Data-driven derivation of equations for the evolution of transport in turbulent flow

C.J. Wareing, A. Roy, S.M. Tobias

School of Mathematics, University of Leeds, Leeds LS2 9JT, UK. *C.J.Wareing@leeds.ac.uk*

Summary

We present preliminary results of DNS of turbulent fluid dynamics coupled with machine learning techniques to derive new equations for the evolution of transport in turbulent flows. We examine Rayleigh-Bénard convective turbulence with the aim to learn the statistics of unresolved scales for turbulent parameterization. Following the approach of Garaud et al. 2010 [1] in order to perform a comparison to their result, we seek a closure model for the transport of entropy and momentum intended for application to rotating stellar convective regions. We use the Dedalus framework [2,3] for spectrally solving differential equations to generate an extended time-series of two-dimensional DNS data at a Rayleigh number of 10^{10}. We then use the data-driven Sparse Identification of Nonlinear Dynamics (SINDy) algorithm [4] to discover the form of the triple correlation terms from the data, ensuring we capture any difference between the bulk and boundary layers. In future work, we intend to apply the same SINDy method to convective rotating turbulence, mean flows and magnetic fields. Further, we intend to repeat the analysis using Bayesian machine learning methods and compare the results to those obtained with SINDy.

Physical Model & Governing Equations

- 2D horizontally-periodic Rayleigh-Bénard convection.
- Non-dimensionalised using the box height and freefall time.
- Rayleigh numbers, $R = 10^6, 10^8$ & 10^{10}. Prandtl number, $P = 1$.
- Stress-free boundary condition at top $z = 0$ and bottom $z = 1$.
- Boussinesq approximation.
- Temperature $\Theta(z=0)=1$, $\Theta(z=1)=0$.

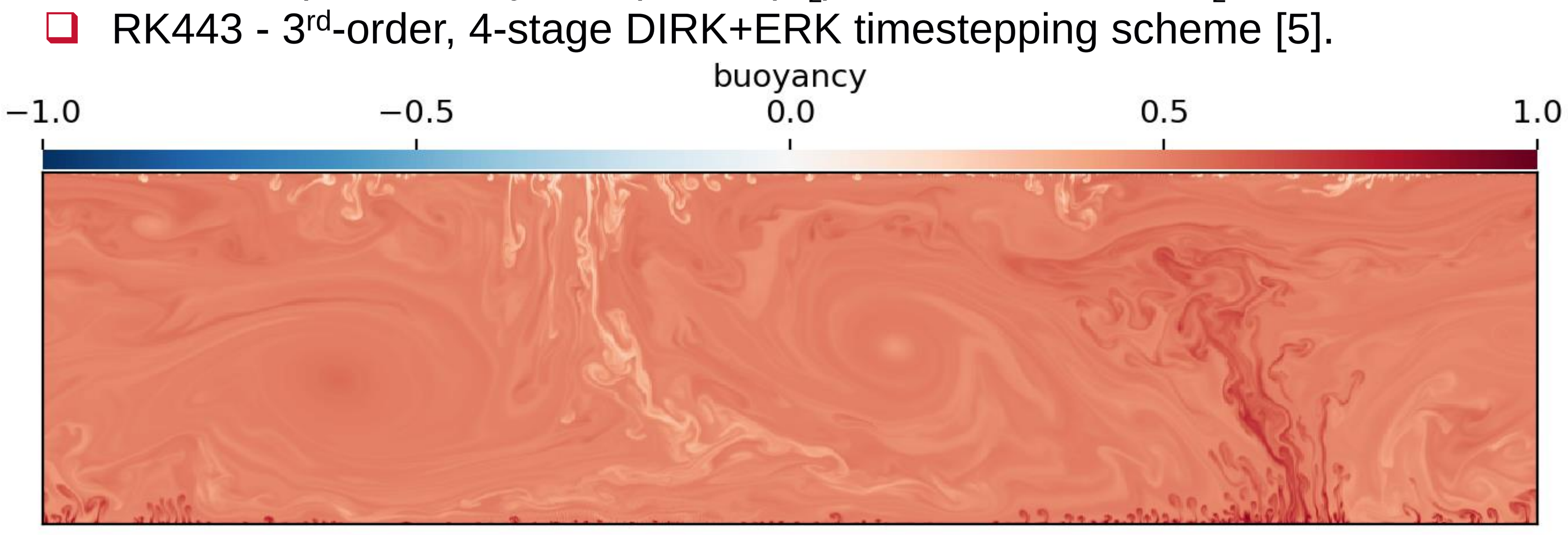
$$\frac{\partial \boldsymbol{u}}{\partial t} + (\boldsymbol{u}\cdot\nabla)\boldsymbol{u} = -\nabla\sigma + \Theta\hat{z} + \sqrt{\frac{P}{R}}\nabla^2\boldsymbol{u}, \quad (1)$$

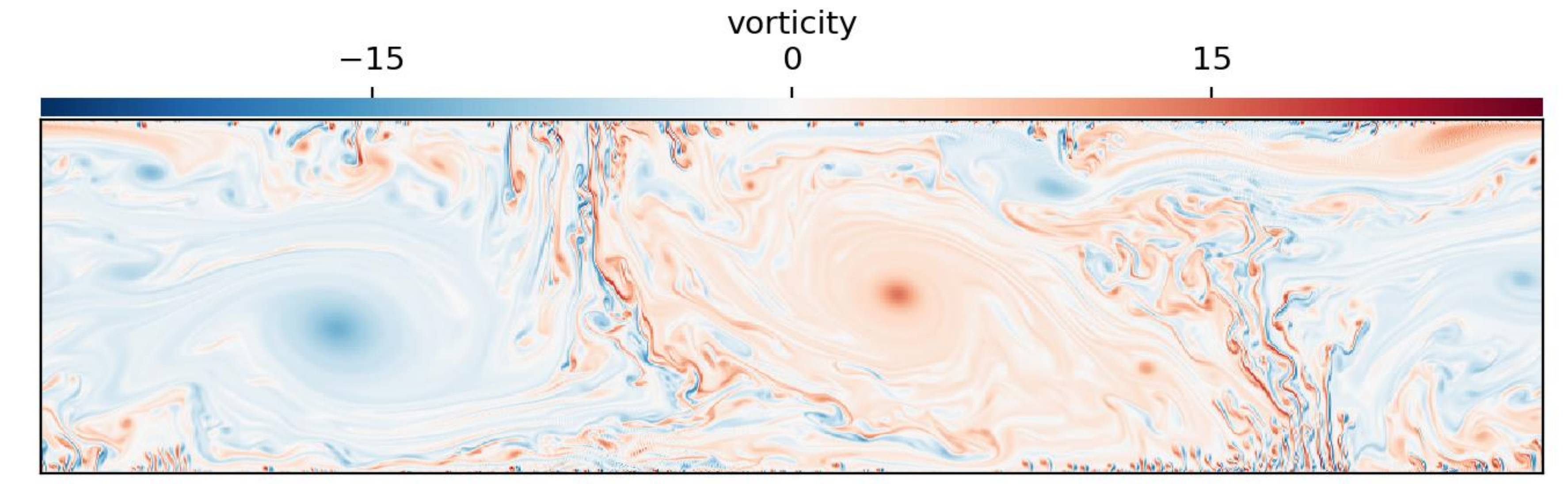
$$\frac{\partial \Theta}{\partial t} + (\boldsymbol{u}\cdot\nabla)\Theta = \frac{1}{\sqrt{PR}}\nabla^2\Theta, \quad (2)$$

$$\nabla\cdot\boldsymbol{u} = 0, \quad (3)$$

Numerical technique

- Dedalus v3 flexible framework for solving PDEs using spectral methods [2].
- Open-source, widely-used, well supported [3].
- 1024 real space Fourier points (N_x) in the x direction, $L_x = 4$.
- 384 real space Chebyshev points (N_z) in the z direction, $L_z = 1$.
- RK443 - 3rd-order, 4-stage DIRK+ERK timestepping scheme [5].





- Snapshot at t=48.5 of resolved 2D convection, Ra = 10^{10}.

Theoretical model

- Garaud et al. [1] retained exact forms of the left-hand sides, developing equations for the triple correlation terms and proposed simple closures for the right-hand sides by splitting mean and fluctuating parts $u_i = \bar{u}_i + u_i'$ with $\bar{u_i'} = 0$, of the form:

$$(\partial_t + \bar{u}_k\partial_k)\bar{R}_{ij} + \bar{R}_{ik}\partial_k\bar{u}_j + \bar{R}_{jk}\partial_k\bar{u}_i + \alpha(\bar{F}_i g_j + \bar{F}_j g_i) - \nu\partial_{kk}\bar{R}_{ij} = -\frac{C_1}{L}\bar{R}^{1/2}\bar{R}_{ij} - \frac{C_2}{L}\bar{R}^{1/2}\left(\bar{R}_{ij} - \frac{1}{3}\bar{R}\delta_{ij}\right) - \nu\frac{C_\nu}{L^2}\bar{R}_{ij}, \quad (4)$$

$$(\partial_t + \bar{u}_j\partial_j)\bar{F}_i + \bar{R}_{ij}\partial_j\bar{\Theta} + \bar{F}_j\partial_j\bar{u}_i + \alpha\bar{Q}g_i - \tfrac{1}{2}(\nu+\kappa)\partial_{jj}\bar{F}_i = -\frac{C_6}{L}\bar{R}^{1/2}\bar{F}_i - \frac{1}{2}(\nu+\kappa)\frac{C_{\nu\kappa}}{L^2}\bar{F}_i, \quad (5)$$

$$(\partial_t + \bar{u}_i\partial_i)\bar{Q} + 2\bar{F}_i\partial_i\bar{\Theta} - \kappa\partial_{ii}\bar{Q} = -\frac{C_7}{L}\bar{R}^{1/2}\bar{Q} - \kappa\frac{C_\kappa}{L^2}\bar{Q}, \quad (6)$$

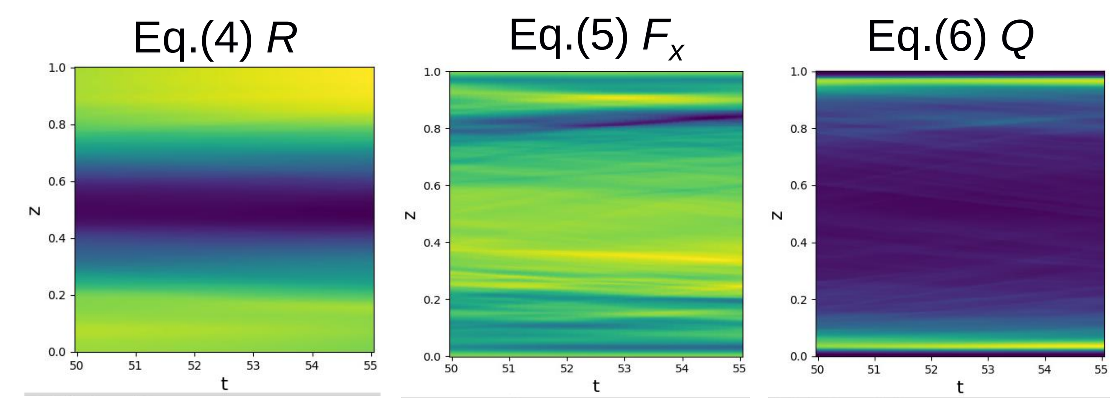
where $\bar{R}_{ij} = \overline{u_i' u_j'}$, $\bar{F}_i = \overline{\Theta' u_i'}$ and $\bar{Q} = \overline{\Theta'^2}$.

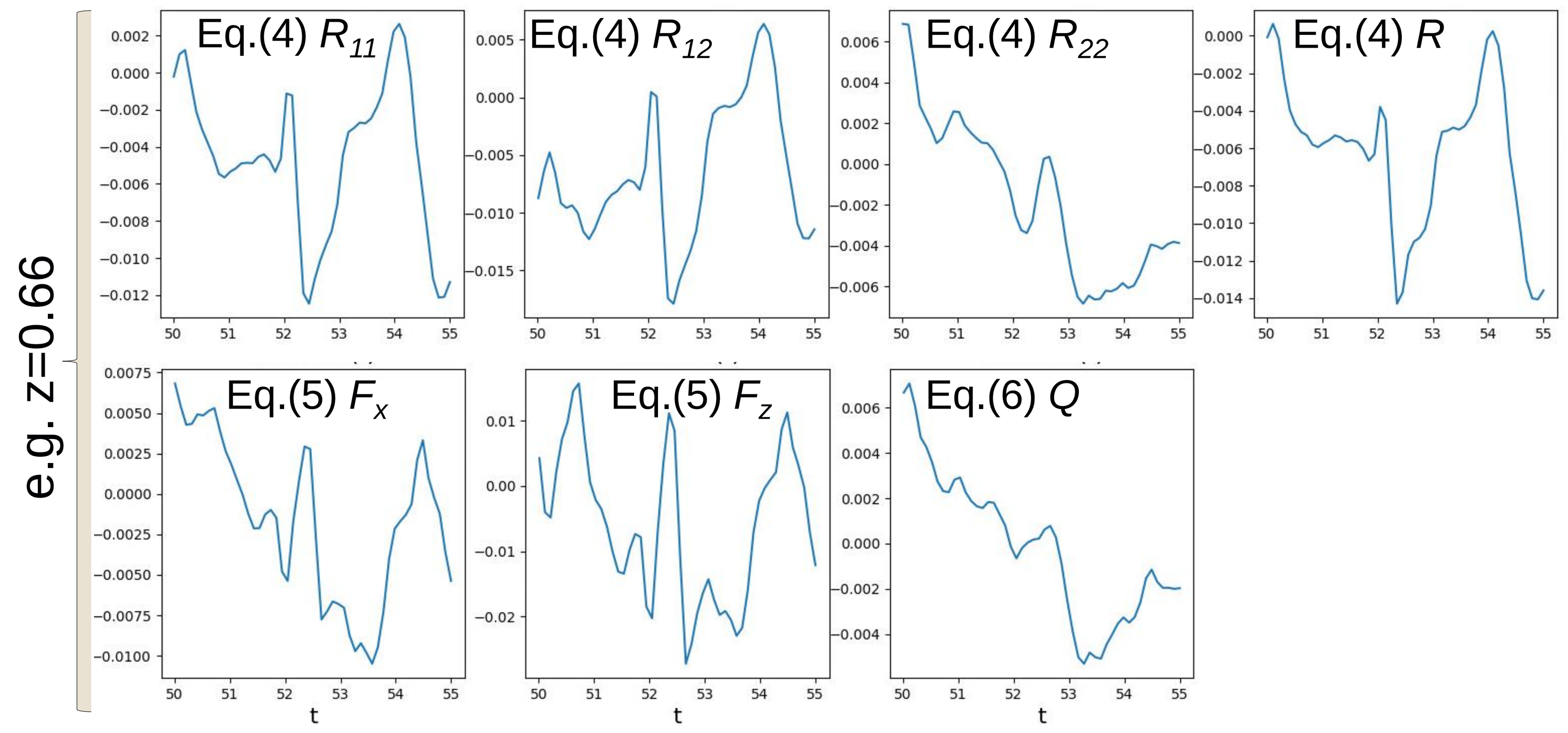
- Assuming variation only with z, they obtain a set of coefficients for 3D Rayleigh-Bénard convection as follows:

$$C_1 \simeq 0.4, \quad C_2 \simeq 0.6, \quad C_\kappa = 2 \pm 0.2.$$
$$C_\nu = 12 \pm 1, \quad C_6 = 1.4 \pm 0.1, \quad (7)$$
$$C_{\nu\kappa} = 6 \pm 0.5, \quad C_7 = 1.4 \pm 0.1.$$

Integrated fluxes

- Integrating along x, we obtain time series of averaged data for:-
  - Temperature, $\bar{\Theta}(z)$
  - x-velocity, $\bar{u}_i(z)$
  - z-velocity, $\bar{u}_j(z)$
  - $\bar{R}(z)$ ($= \Sigma_i \bar{R}_{ii}$); $i = 1,2$
  - $\bar{R}_{ij}(z)$; $i,j = 1,2$
  - $\bar{F}_i(z)$; $i = 1,2$
  - $\bar{Q}(z)$





SINDy

- Discovering governing equations from data by **S**parse **I**dentification of **N**onlinear **Dy**namical (SINDy) systems [4]
- Figure 1 from Brunton et al. [4] gives a demonstrative overview:-

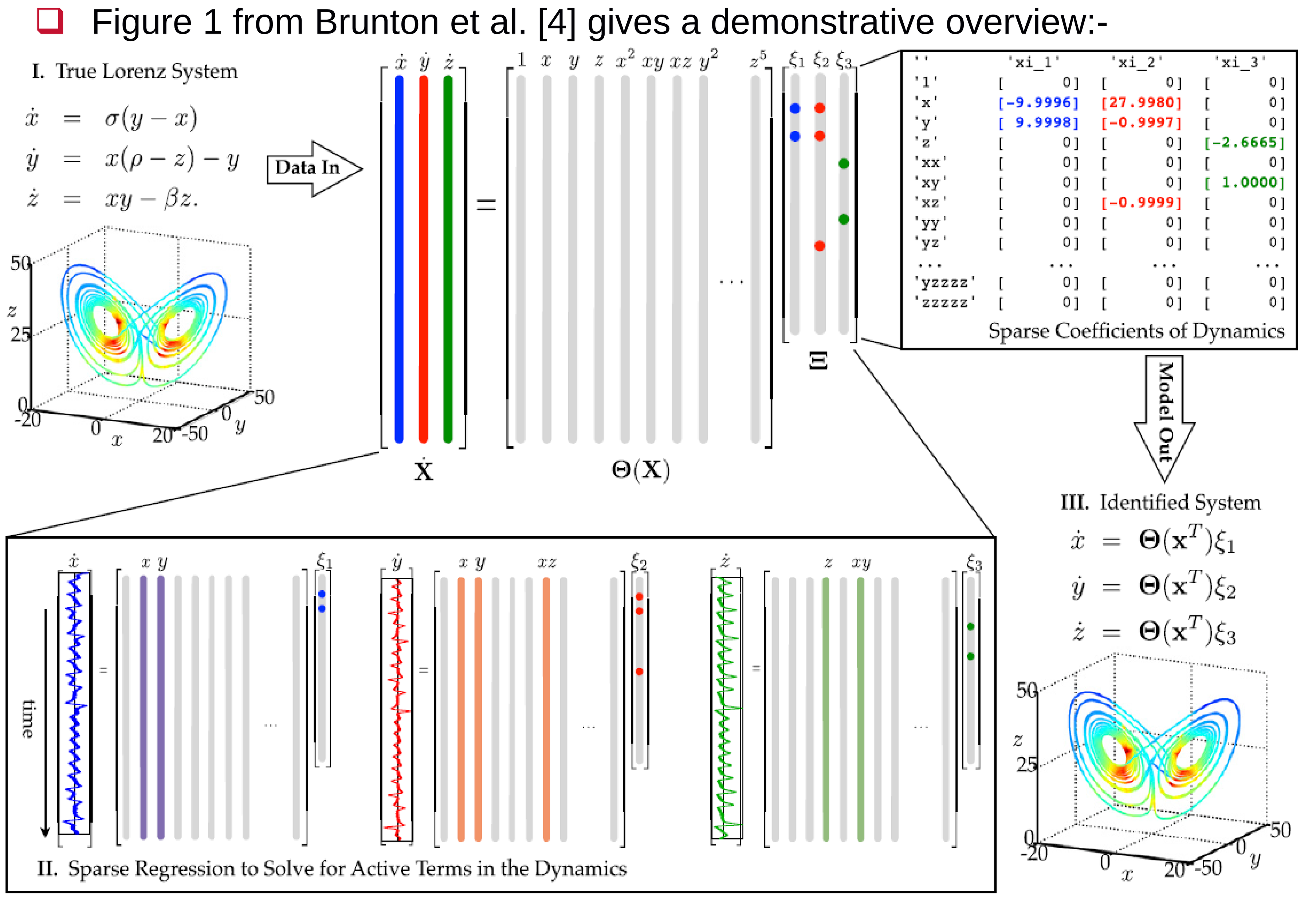


Fig. 1. Schematic of the SINDy algorithm, demonstrated on the Lorenz equations. Data are collected from the system, including a time history of the states X and derivatives $\dot{X}$; the assumption of having $\dot{X}$ is relaxed later. Next, a library of nonlinear functions of the states, $\Theta(X)$, is constructed. This nonlinear feature library is used to find the fewest terms needed to satisfy $\dot{X} = \Theta(X)\Xi$. The few entries in the vectors of Ξ, solved for by sparse regression, denote the relevant terms in the right-hand side of the dynamics. Parameter values are $\sigma = 10, \beta = 8/3, \rho = 28$, $(x_0, y_0, z_0)^T = (-8, 7, 27)^T$. The trajectory on the Lorenz attractor is colored by the adaptive time step required, with red indicating a smaller time step.

Concluding remarks

- Applying SINDy with judicious constraints can identify the correct terms in the equations governing 2D Rayleigh-Bénard convection, e.g.:-

```
Noiseless weak fit:
(R)' = -0.046027 R + 0.000010 R_11 + 0.052413 R^3/2 + 2.000000 Fz + -2.000000 R22Wbar_1 + -2.000000 R12Ubar_1 + -1.000000 WbarR_1
(R11)' = 0.029194 R11R^1/2 + -0.005822 R^3/2 + -2.000000 R12Ubar_1 + -1.000000 WbarR11_1
(R12)' = -0.068491 R12 + -0.009120 R12R^1/2 + 1.000000 Fx + -1.000000 R22Ubar_1 + -1.000000 R12Wbar_1 + -1.000000 WbarR12_1
(R22)' = 0.002815 R22 + -0.027741 R22R^1/2 + 2.000000 Fz + -2.000000 R22Wbar_1 + -1.000000 WbarR22_1
(Fx)' = -0.138289 Fx + 0.000010 Fx_11 + 0.133954 FxR^1/2 + -1.000000 FzUbar_1 + -1.000000 R12Temp_1 + -1.000000 WbarFx_1
(Fz)' = 0.312619 Fz + 0.070370 FzR^1/2 + 1.000000 Q + -1.000000 FzWbar_1 + -1.000000 WbarFz_1 + -1.000000 R22Temp_1
(Q)' = 0.402013 Q + 0.000010 Q_11 + -0.233446 QR^1/2 + -2.000000 FzTemp_1 + -1.000000 WbarQ_1
```

- Preliminary coefficients are different to (7).

  $C_1 \sim 0.05$, $C_2 \sim 0.08$, $C_6 \sim -0.1$, $C_7 \sim 0.2$, $C_\nu \sim 6e3$, $C_{\nu\kappa} \sim 2e4$, $C_\kappa \sim 4e4$

- Future work will look to improve the regression fit and explore the differences.

References:

[1] Garaud, Ogilvie, Miller & Stellmach, MNRAS **407**, 2451-2467, 2010.

[2] Burns, Vasil, Oishi, Lecoanet & Brown, Physical Rev. Res. **2** (2) 023068, 2020.

[3] http://dedalus-project.org

[4] Brunton, Proctor & Kutz, PNAS **113 (15)**, 3932-3937, 2016.

[5] Ascher, Ruuth, Spiteri, Applied Numerical Mathematics **25** (2-3), 151-167, 1997.

Funded by the European Union

European Research Council
Established by the European Commission

Acknowledgements: This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (Grant agreement No. D5S-DLV-786780). The calculations for this paper were performed on the University of Leeds ARC4 facility, hosted and enabled through the ARC HPC resources and support team at the University of Leeds, to whom we extend our grateful thanks.

UNIVERSITY OF LEEDS