19(3):1474–1497, 2021.

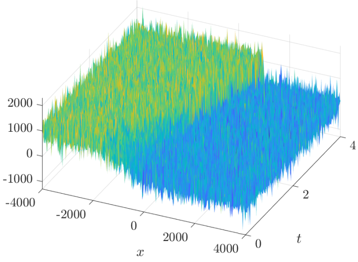
Lorenz 10% noise $\|\widehat{\mathbf{w}} - \mathbf{w}^\star\|_2 \,/\, \|\mathbf{w}^\star\|_2 = 0.0084$



Inviscid Burgers: $\partial_t u = -\partial_x \left(u^2\right)$



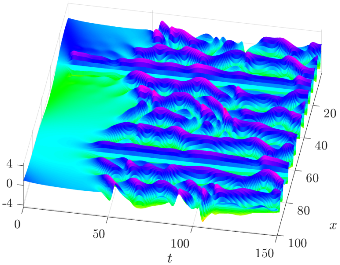
Noise-free



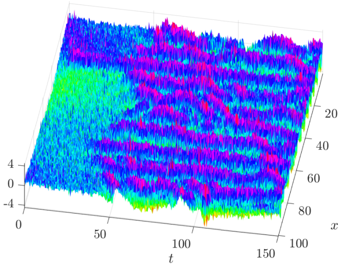
50% Noise

Noise level	Max. Coefficient Error	Identification Rate
0%	4.3×10^{-5}	100%
25%	0.0051	100%
50%	0.012	99.4%
100%	0.025	99%

Kuramoto Sivashinsky $u_t = -(u^2)_x - u_{xx} - u_{xxxx}$



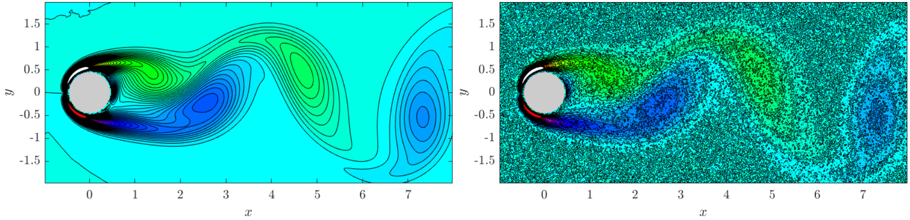
Noise-free



50% Noise

Noise level	Max. Coefficient Error	Identification Rate
0%	8.1×10^{-7}	100%
25%	0.017	100%
50%	0.070	100%
100%	0.31	96.1%

Navier-Stokes $\omega_t = -\nabla \bullet (\omega \mathbf{u}) + 0.01\Delta\omega$

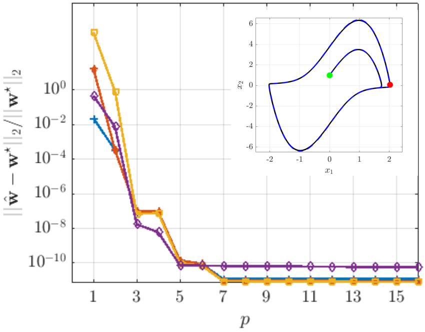


Noise level	Max. Coefficient Error	Identification Rate
0%	5.4×10^{-3}	100%
10%	7.8×10^{-3}	100%
20%	0.014	100%
30%	0.25	83.5%

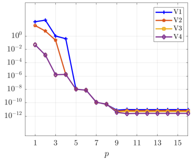
WSINDy

Test Functions

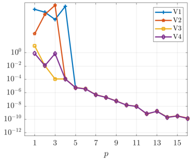
Van der Pol Oscillator ($\Delta t = 0.01$)



Linear 5D ($\Delta t = 0.025$)



Pendulum ($\Delta t = 0.1$)



$$\phi(v; a, p) = \left(1 - \left(\frac{v}{a}\right)^2\right)_+^p$$

- $\phi \in C^{p-1}(\mathbb{R})$, $\text{supp}(\phi) = [-a, a]$
- Trapezoidal rule + large p realizes machine precision:

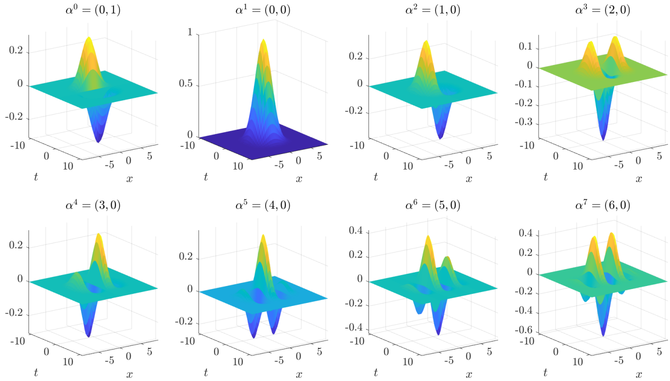
$$\langle \phi', \mathbf{x} \rangle + \langle \phi, \mathbf{F}(\mathbf{x}) \rangle = \mathcal{O}(\Delta t^{p+1})$$

$$(\sigma_{NR} = 0)$$

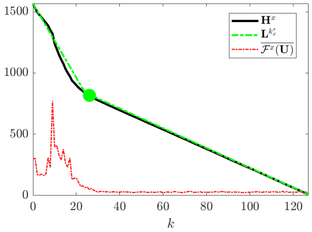
Test Functions

$$\phi(\bullet; a, p) = \left(1 - \left(\frac{\bullet}{a}\right)^2\right)_+^p \tag{4}$$

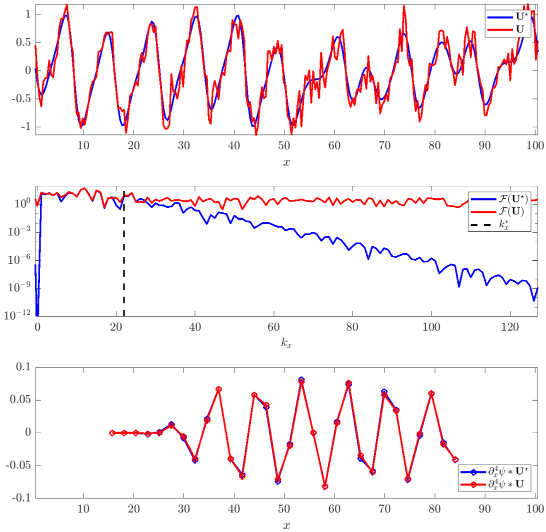
- $\phi \in C^{p-1}(\mathbb{R})$, $\text{supp}\,(\phi) = [-a, a] \to \psi = \prod \phi_i$



Test Functions



1. Find changepoint $\boldsymbol{k^*}$ between *signal-dominated* and *noise-dominated* modes
2. Set $\boldsymbol{k^*}$ to be $\hat{\tau}$ standard deviations into tail of $\hat{\phi}$.
3. Enforce ϕ decays to τ at penultimate gridpoint.



Weak 4th-derivative from KS data with 50% noise.

FFT Speedups

- Reference test function: $\psi_k(\mathbf{x}, t) := \psi(\mathbf{x}_k - \mathbf{x}, t_k - t)$
- Query points: $\mathcal{Q} := \{(\mathbf{x}_k, t_k)\}_{k \in [K]}$
- Convolutional Weak Form:

$$\left(D^{\alpha^0}\psi\right) * u(\mathbf{x}_k, t_k) = \sum_{s,j}^{S,J} \mathbf{w}^\star_{s,j} \left(D^{\alpha^s}\psi\right) * f_j(u)(\mathbf{x}_k, t_k). \tag{5}$$

FFT Speedups

- Reference test function: $\psi_k(\mathbf{x}, t) := \psi(\mathbf{x}_k - \mathbf{x}, t_k - t)$
- Query points: $\mathcal{Q} := \{(\mathbf{x}_k, t_k)\}_{k\in[K]}$
- Convolutional Weak Form:

$$\left(D^{\alpha^0}\psi\right) * u(\mathbf{x}_k, t_k) = \sum_{s,j}^{S,J} \mathbf{w}^\star_{s,j} \left(D^{\alpha^s}\psi\right) * f_j(u)(\mathbf{x}_k, t_k). \tag{5}$$

- Take ψ separable, $\psi(\mathbf{x}, t) = \phi_1(x_1)\cdots\phi_D(x_D)\phi_{D+1}(t)$, use FFT:

$$\Psi^s * f_j(\mathbf{U}) = \mathcal{P}^{\mathcal{Q}}\mathcal{F}^{-1}\left(\mathcal{F}(\Psi^s) \odot \mathcal{F}(f_j(\mathbf{U}))\right), \tag{6}$$

  At-worst $\mathcal{O}(M\log(N))$ complexity for $M = N^{D+1}$ datapoints

- Local vs. Global Test Functions

Sparse Regression

$$\widehat{\mathbf{w}} = \text{argmin}_{\mathbf{w}} \left\|\mathbf{G}\mathbf{w} - \mathbf{b}\right\|_2^2 + \lambda^2 \left\|\mathbf{w}\right\|_0$$

- Sequential thresholding least squares (STLS):
$$\mathbf{w}^{(n+1)} = H_\lambda \left((\mathbf{G}^{(n)})^\dagger \mathbf{b} \right)$$

Sparse Regression

$$\widehat{\mathbf{w}} = \text{argmin}_{\mathbf{w}} \left\|\mathbf{G}\mathbf{w} - \mathbf{b}\right\|_2^2 + \lambda^2 \left\|\mathbf{w}\right\|_0$$

- Sequential thresholding least squares (STLS):

$$\mathbf{w}^{(n+1)} = H_\lambda\left((\mathbf{G}^{(n)})^\dagger \mathbf{b}\right)$$

- Modify to enforce dominant balance (MSTLS):

$$\mathbf{w}^* = H_\lambda\left((\mathbf{G}^{(n)})^\dagger \mathbf{b}\right)$$

$$\mathbf{w}_i^{(n+1)} = 0 \quad \text{if} \quad \frac{\left\|\mathbf{G}_i \mathbf{w}_i^*\right\|_2}{\left\|\mathbf{b}\right\|_2} \notin [\lambda, \lambda^{-1}]$$

Sparse Regression

$$\widehat{\mathbf{w}} = \text{argmin}_{\mathbf{w}} \|\mathbf{G}\mathbf{w} - \mathbf{b}\|_2^2 + \lambda^2 \|\mathbf{w}\|_0$$

- Sequential thresholding least squares (STLS):

$$\mathbf{w}^{(n+1)} = H_\lambda\left((\mathbf{G}^{(n)})^\dagger \mathbf{b}\right)$$

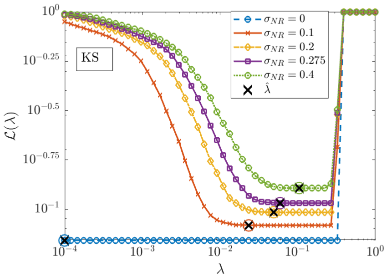
- Modify to enforce dominant balance (MSTLS):

$$\mathbf{w}^* = H_\lambda\left((\mathbf{G}^{(n)})^\dagger \mathbf{b}\right)$$

$$\mathbf{w}_i^{(n+1)} = 0 \quad \text{if} \quad \frac{\|\mathbf{G}_i \mathbf{w}_i^*\|_2}{\|\mathbf{b}\|_2} \notin [\lambda, \lambda^{-1}]$$

- Introduce loss:

$$\mathcal{L}(\lambda) = \frac{\|\mathbf{G}(\mathbf{w}^\lambda - \mathbf{w}^0)\|_2}{\|\mathbf{G}\mathbf{w}^0\|_2} + \frac{\|\mathbf{w}^\lambda\|_0}{\|\mathbf{w}^0\|_0}$$

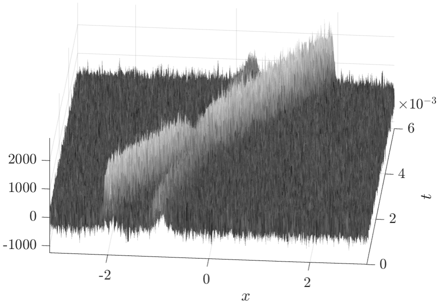


Minimize $\mathcal{L}$ over MSTLS solutions, *finds feasible λ under large noise!*

Rescaling

$$\widetilde{u}(\widetilde{x},\widetilde{t}) := \gamma_u\, u\left(\frac{\widetilde{x}}{\gamma_x}, \frac{\widetilde{t}}{\gamma_t}\right)$$

- Physical laws often involve power laws
  $\implies$ large $\kappa(\mathbf{G})$
- Rescaling is critical for sparse regression
- Choose $\gamma_x, \gamma_t, \gamma_u$ automatically from Θ and ψ to improve $\kappa(\mathbf{G})$



Korteweg-de Vries:

- 50% noise
- $u \sim 10^3$, $\Delta x \sim 1$, $\Delta t \sim 10^{-6}$
- $\kappa(\mathbf{G}) = \mathcal{O}(10^{26})$
- $\kappa(\widetilde{\mathbf{G}}) = \mathcal{O}(10^{6})$

WSINDy Computational Cost

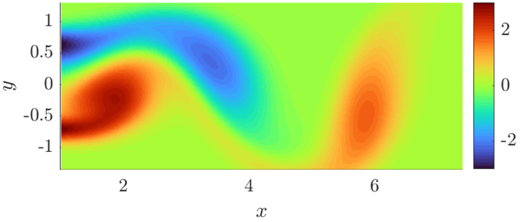
PDE	$\mathbf{U}$ (MB)	$\mathbf{U}$ (dim)	$\mathbf{G}$ (dim)	nz($\mathbf{w}^\star$)	Walltime (s)
IB	0.5	256×256	784×43	1	0.12
KS	0.6	256×301	$1,806 \times 43$	3	0.24
KdV	1.9	400×601	$1,443 \times 43$	2	0.39
NLS	1.0	$2 \times 256 \times 251$	$1,804 \times 190$	6	2.5
NS	**233**	$\mathbf{3 \times 324 \times 149 \times 201}$	$\mathbf{3,872 \times 50}$	**4**	**12**
PM	41	$200 \times 200 \times 128$	$4,608 \times 65$	3	16
SG	85	$129 \times 403 \times 205$	$13,000 \times 73$	3	29
RD	211	$2 \times 256 \times 256 \times 201$	$11,638 \times 181$	14	75

Table: Computations in serial, 8-core Intel i7-2670QM CPU with 2.2 GHz and 8 GB of RAM. (233 MB = 4-min HD Youtube video).

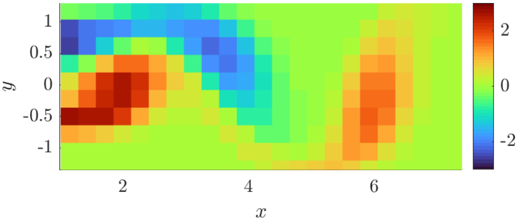
D. A. Messenger and D. M. Bortz. Weak SINDy For Partial Differential Equations. J. Comput. Phys., 443:110525, Oct. 2021.

Data Advantages

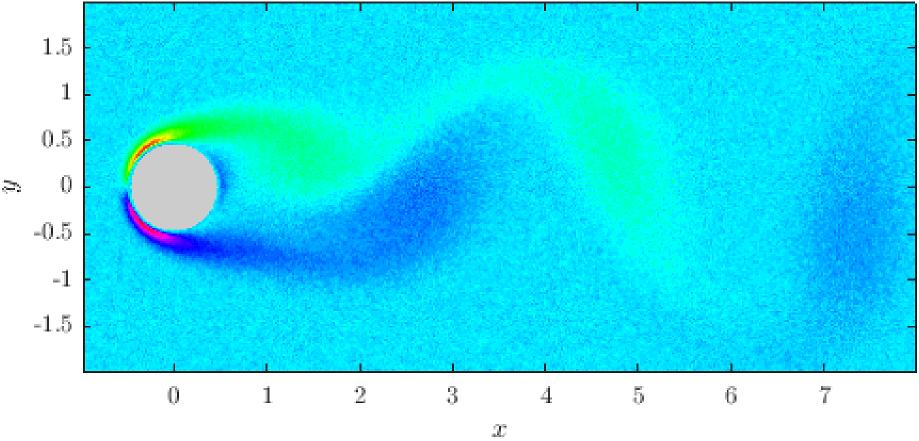
$149 \times 324 \times 201$:
$E_\infty = 3.7 \times 10^{-3}$



$11 \times 24 \times 201$:
$E_\infty = 6.6 \times 10^{-3}$

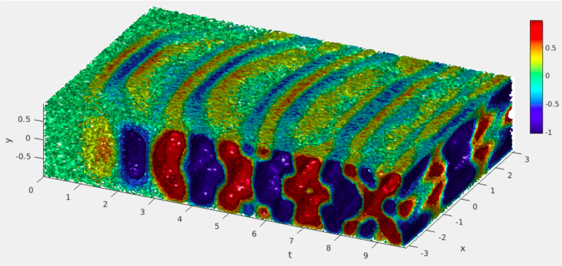


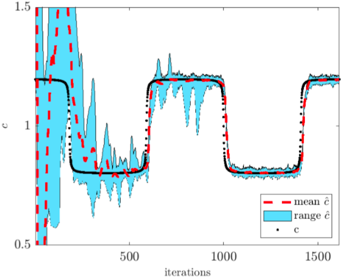
Large-Noise Case



50% noise -> Euler Equations

Online Learning: 2D wave 10% noise, K=25

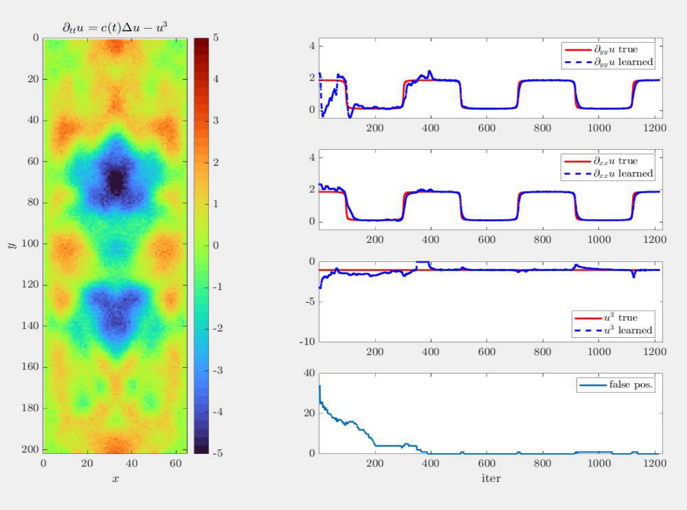




DAM, E. Dall'Anese, & DMB. Online Weak-form Sparse Identification of Partial Differential Equations. In Proc. 3rd MSLS Conf. PMLR 190:241–256. 2022.

Online WSINDy: DAM, Dall'Anese, DMB, MSML, 2022

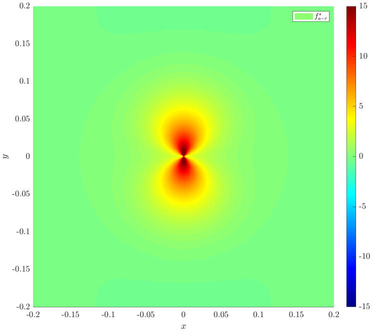
- 30 fr/s (0.033 s/iter)
- 10% noise
- $(\Delta x, \Delta t) = (0.031, 0.024)$
- $K_{\text{mem}} = 15$
- $\dim(\mathbf{G}^{(t)}) = 1548 \times 37$
- $\dim(\mathbf{U}^{(t)}) = 65 \times 202$
- 6.9Mb online storage

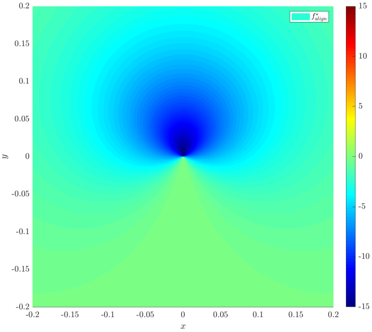


DAM, Wheeler, Liu, & DMB, Learning Anisotropic Interaction Rules from Individual Trajectories in a Heterogeneous Cellular Population. J. R. Soc. Interface, 19(195), Oct. 2022.

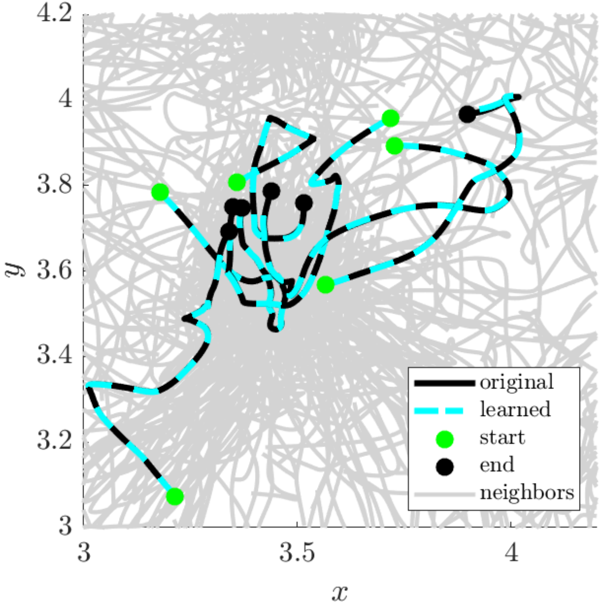
$$\begin{cases} \ddot{x}_i = \frac{1}{N_{\text{tot}}} \sum_{j=1}^{N_{\text{tot}}} f_{\text{a-r}}(|x_i - x_j|, \theta_{ij})(x_i - x_j) \\ \qquad + \frac{1}{N_{\text{tot}}} \sum_{j=1}^{N_{\text{tot}}} f_{\text{align}}(|x_i - x_j|, \theta_{ij})(v_i - v_j) \\ \qquad + \frac{1}{N_{\text{tot}}} \sum_{j=1}^{N_{\text{tot}}} f_{\text{drag}}(|v_i|, \theta_{ij}) v_i + \eta_i. \end{cases} \tag{7}$$

Interacting Particle Systems (Cell Migration)





Learn interaction potentials for migration



Weak form Parameter Estimation

- Minimize the residual

$$\left\|\left[\mathbb{I}_d \otimes (\Phi\Theta(\mathbf{U}))\right]\mathbf{w} + \text{vec}(\Phi(\dot{\mathbf{U}}))\right\|_2^2$$

- If we let

$$\begin{aligned}\mathbf{G} &:= \left[\mathbb{I}_d \otimes (\Phi\Theta(\mathbf{U}))\right], \\ \mathbf{b} &:= \text{-vec}(\Phi(\dot{\mathbf{U}})),\end{aligned}$$

  the OLS solution is

$$\boxed{\widehat{\mathbf{w}} := (\mathbf{G}^T\mathbf{G})^{-1}\mathbf{G}^T\mathbf{B}}$$

Weak form Parameter Estimation

- Minimize the residual

$$\left\| \left[\mathbb{I}_d \otimes (\Phi\Theta(\mathbf{U}))\right] \mathbf{w} + \mathrm{vec}(\Phi(\dot{\mathbf{U}})) \right\|_2^2$$

- If we let

$$\begin{aligned} \mathbf{G} &:= \left[\mathbb{I}_d \otimes (\Phi\Theta(\mathbf{U}))\right], \\ \mathbf{b} &:= \text{-vec}(\Phi(\dot{\mathbf{U}})), \end{aligned}$$

  the OLS solution is

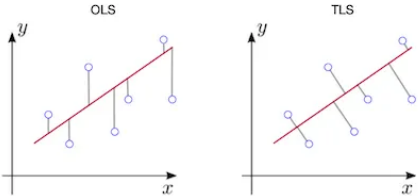
$$\boxed{\widehat{\mathbf{w}} := (\mathbf{G}^T\mathbf{G})^{-1}\mathbf{G}^T\mathbf{B}}$$

- **But, this is a GLS problem:**

$$\left[\mathbb{I}_d \otimes (\Phi\Theta(\mathbf{u}^\star + \boldsymbol{\varepsilon}))\right] \mathbf{w} + \mathrm{vec}\left(\dot{\Phi}(\mathbf{u}^\star + \boldsymbol{\varepsilon})\right)$$

Weak form Parameter Estimation

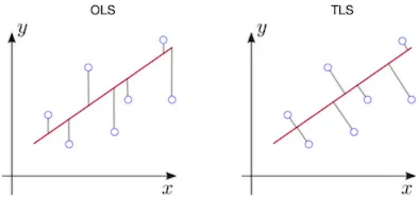
- Specifically, it's an Errors-in-Variables problem:



Ryota Bannai on https://towardsdatascience.com/

Weak form Parameter Estimation

- Specifically, it's an Errors-in-Variables problem:



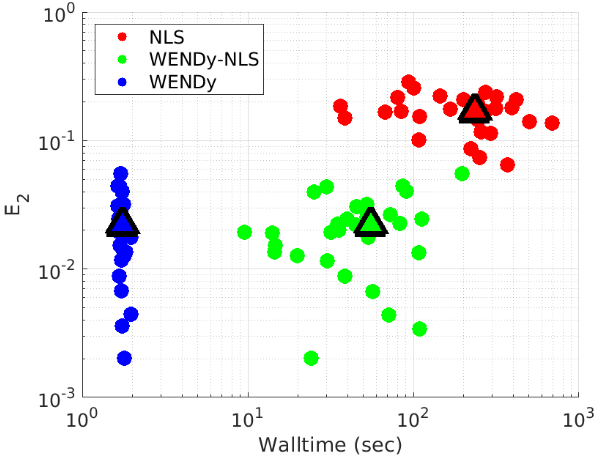
Ryota Bannai on https://towardsdatascience.com/

- **Solve via iterative reweighting of covariance:**

$$\mathbf{G}\mathbf{w}^\star - \mathbf{b} \sim \mathcal{N}(\mathbf{0}, \sigma^2 \mathbf{L}_{\mathbf{w}^\star}(\mathbf{L}_{\mathbf{w}^\star})^T)$$

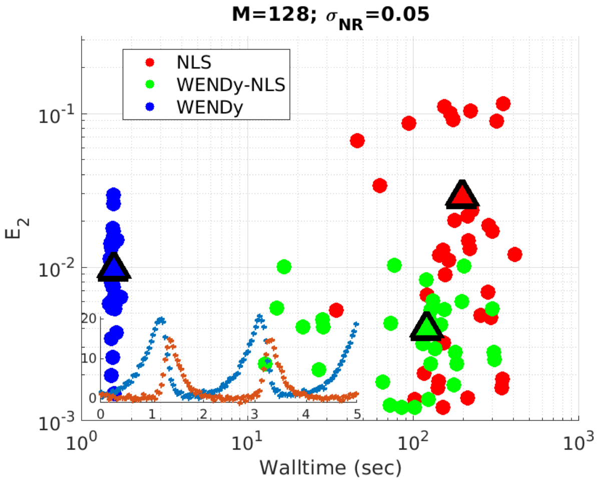
$$\mathbf{L}_\mathbf{w} := [\mathsf{mat}(\mathbf{w})^T \otimes \Phi]\nabla\Theta\mathbf{K} + [\mathbb{I}_d \otimes \dot{\Phi}].$$

Parameter Estimation (FitzHugh-Nagumo)



DMB, DAM, VD, WENDy, Bull. Math. Biol., 2023

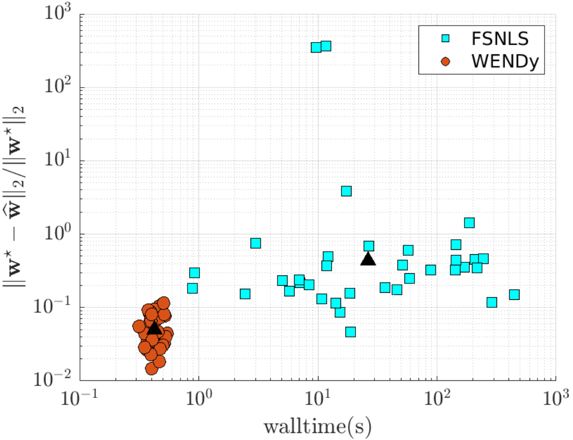
Parameter Estimation (Lotka-Volterra)



DMB, DAM, VD, WENDy, Bull. Math. Biol., 2023

Forthcoming Work

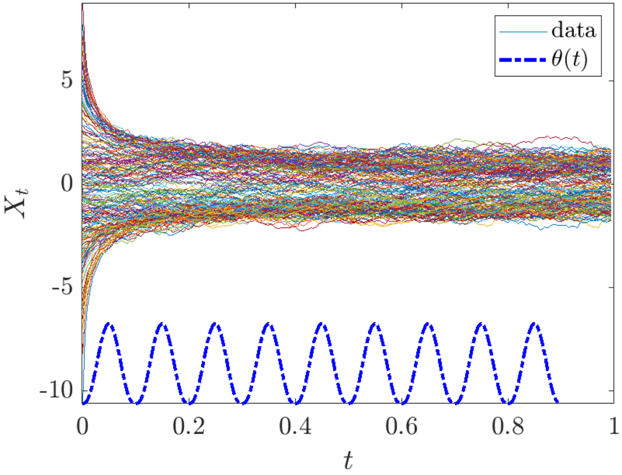
Parameter Estimation (PDE)



Kuramoto-Sivashinsky + 20% noise

128x64 = space x time

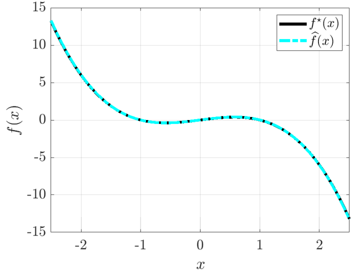
Parameter Estimation (SDE)



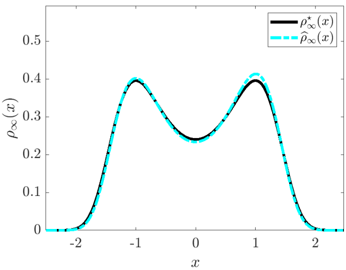
Stochastic Double well potential

200 realizations, 200 timepoints

Parameter Estimation (SDE)

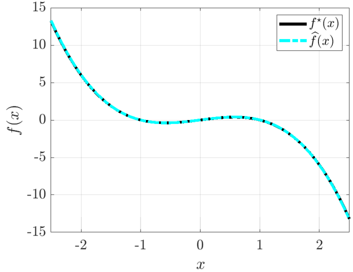


drift function

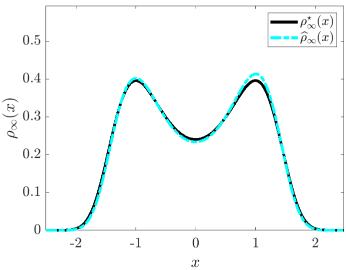


stationary measure

Parameter Estimation (SDE)



drift function



stationary measure

2.5 seconds

$$\widehat{\mathbf{w}}^{(n)} = \texttt{WSINDy}(\mathbf{U}^{(n)}), \qquad \mathbf{U}^{(n)} = u^\star(\mathbf{X}^{(n)}, \mathbf{t}^{(n)}) + \epsilon, \qquad \Delta x^{(n)} \to 0$$

In what sense does $\widehat{\mathbf{w}}^{(n)} \to \mathbf{w}^\star$?

Theorem

Under reasonable assumptions (subGaussian ϵ, poly-trig Θ), there exists a ***critical noise level*** $\sigma_c > 0$ *and a stability tolerance* θ_* *such that for all* $\sigma < \sigma_c$, *all* $\theta \in (0, \theta_*)$, *and sufficiently large* n,

$$\textit{supp}(\widehat{\mathbf{w}}^{(n)}) = \textit{supp}\,(\mathbf{w}^\star) \qquad \textit{and} \qquad \left\|\widehat{\mathbf{w}}^{(n)} - \mathbf{w}^\star\right\|_\infty < C(\theta + \sigma^2) \tag{8}$$

with probability exceeding $1 - 4K(\mathfrak{J}+1)\exp\left(-\frac{c}{2}\,(m_n\theta)^{2/p_{\max}}\right)$ *for any* $\theta > 0$, *where* $m_n = |\textit{supp}\,(\psi) \cap (\mathbf{X}^{(n)}, \mathbf{t}^{(n)})|$.

Classes of (provably) robust equations: $\sigma_c = \infty$ WSINDy recovers the correct form for all noise levels for $p_{ij} \in \{0,1\}, |\alpha^j| \geq 0, |\beta^j| \geq 1, |\gamma^j| \geq 0$:

$$\partial^{\alpha^0} u = \underbrace{\sum_{j=1}^{J_1} w_j^{(1)} \partial^{\alpha^j} \prod_{i=1}^{n} u_i^{p_{ij}}}_{linear+bilinear} + \underbrace{\sum_{i=1}^{n} \sum_{j=1}^{J_2} w_{ij}^{(2)} \partial^{\beta^j} u_i^2}_{quadratic} \tag{9}$$

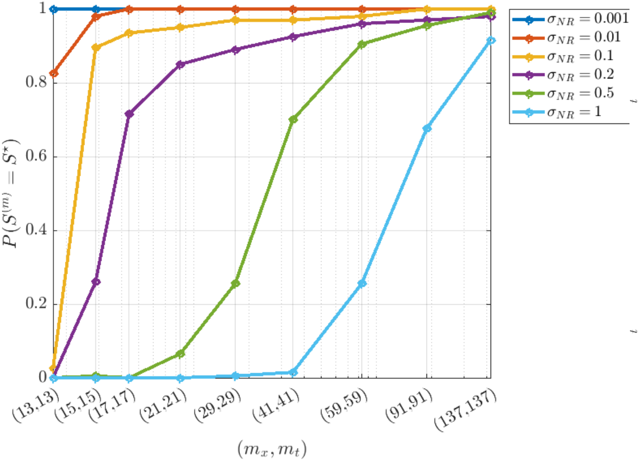
(e.g. $\partial_t \omega = -\nabla \cdot (\omega \mathbf{u}) + \nu \Delta \omega$)

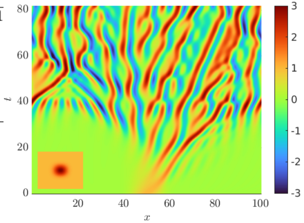
$$\partial^{\alpha^0} u = \underbrace{\sum_{j=1}^{J_1} w_j^{(1)} \partial^{\alpha^j} \prod_{i=1}^{n} u_i^{p_{ij}}}_{linear+bilinear} + \underbrace{\sum_{i=1}^{n} \sum_{j=1}^{J_2} w_{ij}^{(2)} \partial^{\beta^j} u_i^2}_{quadratic} + \underbrace{\sum_{k=1}^{F_1} \nu_k^{(3)} \partial^{\gamma_j} \exp(i\omega_k^T u)}_{trigonometric} \tag{10}$$

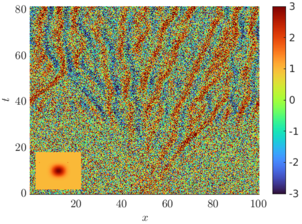
(e.g. $\partial_{tt} u = \Delta u - \sin(\omega u)$)

- (9) also recovers the correct coefficient values
- (10) possibly leads to a bias in the weights

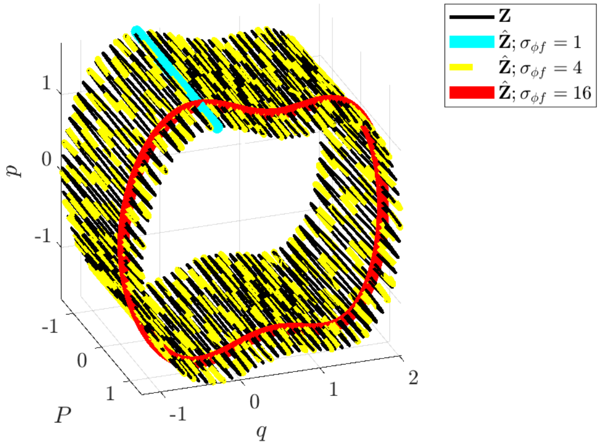
$$\partial_t u = (1)\partial_x^4 u + (0.75)\partial_x^6 u + (-0.5)\partial_x(u^2) + (0.1)\partial_x^3(u^2)$$





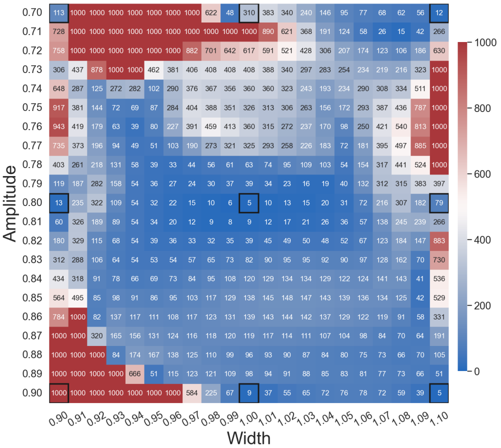


Structure Preserving ROM



DAM, Burby, DMB, Hamiltonian-Respecting ROM for Plasma Physics, SciRep (revised), 2024

Weak Form / Autoencoder ROM

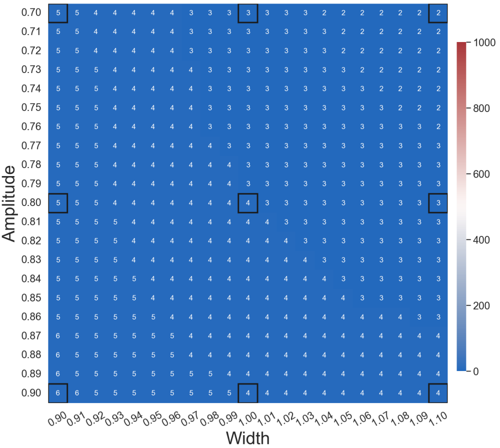


Weak form autoencoder ROM (10% error in data)

2-3 order of magnitude speedup over PDE simulation

Tran et al., CMAME, 2024.

Weak Form / Autoencoder ROM

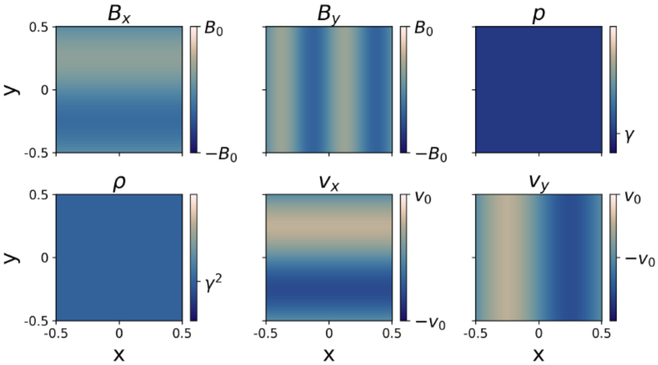


Weak form autoencoder ROM (10% error in data)

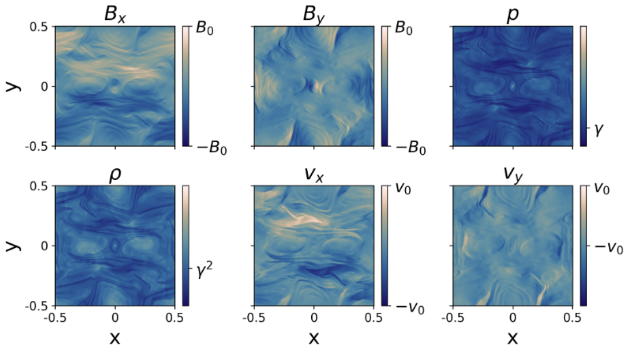
2-3 order of magnitude speedup over PDE simulation

Tran et al., CMAME, 2024.

MHD Results

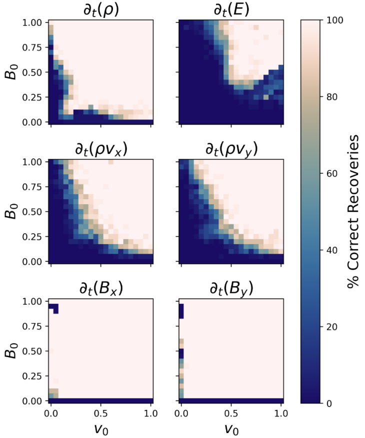


(a) Initial conditions for Orszag-Tang dataset with varied parameters denoted by B_0 for the magnetic field and v_0 for the magnitude of the initial velocity.



(b) Example of final state for Orszag-Tang using the initial conditions shown in panel (a). The boundary conditions are periodic by design so there is no need for a maximum simulation time to limit escape from the simulation.

MHD Results



Vasey et al., MHD + WSINDy, arXiv:2312.05339, 2023

Related Work

- Fasel, Kutz, Brunton, & Brunton. Ensemble-SINDy..., PRSA, 478(2260), 20210904, 2022.
  - E-WSINDy is almost unbeatable!
- Bertsimas & Gurnee, Learning sparse nonlinear dynamics via mixed-integer optimization, Nonlin. Dyn., 2023
  - WSINDy + MIO works great!
- Russo & Laiu, Convergence of weak-SINDy Surrogate Models, SIAM J. App. Dyn. Sys. 2024.
  - WSINDy Reproducing Kernel Hilbert Space
  - WSINDy is a projection operator
- Tang, Liao, Kuske, & Kang, WeakIdent: Weak formulation for Identifying Differential Equations using Narrow-fit and Trimming, J. Comp. Phys., 2023.
  - Narrow fits + Trimming beats sequential least squares

Summary

- WSINDy
  - high computational efficiency
  - ODE, PDE, IPS (1st & 2nd order)
  - robust to large (& frequently ludicrous) noise levels
  - modest data needs / online learning
  - rigorous & theoretically sound
- WENDy
  - accelerating the core of WSINDy
  - avoids numerical approximation of forward problem

Summary: Common Questions

- Need to specify library to prune from

Summary: Common Questions

- Need to specify library to prune from
  - Science informs library: u^{15} needed?

Summary: Common Questions

- Need to specify library to prune from
  - Science informs library: u^{15} needed?
  - Symbolic Regression?

Summary: Common Questions

- Need to specify library to prune from
  - Science informs library: u^{15} needed?
  - Symbolic Regression?
- Integrate by parts

Summary: Common Questions

- Need to specify library to prune from
  - Science informs library: u^{15} needed?
  - Symbolic Regression?
- Integrate by parts
  - Works for an extraordinarily broad class of models
  - differential operators acting on linear combinations of nonlinear functions

Summary: Common Questions

- Need to specify library to prune from
  - Science informs library: u^{15} needed?
  - Symbolic Regression?
- Integrate by parts
  - Works for an extraordinarily broad class of models
  - differential operators acting on linear combinations of nonlinear functions
- Need to know data for all variables in learned equation

Summary: Common Questions

- Need to specify library to prune from
  - Science informs library: u^{15} needed?
  - Symbolic Regression?
- Integrate by parts
  - Works for an extraordinarily broad class of models
  - differential operators acting on linear combinations of nonlinear functions
- Need to know data for all variables in learned equation
  - Differential Algebra

Summary: Common Questions

- Need to specify library to prune from
  - Science informs library: u^{15} needed?
  - Symbolic Regression?
- Integrate by parts
  - Works for an extraordinarily broad class of models
  - differential operators acting on linear combinations of nonlinear functions
- Need to know data for all variables in learned equation
  - Differential Algebra
- Boundary / Initial Conditions identification?

Summary: Common Questions

- Need to specify library to prune from
  - Science informs library: u^{15} needed?
  - Symbolic Regression?
- Integrate by parts
  - Works for an extraordinarily broad class of models
  - differential operators acting on linear combinations of nonlinear functions
- Need to know data for all variables in learned equation
  - Differential Algebra
- Boundary / Initial Conditions identification?
  - Yes, but need different test functions

Acknowledgements

- Funding (as PI)
  - DMS/NIGMS Math Bio Initiative R01GM126559 (2017-2023)
  - NSF MODULUS MCD-2054085 (2021-2024)
  - DOE ASCR DE-SC0023346 (2022-2027)
  - NIH-NIGMS MIRA R35GM149335 (2023-2028)
- John Nardini (PhD 2018)
- Harry Dudley (PhD 2020)
- Jacqui Wentz (PhD 2020)
- Sabina Altus (PhD 2021)
- Lewis Baker (PhD 2021)
- **Dan Messenger (PhD 2022)**
- April Tran (PhD 2026)

github.com/MathBioCU

References (github.com/MathBioCU)

ODE DAM & DMB. WSINDy for ODEs. MMS 2021
doi:10.1137/20M1343166

PDE DAM & DMB. WSINDy for PDEs,. JCP, 2021
doi:10.1016/j.jcp.2021.110525

IPS (order 1) DAM & DMB. WSINDy for SDEs, Physica D, 2022
doi:10.1016/j.physd.2022.133406

Online DAM, Dall'Anese & DMB Online WSINDy MSML 2022
proceedings.mlr.press/v190/a-messenger22a.html

IPS (order 2) DAM, Wheeler, Liu, & DMB. WSINDy for migration, JSRI, 2022
doi:10.1098/rsif.2022.0412

Theory DAM & DMB, Asymptotics..., arXiv:2211.16000

ParEst DMB, DAM, VD, WENDy, BMB, Oct., 2023
doi:10.1007/S11538-023-01208-6

ROM#1 AT, et al., WLaSDI, CMAME, 2024
doi:10.1016/j.cma.2024.116998

ROM#2 DAM, Burby, DMB, Coarse-graining... arXiv:2310.05879

MHD Vasey et al., 2023, WSINDy for MHD arXiv:2312.05339