

Geometrically Analyzing the Equilibria of Parametric Biochemical Networks Admitting Linear Conservation Laws: Case Studies

Changbo Chen and Wenyuan Wu

- ¹ Chongqing Key Laboratory of Automated Reasoning and Cognition,
Chongqing Institute of Green and Intelligent Technology,
Chinese Academy of Sciences, China
² University of Chinese Academy of Sciences

ADG 2023, Belgrade, Serbia
Sep. 22, 2023



中国科学院大学
University of Chinese Academy of Sciences

Outline

- 1 Background
- 2 GeoBlock: A geometric approach for parametric system solving
- 3 Enhancing GeoBlock to handle linear conservation laws

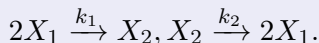
Outline

- 1 Background
- 2 GeoBlock: A geometric approach for parametric system solving
- 3 Enhancing GeoBlock to handle linear conservation laws

Biochemical reaction network

A simple reaction network

- The network consisting of two reactions between two species X_1, X_2 .



- The left hand side is called the reactant.
- The right hand side is called the product.
- The reaction rates are k_1 and k_2 .
- The stoichiometric matrix of the network is:

$$N = \begin{pmatrix} -2 & 2 \\ 1 & -1 \end{pmatrix}$$

The ODE modeling the reaction with mass-action kinetics

$$\begin{pmatrix} \dot{x}_1 \\ \dot{x}_2 \end{pmatrix} = \begin{pmatrix} -2 & 2 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} k_1 x_1^2 \\ k_2 x_2 \end{pmatrix}$$

Biochemical network and parametric system solving

The ODE modeling the biochemical network

- An autonomous ODE $\mathcal{E} : \frac{dX}{dt} = F(U, X)$
- Parameters $U = (U_1, \dots, U_k)$ take values in a finite positive box B_U .
- Unknowns $X = (X_1, \dots, X_m)$ take values in a finite positive box B_X .
- $F(U, X) = (F_1(U, X), \dots, F_m(U, X))$: rational functions.

The parametric system to solve

- Let $S := \wedge U \in B_U \wedge X \in B_X \wedge F = 0$.
- Then S describes the parametric steady states of $\dot{X} = F(U, X)$ in $B_U \times B_X$.

The problem

- Compute a partition of the box B_U into finitely many connected semi-algebraic sets such that above each of which the number of real zeros of S remains unchanged (if it is finite).

Related work: algebraic approaches

- cylindrical algebraic decomposition (CAD)
- comprehensive triangular decomposition of semi-algebraic set
- combining border polynomial or discriminant variety with open CAD
- algebraic methods taking the special structure of reaction networks
- ...

Outline

- 1 Background
- 2 **GeoBlock: A geometric approach for parametric system solving**
- 3 Enhancing GeoBlock to handle linear conservation laws

GeoBlock as a method

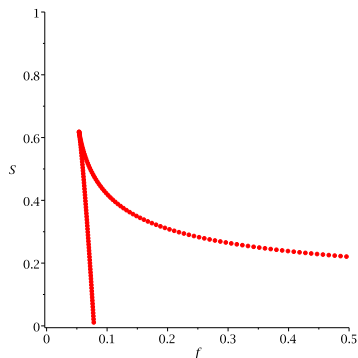
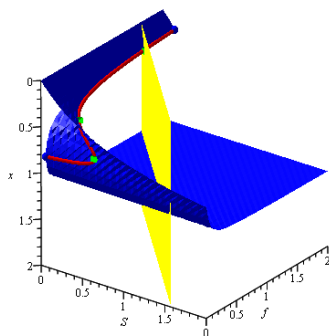
It computes a “border curve”

- It focuses on the bi-parametric case, that is $|U| = 2$.
- It aims to compute a **border curve** \mathcal{C} inside B_U such that the connected components $B_U \setminus \mathcal{C}$ provide the partition in demand.

Main idea of the algorithm

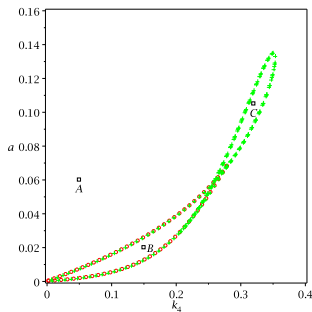
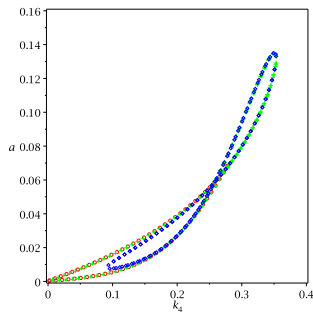
- $S(U, X)$ defines a real algebraic surface in a high-dimensional space.
- Its singular locus and critical locus w.r.t. the projection π (to U) as well its intersection with the boundary of B_X define $\pi^{-1}(\mathcal{C})$.
- To compute $\pi^{-1}(\mathcal{C})$, one firstly computes a witness point per connected component of it and approximates each component by polygonal chains.
- The border \mathcal{C} is then obtained for free by trivial projection.
- It further exploits the block structure of S by combining Tarjan's algorithm for computing strongly connected components of Graphs and Gauss-Jordan elimination.

Illustrating GeoBlock (I): computing a border curve (CASC 2016)



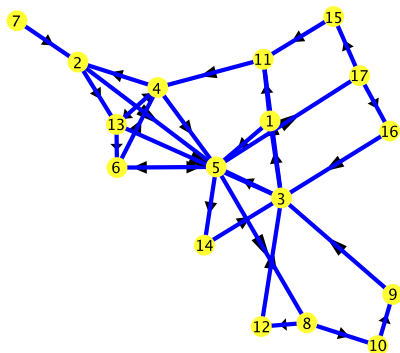
- Left: computing critical locus $\pi^{-1}(\mathcal{C})$ by finding witness points and curve tracing.
- Right: computing a border curve \mathcal{C} by trivial projection.

Illustrating GeoBlock (II): Geometric nature allows producing intrinsic boundaries

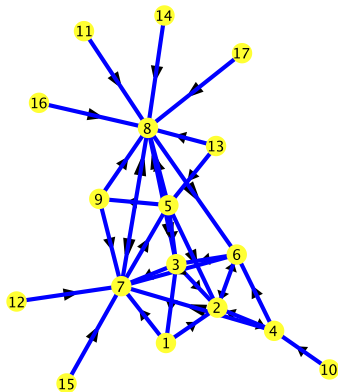


- Left image: boundaries produced by Hurwitz determinants.
- Right image: true fold and Hopf bifurcation boundaries.

Illustrating GeoBlock (III): combining Tarjan's algorithm with Gauss-Jordan elimination allows finding more blocks



Only two blocks if solely by Tarjan's algorithm



Finding nine blocks combining with Gauss-Jordan elimination

GeoBlock as a Maple package

Publicly available at <https://zenodo.org/record/5893837>.

```
> with(GeoBlock);  
[BoxBoundaryPoints, BuildAdjacencyMatrix, ChoosePositive,  
  
ConnectedBlocks, FoldBoundary, FoldDefiningSys,  
  
GaussJordanSolve, GetRealZeros, GridSolve, IneqBoundary,  
  
MixedVolume, PermuteMatrix, PlotCurveWithPoints,  
  
RemoveSelfLoop, RescaleCoeffs, WitnessPoints]
```

General motivation from biology for developing GeoBlock

- Biological systems in practice are sparse and have structures (Busiello, et al., 2017)
- **Sparsity** leads to better robustness and adaptability as well as confers rich structures for a biological network.
- **Power-law degree distribution** (Barabasi, et al., 1999): there are typically many vertices with low degree and a few vertices with high degree and the distribution follows a power-law.
- **Community structure** (Girvan and Newman, 2002): vertices tend to form some communities, inside which the connections are tight while between which the connections are loose.

These statistical properties imply that the biological networks are often organized in **loosely connected modules or clusters** on the whole, which creates opportunities to analyze them with tools from **graph theory**, such as mimicking the loosely connected modules in a biological network by **strongly connected components of a directed graph** in this study.

The development of GeoBlock (with Wenyuan Wu)

Avoiding express swell by curve tracing and numeric projection

A numerical method for computing border curves of bi-parametric real polynomial systems and applications. CASC 2016.

Smaller bifurcation boundaries due to semi-algebraic nature

A numerical method for analyzing the stability of bi-parametric biological systems. SYNASC 2016.

Even less boundaries by utilizing constraints and more stable by recalling

Revealing bistability in neurological disorder models by solving parametric polynomial systems geometrically. AISC 2018.

Relaxing the compactness assumption and exploiting block structure.

A Geometric Approach for Analyzing Parametric Biological Systems by Exploiting Block Triangular Structure (**GeoBlock**). SIAM J. Appl. Dyn. Syst. 21(2): 1573-1596, 2022.

New contribution in this talk

- Efficiently handling linear conservation laws by introducing a bipartite graph to describe the correlations between variables and constraints.

Outline

- 1 Background
- 2 GeoBlock: A geometric approach for parametric system solving
- 3 Enhancing GeoBlock to handle linear conservation laws

Recall the simple reaction network: linear conservation law

The ODE modeling the reaction with mass-action kinetics

$$\begin{pmatrix} \dot{x}_1 \\ \dot{x}_2 \end{pmatrix} = \begin{pmatrix} -2 & 2 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} k_1 x_1^2 \\ k_2 x_2 \end{pmatrix}$$

A simple reaction network

- The stoichiometric matrix of the network is:

$$N = \begin{pmatrix} -2 & 2 \\ 1 & -1 \end{pmatrix}$$

- Suppose that $(a_1, a_2)^T$ generates the left null space of N .
- Further requiring that $a_1 > 0$ and $a_2 > 0$, we can then choose $(a_1, a_2) = (1, 2)$.
- Let $c > 0$, then $x_1 + 2x_2 = c$ is a linear conservation law of the network.
- Note that we have $\dot{x}_1 + 2\dot{x}_2 = 0$ holds.

Handling biochemical networks admitting linear conservation laws (I)

What we have done

- $\dot{X} = F(U, X)$, where F are rational functions
- Let B_U and B_X be finite positive boxes.
- Let $S := \wedge U \in B_U \wedge X \in B_X \wedge F = 0$.
- Compute a partition of the box B_U into finitely many connected semi-algebraic sets such that above each of which the number of real zeros of S remains unchanged (if it is finite).

New challenge

- If the ODE $\dot{X} = F(U, X)$ admits linear conservation laws, then $F(U, X) = 0$ fails to be zero-dimensional for generically chosen values of U .

Handling biochemical networks admitting linear conservation laws (II)

The new formulation

- One first computes some linear independent conservation laws $l_i(X)$, $i = 1, \dots, e$.
- Let $l_i(X) = C_i$, where C_i is a (parametric) constant.
- $C = C_1, \dots, C_e$ take values in a positive box B_C .
- $S' := S \wedge_i l_i(X) = C_i \wedge C \in B_C$ is zero-dimensional for generically chosen values of U and C .
- One solves S' instead of S .

The new problem

- S' may have more equations than unknowns.
- Transform S' into an equivalent square system S'' such that S'' is square and S'' is the easier one to solve among possible transformations.

From over-determined to square

Recall the simple network

- The ODE is:

$$\begin{cases} \dot{x}_1 = f_1(k_1, k_2, x_1, x_2) \\ \dot{x}_2 = f_2(k_1, k_2, x_1, x_2) \end{cases}$$

- The conservation law $\ell := x_1 + 2x_2 = c$.
- The linear conservation law tells that $f_1 + 2f_2 = 0$.
- Thus $\langle f_1, f_2, \ell - c \rangle = \langle f_1, \ell - c \rangle$.

Lemma

- Given a polynomial ODE $\dot{X} = F(U, X)$ with $|F| = |X|$.
- Let $L := \{\ell_i(X) = C_i, i = 1, \dots, e\}$, where $e < |F|$, be a set of independent linear conservation constraints.
- Support the variables in L are partitioned into a set of leading variables X_ℓ and a set of free variables X_f . Accordingly, partition F into F_ℓ and F_f .
- Let $S'' := S' \setminus \{F_\ell = 0\}$. Then S'' is square and $\text{Zero}(S') = \text{Zero}(S'')$.

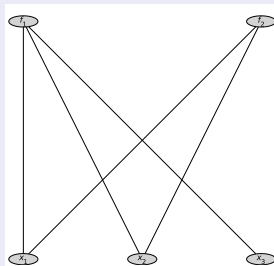
A correlation graph between constraints and their variables

The graph

- Let $F = \{f_1, \dots, f_s\} \subseteq \mathbb{R}[U, X]$ be finite, where $X = \{x_1, \dots, x_m\}$.
- Let $V = X \cup F$.
- Define E as $\{(x_i, f_j) \mid \text{where } x_i \text{ appears in } f_j\}$.

An example

Let $F = \{u_1x_1^2 + u_2x_2 - x_3, x_1^2x_2 + x_2 - 1\}$. Then the associated bipartite graph $G = (V, E)$ is depicted as below:



First example (Biomodels 26, Bradford et al., 2020)

$$\begin{aligned} \dot{x}_1 &= -k_1x_1x_4 - k_{16}x_1x_5 + k_2x_6 + k_{15}x_{11} \\ \dot{x}_2 &= k_3x_6 + k_5x_7 + k_{10}x_9 + k_{13}x_{10} - x_2x_5(k_{11} + k_{12}) - k_4x_2x_4 \\ \dot{x}_3 &= -k_7x_3x_5 + k_6x_7 + k_8x_8 \\ \dot{x}_4 &= x_6(k_2 + k_3) + x_7(k_5 + k_6) - k_1x_1x_4 - k_4x_2x_4 \\ \dot{x}_5 &= k_8x_8 + k_{10}x_9 + k_{13}x_{10} + k_{15}x_{11} - x_2x_5(k_{11} + k_{12}) - k_7x_3x_5 - k_{16}x_1x_5 \\ \dot{x}_6 &= k_1x_1x_4 - x_6(k_2 + k_3) \\ \dot{x}_7 &= k_4x_2x_4 - x_7(k_5 + k_6) \\ \dot{x}_8 &= k_7x_3x_5 - x_8(k_8 + k_9) \\ \dot{x}_9 &= k_{11}x_2x_5 + k_9x_8 - k_{10}x_9 \\ \dot{x}_{10} &= k_{12}x_2x_5 - x_{10}(k_{13} + k_{14}) \\ \dot{x}_{11} &= k_{16}x_1x_5 + k_{14}x_{10} - k_{15}x_{11}, \end{aligned} \tag{1}$$

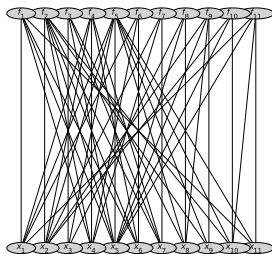
Linear conservation law

$$\dot{\ell}_1 := \dot{x}_5 + \dot{x}_8 + \dot{x}_9 + \dot{x}_{10} + \dot{x}_{11} = 0 \tag{2}$$

The new constraints brought by conservation laws

$$\begin{aligned} \ell_1 &:= x_5 + x_8 + x_9 + x_{10} + x_{11} = k_{17} \\ \ell_2 &:= x_4 + x_6 + x_7 = k_{18} \\ \ell_3 &:= x_1 + x_2 + x_3 + x_6 + x_7 + x_8 + x_9 + x_{10} + x_{11} = k_{19}. \end{aligned} \tag{3}$$

The associated bipartite graph of F



Choosing a set of variables and removing corresponding equations

- The variables must all be leading ones.
- The variables having the least number of neighbors.

$[1, 5], [2, 6], [3, 3], [4, 5], [5, 8], [6, 4], [7, 4], [8, 4], [9, 3], [10, 4], [11, 3].$

- The variable not appearing in nonlinear terms.

A possible set is $\{x_3, x_6, x_9\}$, implying $\{f_3, f_6, f_9\}$ can be removed from F .

The result of applying GeoBlock on S''

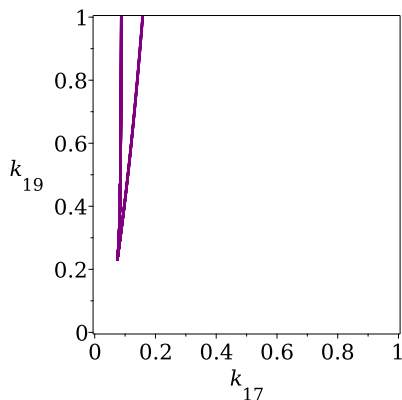


Figure: The computed border curve.

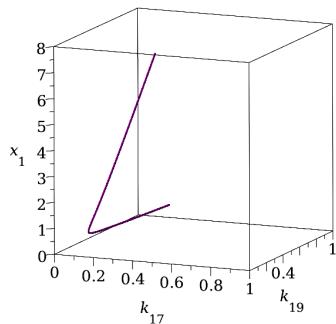


Figure: The preimage of the border curve in the space (k_{17}, k_{19}, x_1) .

Comparison with the work of Bradford et al., 2020

The right one is obtained by combining different algebraic approaches, such as cylindrical algebraic decomposition, real triangular decomposition, virtual term substitution and parametric Gauss elimination.

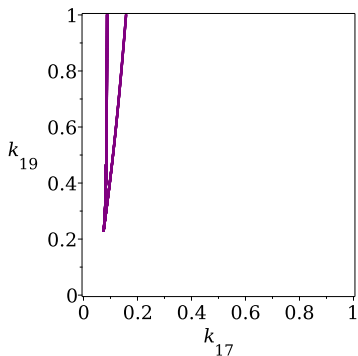


Figure: The computed border curve (rescaled to $[0, 1] \times [0, 1]$).

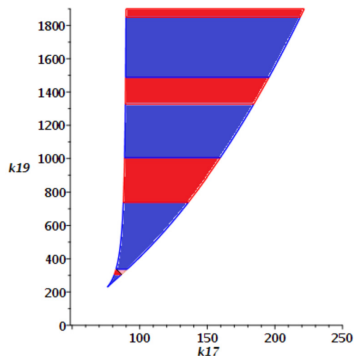


Figure: Visualization of the cells obtained by algebraic approaches.

Second example (16 variables, 2 parameters)

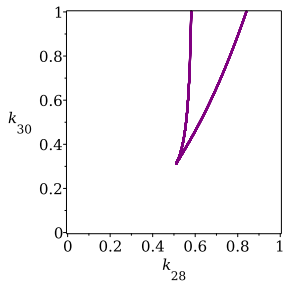


Figure: The computed border curve for the second network

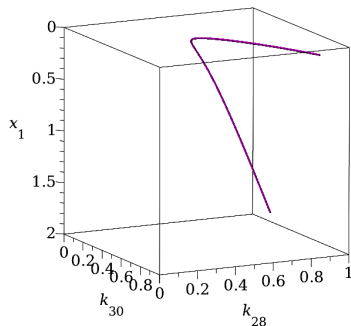


Figure: The preimage of the border curve in the space (k_{28}, k_{30}, x_1)

Comparison with the work of Bradford et al., 2020

Algebraic approaches fail to compute the cell decomposition and the right one is obtained by grid sampling.

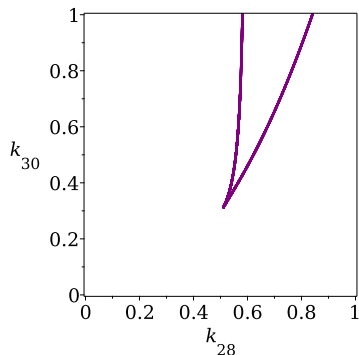


Figure: The computed border curve (rescaled to $[0, 1] \times [0, 1]$).

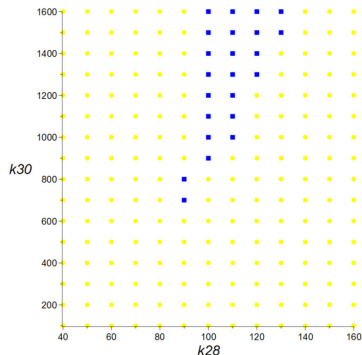


Figure: Grid sampling.

Conclusion

- We introduce a bipartite graph to describe the correlations between variables and constraints.
- We show how to adapt GeoBlock to handle biochemical networks admitting linear conservation laws.
- We show how the bipartite graph helps GeoBlock to choose right systems to solve.
- We apply the geometric method to analyze the equilibria of two biochemical networks challenging for algebraic approaches.