Parameter-free statistical model invalidation for biochemical reaction networks

Kenneth L. Ho (Stanford)

Joint work with Heather Harrington (Oxford)

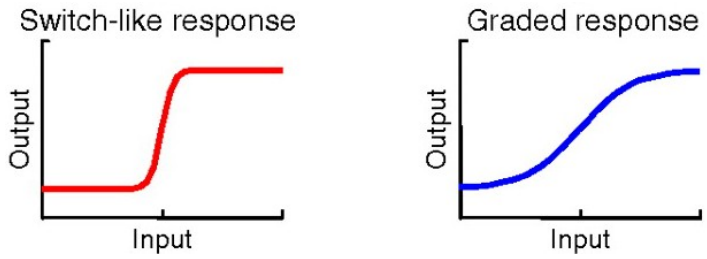
Theranos, Apr. 2015

Motivation

- **Model selection**: observed data, multiple models; which model is 'best'?
- Example (sequential phosphorylation):

$$\begin{tikzcd} S_0 \arrow[r, "\kappa_{01}"] \arrow[rr, bend right, "\kappa_{02}"'] & S_1 \arrow[r, "\kappa_{12}"] & S_2 \end{tikzcd}$$

- Two models: distributive ($\kappa_{02} = 0$), processive ($\kappa_{02} > 0$)



- Closely related to **model invalidation**

[Aoki/Yamada/Kunida/Yasuda/Matsuda]

Problem setting

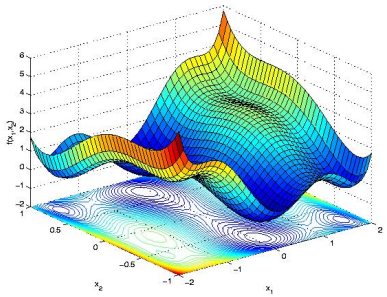
- Data x, model $f = f(\kappa)$ with parameters κ
- How to tell if model is incompatible with data?
- **Known parameters**: compute $\hat{x} = f(\kappa)$ and check $\|x - \hat{x}\|$
- **Unknown parameters**: fit parameters and check best-case error

Problem setting

- Data x, model $f = f(\kappa)$ with parameters κ
- How to tell if model is incompatible with data?
- **Known parameters**: compute $\hat{x} = f(\kappa)$ and check $\|x - \hat{x}\|$
- **Unknown parameters**: fit parameters and check best-case error

Parameter problem: in biology, parameters are hardly ever known

- Technical limitations, uncertainties, etc.
- Partial data: experimentally inaccessible species
- Nonlinear, high-dimensional optimization often required

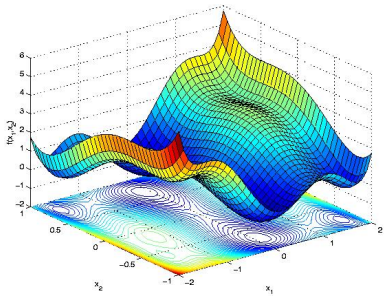


Problem setting

- Data x, model $f = f(\kappa)$ with parameters κ
- How to tell if model is incompatible with data?
- **Known parameters**: compute $\hat{x} = f(\kappa)$ and check $\|x - \hat{x}\|$
- **Unknown parameters**: fit parameters and check best-case error

Parameter problem: in biology, parameters are hardly ever known

- Technical limitations, uncertainties, etc.
- Partial data: experimentally inaccessible species
- Nonlinear, high-dimensional optimization often required



What can be done?

Previous work

- Optimization 'tricks': random seeding, simulated annealing, etc.

Previous work

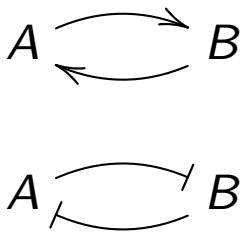
- Optimization 'tricks': random seeding, simulated annealing, etc.
- Convex relaxation + SDP: lower bound for best-case error
  - Polynomial-time search through parameter space

Previous work

- Optimization 'tricks': random seeding, simulated annealing, etc.
- Convex relaxation + SDP: lower bound for best-case error
  - Polynomial-time search through parameter space

This talk: (quantitative) parameter-free methods

- Like SDP but no dependence on parameters
- Based only on model structure/topology
- Not necessarily 'better' but new framework

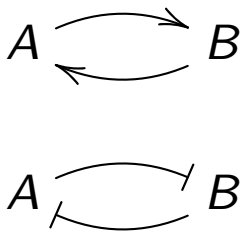


Previous work

- Optimization 'tricks': random seeding, simulated annealing, etc.
- Convex relaxation + SDP: lower bound for best-case error
  - Polynomial-time search through parameter space

This talk: (quantitative) parameter-free methods

- Like SDP but no dependence on parameters
- Based only on model structure/topology
- Not necessarily 'better' but new framework



Philosophically related (qualitative):

- Chemical reaction network theory
- Stoichiometric network analysis

[Clarke, Feinberg, Horn, Jackson, ...]

Chemical reaction networks

- Reactions: $$\sum_{j=1}^{N} r_{ij} X_j \xrightarrow{\kappa_i} \sum_{j=1}^{N} p_{ij} X_j, \qquad i = 1, \ldots, R$$
- Mass-action dynamics: $$\dot{x}_j = \sum_{i=1}^{R} \kappa_i \left(p_{ij} - r_{ij}\right) \prod_{k=1}^{N} x_k^{r_{ik}}, \quad j = 1, \ldots, N$$

Chemical reaction networks

- Reactions: $$\sum_{j=1}^{N} r_{ij} X_j \xrightarrow{\kappa_i} \sum_{j=1}^{N} p_{ij} X_j, \qquad i = 1, \ldots, R$$
- Mass-action dynamics: $$\dot{x}_j = \sum_{i=1}^{R} \kappa_i \left(p_{ij} - r_{ij}\right) \prod_{k=1}^{N} x_k^{r_{ik}}, \quad j = 1, \ldots, N$$

Example:

$$X_1 + X_2 \xrightarrow{\kappa_1} 2X_1$$

$$3X_1 \xrightarrow{\kappa_2} X_2 + 2X_3$$

$$\begin{tikzcd} & X_1 \\ X_3 \arrow[ur, "\kappa_3"] \arrow[dr, "\kappa_4"'] & \\ & X_2 \end{tikzcd}$$

$$\dot{x}_1 = \kappa_1 x_1 x_2 - 3\kappa_2 x_1^3 + \kappa_3 x_3$$

$$\dot{x}_2 = -\kappa_1 x_1 x_2 + \kappa_2 x_1^3 + \kappa_4 x_3$$

$$\dot{x}_3 = 2\kappa_2 x_1^3 - \left(\kappa_3 + \kappa_4\right) x_3$$

Chemical reaction networks

- Reactions: $$\sum_{j=1}^{N} r_{ij} X_j \xrightarrow{\kappa_i} \sum_{j=1}^{N} p_{ij} X_j, \qquad i = 1, \ldots, R$$
- Mass-action dynamics: $$\dot{x}_j = \sum_{i=1}^{R} \kappa_i \left(p_{ij} - r_{ij}\right) \prod_{k=1}^{N} x_k^{r_{ik}}, \quad j = 1, \ldots, N$$

Example: $$X_1 + X_2 \xrightarrow{\kappa_1} 2X_1$$

$$3X_1 \xrightarrow{\kappa_2} X_2 + 2X_3$$

$$\begin{tikzcd} & X_1 \\ X_3 \arrow[ur, "\kappa_3"] \arrow[dr, "\kappa_4"'] & \\ & X_2 \end{tikzcd}$$

$$\dot{x}_1 = \kappa_1 x_1 x_2 - 3\kappa_2 x_1^3 + \kappa_3 x_3$$

$$\dot{x}_2 = -\kappa_1 x_1 x_2 + \kappa_2 x_1^3 + \kappa_4 x_3$$

$$\dot{x}_3 = 2\kappa_2 x_1^3 - (\kappa_3 + \kappa_4) x_3$$

Chemical reaction networks

- Reactions: $$\sum_{j=1}^{N} r_{ij} X_j \xrightarrow{\kappa_i} \sum_{j=1}^{N} p_{ij} X_j, \qquad i = 1, \ldots, R$$
- Mass-action dynamics: $$\dot{x}_j = \sum_{i=1}^{R} \kappa_i \left(p_{ij} - r_{ij}\right) \prod_{k=1}^{N} x_k^{r_{ik}}, \quad j = 1, \ldots, N$$

Example: $$X_1 + X_2 \xrightarrow{\kappa_1} 2X_1$$

$$3X_1 \xrightarrow{\kappa_2} X_2 + 2X_3$$

$$\begin{tikzcd} & X_1 \\ X_3 \arrow[ur, "\kappa_3"] \arrow[dr, "\kappa_4"'] & \\ & X_2 \end{tikzcd}$$

$$\dot{x}_1 = \kappa_1 x_1 x_2 - 3\kappa_2 x_1^3 + \kappa_3 x_3$$

$$\dot{x}_2 = -\kappa_1 x_1 x_2 + \kappa_2 x_1^3 + \kappa_4 x_3$$

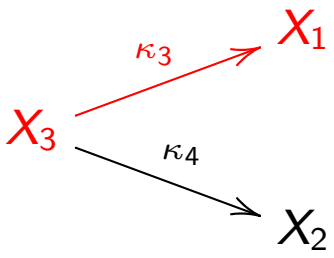
$$\dot{x}_3 = 2\kappa_2 x_1^3 - (\kappa_3 + \kappa_4) x_3$$

Chemical reaction networks

- Reactions: $$\sum_{j=1}^{N} r_{ij} X_j \xrightarrow{\kappa_i} \sum_{j=1}^{N} p_{ij} X_j, \qquad i = 1, \ldots, R$$
- Mass-action dynamics: $$\dot{x}_j = \sum_{i=1}^{R} \kappa_i \left(p_{ij} - r_{ij}\right) \prod_{k=1}^{N} x_k^{r_{ik}}, \quad j = 1, \ldots, N$$

Example: $$X_1 + X_2 \xrightarrow{\kappa_1} 2X_1$$

$$3X_1 \xrightarrow{\kappa_2} X_2 + 2X_3$$



$$\dot{x}_1 = \kappa_1 x_1 x_2 - 3\kappa_2 x_1^3 + \kappa_3 x_3$$

$$\dot{x}_2 = -\kappa_1 x_1 x_2 + \kappa_2 x_1^3 + \kappa_4 x_3$$

$$\dot{x}_3 = 2\kappa_2 x_1^3 - (\kappa_3 + \kappa_4) x_3$$

Chemical reaction networks

- Reactions: $$\sum_{j=1}^{N} r_{ij} X_j \xrightarrow{\kappa_i} \sum_{j=1}^{N} p_{ij} X_j, \qquad i = 1, \ldots, R$$
- Mass-action dynamics: $$\dot{x}_j = \sum_{i=1}^{R} \kappa_i \left(p_{ij} - r_{ij}\right) \prod_{k=1}^{N} x_k^{r_{ik}}, \quad j = 1, \ldots, N$$

Example:

$$X_1 + X_2 \xrightarrow{\kappa_1} 2X_1$$

$$3X_1 \xrightarrow{\kappa_2} X_2 + 2X_3$$

$$\begin{tikzcd} & X_1 \\ X_3 \arrow[ur, "\kappa_3"] \arrow[dr, "\kappa_4"'] & \\ & X_2 \end{tikzcd}$$

$$\dot{x}_1 = \kappa_1 x_1 x_2 - 3\kappa_2 x_1^3 + \kappa_3 x_3$$

$$\dot{x}_2 = -\kappa_1 x_1 x_2 + \kappa_2 x_1^3 + \kappa_4 x_3$$

$$\dot{x}_3 = 2\kappa_2 x_1^3 - (\kappa_3 + \kappa_4) x_3$$

Model compatibility

- ODE: quantitative test for model compatibility

Example:

$$\dot{x}_1 = \kappa_1 x_1 x_2 - 3\kappa_2 x_1^3 + \kappa_3 x_3$$
$$\dot{x}_2 = -\kappa_1 x_1 x_2 + \kappa_2 x_1^3 + \kappa_4 x_3$$
$$\dot{x}_3 = 2\kappa_2 x_1^3 - (\kappa_3 + \kappa_4) x_3$$

Model compatibility

- ODE: quantitative test for model compatibility
- **Problem**: partial data

Example:

$$\dot{x}_1 = \kappa_1 x_1 x_2 - 3\kappa_2 x_1^3 + \kappa_3 x_3$$
$$\dot{x}_2 = -\kappa_1 x_1 x_2 + \kappa_2 x_1^3 + \kappa_4 x_3$$
$$\dot{x}_3 = 2\kappa_2 x_1^3 - (\kappa_3 + \kappa_4) x_3$$

Model compatibility

- ODE: quantitative test for model compatibility
- **Problem**: partial data
  - Derivatives unreliable; assume steady state

Example:

$$0 = \dot{x}_1 = \kappa_1 x_1 x_2 - 3\kappa_2 x_1^3 + \kappa_3 x_3$$
$$0 = \dot{x}_2 = -\kappa_1 x_1 x_2 + \kappa_2 x_1^3 + \kappa_4 x_3$$
$$0 = \dot{x}_3 = 2\kappa_2 x_1^3 - (\kappa_3 + \kappa_4) x_3$$

Model compatibility

- ODE: quantitative test for model compatibility
- **Problem**: partial data
  - Derivatives unreliable; assume steady state
  - Eliminate experimentally inaccessible species

Example:

$$0 = \dot{x}_1 = \kappa_1 x_1 x_2 - 3\kappa_2 x_1^3 + \kappa_3 x_3$$

$$0 = \dot{x}_2 = -\kappa_1 x_1 x_2 + \kappa_2 x_1^3 + \kappa_4 x_3$$

$$0 = \dot{x}_3 = 2\kappa_2 x_1^3 - (\kappa_3 + \kappa_4) x_3$$

Model compatibility

- ODE: quantitative test for model compatibility
- **Problem**: partial data
  - Derivatives unreliable; assume steady state
  - Eliminate experimentally inaccessible species

Example:

$$0 = \dot{x}_1 = \kappa_1 x_1 x_2 - 3\kappa_2 x_1^3 + \kappa_3 x_3$$
$$0 = \dot{x}_2 = -\kappa_1 x_1 x_2 + \kappa_2 x_1^3 + \kappa_4 x_3$$
$$0 = \dot{x}_3 = 2\kappa_2 x_1^3 - (\kappa_3 + \kappa_4) x_3$$

- Observe x_1, x_2; eliminate $x_3 = \frac{2\kappa_2 x_1^3}{\kappa_3 + \kappa_4}$

Model compatibility

- ODE: quantitative test for model compatibility
- **Problem**: partial data
  - Derivatives unreliable; assume steady state
  - Eliminate experimentally inaccessible species

Example:

$$0 = \kappa_1 x_1 x_2 + \left(\frac{2\kappa_3}{\kappa_3 + \kappa_4} - 3 \right) \kappa_2 x_1^3$$

$$0 = -\kappa_1 x_1 x_2 + \left(\frac{2\kappa_4}{\kappa_3 + \kappa_4} + 1 \right) \kappa_2 x_1^3$$

- Observe x_1, x_2; eliminate $x_3 = \frac{2\kappa_2 x_1^3}{\kappa_3 + \kappa_4} \implies$ steady-state invariants

Model compatibility

- ODE: quantitative test for model compatibility
- **Problem**: partial data
  - Derivatives unreliable; assume steady state
  - Eliminate experimentally inaccessible species

Example:

$$0 = \kappa_1 x_1 x_2 + \left(\frac{2\kappa_3}{\kappa_3 + \kappa_4} - 3 \right) \kappa_2 x_1^3$$

$$0 = -\kappa_1 x_1 x_2 + \left(\frac{2\kappa_4}{\kappa_3 + \kappa_4} + 1 \right) \kappa_2 x_1^3$$

- Observe x_1, x_2; eliminate $x_3 = \frac{2\kappa_2 x_1^3}{\kappa_3 + \kappa_4} \implies$ steady-state invariants
- In general, use computational algebraic geometry (Gröbner bases):

$$0 = \sum_{i=1}^{n} \alpha_i(\kappa) \varphi_i(x)$$

[Manrai/Gunawardena]

Parameter-free invalidation

$$\sum_{i=1}^{n} \alpha_i(\kappa)\varphi_i(x^{(1)}) = 0$$

$$\vdots$$

$$\sum_{i=1}^{n} \alpha_i(\kappa)\varphi_i(x^{(m)}) = 0$$

[Manrai/Gunawardena]

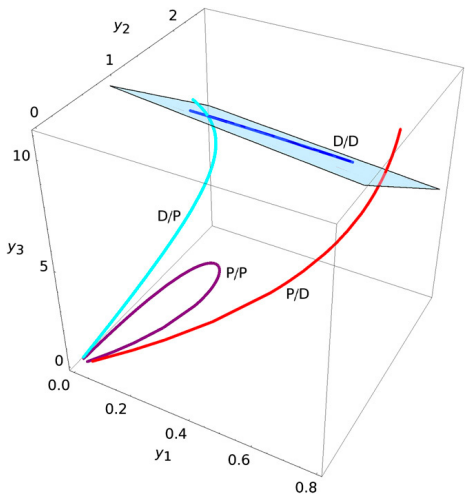
Parameter-free invalidation

$$\sum_{i=1}^{n} \alpha_i(\kappa)\varphi_i(x^{(1)}) = 0$$

$$\vdots$$

$$\sum_{i=1}^{n} \alpha_i(\kappa)\varphi_i(x^{(m)}) = 0$$

- Let $y^{(k)} = (\varphi_1(x^{(k)}), \ldots, \varphi_n(x^{(k)})) \in \mathbb{R}^n$
- **Geometry**: $y^{(1)}, \ldots, y^{(m)}$ are coplanar
- Depends only on data, hence parameter-free

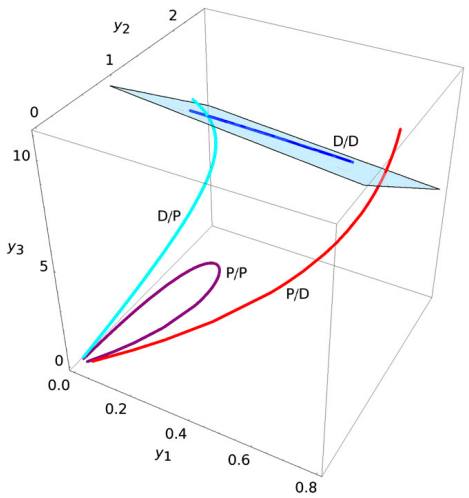


[Manrai/Gunawardena]

Parameter-free invalidation

$$\sum_{i=1}^{n} \alpha_i(\kappa)\varphi_i(x^{(1)}) = 0$$
$$\vdots$$
$$\sum_{i=1}^{n} \alpha_i(\kappa)\varphi_i(x^{(m)}) = 0$$

- Let $y^{(k)} = (\varphi_1(x^{(k)}), \ldots, \varphi_n(x^{(k)})) \in \mathbb{R}^n$
- **Geometry**: $y^{(1)}, \ldots, y^{(m)}$ are coplanar
- Depends only on data, hence parameter-free



Linear algebra: $Y\alpha = 0$

- Compatible $\implies \exists \alpha \in \text{null}(Y) \implies \dim(\text{null}(Y)) > 0$
- Compute SVD, reject if $\sigma_{\min}(Y) > 0$

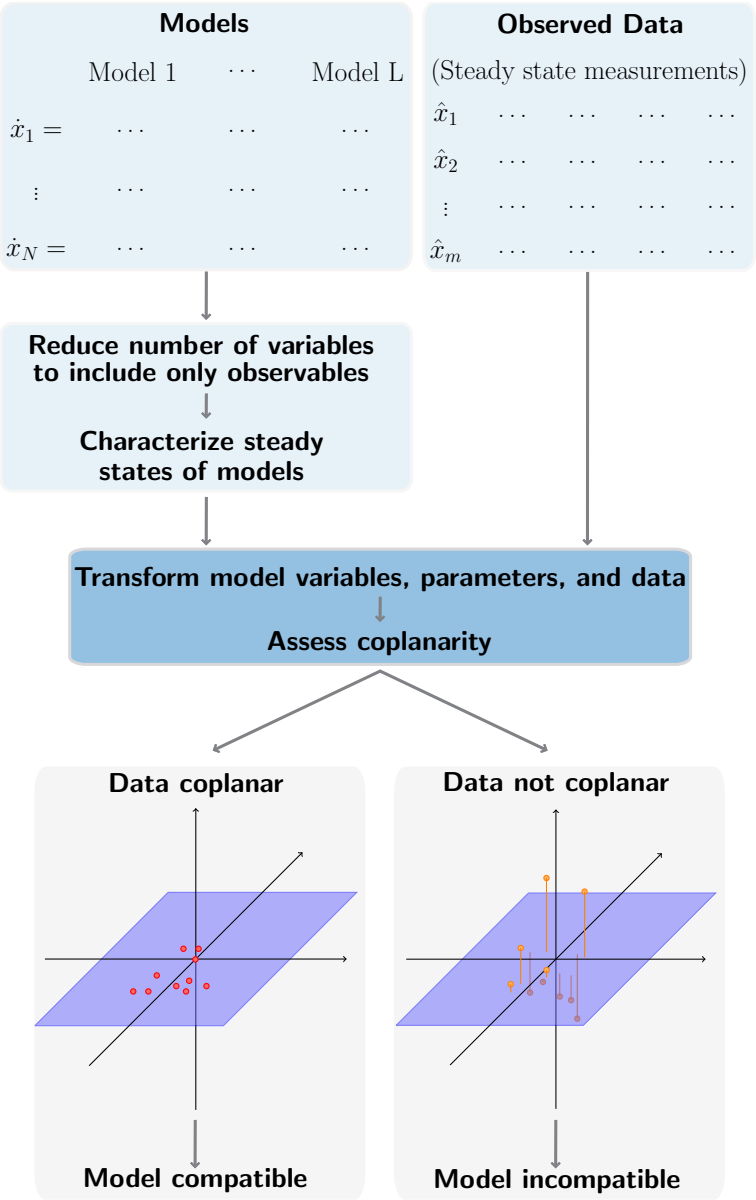
[Harrington/Ho/Thorne/Stumpf, Manrai/Gunawardena]

Statistical rejection

- **Null hypothesis**: $\sigma_{\min}(Y) = \min_{\|\alpha\|=1} \|Y\alpha\| = 0 \implies \exists \alpha$ such that $Y\alpha = 0$
- Assume Gaussian noise in $x^{(k)}$, estimate noise in $y^{(k)} = \varphi(x^{(k)})$
- To first-order, Gaussian noise in $Y \implies z = Y\alpha$ Gaussian
- Rescale rows: $DY \implies (Dz)_i \sim \mathcal{N}(0, \sigma_i^2)$, $\sigma_i^2 \leq 1$
- Tail bound: $\Pr(\sigma_{\min}(DY) > t) \leq \Pr(\|Dz\| > t) \leq \Pr(\chi_m > t)$
- Other bounds possible: Weyl, Wielandt-Hoffman, concentration of measure, etc.

[Harrington/Ho/Thorne/Stumpf]

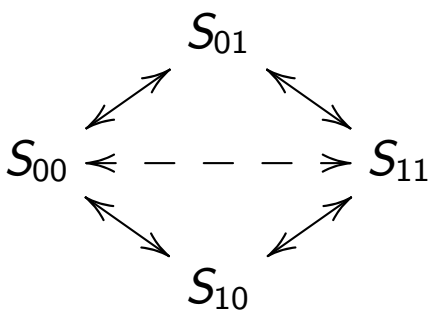
Algorithm



- Test coplanarity for each invariant
- Reject model if any invariant fails
- **Main costs**: elimination, $\sigma_{\min}(\mathbb{R}^{m\times n})$

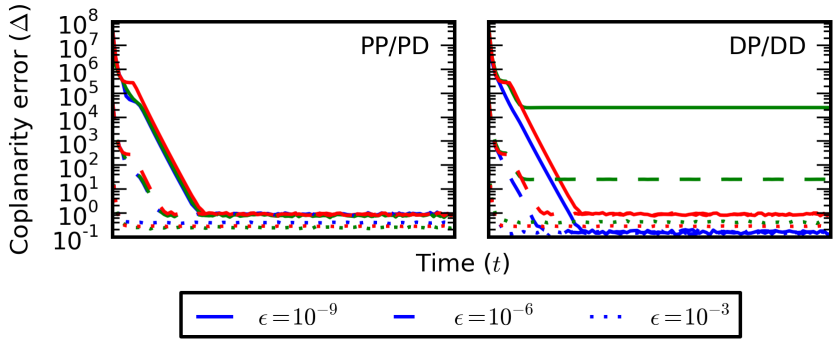
[Harrington/Ho/Thorne/Stumpf]

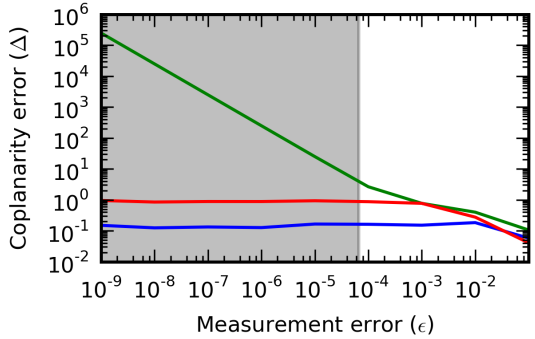
Example: two-site phosphorylation



- Kinase/phosphatase: distributive/processive
- Four models: PP, PD, DP, DD
- 12 species, 22 parameters

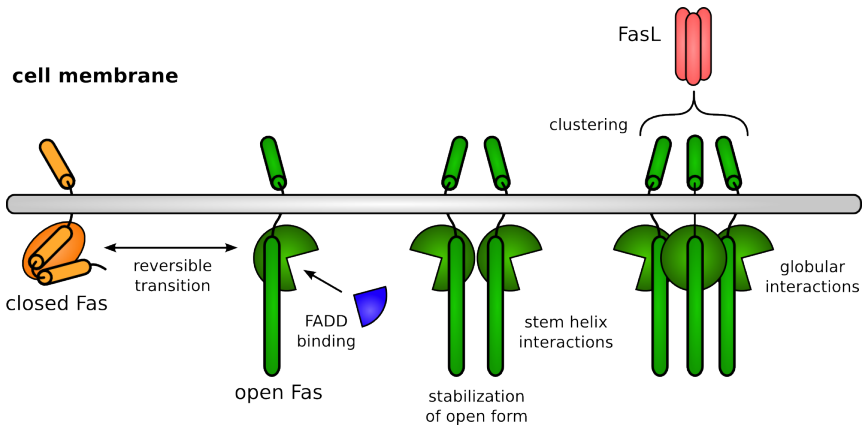
- Variable ordering: $(ks_{00}, ks_{01}, ks_{10}, fs_{01}, fs_{10}, fs_{11}, k, f, \overbrace{s_{00}, s_{01}, s_{10}, s_{11}}^{\text{obs}})$
- Kinase is discriminative, can reject DP/DD models on basis of PP data
- Can discriminate instead on phosphatase by reversing variable ordering





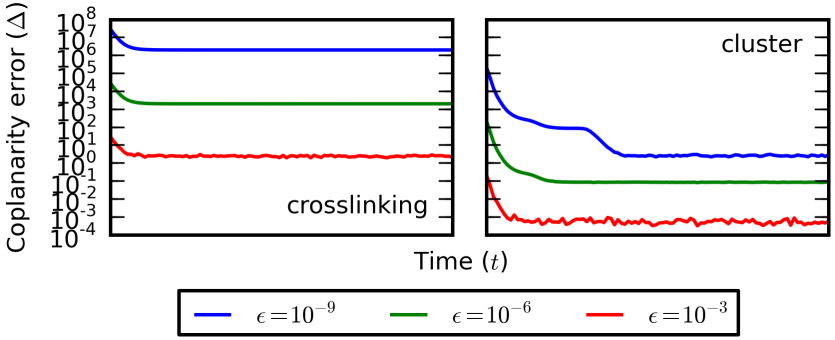
[Harrington/Ho/Thorne/Stumpf, Manrai/Gunawardena]

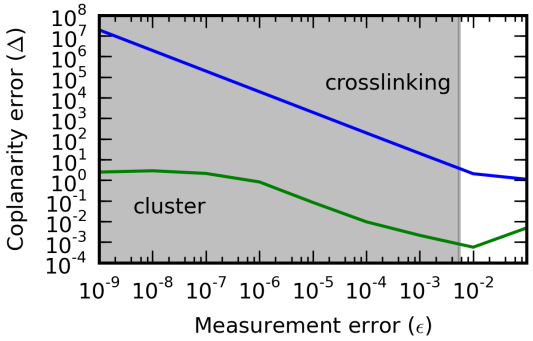
Example: cell death signaling



- Extrinsic pathway
- FasL/Fas interactions
- Measure activated Fas

- Crosslinking model: sequential Fas recruitment (8 species, 2 parameters)
- Cluster model: scaffold for Fas clustering, bistable (6 species, 9 parameters)
- Can reject crosslinking model from cluster data





[Harrington/Ho/Thorne/Stumpf, Ho/Harrington, Lai/Jackson]

Summary

- Parameter-free statistical model invalidation
- Detection of non-parametric linear structure
- Very efficient once invariants have been computed

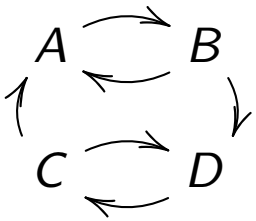
Broader perspective: parameter-independent model properties

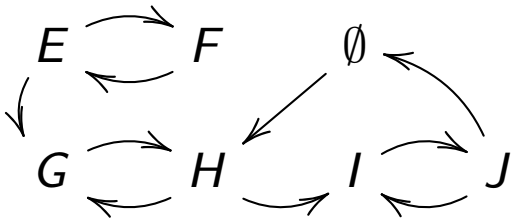
- Structure, topology, robustness, modularity
- Algebraic systems biology

Limitations: nonlinear elimination, steady-state data, necessary but not sufficient

Extension: complex-linear networks

- CRNs: nonlinear ODEs $\implies$ nonlinear elimination
- Fundamental insight of CRNT: hidden linearity
- **Complex-balanced networks**: underlying Laplacian dynamics
- Study properties of Laplacian matrices
- Steady state: kernel = zero + constant + rank-one
- No elimination: decomposition based on graph connectivity
- Test σ_1, σ_2, etc.

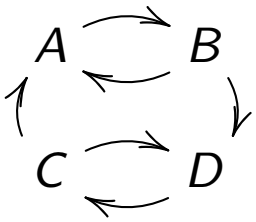


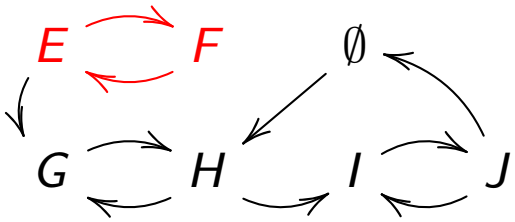


[Harrington/Ho]

Extension: complex-linear networks

- CRNs: nonlinear ODEs $\implies$ nonlinear elimination
- Fundamental insight of CRNT: hidden linearity
- **Complex-balanced networks**: underlying Laplacian dynamics
- Study properties of Laplacian matrices
- Steady state: kernel = zero + constant + rank-one
- No elimination: decomposition based on graph connectivity
- Test σ_1, σ_2, etc.

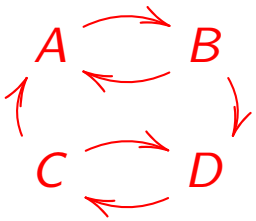


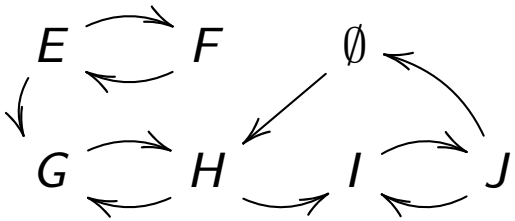


[Harrington/Ho]

Extension: complex-linear networks

- CRNs: nonlinear ODEs $\implies$ nonlinear elimination
- Fundamental insight of CRNT: hidden linearity
- **Complex-balanced networks**: underlying Laplacian dynamics
- Study properties of Laplacian matrices
- Steady state: kernel = zero + constant + rank-one
- No elimination: decomposition based on graph connectivity
- Test σ_1, σ_2, etc.

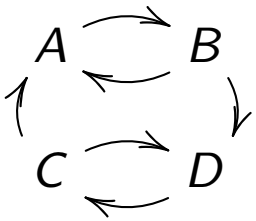


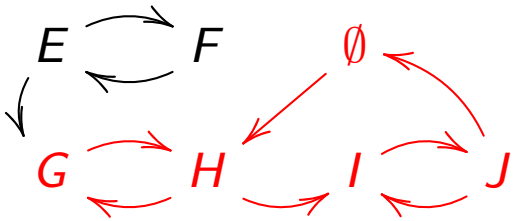


[Harrington/Ho]

Extension: complex-linear networks

- CRNs: nonlinear ODEs $\implies$ nonlinear elimination
- Fundamental insight of CRNT: hidden linearity
- **Complex-balanced networks**: underlying Laplacian dynamics
- Study properties of Laplacian matrices
- Steady state: kernel = zero + constant + rank-one
- No elimination: decomposition based on graph connectivity
- Test σ_1, σ_2, etc.





[Harrington/Ho]

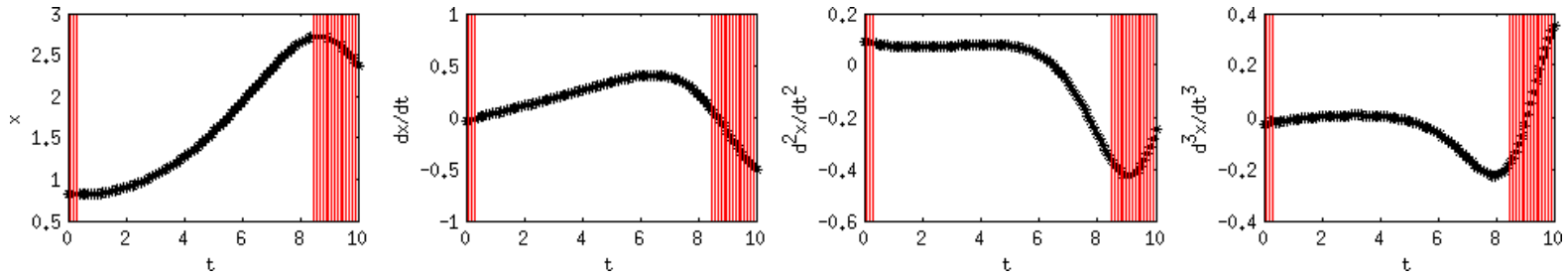
Extension: time course data

- Differential elimination: dynamical invariants involving derivatives
- **Example**: Lotka-Volterra

$$\dot{x} = ax - bxy$$
$$\dot{y} = -cy + dxy$$

$$\Longrightarrow \qquad acx^2 + ax\dot{x} - bcx^3 - bx^2\dot{x} = -\dot{x}^2 + x\ddot{x}$$

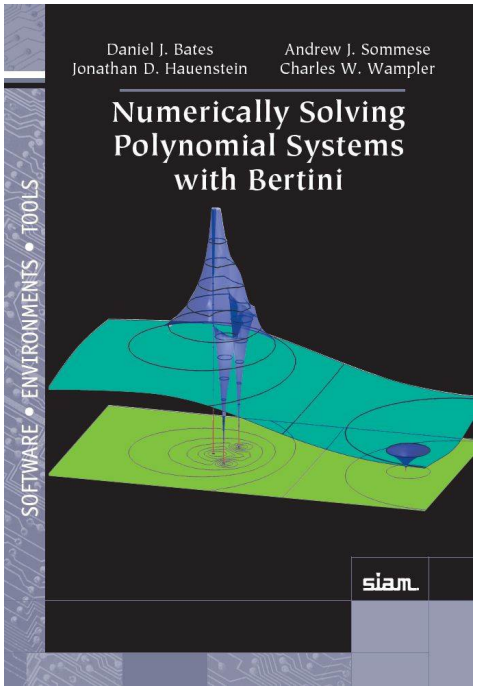
- Estimate derivatives using Gaussian processes



[Harrington/Ho/Meshkat]

Extension: numerical algebraic geometry

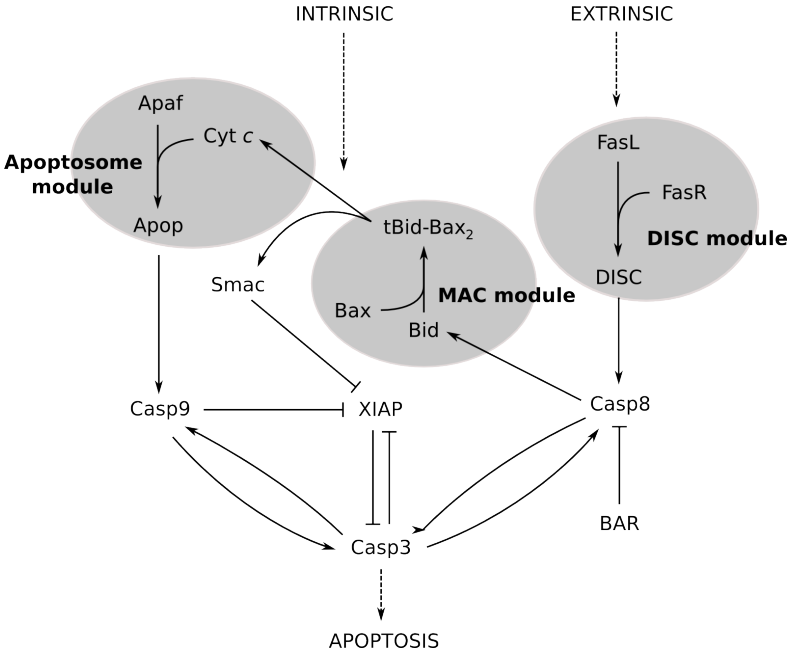
- **Goal**: solve global nonlinear optimization
- Exploit polynomial structure, numerical algebraic geometry
- Find closest intersection between model and data varieties
- Maximum-likelihood parameter estimation, model invalidation, model selection

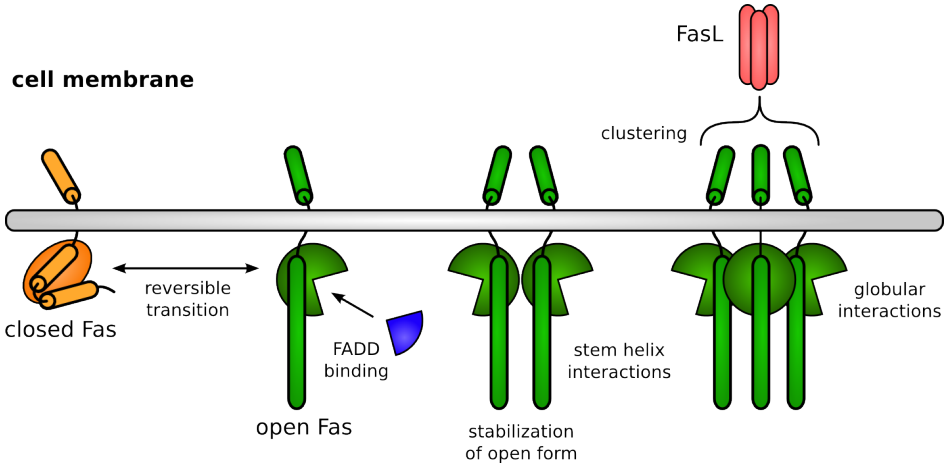


[Gross/Davis/Ho/Bates/Harrington]

Other work

- Mathematical modeling of cell signaling networks
- Automated all-atom protein crystal structure refinement
- Fast multipole methods, direct solvers, matrix factorizations

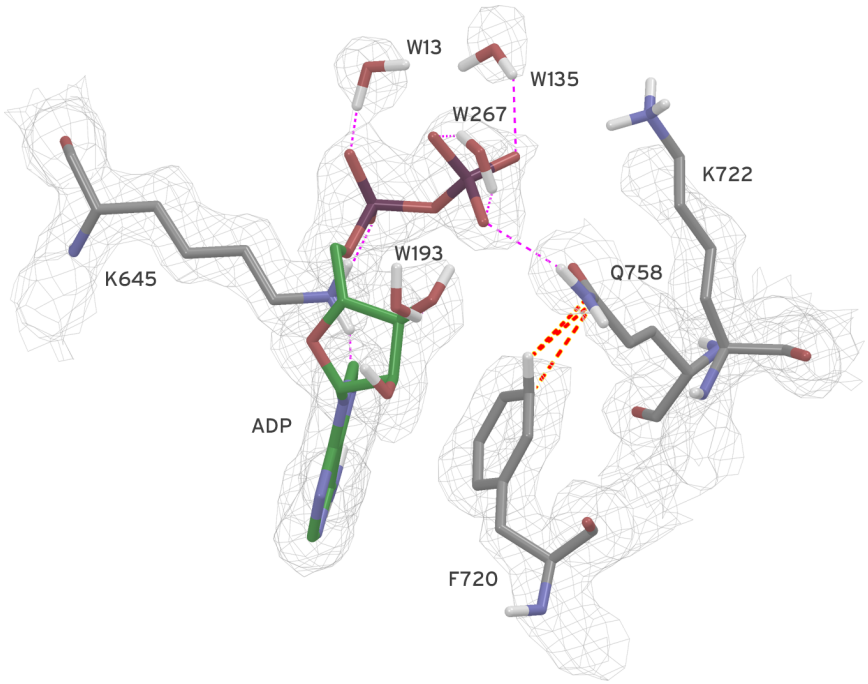


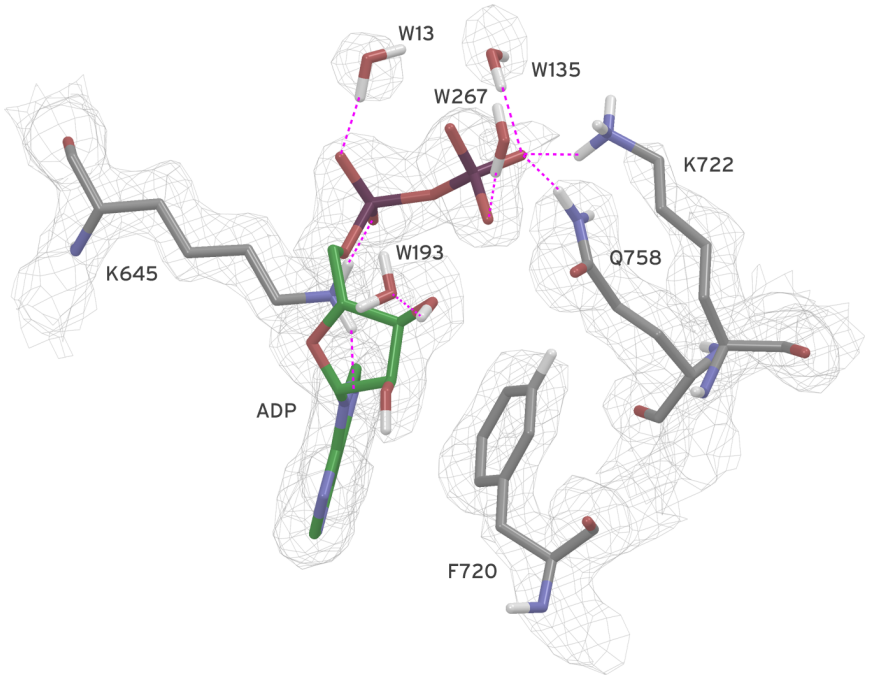


[Harrington/Ho/Ghosh/Tung, Ho/Harrington]

Other work

- Mathematical modeling of cell signaling networks
- Automated all-atom protein crystal structure refinement
- Fast multipole methods, direct solvers, matrix factorizations

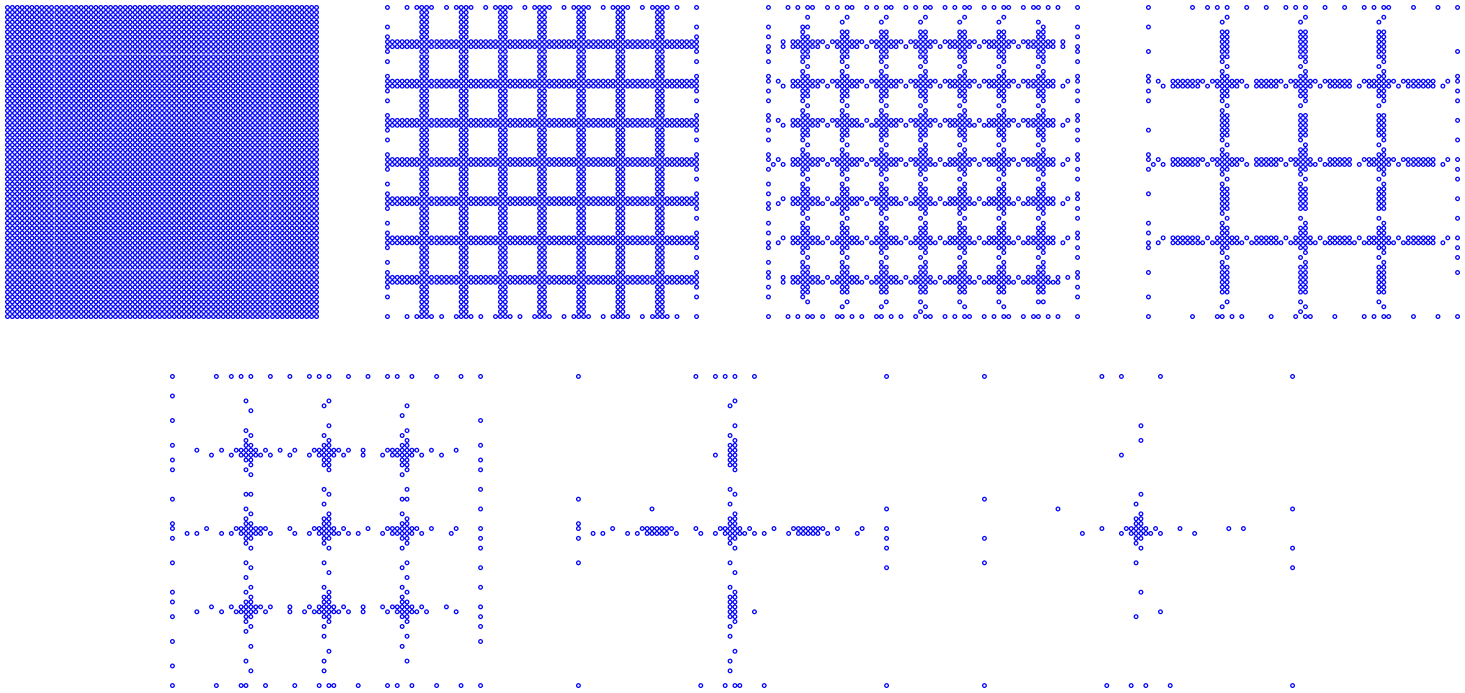




[Bell/Ho/Farid]

Other work

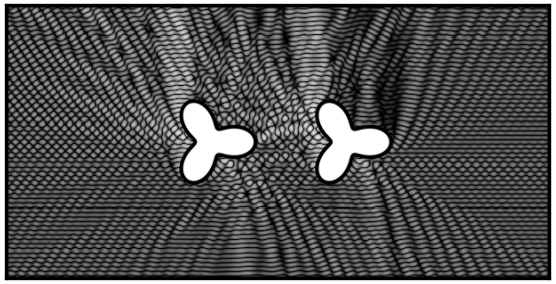
- Mathematical modeling of cell signaling networks
- Automated all-atom protein crystal structure refinement
- Fast multipole methods, direct solvers, matrix factorizations

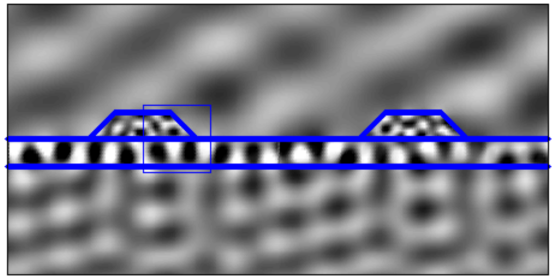


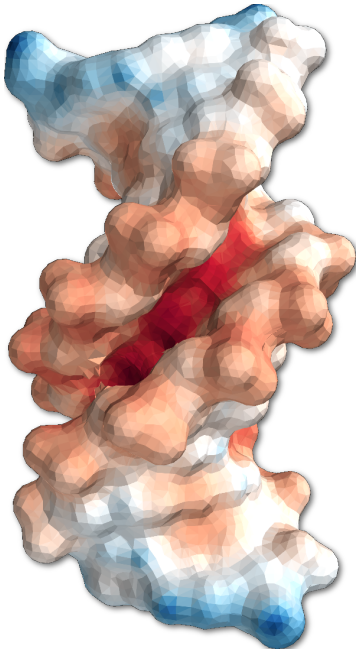
[Greengard/Ho/Lee, Ho/Greengard, Ho/Ying, Li/Yang/Martin/Ho/Ying, Minden/Damle/Ho/Ying]

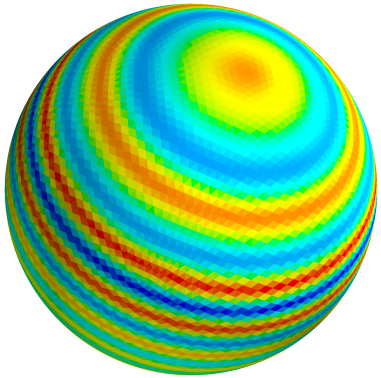
Other work

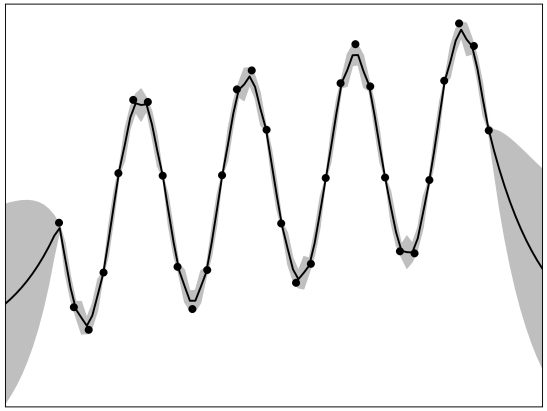
- Mathematical modeling of cell signaling networks
- Automated all-atom protein crystal structure refinement
- Fast multipole methods, direct solvers, matrix factorizations











[Greengard/Ho/Lee, Ho/Greengard, Ho/Ying, Li/Yang/Martin/Ho/Ying, Minden/Damle/Ho/Ying]

References

Parameter-free invalidation:

- E. Gross, B. Davis, K.L. Ho, D.J. Bates, H.A. Harrington. Numerical algebraic geometry for model selection. In preparation.
- H.A. Harrington, K.L. Ho. Parameter-free statistical model invalidation for weakly complex-balanced chemical reaction networks. In preparation.
- H.A. Harrington, K.L. Ho, N. Meshkat. Model rejection using differential algebra techniques. In preparation.
- H.A. Harrington, K.L. Ho, T. Thorne, M.P.H. Stumpf. Parameter-free model discrimination criterion based on steady-state coplanarity. **Proc. Natl. Acad. Sci. U.S.A.** 109 (39): 15746–15751, 2012.

Other:

- J.A. Bell, K.L. Ho, R. Farid. Significant reduction in errors associated with nonbonded contacts in protein crystal structures: automated all-atom refinement with *PrimeX*. **Acta Cryst.** D68: 935–952, 2012.
- L. Greengard, K.L. Ho, J.-Y. Lee. A fast direct solver for scattering from periodic structures with multiple material interfaces in two dimensions. **J. Comput. Phys.** 258: 738–751, 2014.
- H.A. Harrington, K.L. Ho, S. Ghosh, KC Tung. Construction and analysis of a modular model of caspase activation in apoptosis. **Theor. Biol. Med. Model.** 5: 26, 2008.
- K.L. Ho. Fast direct methods for molecular electrostatics. Ph.D. thesis, New York University, 2012.
- K.L. Ho, L. Greengard. A fast direct solver for structured linear systems by recursive skeletonization. **SIAM J. Sci. Comput.** 34 (5): A2507–A2532, 2012.
- K.L. Ho, L. Greengard. A fast semidirect least squares algorithm for hierarchically block separable matrices. **SIAM J. Matrix Anal. Appl.** 35 (2): 725–748, 2014.
- K.L. Ho, H.A. Harrington. Bistability in apoptosis by receptor clustering. **PLoS Comput. Biol.** 6 (10): e1000956, 2010.
- K.L. Ho, L. Ying. Hierarchical interpolative factorization for elliptic operators: differential equations. Preprint, arXiv:1307.2895 [math.NA], 2013. To appear, **Comm. Pure Appl. Math.**
- K.L. Ho, L. Ying. Hierarchical interpolative factorization for elliptic operators: integral equations. Preprint, arXiv:1307.2666 [math.NA], 2013. To appear, **Comm. Pure Appl. Math.**
- Y. Li, H. Yang, E. Martin, K. Ho, L. Ying. Butterfly factorization. Preprint, arXiv:1502.01379 [math.NA], 2015.
- V. Minden, A. Damle, K.L. Ho, L. Ying. A technique for updating hierarchical factorizations of integral operators. Preprint, arXiv:1411.5706 [math.NA], 2014.