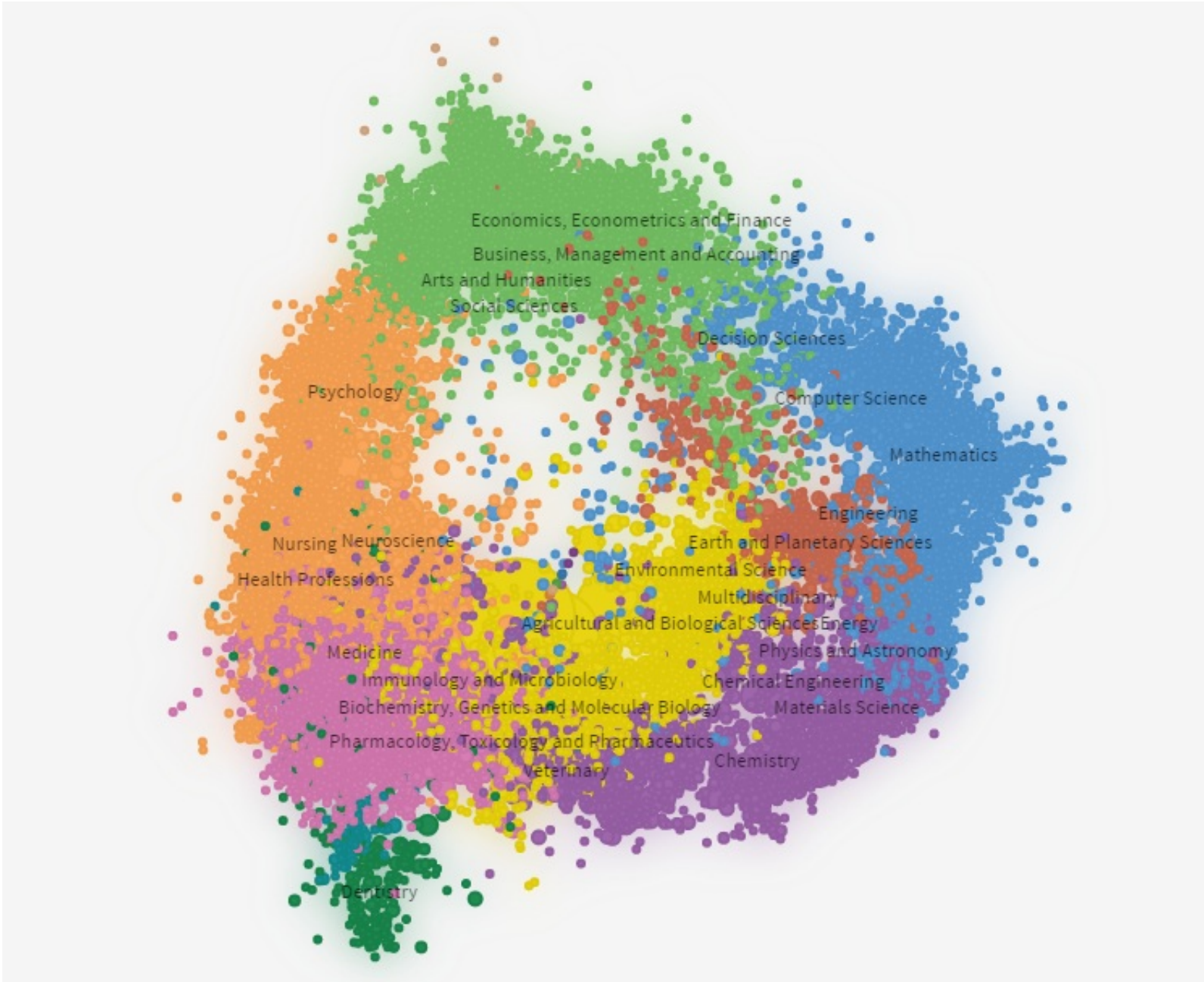
Faces of entropy: from materials science to biology and quantum information science

Maria Emelianenko

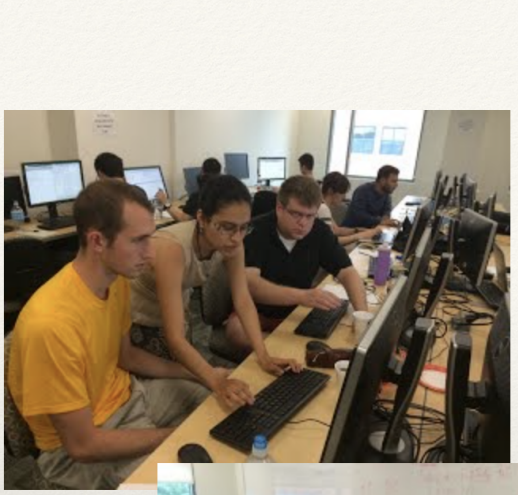
Dept. of Mathematical Sciences and
Quantum Science and Engineering Center
George Mason University

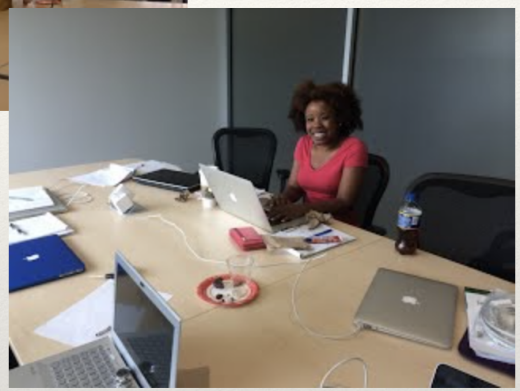
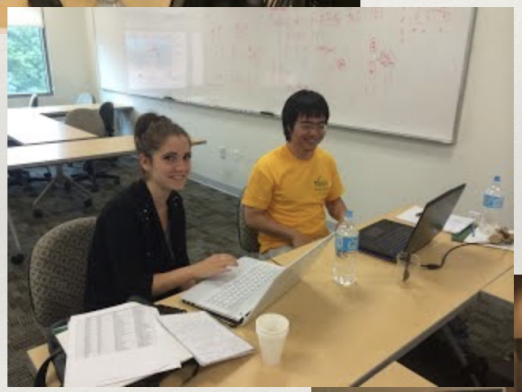
PSU symposium, Nov 14, 2023

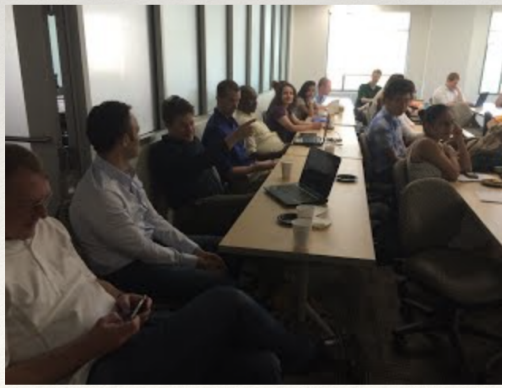
Boundaries between scientific domains



Crossing domain boundaries: Mason Modeling Days



Collaborators

- Materials modeling and simulation:
  joint with D. Kinderlehrer (CMU), Y. Epshteyn (U. Utah), K. Salemeruiz (U, Luxembourg), M. Darehbidi (GMU, Mech. Engr.), C. Torres (UTFSM, Chile), D. Golovaty (U. Akron), D. Torrejon (Blacksky)
- Adaptive biological networks:
  joint with T. Oellerich (GMU, Math), M. Pierobon (GMU, Bio), L. Liotta (GMU, Bio), R. Araujo (U. Queensland, Australia)
- Quantum computing:
  joint with M. Jarret (GMU, Math), R. Turner (GMU, Math), M. Tian (GMU, Physics), X. Gitiaux (GMU, CS), I. Morris (GMU, Physics), E. Galvao (Portugal)

Materials and biological projects supported by NSF CAREER grant DMS-1056821, Simons Foundation grant and GMU Provost MDR grant. QC projects supported by QSEC seed grants.

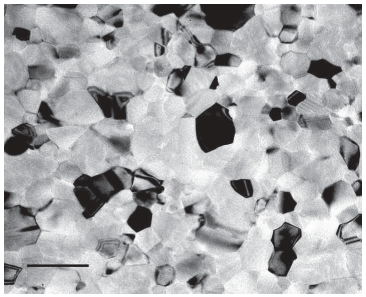
SIMONS FOUNDATION

Outline of the talk

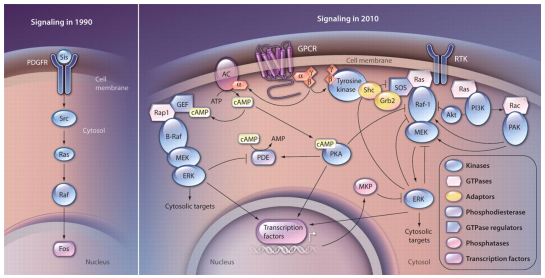
1. Concept of graph entropy: mathematical definition
2. Concept of random walks and generalized master equations
3. Concept of data-driven dynamics identification without regularization
4. Crossing boundary into materials science
5. Crossing boundary into biology
6. Crossing boundary into quantum computing

Ubiquitous networks

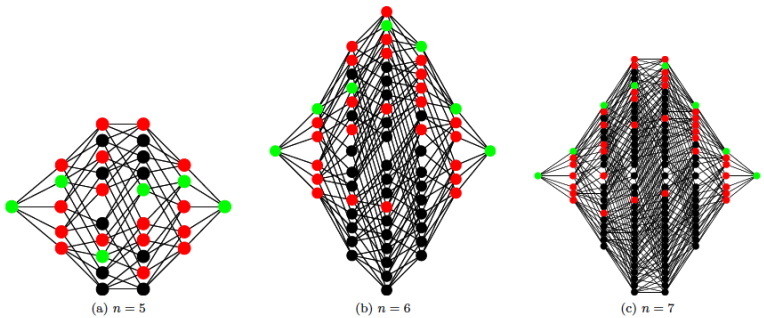
- Materials networks



- Biological networks



- Quantum graphs



Characterizing networks

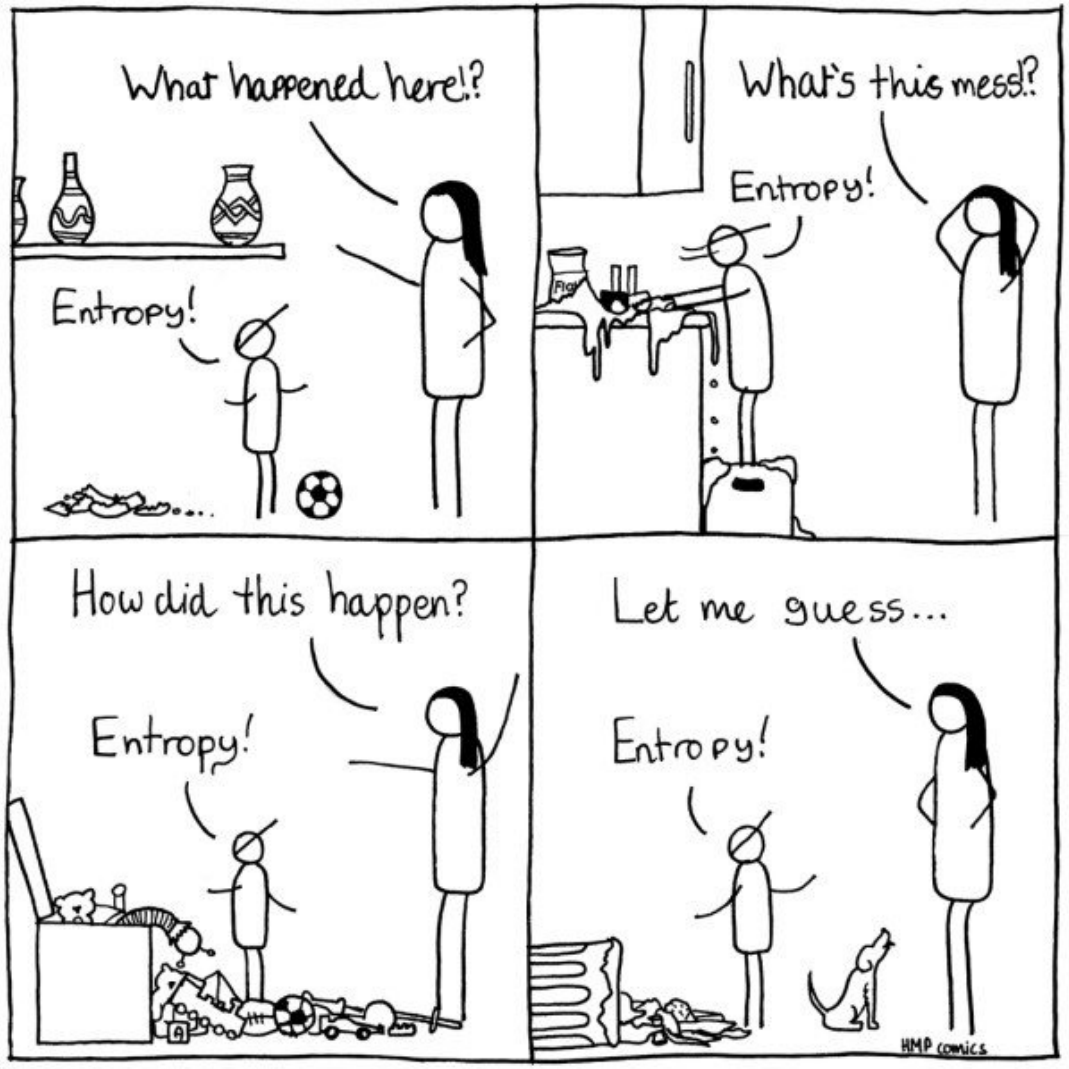
1. Static properties: node statistics, connectivity, graph entropy, centrality measures, distance between graphs, max-cut etc
2. Dynamics: random walks on graphs, graph flows etc.

Characterizing networks

1. Static properties: node statistics, connectivity, graph entropy, centrality measures, distance between graphs, max-cut etc
2. Dynamics: random walks on graphs, graph flows etc.

We will talk about entropy first.

What is entropy?



This is why we don't teach our children about entropy until much later...

Faces of entropy

Entropy = measure of disorder. But what does it mean precisely?

It depends on the context. Three faces of entropy:

- thermodynamic entropy, viewed as a continuum field,
- entropy of mixing,
- configurational entropy/ entropy of information/ network entropy

We will talk about these in the context of materials, biology and QC applications.

Configurational entropy

The math entropy (configurational entropy), which is the negative of the Shannon Entropy, is

$$\Phi(p) = \sum_{k=1\ldots r} p_k \log p_k, \text{ where } \sum_{k=1\ldots r} p_k = 1, \ p_k \geqq 0.$$

We always take $0 \log 0 = 0$. This entropy is supposed to suggest the randomness present in a random variable ω defined on the integers $\mathbf{Z}$ with values in the events $X = \{x_1, \ldots x_r\}$ having

$$\text{Prob}(\omega_i = x_k) = p_k, \ k = 1, \ldots, r.$$

Now since

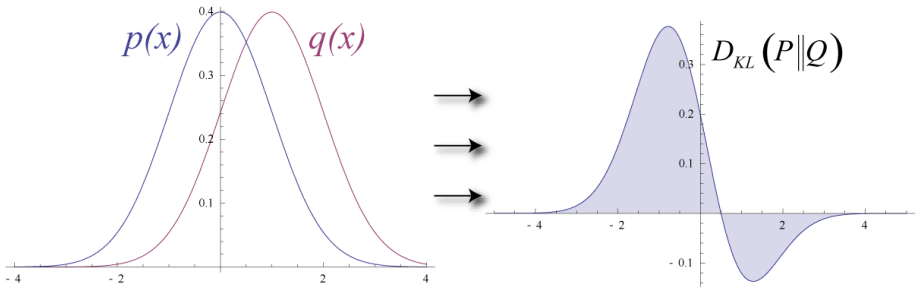
$$\log \frac{1}{r} \leqq \Phi(p) \leqq 0,$$

1. The minimum of Φ is achieved for the uniform distribution $p_k = 1/r$ for all k: most random in the sense that the distribution does not distinguish among the events x_k
2. Max is achieved for a distribution where one of the $p_k = 1$ and the remainder are 0: the least random in the sense that it says that x_k occurs and none other does.

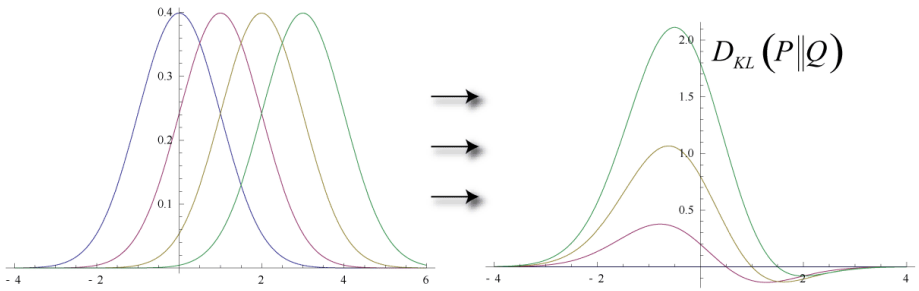
Kullback-Leibler relative entropy

Kullback-Leibler Relative Entropy, defined to be

$$\Phi(p\|q) = \sum_{i=1,\ldots,r} p_i \log \frac{p_i}{q_i}, \ p, q \text{ probability vectors.}$$



Original Gaussian PDF's KL Area to be Integrated



Properties:

1. $\Phi(p\|q) \geqq 0$
2. $\Phi(p^{(n)}\|p^{eq})$ is decreasing for any probability matrix P and sequence of iterates $p^{(n)}$, that is, for any Markov Chain
3. The diffusion equation is the Euler-Lagrange equation for the entropy.

Characterizing networks

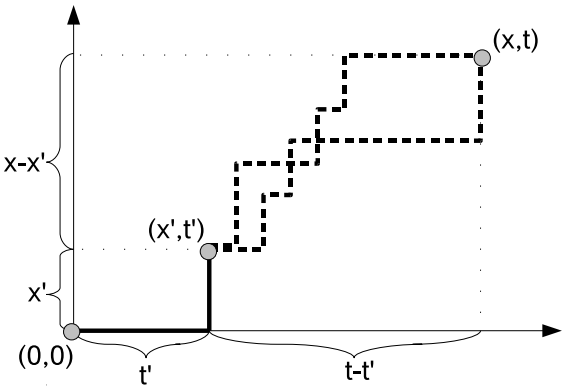
1. Static properties: node statistics, connectivity, graph entropy, centrality measures, distance between graphs, max-cut etc
2. Dynamics: random walks on graphs, graph flows etc.

Continuous Time Random Walks

The walker moves according to two parameters:
(1) Waiting times with pdf $w(t)$ (2) Jump sizes with pdf $\mu(x)$.

Chapman-Kolmogorov type master equation:

$$p(s,t) = p_0(s)\psi(t) + \int\int p(s-s',t-t')\mu(s')w(t')dt'ds'$$



Equivalent form in terms of the memory kernel $\hat{\Phi}(u) = \dfrac{1-\hat{w}(u)}{u\hat{w}(u)}$:

$$\int \Phi(t-t')\frac{\partial}{\partial t}p(x,t')dt' = \int [p(x-x',t) - p(x,t)]\mu(x')dx'$$

becomes a Montroll-Weiss equation in Fourier-Laplace space

$$\hat{p}(k,u) = \frac{\hat{p}_0(k)(1-\hat{w}(u))}{u(1-\hat{w}(u)\hat{\mu}(k))}$$

Uniform approach to fractional and regular dynamics

- $\Phi(t) = c\delta(t)$: Markov process, no memory

$$\frac{\partial}{\partial t} p(x,t) = \lambda \int [p(x-x',t) - p(x,t)]\mu(x')dx'$$

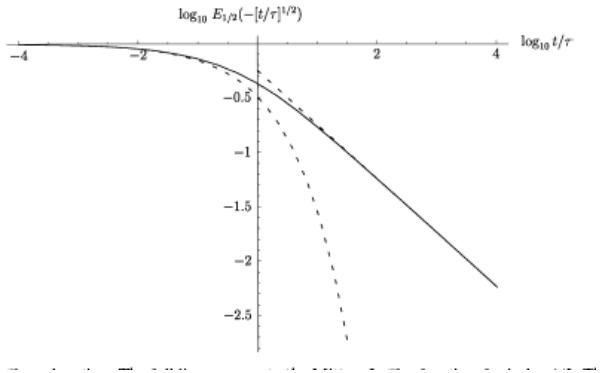
- $\Phi(t) \neq c\delta(t)$: non-Markov process with memory kernel $\Phi(t)$.
  For the choice of $\hat{\Phi}(u) = \frac{1}{\lambda} u^{\beta-1}$

$$\frac{\partial^\beta}{\partial t^\beta} p(x,t) = \lambda \int [p(x-x',t) - p(x,t)]\mu(x')dx'$$

Hence the process is Markov only in the case of exponential waiting times.

Notations:

- $E_\beta = \sum_{n=0}^{\infty} \frac{z^n}{\Gamma(\beta n + 1)}$ is the Mittag-Leffler function, which interpolates between the stretched exponential form and long-time inverse power law behavior.



- Caputo fractional derivative definition

$$\frac{d^\beta}{dt^\beta} f(t) = \frac{1}{\Gamma(1-\beta)} \int_0^t \frac{f'(\tau)}{(t-\tau)^\beta} d\tau,$$

for which

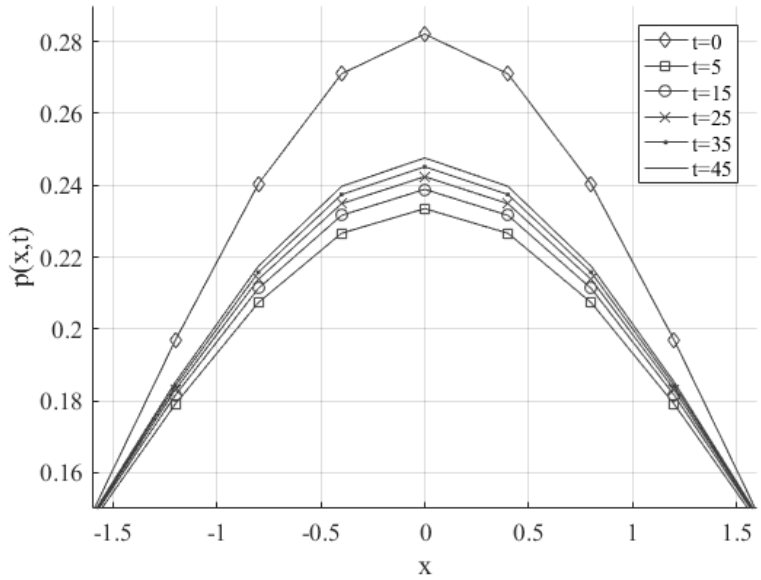
$$\mathcal{L}\left[\frac{d^\beta}{dt^\beta} f(t)\right] = u^\beta \hat{f}(u) - u^{\beta-1} f(0).$$

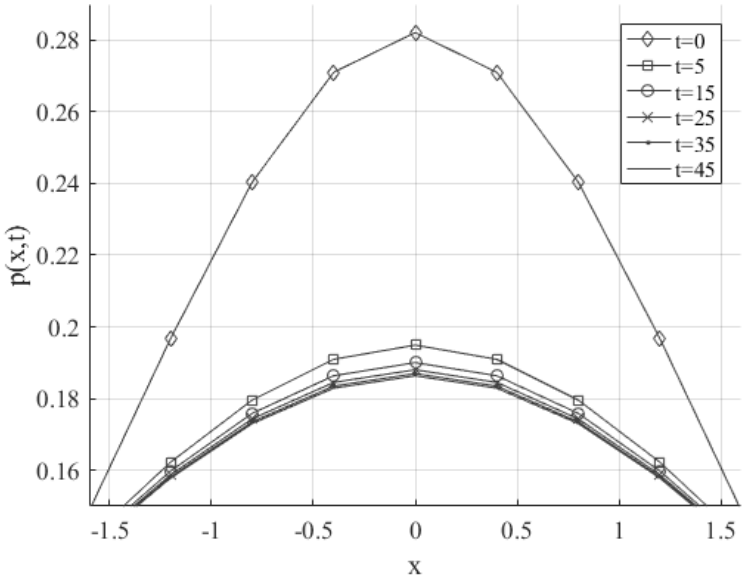
Some results on generalized master equations

[Torrejon, Emelianenko: SIAP 2018]

Jump sizes	Inter-arrival times
	Homogeneous Poisson case, arrival rate λ
$\boldsymbol{\mu(r)}$	$\frac{\partial}{\partial t}p(x,t) = \lambda \int_{\mathbb{R}} \mu(r)[p(x-r,t) - p(x,t)]dr$
$\boldsymbol{\mu(r,t)}$	$\frac{\partial}{\partial t}p(x,t) = \lambda \int_{\mathbb{R}} \mu(r,t)[p(x-r,t) - p(x,t)]dr$
$\boldsymbol{\mu(r,t_0)}$	$\begin{cases} \bar{p}(x,t) = p(x,t) + \frac{1}{\lambda}\frac{\partial}{\partial t}p(x,t) \\ \frac{\partial}{\partial t}\bar{p}(x,t) = \lambda \int_{\mathbb{R}} \mu(r,t)[\bar{p}(x-r,t) - \bar{p}(x,t)]dr \end{cases}$
	Fractional case, $\boldsymbol{w(s) = -\frac{d}{ds}E_\beta(-s^\beta)}$
$\boldsymbol{\mu(r)}$	$\partial_t^\beta p(x,t) = \int_{\mathbb{R}} \mu(r)[p(x-r,t) - p(x,t)]dr$
$\boldsymbol{\mu(r,t)}$	$\partial_t^\beta p(x,t) = \int_{\mathbb{R}} \mu(r,t)[p(x-r,t) - p(x,t)]dr$
$\boldsymbol{\mu(r,t_0)}$	$\begin{cases} \bar{p}(x,t) = \frac{\partial}{\partial t}p(x,t) + \partial_t^{1-\beta}p(x,t) + p(x,0)\Phi_\beta(t) \\ D_t^\beta \bar{p}(x,t) = \int_{\mathbb{R}} \mu(r,t)[\bar{p}(x-r,t) - \bar{p}(x,t)]dr \end{cases}$

Comparison of the evolution of instantaneous fractional CTRW (left) against delayed fractional CTRW (right) with jump kernel $\mu_1(r, t; 0)$, starting from the same Gaussian distribution.





Left graph keeps diffusing, while the right one undergoes a reversal in its trend (it goes down initially then starts growing).

Data-Driven Dynamics Identification Methods

- Dynamic Mode Decomposition (DMD) [1,2]
- Sparse Identification of Nonlinear Dynamics (SINDy) [2,3] and variations
- Koopman Operator Theory[1,2]
- Conditional Gradient-based Identification of Nonlinear Dynamics (CINDy)[4]
- Physics-informed neural networks [5]

---

[1] J. Nathan Kutz et. al., "Dynamic Mode Decomposition: Data-Driven Modeling of Complex Systems", SIAM, 2016.

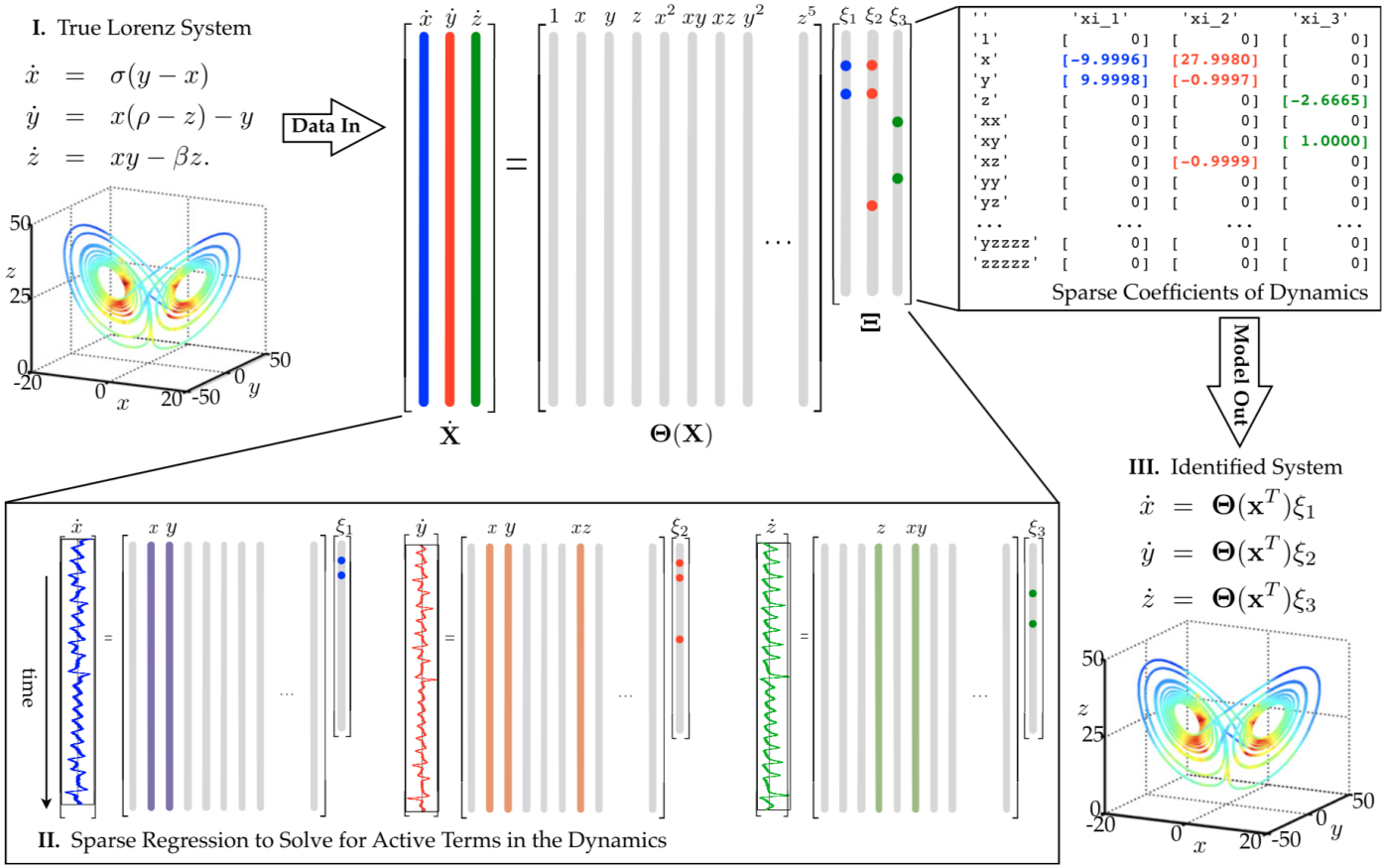
[2] Steven L. Brunton, J. Nathan Kutz, "Data-driven science and engineering", Cambridge University Press, 2019.

[3] S. L. Brunton et. al., "Discovering governing equations from data by sparse identification of nonlinear dynamical systems", Proceedings of the National Academy of Sciences, vol. 113, no. 15, 2016.

[4] A. Carderera et. al. "Cindy: Conditional gradient-based identification of non-linear dynamics – noise-robust recovery", 2021.

[5] Raissi et al, 2019.

Learning models from data



[Brunton et al, "Discovering governing equations from data by sparse identification of nonlinear dynamical systems", PNAS, 2016]

Modified Approach[6]

Consider the dynamical system

$$\frac{d}{dt}\mathbf{x} = \mathbf{f}(\mathbf{x})$$

[6] T. Oellerich, M. Emelianenko, "Learning biological network dynamics from data", 2023

Modified Approach[6]

Consider the dynamical system

$$\frac{d}{dt}\mathbf{x} = \mathbf{f}(\mathbf{x})$$

Let $\mathbf{X} = \begin{bmatrix}\mathbf{x}(t_1) & \dots & \mathbf{x}(t_m)\end{bmatrix}^T$ be the time series data collected from the system. Then, we can develop a library of candidate nonlinear functions $\mathbf{\Theta}(\mathbf{X})$ constructed from the data in $\mathbf{X}$:

$$\mathbf{\Theta}(\mathbf{X}) = \begin{bmatrix}\mathbf{1} & \mathbf{X} & \mathbf{X}^2 & \dots & \mathbf{X}^d & \dots & \sin(\mathbf{X}) & \dots\end{bmatrix}$$

Now consider an approximation of $\mathbf{f}(\mathbf{x})$ using a generalized linear model:

$$\mathbf{f}_k(\mathbf{x}) \approx \mathbf{C}(\mathbf{x})\boldsymbol{\omega}_k \qquad \mathbf{C}(\mathbf{x}) = \begin{bmatrix}\mathbf{\Theta}(\mathbf{x}) \\ -\mathbf{\Theta}(\mathbf{x})\end{bmatrix}$$

where $\boldsymbol{\omega}_k$ is ≥ 0.

[6] T. Oellerich, M. Emelianenko, "Learning biological network dynamics from data", 2023

Non-Negative Least Squares (NNLS)

In order to find the optimal $\boldsymbol{\omega}_k$, we will utilize the time series data in the **Non-negative Least Squares (NNLS)** algorithm[7] as follows:

$$\boldsymbol{\omega}_k = \underset{\boldsymbol{\omega}_k^{'} \geq 0}{\arg\min} \left\| \begin{bmatrix} \boldsymbol{\Theta}(\mathbf{X}) \\ -\boldsymbol{\Theta}(\mathbf{X}) \end{bmatrix} \boldsymbol{\omega}_k^{'} - \dot{\mathbf{X}}_k \right\|_2$$

where the top entries of $\boldsymbol{\omega}_k$ will correspond to positive coefficients in the recovered dynamics and the bottom entries are the negative. Under certain conditions, this problem is effectively related to

$$\boldsymbol{\xi}_k = \underset{\boldsymbol{\xi}_k^{'} \geq 0}{\arg\min} \|\dot{\mathbf{X}}_k - \boldsymbol{\Theta}(\mathbf{X})\boldsymbol{\xi}_k^{'}\|_2 + \alpha \mathbf{I}^T \boldsymbol{\xi}_k^{'}$$

[7] Slawski, Hein (2013): Consistency and sparse recovery without regularization via NNLS

Non-Negative Least Squares (NNLS)

In order to find the optimal $\boldsymbol{\omega}_k$, we will utilize the time series data in the **Non-negative Least Squares (NNLS)** algorithm[7] as follows:

$$\boldsymbol{\omega}_k = \arg\min_{\boldsymbol{\omega}_k' \geq 0} \left\| \begin{bmatrix} \boldsymbol{\Theta}(\mathbf{X}) \\ -\boldsymbol{\Theta}(\mathbf{X}) \end{bmatrix} \boldsymbol{\omega}_k' - \dot{\mathbf{X}}_k \right\|_2$$

where the top entries of $\boldsymbol{\omega}_k$ will correspond to positive coefficients in the recovered dynamics and the bottom entries are the negative. Under certain conditions, this problem is effectively related to

$$\boldsymbol{\xi}_k = \arg\min_{\boldsymbol{\xi}_k' \geq 0} \|\dot{\mathbf{X}}_k - \boldsymbol{\Theta}(\mathbf{X})\boldsymbol{\xi}_k'\|_2 + \alpha \mathbf{I}^T \boldsymbol{\xi}_k'$$

Comparatively, SINDy solves the optimization problem:

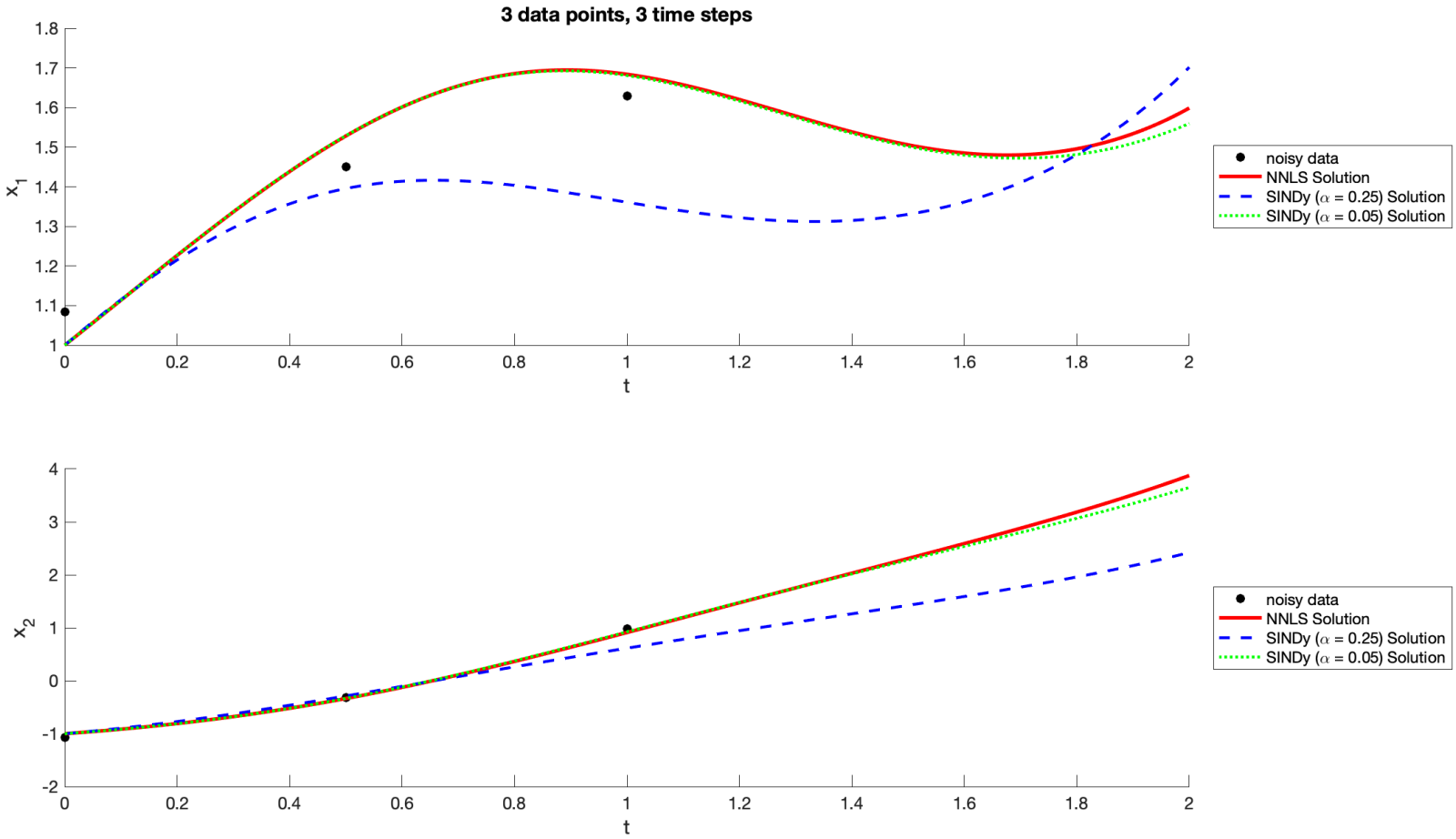
$$\boldsymbol{\xi}_k = \arg\min_{\boldsymbol{\xi}_k'} \|\dot{\mathbf{X}}_k - \boldsymbol{\Theta}(\mathbf{X})\boldsymbol{\xi}_k'\|_2 + \alpha \|\boldsymbol{\xi}_k'\|_1$$

where α weights the sparsity constraint.
Can be solved via active set method, gradient descent, etc.

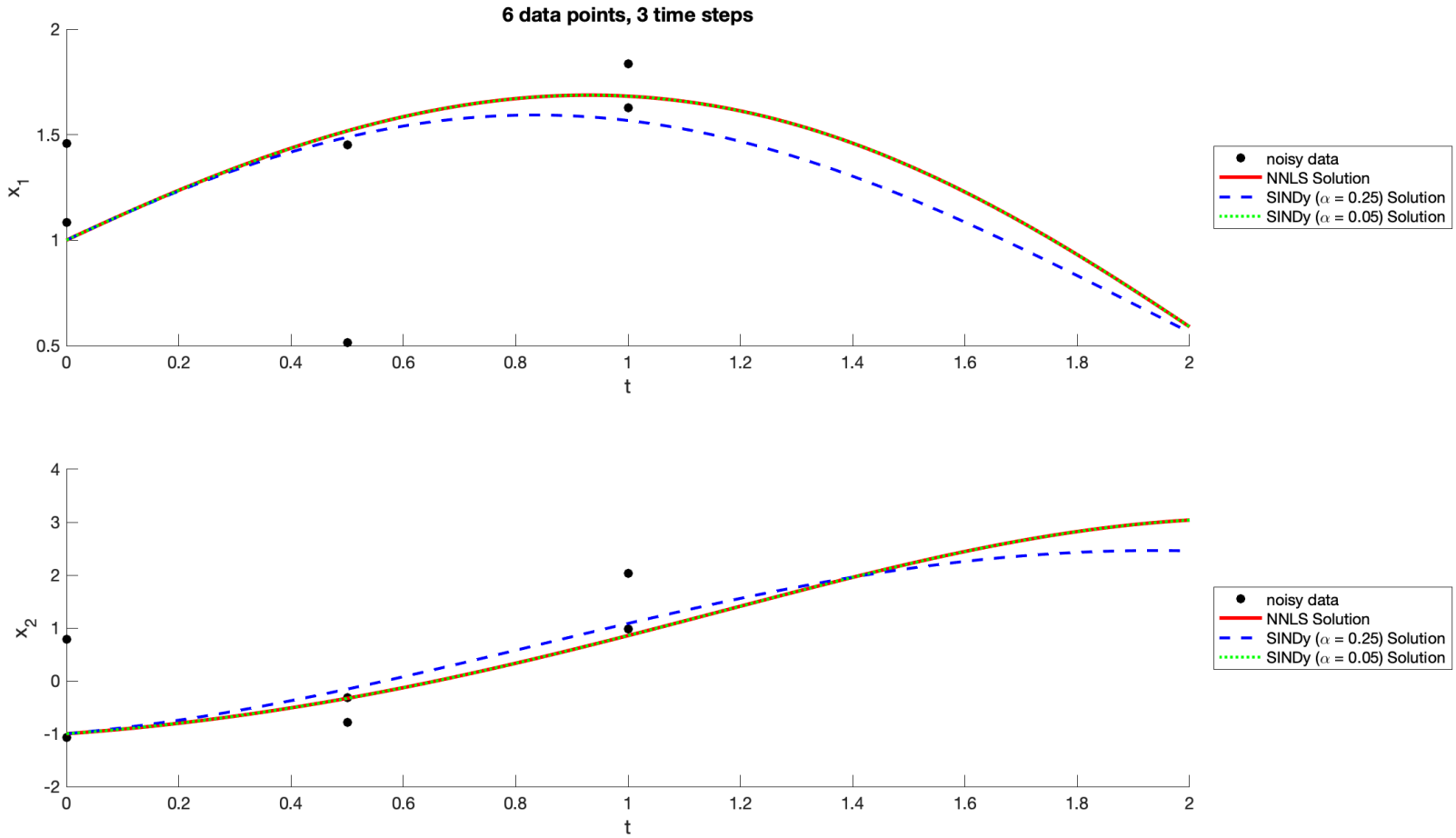
[7] Slawski, Hein (2013): Consistency and sparse recovery without regularization via NNLS

Comparing NNLS and SINDy

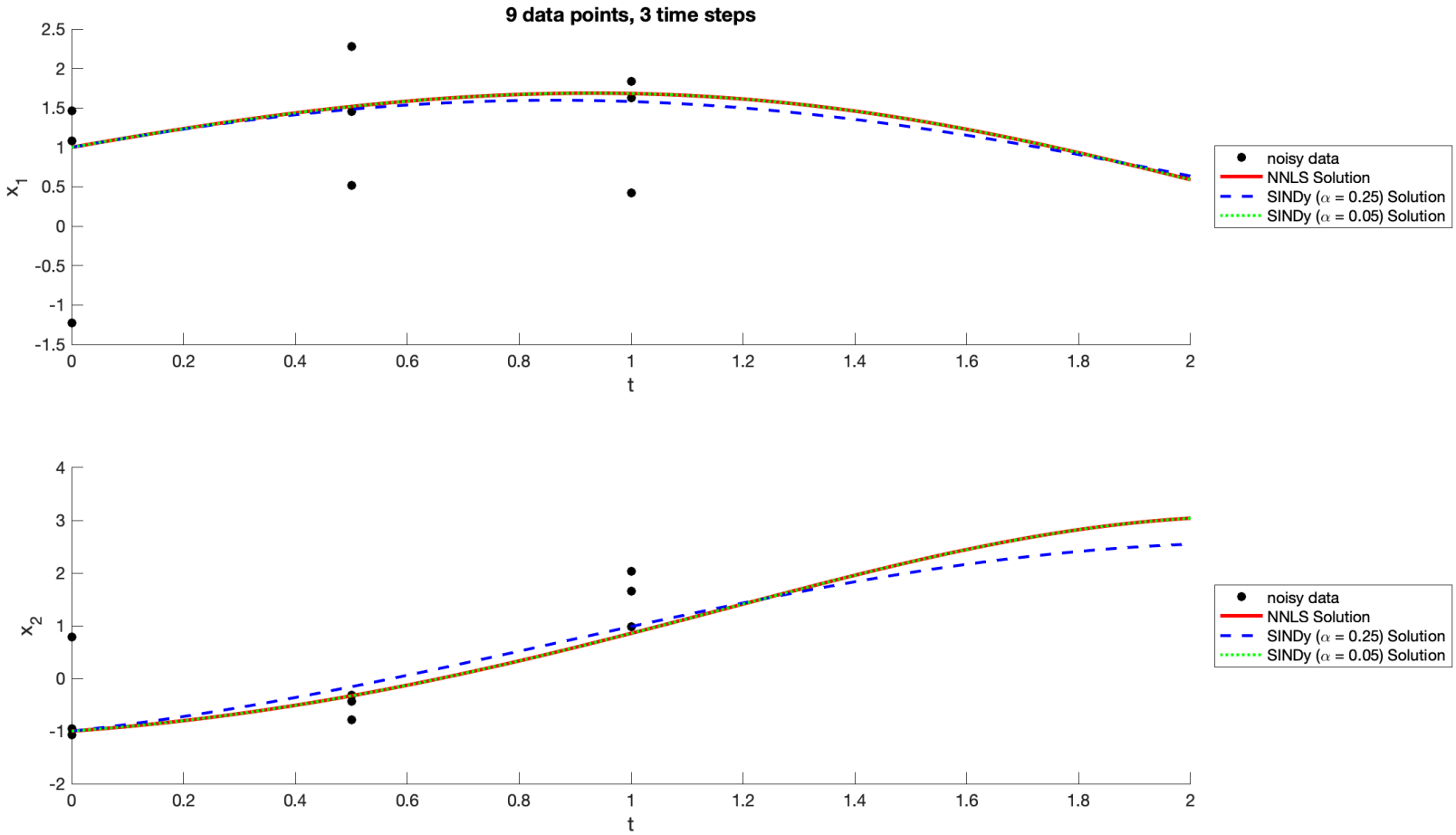
$$\frac{dx_1}{dt} = 0.5 - 0.75x_2 \qquad \frac{dx_2}{dt} = -1 + 2x_1 + 0.25x_1x_2$$



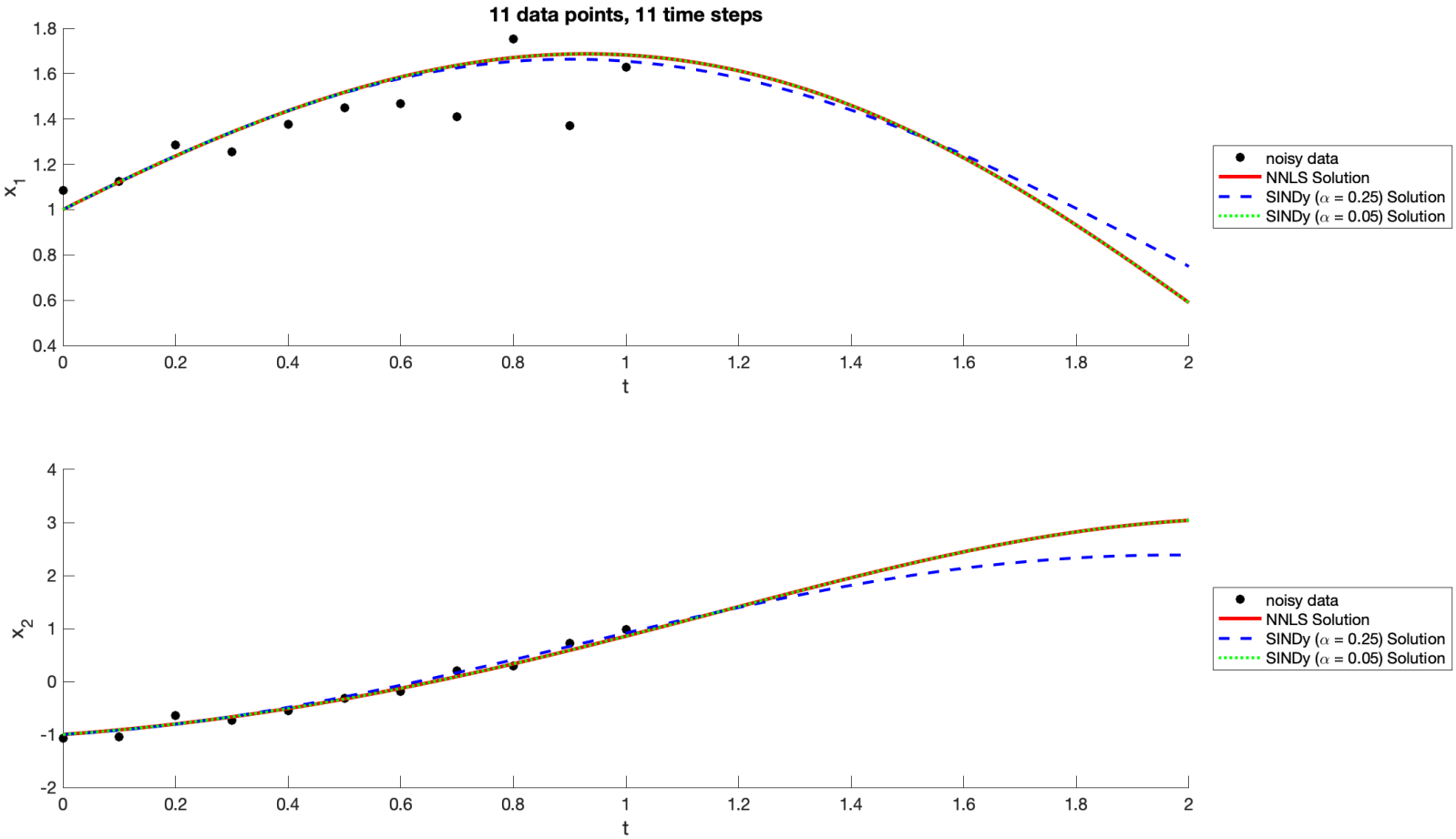
$$\frac{dx_1}{dt} = 0.5 - 0.75x_2 \qquad \frac{dx_2}{dt} = -1 + 2x_1 + 0.25x_1x_2$$



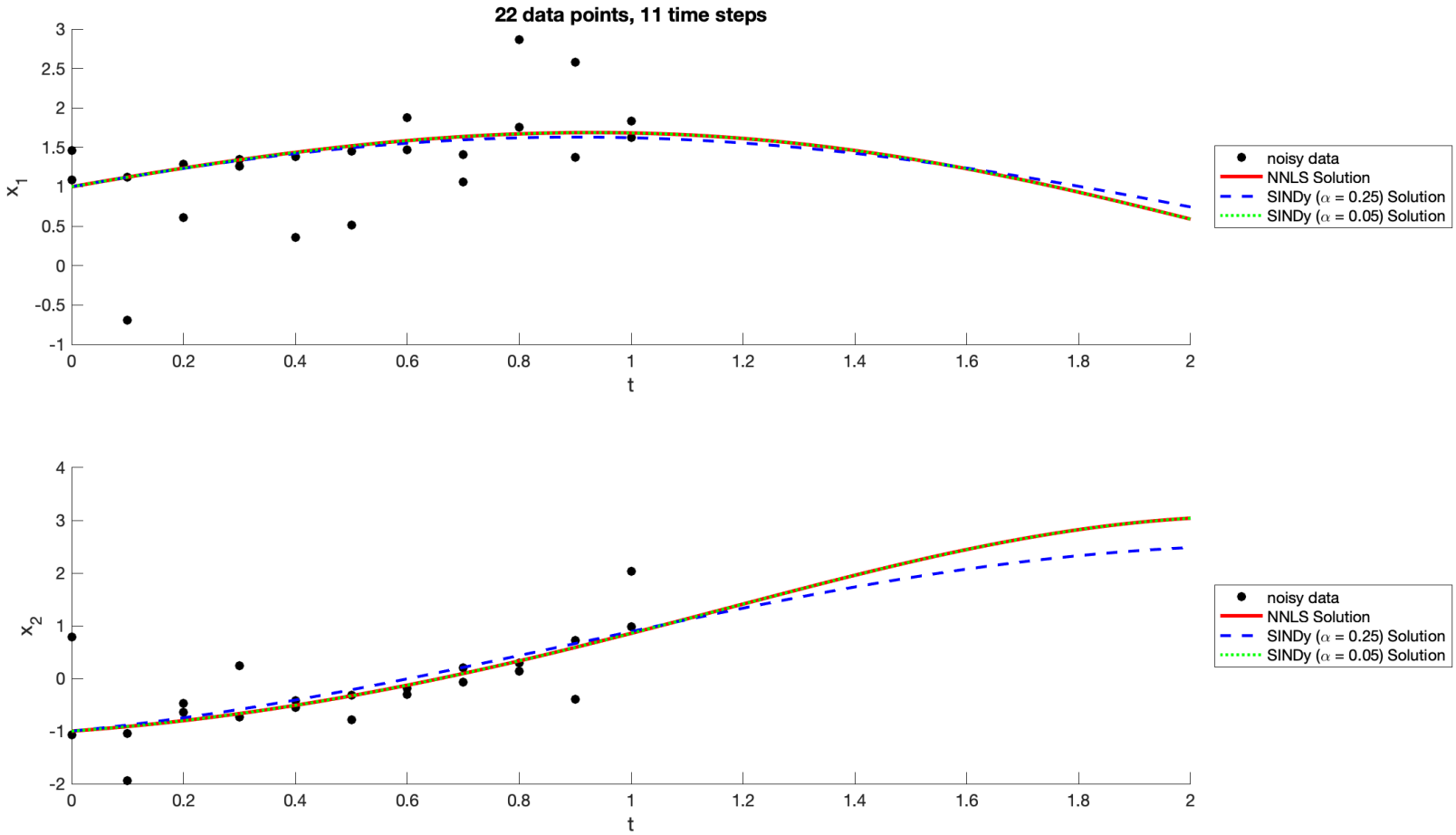
$$\frac{dx_1}{dt} = 0.5 - 0.75x_2 \qquad \frac{dx_2}{dt} = -1 + 2x_1 + 0.25x_1x_2$$



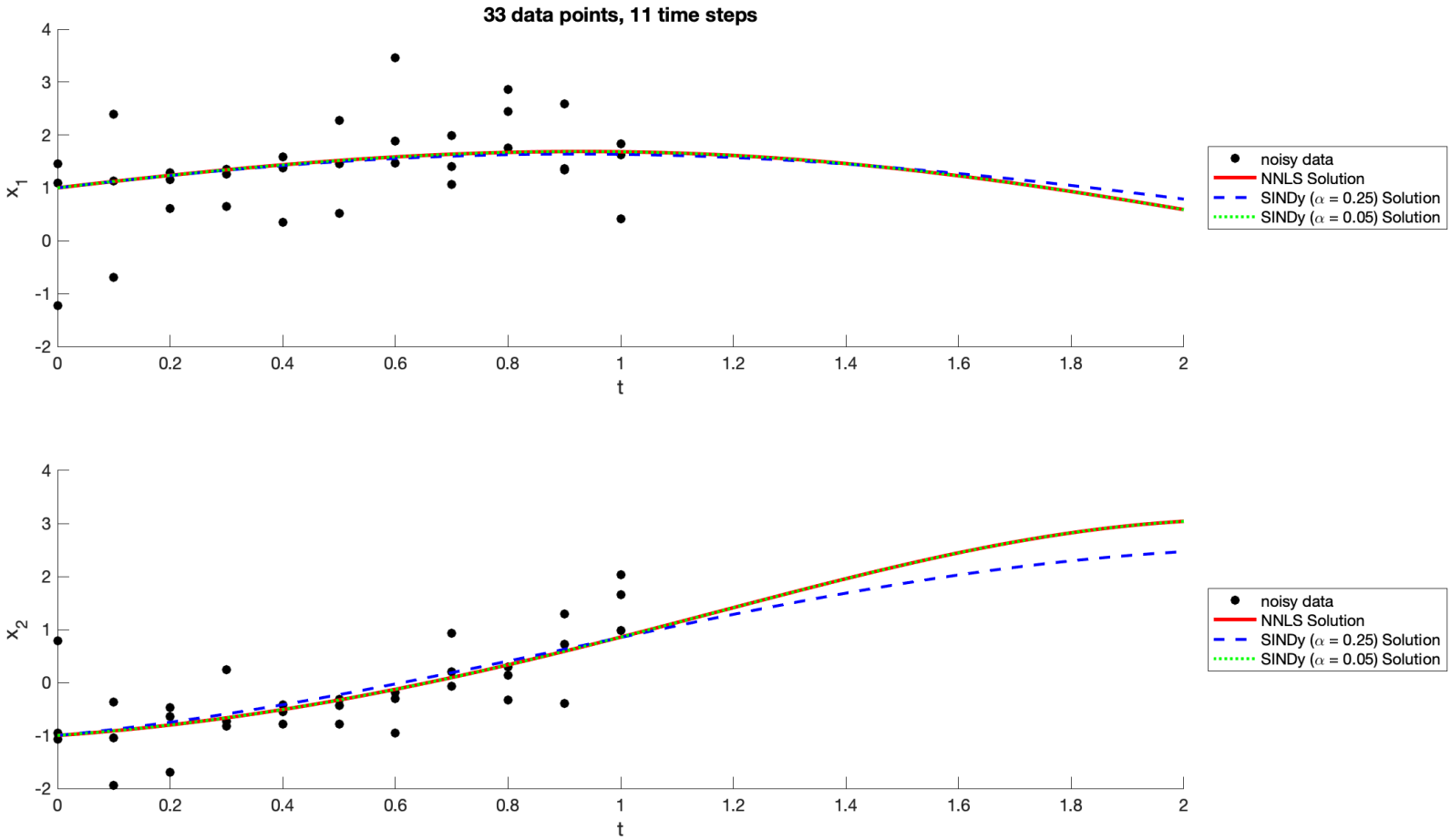
$$\frac{dx_1}{dt} = 0.5 - 0.75x_2 \qquad \frac{dx_2}{dt} = -1 + 2x_1 + 0.25x_1x_2$$



$$\frac{dx_1}{dt} = 0.5 - 0.75x_2 \qquad \frac{dx_2}{dt} = -1 + 2x_1 + 0.25x_1x_2$$

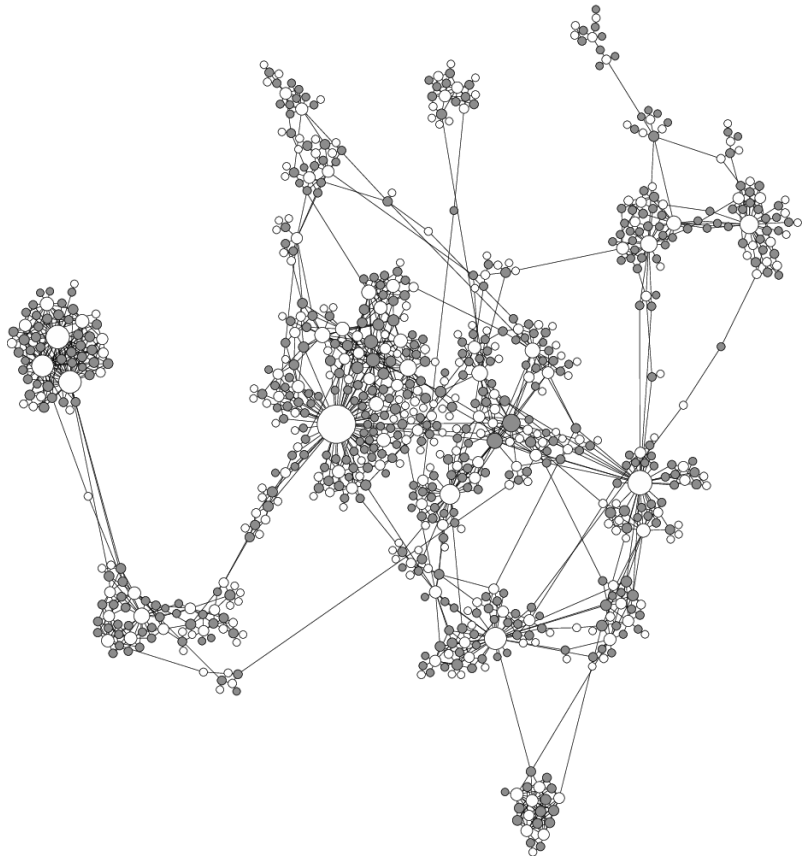


$$\frac{dx_1}{dt} = 0.5 - 0.75x_2 \qquad \frac{dx_2}{dt} = -1 + 2x_1 + 0.25x_1x_2$$



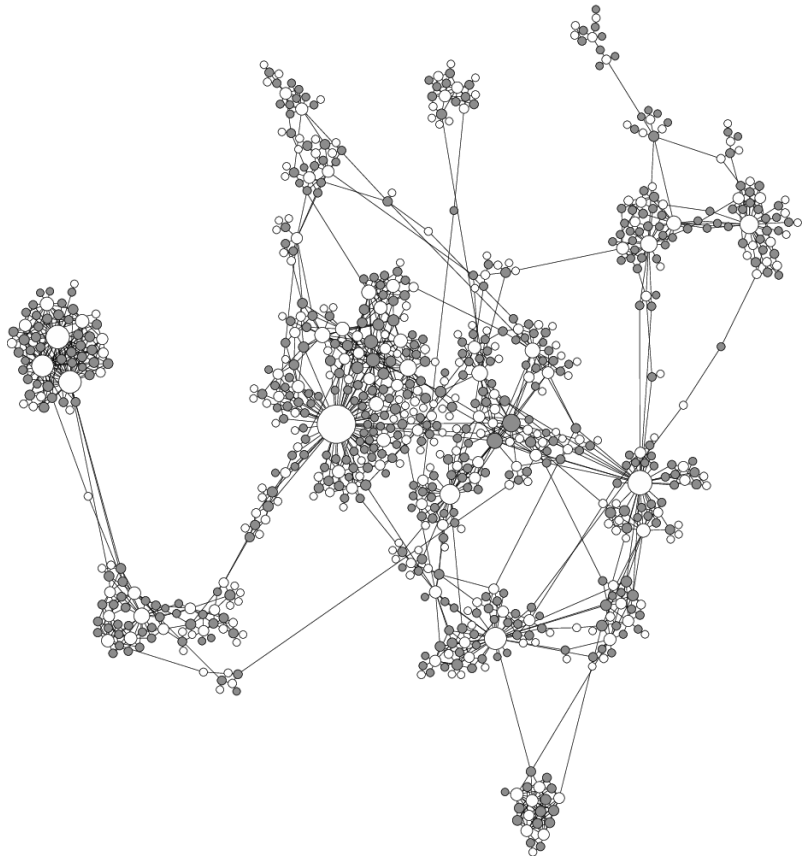
Mathematical concepts so far

1. Entropy as a tool for characterizing disorder
2. Random walks as a tool for characterizing dynamics
3. Data-driven sparse recovery of dynamics without regularization



Mathematical concepts so far

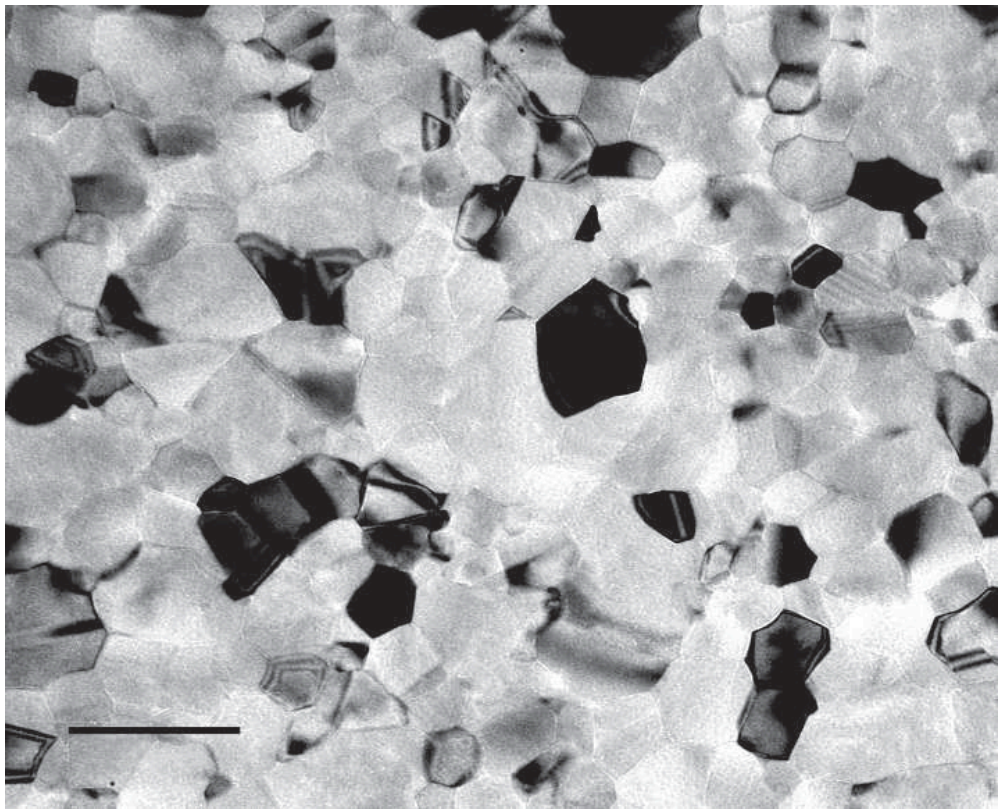
1. Entropy as a tool for characterizing disorder
2. Random walks as a tool for characterizing dynamics
3. Data-driven sparse recovery of dynamics without regularization

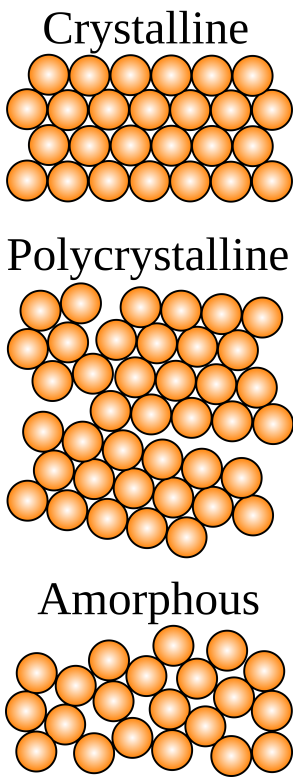


Now let's step into the world of applications!

Crossing into materials science

Polycrystalline materials are solids that are composed of many crystallites (grains) of varying size and orientation. The variation in directions is called texture and can be random or directed. Microstructure is the collection of grains with associated orientations.





Disorder in microstructures

Joint work with K. Saleme Ruiz [MSMSE 2018]

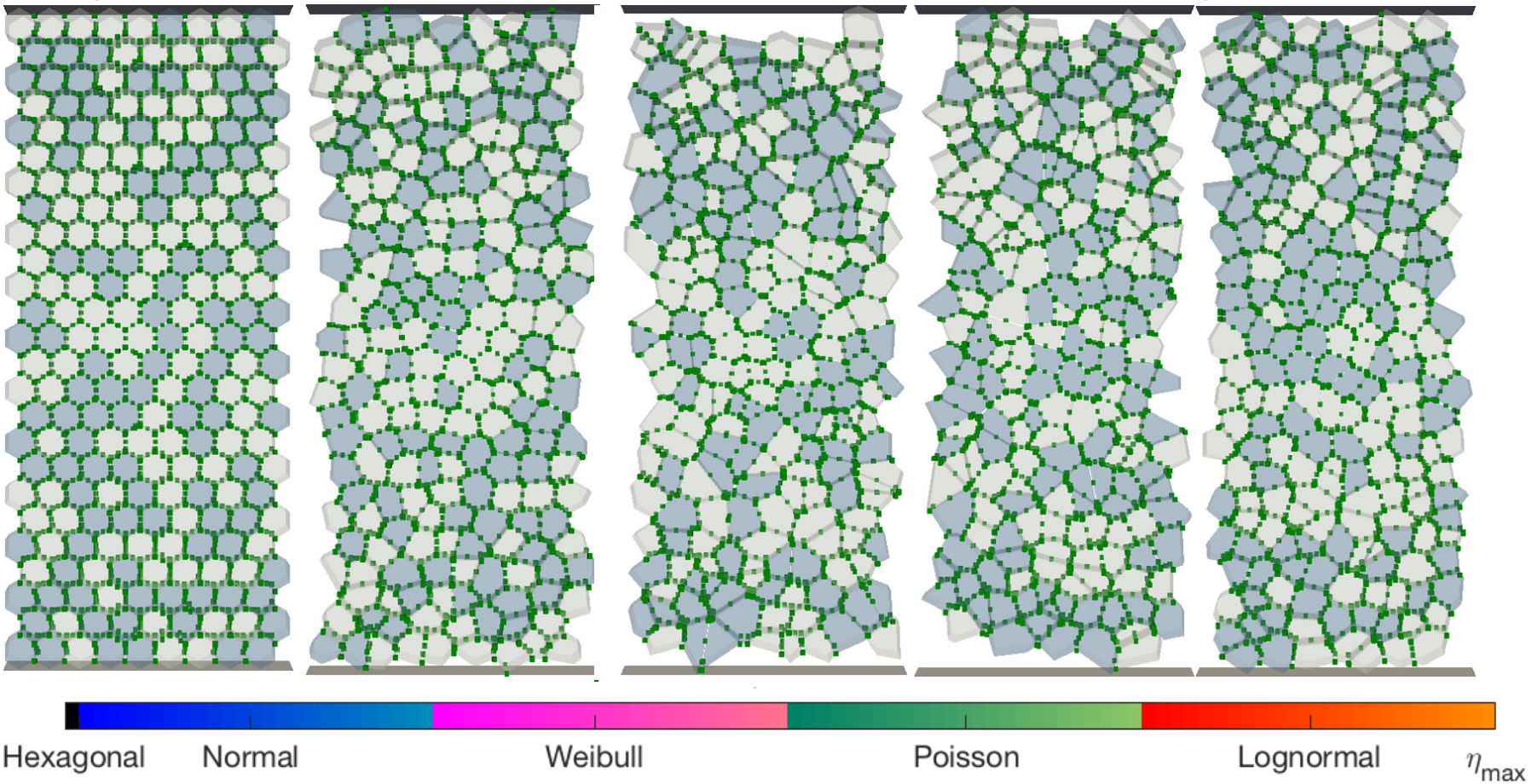
Entropy can be used to characterize microstructure disorder:

$$\eta_S(X) = -\sum_{k=1}^{K} p^k \ln(p^k)$$

$p(X)$: probability density of a certain random variable X.
X: areas, number of sides, orientations etc.

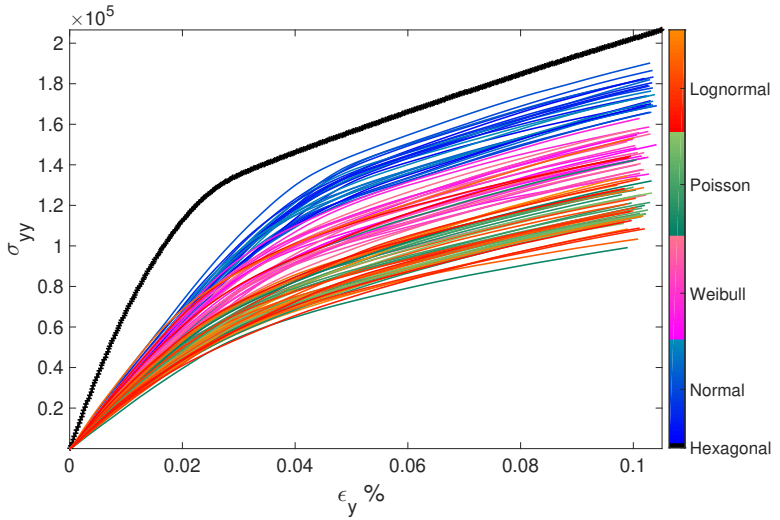
Disorder in microstructures

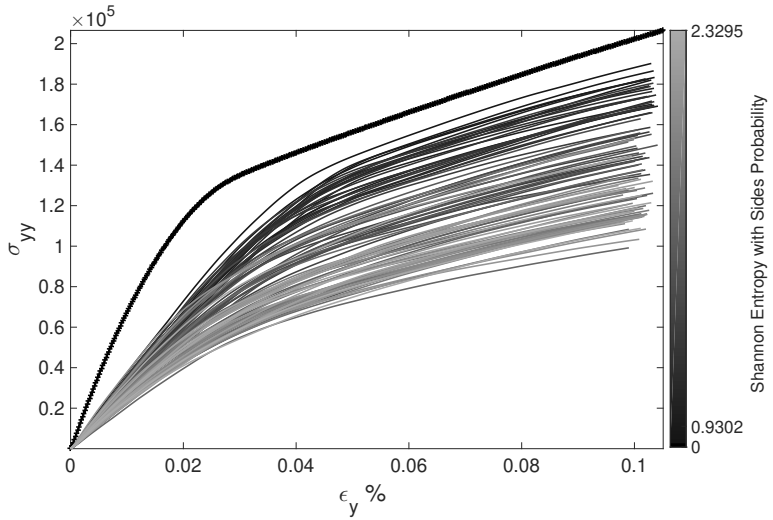
Types of random perturbation of microstructure:
hexagonal, Normal-Voronoi, Weibull-Voronoi, Poisson-Voronoi, Lognormal-Voronoi.



Mechanical response

Discrete Element Method (DEM) code was used to calculate stress-strain curves in a quasi-static uniaxial compression loading test. 25 samples of each type of perturbation were considered.





Darker curves have smaller Shannon side-based entropy.

Observation:

Better elastic properties are linked to lower **side-based Shannon entropy**.

Types of configurational entropy

- Renyi entropy:

$$\eta_\beta(X) = \frac{1}{1-\beta} \ln \sum_{k=1}^{K} (p^k)^\beta$$

- Shannon entropy:

$$\eta_S(X) = -\sum_{k=1}^{K} p^k \ln(p^k) = \lim_{\beta \to 1} \eta_\beta(X)$$

What is the effect of parameter β on the microstructure characterization?

Comparing microstructures

Jensen-Renyi divergence: distance between microstructures X and Y:

$$Dist_{JR}(X,Y) = \sqrt{\eta_\beta(p_Z) - \frac{1}{2}\eta_\beta(p_X) - \frac{1}{2}\eta_\beta(p_Y)}$$

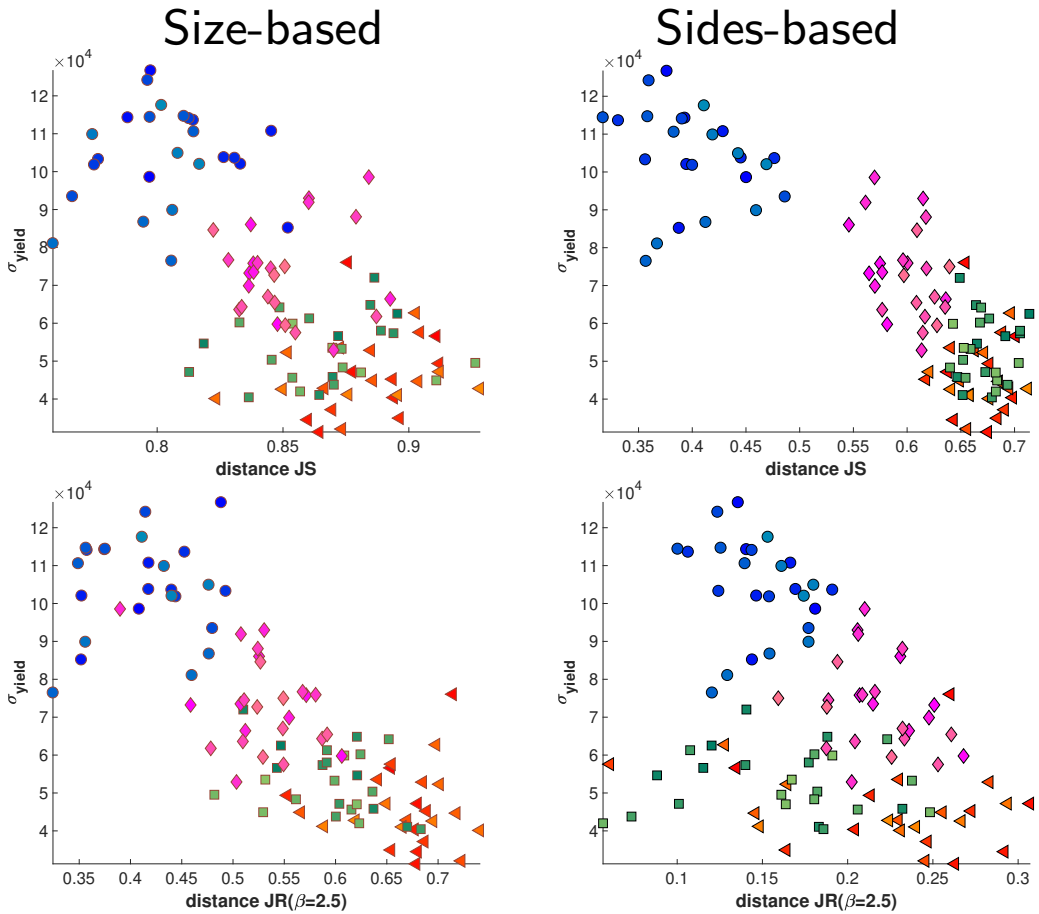
where $p_Z = \frac{1}{2}p_X + \frac{1}{2}p_Y$.
We consider 2 types of statistics p_X:

1. Size-based entropy (X is the grain area)
2. Side-based entropy (X is the number of sides)

Yield stress shows a clear difference between side- and size- based entropies.

$\beta = 1$
better
for sides

$\beta = 2.5$
better
for sizes

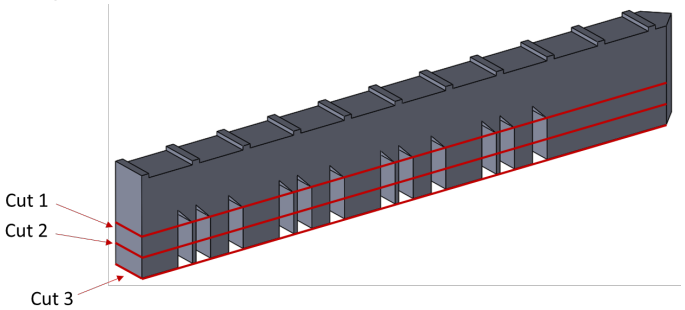


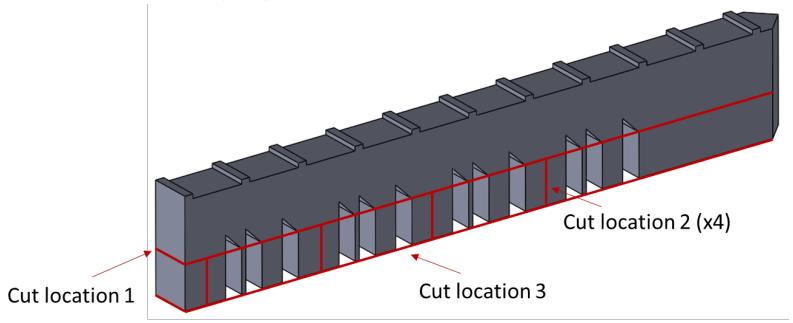
Observation:

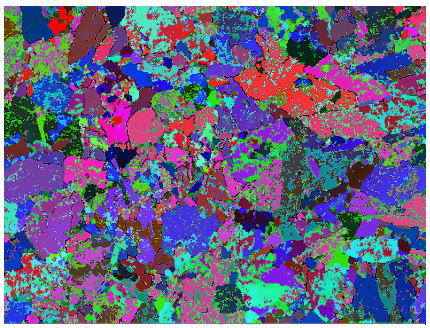
For **size-based entropy measures**, $\beta = 2.5$ better correlates with elastic response than $\beta = 1$.

Microstructure for different configurations

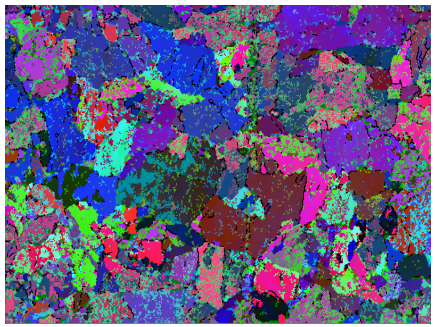
NIST AM Benchmark LPBF test data for 15-5 stainless steel



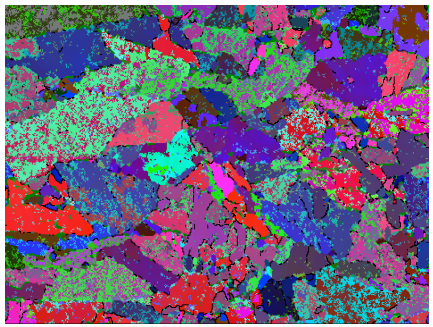




Leg 7 (thick)
Longitudinal orientation

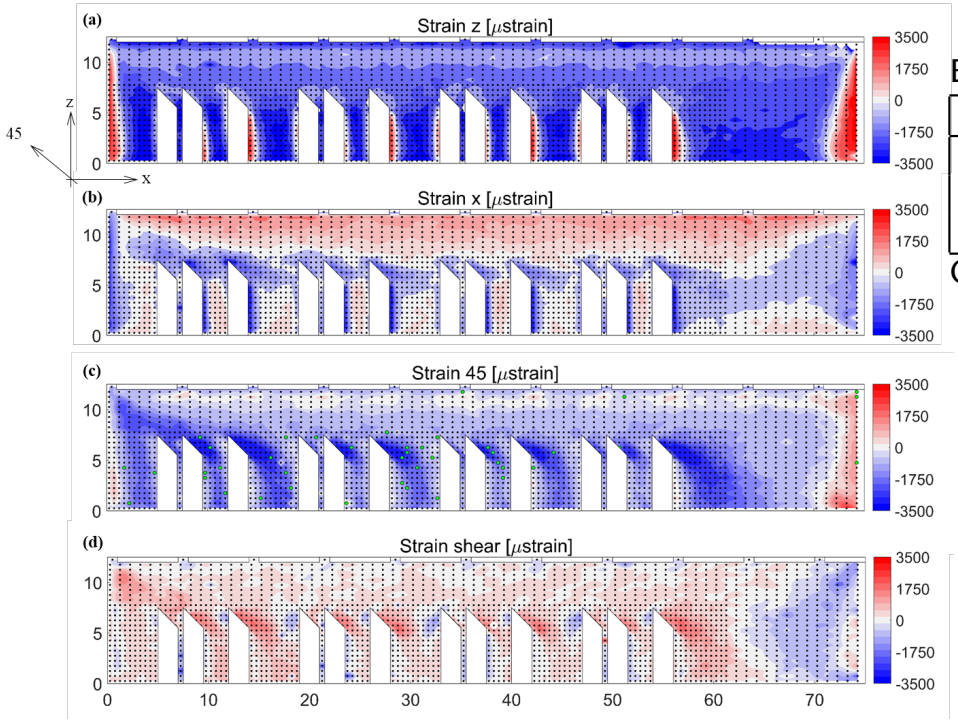


Leg 7 (thick)
Transverse orientation



Leg 8 (thin)
Transverse orientation

AM samples: residual strain and entropy



Entropy measured from NIST benchmark samples:

Sample	η roundness	η circularity
L7 (thick) LONG	2.1162	2.2452
L7 (thick) TRANS	2.1819	2.2580
L8 (thin) TRANS	2.1499	2.2375

Open questions:

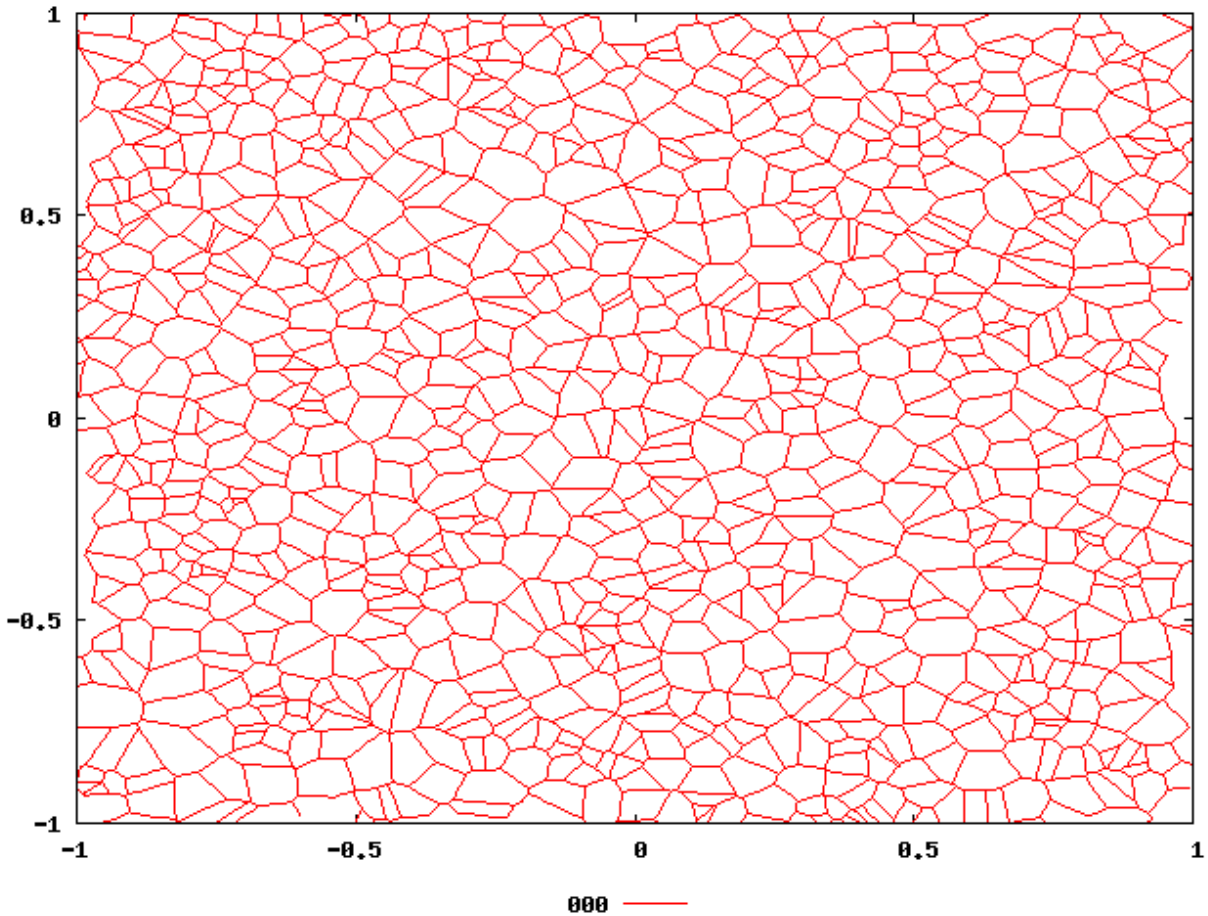
- Which microstructure properties best correlate with residual stresses and strains: size/shape distributions, entropies?
- How does the melt pool geometry/characteristics affect stress/strain distribution?
- How sensitive are entropy calculations to the quality of the segmentation?

Entropy for smart materials design

1. Different types of microstructure entropies correlates with different materials properties
2. We can use this concept to reverse engineer microstructures with desired mechanical response

Grain growth

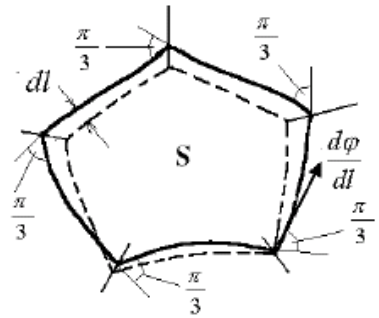
Grain growth is the increase in size of grains (crystallites) in a material at high temperature. This occurs when recovery and recrystallization are complete and further reduction in the internal energy can only be achieved by reducing the total area of grain boundary.

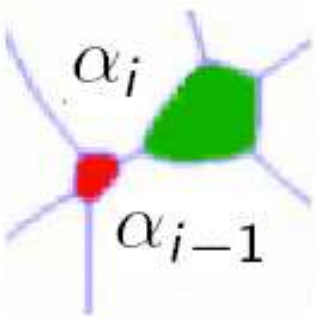


Texture modeling in 2D

Types of collective variables:

1. number of sides of a grain (n)
2. relative area of a grain (S)
3. dihedral angle (ϕ)
4. misorientation ($\beta = \Delta\alpha$)

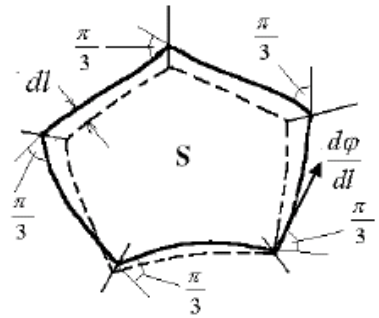


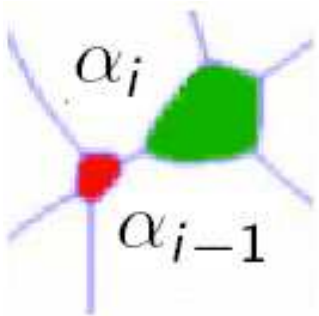


Texture modeling in 2D

Types of collective variables:

1. number of sides of a grain (n)
2. relative area of a grain (S)
3. dihedral angle (ϕ)
4. misorientation ($\beta = \Delta\alpha$)





Here we focus on **texture**, so relevant distributions are:

- $\rho_n(\beta, t)$: MDF, misorientation distribution function (number-weighted pdf)
- $\rho_l(\beta, t)$: GBCD, grain boundary character distribution (length-weighted pdf)

What are the mechanisms that control the development of these distributions?

Entropy theory of materials texture

(Joint work with D. Kinderlehrer, Y. Epshteyn, E. Eggeling, K. Barmak, S. Ta'asan, R. Sharp)
[Barmak et al 2011-2015]

The total surface energy of the system: $E[\rho] = \int \gamma(\beta)\rho(\beta, t)\, d\beta$.

Add a configurational entropy term to get the free energy:

$$F_\sigma[\rho] = E[\rho] + \sigma \int \rho(\beta) \log(\rho(\beta))\, d\beta,$$

where σ is a temperature-like parameter.

Entropy model

Mass transport paradigm (Benamou and Bernier, 2000): Given two probability densities f and f^*, Wasserstein distance $d(f, f^*)$ between them is given by

$$\frac{1}{\tau} d(f, f^*)^2 = \inf \int_0^\tau \int_\Omega v^2 f d\psi dt$$

over deformation paths $f(\psi, t)$ subject to

$$f_t + (cf)_\psi = 0, \qquad f(\psi, 0) = f^*(\psi), f(\psi, \tau) = f(\psi)$$

Variational principle:

$$\frac{\mu}{2\tau} d(\rho, \rho^*)^2 + F_\lambda(\rho) = \inf_\eta \{ \frac{\mu}{2\tau} d(\eta, \rho^*)^2 + F_\lambda(\eta) \}$$

Can be solved using implicit scheme obtained by setting $\rho^* = \rho^{k-1}, \rho^k = \rho$.
We can recover GBCD ρ as

$$\rho(\alpha, t) = \lim_{\tau \to 0} \rho^\tau(\alpha, t)$$

Hence it satisfies Fokker-Planck equation [Jordan,Kinderlehrer,Otto 1998]:

$$\frac{\partial \rho}{\partial t} = \frac{\partial}{\partial \alpha} (\lambda \frac{\partial \rho}{\partial \alpha} + \frac{d\gamma}{d\alpha} \rho)$$

Entropy-based model validation

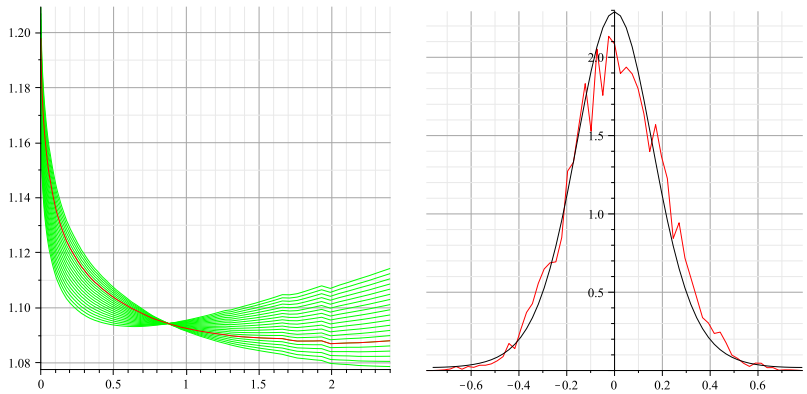
Minimization of F_λ gives a Boltzmann distribution (stationary solution to Fokker-Planck): $\rho_\lambda(\alpha) = e^{-\gamma(\alpha)/\lambda}$, where $Z_\lambda = \int_\Omega e^{-\gamma(\alpha)/\lambda}$.

Question: does it match the distribution obtained in simulation?
Relative entropy approach: Seek λ s.t. relative entropy $\Phi_\lambda(\eta) = F_\lambda(\eta) + \lambda \log Z_\lambda$ exhibits an exponential decay.

Entropy-based model validation

Minimization of F_λ gives a Boltzmann distribution (stationary solution to Fokker-Planck): $\rho_\lambda(\alpha) = e^{-\gamma(\alpha)/\lambda}$, where $Z_\lambda = \int_\Omega e^{-\gamma(\alpha)/\lambda}$.

Question: does it match the distribution obtained in simulation?
Relative entropy approach: Seek λ s.t. relative entropy $\Phi_\lambda(\eta) = F_\lambda(\eta) + \lambda \log Z_\lambda$ exhibits an exponential decay.



(1) There is a unique such λ
(2) The distribution for this choice of λ precisely matches the simulation result!
[Barmak, Eggeling, E, Epshteyn, Kinderlehrer, Taasan, Sharp, PRB (2011)]

Entropy theory for texture in 2d

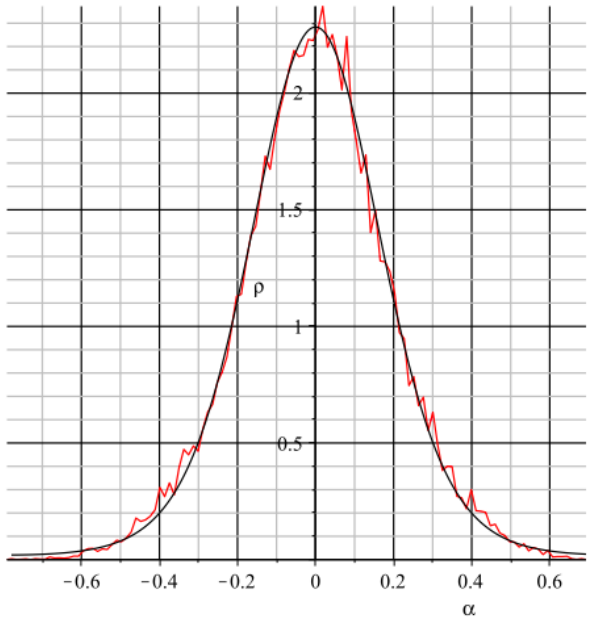
Given interfacial energy $\gamma(\alpha)$, we compare Boltzmann distribution for the special value of parameter λ with simulation:

- Compute relative entropy - it is exponentially decaying for a unique value of λ
- Compute Boltzmann distribution for the same λ
- Compare with the stationary GBCD

Entropy theory for texture in 2d

Given interfacial energy $\gamma(\alpha)$, we compare Boltzmann distribution for the special value of parameter λ with simulation:

- Compute relative entropy - it is exponentially decaying for a unique value of λ
- Compute Boltzmann distribution for the same λ
- Compare with the stationary GBCD



Compelling evidence: this method yields a good fit in 2d.

Random walk model

What we know:

1. Coarsening process contains discontinuities (jumps) created by topological reconfigurations
2. The rate with which these events occur (arrival rate λ) goes down over time
3. Knowing the arrival rate, we can write the evolution equation for the mesoscale statistics

Idea:

- Use non-homogeneous Poisson-type jump process to describe evolution of the statistics

Random walk model

(Joint work with D. Torrejon, D. Golovaty [J. Ellip. Parab. Eq. 2016])

Master equation for the random walk in misorientation space:

$$\frac{\partial \rho(\beta, t)}{\partial t} = \lambda(t) \int (\rho(\beta - \Delta\beta, t) - \rho(\beta))\mu(\Delta\beta, t) d(\Delta\beta)$$

Parameters: jump kernel and waiting times distribution (estimated from simulations).

Results:
(+) Able to capture correct evolution in 1D,
($\pm$) Reproduces 2D statistics up to a certain stage, fails to capture stagnation
($-$) Needs to have the jump and waiting times pdfs

Hierarchy of mesoscale models:

- Entropy model: energy gradient flow description (Kinderlehrer, E., Epshteyn, Liu et al)
  Works for certain types of energy, one parameter σ

$$\frac{\partial \rho}{\partial t} = \frac{\partial}{\partial \beta}(\sigma \frac{\partial \rho}{\partial \beta} + \frac{d\gamma}{d\beta}\rho)$$

- Kinetic model: most granular description of collisions (E., Yegorov) Works for arbitrary energy, parameter λ

$$\rho_n(\beta, t) + N(t)\frac{\partial \rho_n(\beta, t)}{\partial N(t)} = -\frac{\lambda(t)}{3}(P^{s,+}(\beta, t) - P^{s,-}(\beta, t)) + \frac{1}{3}P_1^d(\beta, t) + \frac{2}{3}P_2^d(\beta, t)$$

- Random walk model: mean field approximation of collisions (E., Golovaty, Torrejon et al)
  Works with arbitrary energy, parameters λ and μ

$$\frac{\partial \rho(\beta, t)}{\partial t} = \lambda(t) \int (\rho(\beta - \Delta\beta, t) - \rho(\beta))\mu(\Delta\beta, t)d(\Delta\beta)$$

Hierarchy of mesoscale models:

- Entropy model: energy gradient flow description (Kinderlehrer, E., Epshteyn, Liu et al)
  Works for certain types of energy, one parameter σ

$$\frac{\partial \rho}{\partial t} = \frac{\partial}{\partial \beta}(\sigma \frac{\partial \rho}{\partial \beta} + \frac{d\gamma}{d\beta}\rho)$$

- Kinetic model: most granular description of collisions (E., Yegorov) Works for arbitrary energy, parameter λ

$$\rho_n(\beta,t)+N(t)\frac{\partial \rho_n(\beta,t)}{\partial N(t)} = -\frac{\lambda(t)}{3}(P^{s,+}(\beta,t)-P^{s,-}(\beta,t))+\frac{1}{3}P_1^d(\beta,t)+\frac{2}{3}P_2^d(\beta,t)$$

- Random walk model: mean field approximation of collisions (E., Golovaty, Torrejon et al)
  Works with arbitrary energy, parameters λ and μ

$$\frac{\partial \rho(\beta,t)}{\partial t} = \lambda(t)\int (\rho(\beta-\Delta\beta,t)-\rho(\beta))\mu(\Delta\beta,t)d(\Delta\beta)$$

Remaining challenges:

- Formal derivation of the continuum limits for granular models
- Generalization to 3D
- Can we account for changes during AM processing?

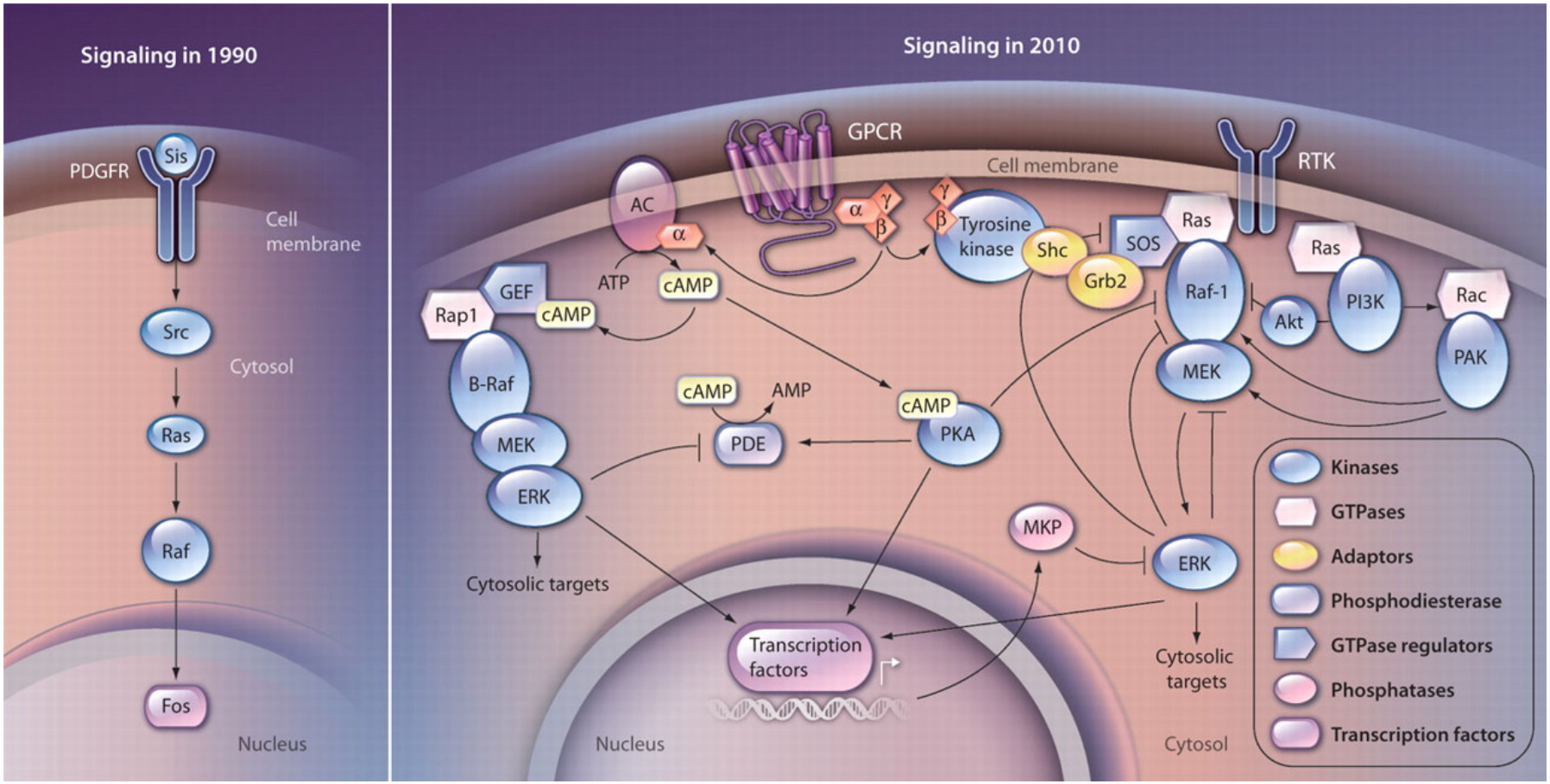
Entropy and random walk theories in polycrystals

1. Allow to describe evolution of the microstructure
2. Links microstructure to mechanical properties

Many questions remain about the mechanisms responsible for this behavior.

Simulations are extremely time consuming: see our GPU-based version in [Sazo 2019].

Crossing into biology



Protein interaction network

Biological networks are graph structures.
For instance, protein interaction network may be represented as a weighted graph with vertices representing proteins and with weights

$$p_{ij} = \frac{w_{ij}}{\sum_k w_{ik}},$$

where $w_{ij} = \frac{1}{2}(1 + C_{ij})$ are the transformed Pearson correlation coefficient C_{ij} of gene expression between genes i and j in the same phenotype.
Probability flux between nodes i and j is given by

$$E_{ij} = \gamma \sum_{L=1}^{\infty} \alpha_L p_{ij}^L,$$

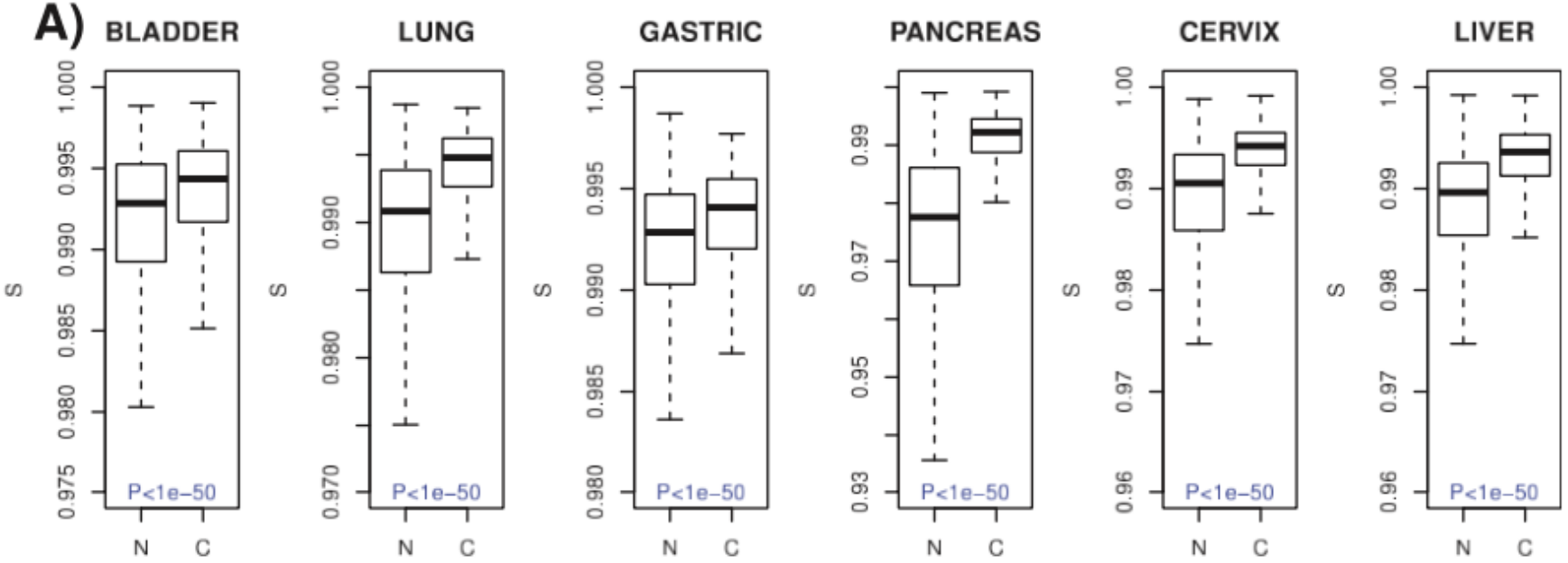
where P^L denotes flux on a path of length L.
Network entropy of a node i is:

$$S_i = -\frac{1}{\log k_i} \sum_j p_{ij} \log p_{ij},$$

where k_i is the degree of gene i, assumed to be ≥ 2.

Cancer increases entropy

Below local entropy distributions are shown for high-degree nodes.



[West et al, Sci Rep 2012]

Question: entropy grows as a consequence of the loss of local connectivity. Which nodes are critical in this process? Can we target them in drug therapy?

Suppose we have a network comprised of N nodes, labeled $\vec{P} = [P_1, P_2, \ldots, P_N]^T \in P$. Consider a time dependent input being applied to this system, denoted $\vec{U}(t) = \vec{U} \in \mathbf{U}$. The resulting dynamical system has the form:

$$\vec{P'} = \vec{F}(\vec{U}, \vec{P}) \tag{1}$$

where f_i models the set of activation/de-activation reactions for the respective node. We will use $\vec{F} = [f_1, f_2 \ldots f_N]^T$ to denote the vector of right hand sides and $h(\vec{P}) = \sum_{i=1}^{N} h_i P_i$ as the output.
Let us consider the first case, $\det(D_{\vec{P}}\vec{F}) \neq 0$.

Theorem[Araujo, Nature Comm. 2018]

A biochemical system for which $\det(D_{\vec{P}}\vec{F}) \neq 0$ at the system's steady state will exhibit homeostasis if

$$hC^T D_{\vec{U}}\vec{F} = 0$$

at steady state. Here $h = [h_1, ..., h_N]$ is the vector of coefficients which determine the weighted contribution from each node to the total output of the system and C denotes the cofactor matrix of $D_{\vec{P}}\vec{F}$.

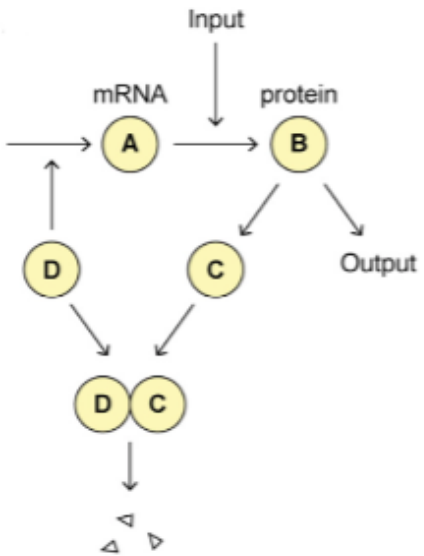
Adaptation condition generalized to singular systems

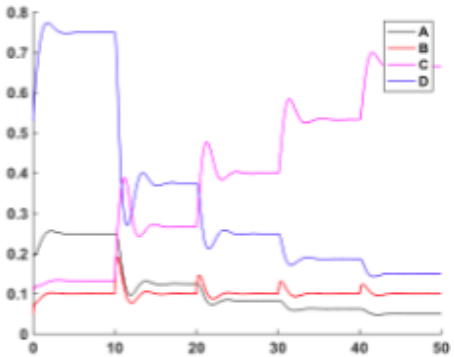
Theorem [Oellerich, Emelianenko, Araujo, Liotta, 2020]

A biochemical system for which $\det(D_{\vec{P}}\vec{F}) = 0$ at the system's steady state will exhibit homeostasis if

$$hV_1\Sigma_1^{-1}W_1^T D_{\vec{U}}\vec{F} = 0$$

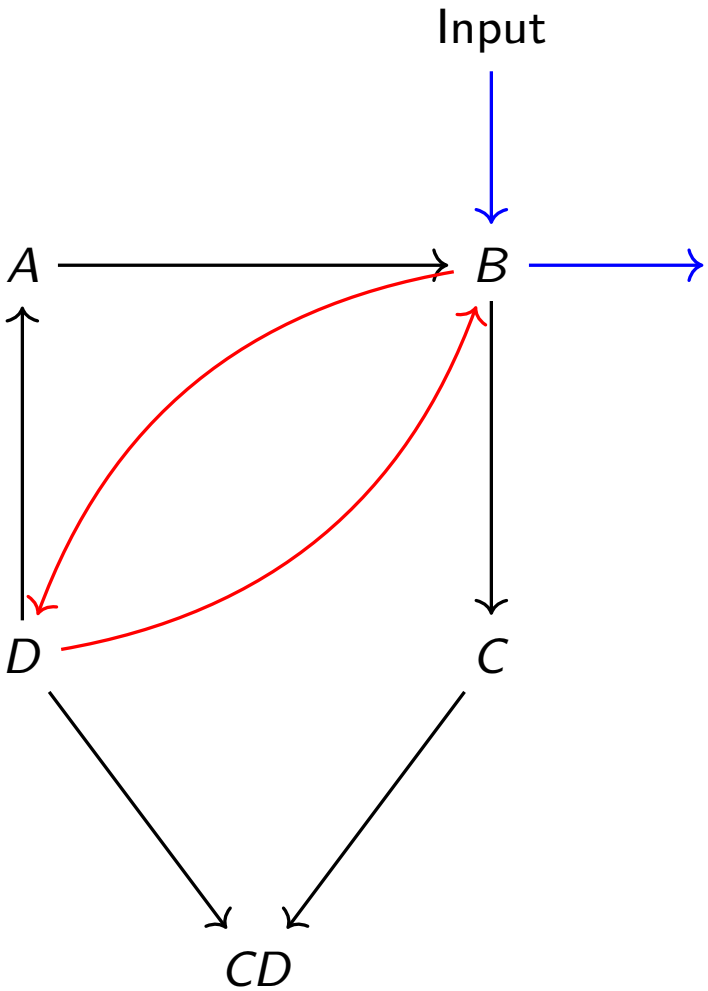
at the system's steady state. Here, V_1, Σ_1, and W_1 are derived from the full singular value decomposition of $D_{\vec{P}}\vec{F}$.





Adaptation is fragile

Hypothesis: Adding certain connections to the network leads to the loss of adaptive behavior.



$$\begin{cases} f_A &= \frac{dA}{dt} = k_1 D - k_2 A \\ f_B &= \frac{dB}{dt} = k_3 IA - k_4 B + a_1 D \\ f_C &= \frac{dC}{dt} = k_4 B - k_5 C \cdot D \\ f_D &= \frac{dD}{dt} = k_6 - k_5 C \cdot D + a_2 B \end{cases} \tag{2}$$

This system fails to adapt.

Network entropy in biology

Entropy plays a big role in the study of biological networks:

1. Cancer mutations alter network connectivity
2. Entropy is one of the hallmark features of the biological network that may be used for predicting diseases and detecting anomalies in the data
3. Network entropy and thermodynamic entropy descriptions need to be explored
4. Connection with thermodynamics and gradient flow theory is the key to uncovering adaptation mechanisms

Discovering Dynamics for Biological Systems

We want to:

- Recover the dynamics from biological data
- Recreate network structure from dynamics

Known problems:

- Most ML methods do not work well with small noisy datasets
- Hyperparameters and regularization parameters are hard to work with
- Rational functions are typically present in biological models (Michaelis–Menten kinetics, etc.)
- May not have access to derivative information or it may not be accurate enough

Discovering Dynamics for Biological Systems

We want to:

- Recover the dynamics from biological data
- Recreate network structure from dynamics

Known problems:

- Most ML methods do not work well with small noisy datasets
- Hyperparameters and regularization parameters are hard to work with
- Rational functions are typically present in biological models (Michaelis–Menten kinetics, etc.)
- May not have access to derivative information or it may not be accurate enough

Interested in certain properties of dynamics:

- Conservation Laws ([Oellerich/E.., 2023])
- Special network structures, such as compounds in the network [Oellerich/E.., 2021]
- Adaptation properties ([Tang et. al., 2016], [Araujo et. al., 2018], [Oellerich et. al., 2021])

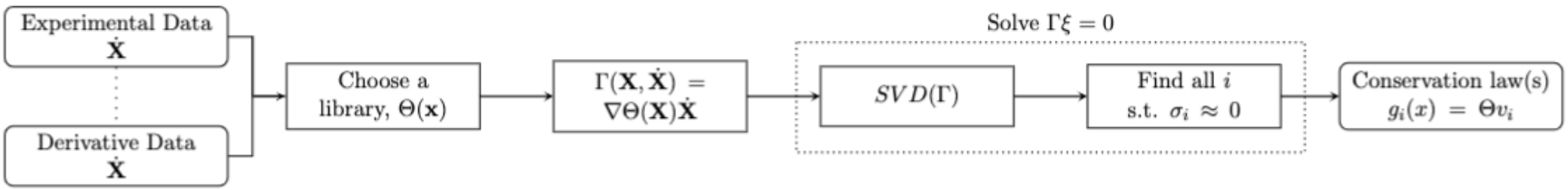
Workflow for identifying conservation laws from data

$$\dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}), \qquad g(\mathbf{x}) = C \approx \Theta(\mathbf{x})^T \boldsymbol{\xi} \tag{3}$$

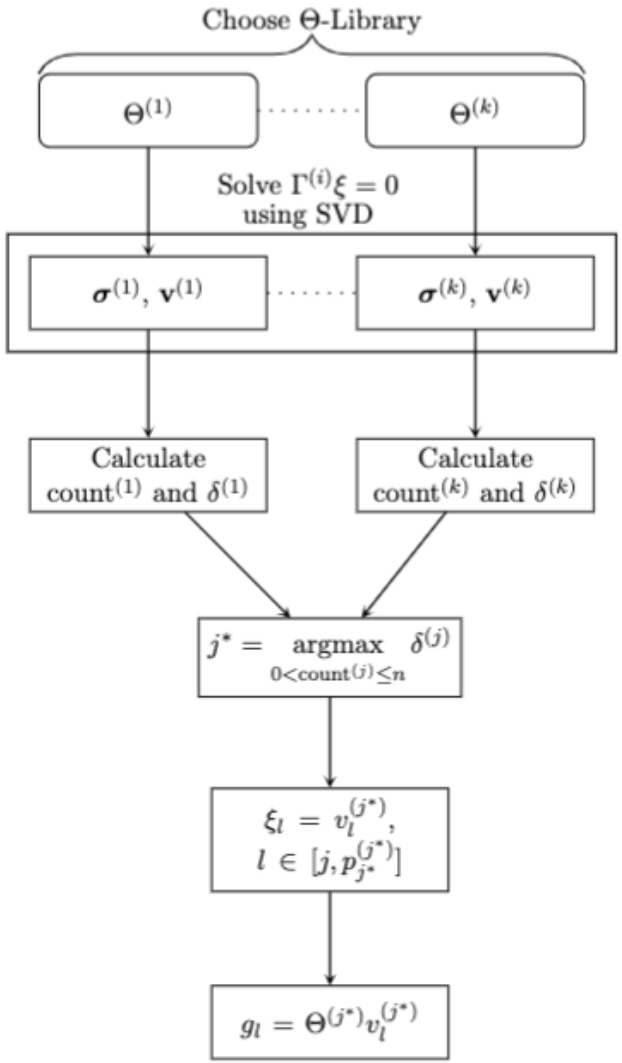
where

$$\Theta(\mathbf{x}) = [\boldsymbol{\theta}_1(\mathbf{x}), \boldsymbol{\theta}_2(\mathbf{x}) \dots \boldsymbol{\theta}_p(\mathbf{x})]^T \in \mathbb{R}^{p \times 1}, \qquad \boldsymbol{\xi} = [\xi_1, \dots \xi_p]^T \in \mathbb{R}^{p \times 1} \tag{4}$$

$$0 = \frac{d}{dt} C = \frac{d}{dt} g(\mathbf{x}) = \left(\frac{d\Theta}{dt} \right)^T \boldsymbol{\xi} = \Gamma(\mathbf{x}, \dot{\mathbf{x}}) \boldsymbol{\xi} \Rightarrow \Gamma(\mathbf{X}, \dot{\mathbf{X}})) \boldsymbol{\xi} = 0 \tag{5}$$



Workflow for optimizing the choice of the library using data

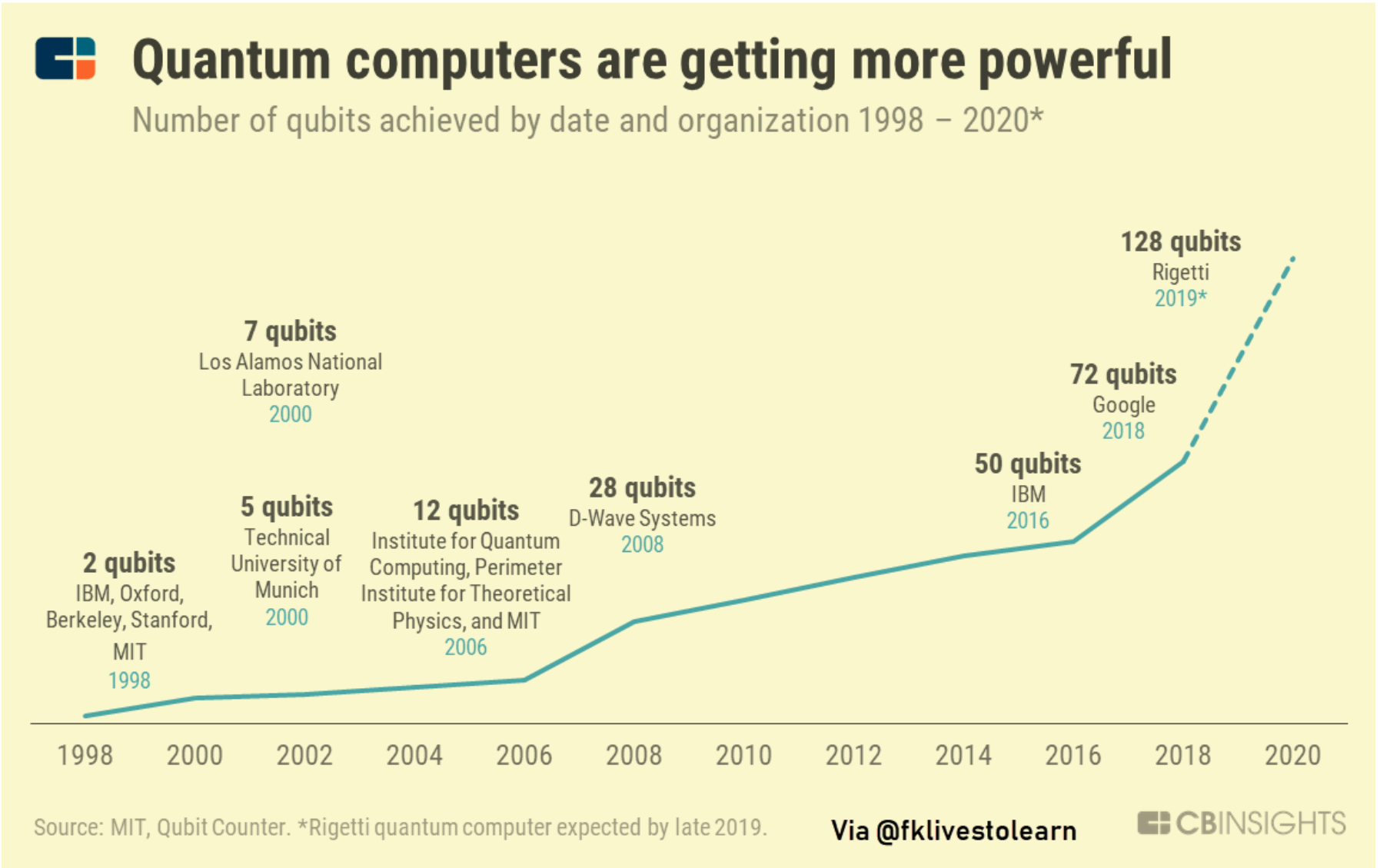


Identifying the right library

Poly Order	Sine	Ln	len(Θ)	δ	count	$\|\Gamma\xi_{cons}\|_2$
1.0	0	0	3	NaN	0	NaN
2.0	0	0	9	NaN	0	NaN
3.0	0	0	19.0	1.8648e-12	5.0	4.3008e-13
1.0	0	1.0	6.0	5.4734e-3	1.0	7.786e-14
2.0	0	1.0	12.0	1.6257e-11	1.0	2.0163e-15
3.0	0	1.0	22.0	1.0229e-12	6.0	3.4132e-13
1.0	1.0	0	15.0	4.257e-12	1.0	3.7476e-13
2.0	1.0	0	21.0	9.6393e-13	6.0	7.0661e-13
3.0	1.0	0	31.0	6.9411e-13	13.0	3.1455e-13
1.0	1.0	1.0	18.0	1.1416e-12	2.0	3.2407e-13
2.0	1.0	1.0	24.0	2.32e-12	8.0	5.2193e-13
3.0	1.0	1.0	34.0	5.8818e-13	16.0	7.3631e-13

$$\begin{aligned}
\dot{x} &= -k_1 x + k_2 y S \\
\dot{y} &= k_1 x - k_2 y S - k_3 y \\
\dot{S} &= -k_2 y S \\
\frac{d}{dt}(x + y - \frac{k_3}{k_2}\ln(S)) &= 0
\end{aligned}$$

Crossing into quantum computing



Quantum computing concepts

Computation with Qubits

How does the use of qubits affect computation?

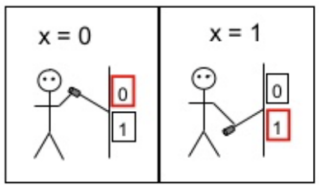
Classical Computation

Data unit: bit

● = '1' ○ = '0'

Valid states:

x = '0' or '1'

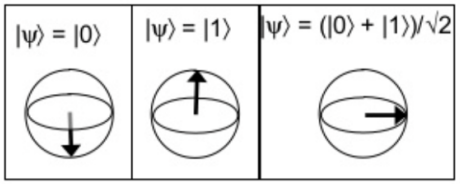


Quantum Computation

Data unit: qubit

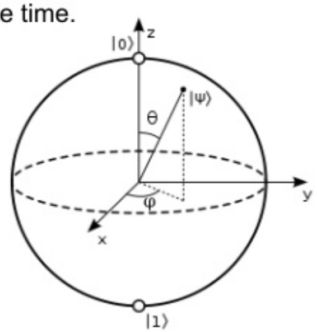
Valid states:

$|\psi\rangle = c_1|0\rangle + c_2|1\rangle$



Quantum Phenomenon: Superposition and Entanglement

- Superposition
  - Superposition is the ability of a quantum system to be in multiple states at the same time.



- Entanglement
  - Multiple particles are associated in such a way that measurement of one quantum state of one particle is determined by the measurement of the state of another particle.

Quantum information science notions

In classical world (configurational entropy):

$$S(p) = \sum_{k=1\ldots r} p_k \ln p_k, \text{ where } \sum_{k=1\ldots r} p_k = 1, \; p_k \geqq 0. \tag{6}$$

In quantum statistical mechanics (von Neumann entanglement entropy):

$$S = -\mathbf{tr}[\rho \ln \rho],$$

where ρ is the density matrix. Using eigendecomposition of the form $\rho = \sum_j \eta_j |j\rangle\langle j|$,

$$S = -\sum \eta_j \ln \eta_j$$

In this context, entropy measures departure of the system from pure state (amount of mixing of the quantum state).

Quantum vs. classical information science

In classical world (Renyi entropy):

$$\eta_\beta(X) = \frac{1}{1-\beta} \ln \sum_{k=1}^{K} (p^k)^\beta$$

Shannon entropy is the limit of Renyi entropy:

$$\eta_S(X) = -\sum_{k=1}^{K} p^k \ln(p^k) = \lim_{\beta \to 1} \eta_\beta(X)$$

In quantum world (von Neumann entanglement entropy):

$$S_\beta(\rho_A) = \frac{1}{1-\alpha} \ln \mathbf{tr}(\rho_A^\beta)$$

von Neumann entropy is the limit of Renyi entropy:

$$S = -\sum \eta_j \ln \eta_j = \lim_{\beta \to 1} S_\beta$$

Quantum circuits for the multi-state SWAP test

How does one implement entropy/overlap calculation on an actual quantum computer?

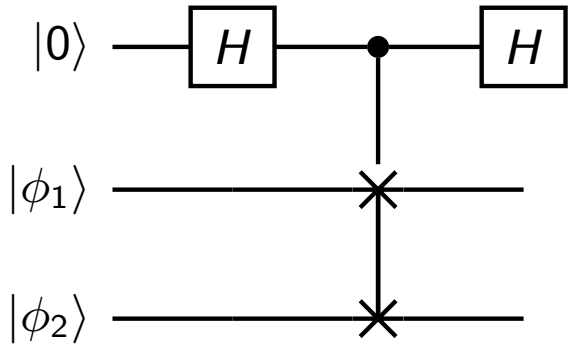


Figure: SWAP test for 2 states.

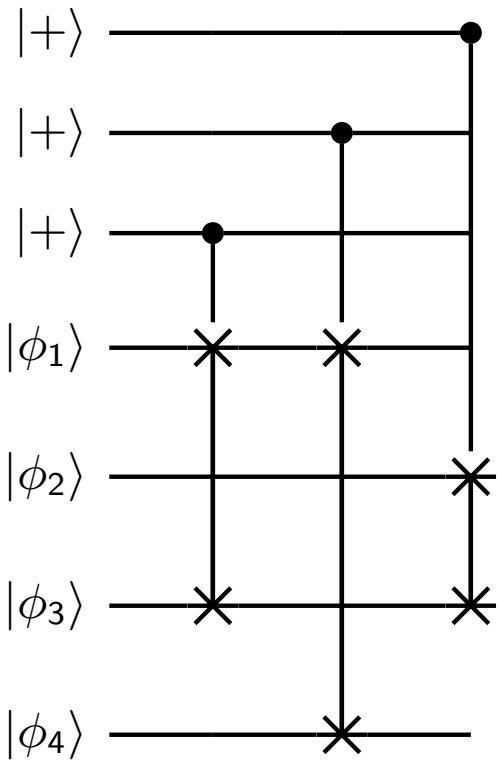
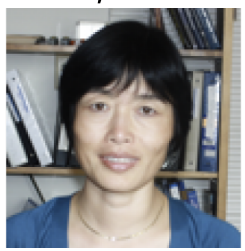
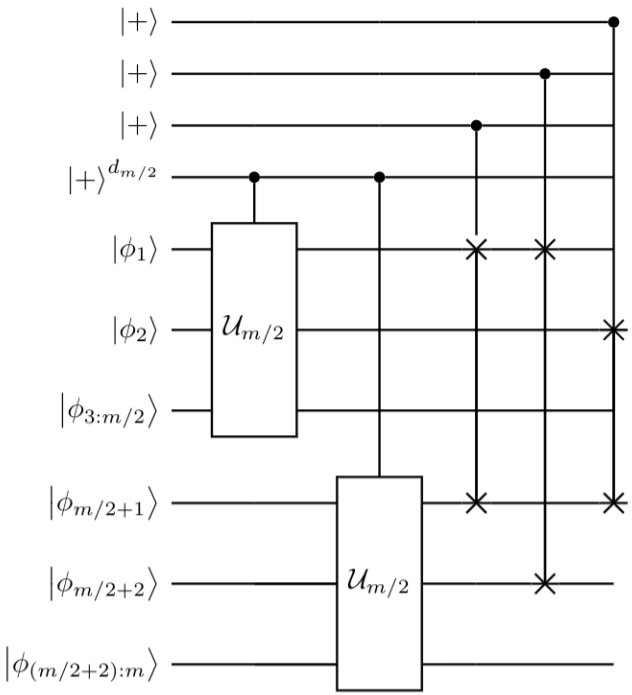


Figure: SWAP test for 4 states.

Recursive construction for n states

Joint work with M. Tian, X. Gitiaux, I. Morris





Theorem (Gitiaux, Morris, Emelianenko, Tian 2021)

Let $n = m(m-1)$. Then, there exists a unitary $\mathcal{U}_m$ that maps

$$\mathcal{U}_m : \otimes |0\rangle^{d_m} \otimes_{i=1}^{m} |\phi_i\rangle \to \sum_{ij=1}^{n} |ij\rangle \, |\phi_i \phi_j G_{ij}\rangle \, ,$$

where G_{ij} is a garbage state and $d_m = O(\ln m)$ is the number of ancillaries. $\mathcal{U}$ can be computed by a quantum circuit with $O(m)$ controlled swap gates. In comparison, pairwise implementation needs $O(m)$ ancillaries, $O(n)$ controlled SWAPs and is destructive.

Random walks on graphs

Classical random walk on a graph: a Markov process with infinitesimal generator matrix M with jumping rate γ, defined as

$$M_{ab} = \begin{cases} -\gamma, & a \neq b, a \text{ and } b \text{ connected by an edge} \\ 0, & a \neq b, a \text{ and } b \text{ not connected by an edge} \\ -k\gamma, & a = b, \text{where } k \text{ is the degree of } a \end{cases}$$

Evolution is given by

$$\frac{dp_a(t)}{dt} = -\sum_b M_{ab} p_b(t),$$

where $p_a(t)$ is the probability of being at vertex a at time t.

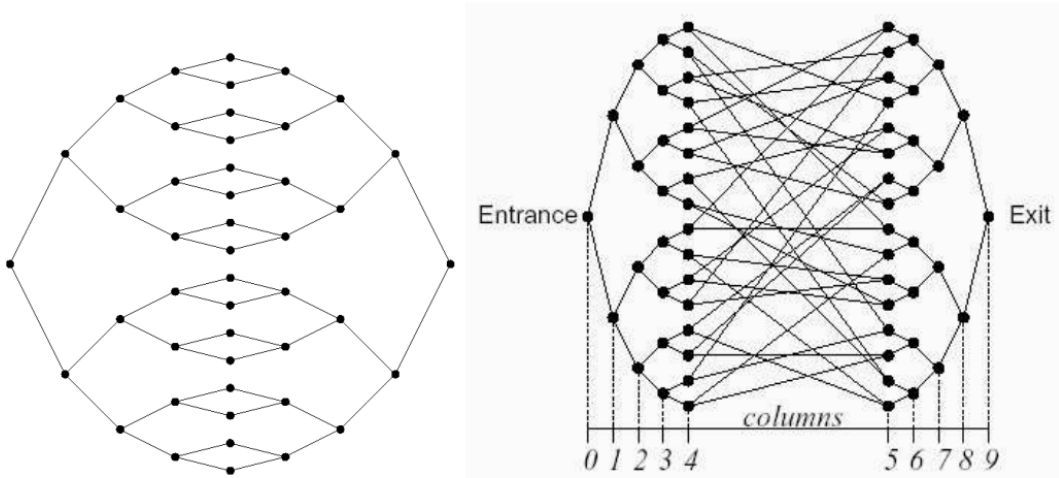
Quantum walk satisfies Schrödinger equation

$$i\frac{d}{dt}\langle a|\psi(t)\rangle = -\sum_b \langle a|b\rangle \langle b|\psi(t)\rangle$$

with a Hamiltonian H, where $\langle a|\psi(t)\rangle$ is the amplitude of being at vertex a at time t. An analogue of the above classical walk: quantum Hamiltonian with matrix elements

Quantum advantage of quantum walk

Example: glued tree graph with $2^{n+1} + 2^n - 2$ vertices [Childs et al, 2002]:



Classical walker needs exponential time to get from 'entrance' to 'exit': number of vertices in column n contains 2^n vertices
Quantum walker is able to choose all directions simultaneously: reduces to $2n + 1$-dimensional subspace of the original space (linear time).

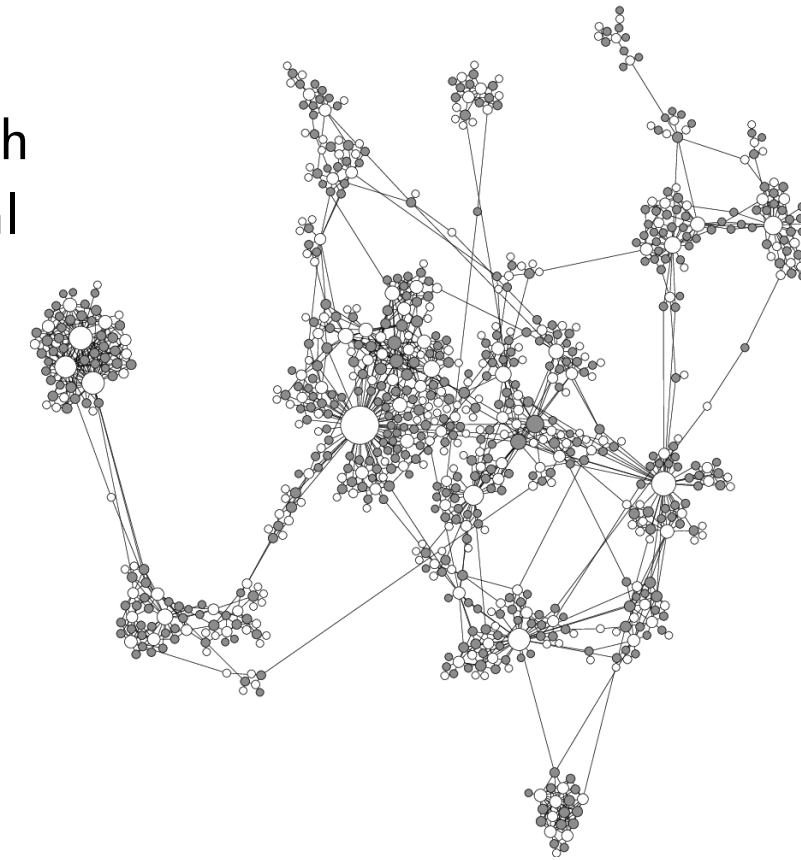
Thinking conceptually helps transcend boundaries

Concept of entropy

1. Entropy is a measure of disorder and is used in many contexts
2. Different entropy forms may correlate with different aspects of the physical/biological processes
3. Quantum entanglement can be measured using entropy, but circuit optimization may be nontrivial

Concept of random walks

1. Can help characterize dynamics
2. Can be quantum and fractional
3. Quantum walk may produce exponential speedup comparing to classical case



The search for hidden mechanisms and connections is far from complete.