# Polynomial system solving in applications

Julia Lindberg

University of Wisconsin-Madison

LAAS BrainPOP Seminar

# Motivation

- Polynomial systems arise naturally in many applications

# Motivation

- Polynomial systems arise naturally in many applications
- Solutions to these systems often have physical meaning

# Motivation

- Polynomial systems arise naturally in many applications
- Solutions to these systems often have physical meaning
- This talk will outline techniques to solve these systems

# Motivation

- Polynomial systems arise naturally in many applications
- Solutions to these systems often have physical meaning
- This talk will outline techniques to solve these systems
- We will consider three examples
  1. Lagrange systems of polynomial optimization problems
  2. Gaussian mixture moment systems
  3. Power flow equations

# Set up

- Consider the variety defined by polynomials $f_1, \ldots, f_n \in \mathbb{R}[x_1, \ldots, x_n]$

$$\mathcal{V}(F) = \{x \in \mathbb{C}^n : f_1(x) = 0, \ldots, f_n(x) = 0\}$$

where $\dim(\mathcal{V}(F)) = 0$

# Set up

- Consider the variety defined by polynomials $f_1, \dots, f_n \in \mathbb{R}[x_1, \dots, x_n]$

$$\mathcal{V}(F) = \{x \in \mathbb{C}^n : f_1(x) = 0, \dots, f_n(x) = 0\}$$

  where $\dim(\mathcal{V}(F)) = 0$

- **Goal:** Find all points in $\mathcal{V}(F)$

# Table of Contents

# Method 1
## Homotopy continuation

- Idea: Solving most polynomial systems is hard, but some are easy

# Method 1

## Homotopy continuation

- Idea: Solving most polynomial systems is hard, but some are easy

$$H_T = \begin{cases} 2(x_2 y_1 - x_1 y_2) + 3y_1 = 0 \\ 2(x_1 y_2 - x_2 y_1) + 4y_2 = 0 \\ x_1^2 + y_1^2 = 1 \\ x_2^2 + y_2^2 = 1 \end{cases} \qquad H_S = \begin{cases} x_1^2 = 1 \\ x_2^2 = 1 \\ y_1^2 = 1 \\ y_2^2 = 1 \end{cases}$$

# Method 1

## Homotopy continuation

- Idea: Solving most polynomial systems is hard, but some are easy

$$H_T = \begin{cases} 2(x_2y_1 - x_1y_2) + 3y_1 = 0 \\ 2(x_1y_2 - x_2y_1) + 4y_2 = 0 \\ x_1^2 + y_1^2 = 1 \\ x_2^2 + y_2^2 = 1 \end{cases} \qquad H_S = \begin{cases} x_1^2 = 1 \\ x_2^2 = 1 \\ y_1^2 = 1 \\ y_2^2 = 1 \end{cases}$$

- Can I map my solutions from $H_S$ to $H_T$?

# Method 1

## Homotopy continuation

- Idea: Solving most polynomial systems is hard, but some are easy

$$H_T = \begin{cases} 2(x_2y_1 - x_1y_2) + 3y_1 = 0 \\ 2(x_1y_2 - x_2y_1) + 4y_2 = 0 \\ x_1^2 + y_1^2 = 1 \\ x_2^2 + y_2^2 = 1 \end{cases} \qquad H_S = \begin{cases} x_1^2 = 1 \\ x_2^2 = 1 \\ y_1^2 = 1 \\ y_2^2 = 1 \end{cases}$$

- Can I map my solutions from $H_S$ to $H_T$?
- Define $H_t := (1-t)H_S + tH_T$ and compute $H_t$ as $t \to 1$
  - Called following *homotopy paths*

- Idea: Solving most polynomial systems is hard, but some are easy

$$H_T = \begin{cases} 2(x_2y_1 - x_1y_2) + 3y_1 = 0 \\ 2(x_1y_2 - x_2y_1) + 4y_2 = 0 \\ x_1^2 + y_1^2 = 1 \\ x_2^2 + y_2^2 = 1 \end{cases} \qquad H_S = \begin{cases} x_1^2 = 1 \\ x_2^2 = 1 \\ y_1^2 = 1 \\ y_2^2 = 1 \end{cases}$$

- Can I map my solutions from $H_S$ to $H_T$?
- Define $H_t := (1-t)H_S + tH_T$ and compute $H_t$ as $t \to 1$
    - Called following *homotopy paths*
- Typically use predictor-corrector methods
    - Predict: Take step along tangent direction at a point
    - Correct: Use Newton's method

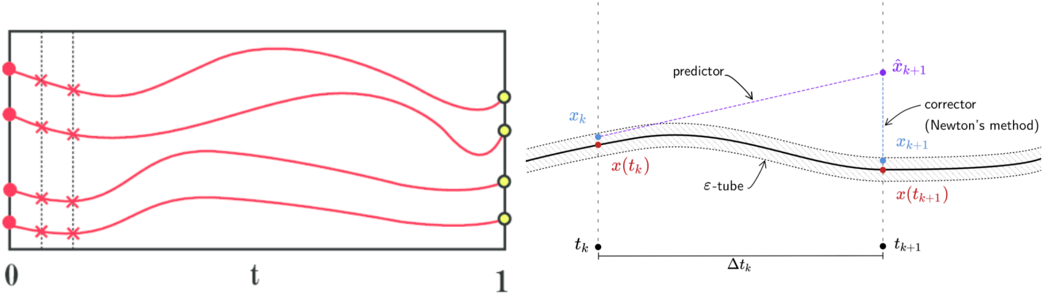



Figure: The homotopy $H_t = (1-t)H_S + tH_T$ (left)[KW14] and a predictor corrector step (right) [BT18]

# Homotopy continuation

## Start system

- Want to pick a start system, $H_S$, such that
  1. The solutions of $H_S$ are easy to find
  2. The number of solutions to $H_S \approx$ the number of solutions to $H_T$

- Want to pick a start system, $H_S$, such that
  1. The solutions of $H_S$ are easy to find
  2. The number of solutions to $H_S \approx$ the number of solutions to $H_T$

- Need to know $|\mathcal{V}(F)|$

**Theorem (Bezout)**

$|\mathcal{V}(F)| \leq d_1 \cdots d_n \text{ where } d_i = \deg(f_i)$

**Theorem (Bezout)**

$|\mathcal{V}(F)| \leq d_1 \cdots d_n$ *where* $d_i = \deg(f_i)$

- If $|\mathcal{V}(F)| \approx d_1 \cdots d_n$ then a **total degree** start system is suitable. i.e.

$$H_S = \langle x_1^{d_1} - 1, \ldots, x_n^{d_n} - 1 \rangle$$

**Theorem (Bezout)**

$|\mathcal{V}(F)| \leq d_1 \cdots d_n$ *where* $d_i = \deg(f_i)$

- If $|\mathcal{V}(F)| \approx d_1 \cdots d_n$ then a **total degree** start system is suitable. i.e.

$$H_S = \langle x_1^{d_1} - 1, \dots, x_n^{d_n} - 1 \rangle$$

- This bound is **generically** tight but can be a strict upper bound when $f_i$ are sparse

# Newton Polytopes

- The *monomial support* of $f$ is the set of exponent vectors of the monomials of $f$

# Newton Polytopes

- The *monomial support* of $f$ is the set of exponent vectors of the monomials of $f$
- Ex. $f(x, y) = 1 + 2x - 3y + xy$ has monomial support

$$\{(0,0),(1,0),(0,1),(1,1)\}$$

# Newton Polytopes

- The *monomial support* of $f$ is the set of exponent vectors of the monomials of $f$
- Ex. $f(x, y) = 1 + 2x - 3y + xy$ has monomial support

$$\{(0, 0), (1, 0), (0, 1), (1, 1)\}$$

- The *Newton polytope* of $f$ is the convex hull of its monomial support

# Newton Polytopes

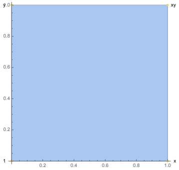
- The *monomial support* of $f$ is the set of exponent vectors of the monomials of $f$

- Ex. $f(x, y) = 1 + 2x - 3y + xy$ has monomial support

$$\{(0,0),(1,0),(0,1),(1,1)\}$$

- The *Newton polytope* of $f$ is the convex hull of its monomial support

- Ex. If $f(x, y) = 1 + 2x - 3y + xy$, then $\mathrm{Newt}(f) = \mathrm{Conv}\{(0,0),(1,0),(0,1),(1,1)\}$



# Mixed Volume

- Given $k$ polytopes, $P_1, \ldots, P_k$ in $\mathbb{R}^n$ their *mixed volume* is

$$\mathrm{MVol}(P_1, \ldots, P_k) = \sum_{J \subseteq [k]} (-1)^{k-|J|} \cdot \mathrm{Vol}_n(P_J)$$

  where $P_J = P_{i_1} + \ldots + P_{i_{|J|}}$ and $J = (i_1, \ldots, i_{|J|})$

# Mixed Volume

- Given $k$ polytopes, $P_1, \ldots, P_k$ in $\mathbb{R}^n$ their *mixed volume* is

$$\mathrm{MVol}(P_1, \ldots, P_k) = \sum_{J \subseteq [k]} (-1)^{k-|J|} \cdot \mathsf{Vol}_n(P_J)$$

  where $P_J = P_{i_1} + \ldots + P_{i_{|J|}}$ and $J = (i_1, \ldots, i_{|J|})$

- Ex. $k = n = 2$,

$$\mathrm{MVol}(P_1, P_2) = \mathrm{Area}(P_1 + P_2) - \mathrm{Area}(P_1) - \mathrm{Area}(P_2)$$

# Start system

## Polyhedral

**Theorem (BKK Bound [Ber75, Kho78, Kou76])**

$|\mathcal{V}(F)| \leq \mathrm{MVol}(\mathrm{Newt}(f_1), \ldots, \mathrm{Newt}(f_n))$

**Theorem (BKK Bound [Ber75, Kho78, Kou76])**

$|\mathcal{V}(F)| \leq \mathrm{MVol}(\mathrm{Newt}(f_1), \ldots, \mathrm{Newt}(f_n))$

- If $\mathrm{MVol}(\mathrm{Newt}(f_1), \ldots, \mathrm{Newt}(f_n)) \ll d_1 \cdots d_n$ then a **polyhedral** start system is suitable

**Theorem (BKK Bound [Ber75, Kho78, Kou76])**

$|\mathcal{V}(F)| \leq \mathrm{MVol}(\mathrm{Newt}(f_1), \ldots, \mathrm{Newt}(f_n))$

- If $\mathrm{MVol}(\mathrm{Newt}(f_1), \ldots, \mathrm{Newt}(f_n)) \ll d_1 \cdots d_n$ then a **polyhedral** start system is suitable
- There exists an algorithm that finds this binomial start system [HS95]

**Theorem (BKK Bound [Ber75, Kho78, Kou76])**

$|\mathcal{V}(F)| \leq \mathrm{MVol}(\mathrm{Newt}(f_1), \ldots, \mathrm{Newt}(f_n))$

- If $\mathrm{MVol}(\mathrm{Newt}(f_1), \ldots, \mathrm{Newt}(f_n)) \ll d_1 \cdots d_n$ then a **polyhedral** start system is suitable
- There exists an algorithm that finds this binomial start system [HS95]
- In general, not easy to compute the mixed volume (#P hard)

# Table of Contents

# Example 1

## Critical points of polynomial optimization problems

- Want to solve polynomial optimization problem of the form

$$\min_{x \in \mathbb{R}^n} g(x) \quad \text{subject to} \quad f_i(x) = 0, \; i \in [m]$$

# Example 1

## Critical points of polynomial optimization problems

- Want to solve polynomial optimization problem of the form

$$\min_{x \in \mathbb{R}^n} g(x) \quad \text{subject to} \quad f_i(x) = 0, \; i \in [m]$$

- Consider Lagrangian

$$\mathcal{L}(x, \lambda) = g(x) - \sum_{i=1}^{m} \lambda_i f_i(x)$$

# Example 1

## Critical points of polynomial optimization problems

- Want to solve polynomial optimization problem of the form

$$\min_{x \in \mathbb{R}^n} \; g(x) \quad \text{subject to} \quad f_i(x) = 0, \;\; i \in [m]$$

- Consider Lagrangian

$$\mathcal{L}(x, \lambda) = g(x) - \sum_{i=1}^{m} \lambda_i f_i(x)$$

- Smooth critical points = solutions to polynomial system

$$\ell_i = \frac{\partial \mathcal{L}}{\partial x_i} = 0$$
$$f_i = \frac{\partial \mathcal{L}}{\partial \lambda_i} = 0$$

**Theorem (L., Nicholson, Rodriguez, Wang [LNRW21])**

*For generic $f_i, g$ the number of smooth critical points to*

$$\min_{x \in \mathbb{R}^n} g(x) \quad \textit{subject to} \quad f_i(x) = 0, \ i \in [m]$$

*is equal to* $\mathrm{MVol}(\ell_1, \ldots, \ell_n, f_1, \ldots, f_m)$. *Moreover, the number of critical points is dictated by* $\mathrm{Newt}(f_1), \ldots, \mathrm{Newt}(f_m), \mathrm{Newt}(g)$.

**Theorem (L., Nicholson, Rodriguez, Wang [LNRW21])**

*For generic $f_i, g$ the number of smooth critical points to*

$$\min_{x\in\mathbb{R}^n} g(x) \quad \textit{subject to} \quad f_i(x) = 0,\ i \in [m]$$

*is equal to* $\mathrm{MVol}(\ell_1, \ldots, \ell_n, f_1, \ldots, f_m)$. *Moreover, the number of critical points is dictated by* $\mathrm{Newt}(f_1), \ldots, \mathrm{Newt}(f_m), \mathrm{Newt}(g)$.

- **Implication 1:** Polyhedral start systems will efficiently find all complex critical points for polynomial optimization problems

**Theorem (L., Nicholson, Rodriguez, Wang [LNRW21])**

*For generic $f_i, g$ the number of smooth critical points to*

$$\min_{x \in \mathbb{R}^n} g(x) \quad \textit{subject to} \quad f_i(x) = 0, \; i \in [m]$$

*is equal to* $\mathrm{MVol}(\ell_1, \ldots, \ell_n, f_1, \ldots, f_m)$. *Moreover, the number of critical points is dictated by* $\mathrm{Newt}(f_1), \ldots, \mathrm{Newt}(f_m), \mathrm{Newt}(g)$.

- **Implication 1:** Polyhedral start systems will efficiently find all complex critical points for polynomial optimization problems
- **Implication 2:** Only consider monomials corresponding to vertices of Newton polytopes

# Example 1

## Critical points of polynomial optimization problems

- Consider

$$\min_{x\in\mathbb{R}^3} 3x_1 - x_2 + 2x_3 \quad \text{subject to} \quad x_1^2 - x_1x_2 + x_1x_3 - 2x_2^2 + 3x_2x_3 + 4x_3^2 - x_1 + 2x_2 - x_3 - 1 = 0$$

# Example 1

## Critical points of polynomial optimization problems

- Consider

$$\min_{x\in\mathbb{R}^3} 3x_1 - x_2 + 2x_3 \quad \text{subject to} \quad x_1^2 - x_1x_2 + x_1x_3 - 2x_2^2 + 3x_2x_3 + 4x_3^2 - x_1 + 2x_2 - x_3 - 1 = 0$$

- The Lagrange system is

$$\begin{aligned}
\ell_1 &= 3 - \lambda(2x_1 - x_2 + x_3 - 1) \\
\ell_2 &= -1 - \lambda(-x_1 - 4x_2 + 3x_3 + 2) \\
\ell_3 &= 2 - \lambda(x_1 + 3x_2 + 8x_3 - 1) \\
f &= x_1^2 - x_1x_2 + x_1x_3 - 2x_2^2 + 3x_2x_3 + 4x_3^2 - x_1 + 2x_2 - x_3 - 1
\end{aligned}$$

Using a total degree start system, need to track 16 paths

# Example 1

## Critical points of polynomial optimization problems

- Consider

$$\min_{x\in\mathbb{R}^3} 3x_1 - x_2 + 2x_3 \quad \text{subject to} \quad x_1^2 - x_1x_2 + x_1x_3 - 2x_2^2 + 3x_2x_3 + 4x_3^2 - x_1 + 2x_2 - x_3 - 1 = 0$$

- The Lagrange system is

$$\begin{aligned} \ell_1 &= 3 - \lambda(2x_1 - x_2 + x_3 - 1) \\ \ell_2 &= -1 - \lambda(-x_1 - 4x_2 + 3x_3 + 2) \\ \ell_3 &= 2 - \lambda(x_1 + 3x_2 + 8x_3 - 1) \\ f &= x_1^2 - x_1x_2 + x_1x_3 - 2x_2^2 + 3x_2x_3 + 4x_3^2 - x_1 + 2x_2 - x_3 - 1 \end{aligned}$$

Using a total degree start system, need to track 16 paths

- I claim there are only 2 critical points

# Example 1

## Critical points of polynomial optimization problems

- Consider the same optimization problem restricted to the monomials corresponding to vertices of each Newton polytope:

$$\min_{x \in \mathbb{R}^3} 3x_1 - x_2 + 2x_3 \quad \text{subject to} \quad x_1^2 - 2x_2^2 + 4x_3^2 - 1 = 0$$

# Example 1

## Critical points of polynomial optimization problems

- Consider the same optimization problem restricted to the monomials corresponding to vertices of each Newton polytope:

$$\min_{x \in \mathbb{R}^3} 3x_1 - x_2 + 2x_3 \quad \text{subject to} \quad x_1^2 - 2x_2^2 + 4x_3^2 - 1 = 0$$

- The Lagrange system is

$$\begin{aligned} \ell_1 &= 3 - 2\lambda x_1 \\ \ell_2 &= -1 + 4\lambda x_2 \\ \ell_3 &= 2 - 8\lambda x_3 \\ f &= x_1^2 - 2x_2^2 + 4x_3^2 - 1 \end{aligned}$$

# Example 1

## Critical points of polynomial optimization problems

- Consider the same optimization problem restricted to the monomials corresponding to vertices of each Newton polytope:

$$\min_{x\in\mathbb{R}^3} 3x_1 - x_2 + 2x_3 \quad \text{subject to} \quad x_1^2 - 2x_2^2 + 4x_3^2 - 1 = 0$$

- The Lagrange system is

$$\begin{aligned} \ell_1 &= 3 - 2\lambda x_1 \\ \ell_2 &= -1 + 4\lambda x_2 \\ \ell_3 &= 2 - 8\lambda x_3 \\ f &= x_1^2 - 2x_2^2 + 4x_3^2 - 1 \end{aligned}$$

- Can explicitly solve this system and get 2 complex solutions, then use this as a start system for the general problem

# Example 1

## Critical points of polynomial optimization problems

- Consider the same optimization problem restricted to the monomials corresponding to vertices of each Newton polytope:

$$\min_{x \in \mathbb{R}^3} 3x_1 - x_2 + 2x_3 \quad \text{subject to} \quad x_1^2 - 2x_2^2 + 4x_3^2 - 1 = 0$$

- The Lagrange system is

$$\begin{aligned} \ell_1 &= 3 - 2\lambda x_1 \\ \ell_2 &= -1 + 4\lambda x_2 \\ \ell_3 &= 2 - 8\lambda x_3 \\ f &= x_1^2 - 2x_2^2 + 4x_3^2 - 1 \end{aligned}$$

- Can explicitly solve this system and get 2 complex solutions, then use this as a start system for the general problem
- This generalizes to minimizing a linear function over any hypersurface

# Table of Contents

# Example 2

## Gaussian mixture moment systems

- Classical problem in statistics is density estimation: “Given $N$ samples from a density $p$, can I estimate $p$?”

# Example 2

## Gaussian mixture moment systems

- Classical problem in statistics is density estimation: "Given $N$ samples from a density $p$, can I estimate $p$?"

- Gaussian mixture models are a popular family to consider since they universally approximate smooth densities

# Example 2

## Gaussian mixture moment systems

- Classical problem in statistics is density estimation: "Given $N$ samples from a density $p$, can I estimate $p$?"
- Gaussian mixture models are a popular family to consider since they universally approximate smooth densities
- **Problem:** Given $N$ samples distributed as mixture of $k$ Gaussians, recover mean $\mu_i$, variance, $\sigma_i^2$ and mixing coefficient, $\lambda_i$ of each component

# Example 2

## Gaussian mixture moment systems

- Classical problem in statistics is density estimation: "Given $N$ samples from a density $p$, can I estimate $p$?"
- Gaussian mixture models are a popular family to consider since they universally approximate smooth densities
- **Problem:** Given $N$ samples distributed as mixture of $k$ Gaussians, recover mean $\mu_i$, variance, $\sigma_i^2$ and mixing coefficient, $\lambda_i$ of each component
- Many techniques to do this, consider **method of moments**

# Moments

## Definition

Let $f : \mathbb{R} \to \mathbb{R}$ be the probability density function of a random variable $X$. For $i \geq 0$, the $i-th$ *moment of* $X$ is

$$m_i = \mathbb{E}[X^i] = \int_{\mathbb{R}} x^i f(x) dx.$$

# Moments

## Definition

Let $f : \mathbb{R} \to \mathbb{R}$ be the probability density function of a random variable $X$. For $i \geq 0$, the *$i-th$ moment of $X$* is

$$m_i = \mathbb{E}[X^i] = \int_{\mathbb{R}} x^i f(x) dx.$$

- For parameterized distributions, moments are functions of parameters

# Moments

## Definition

Let $f : \mathbb{R} \to \mathbb{R}$ be the probability density function of a random variable $X$. For $i \geq 0$, the $i-th$ *moment of* $X$ is

$$m_i = \mathbb{E}[X^i] = \int_{\mathbb{R}} x^i f(x) dx.$$

- For parameterized distributions, moments are functions of parameters
- Ex. Suppose $X \sim \text{Unif}[a, b]$ where $-\infty < a < b < \infty$

# Moments

## Definition

Let $f : \mathbb{R} \to \mathbb{R}$ be the probability density function of a random variable $X$. For $i \geq 0$, the *$i-th$ moment of $X$* is

$$m_i = \mathbb{E}[X^i] = \int_{\mathbb{R}} x^i f(x) dx.$$

- For parameterized distributions, moments are functions of parameters
- Ex. Suppose $X \sim \text{Unif}[a, b]$ where $-\infty < a < b < \infty$
- The first few moments are:

$$m_1 = \frac{1}{2}(a + b)$$
$$m_2 = \frac{1}{3}(a^2 + 2ab + b^2)$$
$$m_3 = \frac{1}{4}(b^3 + ab^2 + a^2 b + a^3)$$

- Consider a statistical model with $n$ unknown parameters, $\theta = (\theta_1, \dots, \theta_n)$ and the moments up to order $n$ as functions of $\theta$

$$m_1 = g_1(\theta), \dots, m_n = g_n(\theta)$$

and samples $y_1, \dots, y_N$

# Method of Moments

## Procedure

- Consider a statistical model with $n$ unknown parameters, $\theta = (\theta_1, \dots, \theta_n)$ and the moments up to order $n$ as functions of $\theta$

$$m_1 = g_1(\theta), \dots, m_n = g_n(\theta)$$

  and samples $y_1, \dots, y_N$
- **Method of Moments:**
  1. Compute sample moments

$$\overline{m}_i = \frac{1}{N} \sum_{j=1}^{N} y_j^i$$

# Method of Moments

## Procedure

- Consider a statistical model with $n$ unknown parameters, $\theta = (\theta_1, \ldots, \theta_n)$ and the moments up to order $n$ as functions of $\theta$

$$m_1 = g_1(\theta), \ldots, m_n = g_n(\theta)$$

  and samples $y_1, \ldots, y_N$

- **Method of Moments:**
  1. Compute sample moments

$$\overline{m}_i = \frac{1}{N} \sum_{j=1}^{N} y_j^i$$

  2. Solve $g_i(\theta) = \overline{m}_i$ for $i = 1, \ldots, n$ to recover parameters

- The moments of the Gaussian distributions are $M_0(\mu, \sigma^2) = 1$, $M_1(\mu, \sigma^2) = \mu$,

$$M_\ell(\mu, \sigma^2) = \mu M_{\ell-1} + (\ell - 1)\sigma^2 M_{\ell-2}, \qquad \ell \geq 2$$

- The moments of the Gaussian distributions are $M_0(\mu, \sigma^2) = 1$, $M_1(\mu, \sigma^2) = \mu$,

$$M_\ell(\mu, \sigma^2) = \mu M_{\ell-1} + (\ell - 1)\sigma^2 M_{\ell-2}, \qquad \ell \geq 2$$

- The moments of mixtures of $k$ Gaussians are

$$m_\ell = \sum_{i=1}^{k} \lambda_i M_\ell(\mu_i, \sigma_i^2), \qquad \ell \geq 0$$

# Method of Moments

## Gaussian Mixture Models

- The moments of the Gaussian distributions are $M_0(\mu, \sigma^2) = 1$, $M_1(\mu, \sigma^2) = \mu$,

$$M_\ell(\mu, \sigma^2) = \mu M_{\ell-1} + (\ell - 1)\sigma^2 M_{\ell-2}, \qquad \ell \geq 2$$

- The moments of mixtures of $k$ Gaussians are

$$m_\ell = \sum_{i=1}^{k} \lambda_i M_\ell(\mu_i, \sigma_i^2), \qquad \ell \geq 0$$

- There are three special cases of Gaussian mixture models commonly studied in the statistics literature:
  1. The mixing coefficients are known
  2. The mixing coefficients are known and the variances are equal
  3. Only the means are unknown

# Gaussian Mixture Models

**Theorem (L., Améndola, Rodriguez [LAR21])**

*In all cases, Gaussian mixture models are algebraically identifiable using moment equations of lowest degree. Moreover, the mixed volume of each of set of equations is given below.*

| | Known mixing coefficients | Known mixing coefficients + equal variances | Unknown means |
|---|---|---|---|
| Moment equations | $m_1, \ldots, m_{2k}$ | $m_1, \ldots, m_{k+1}$ | $m_1, \ldots, m_k$ |
| Unknowns | $\mu_i, \sigma_i^2$ | $\mu_i, \sigma^2$ | $\mu_i$ |
| Mixed volume | $(2k-1)!!k!$ | $\frac{(k+1)!}{2}$ | $k!$ |
| Mixed volume tight | Yes for $k \leq 8$ | Yes for $k \leq 8$ | Yes |

| | Mixed Volume | Bezout Bound |
|---|---|---|
| Known mixing coefficients | $(2k-1)!!k!$ | $(2k)!$ |
| Known mixing coefficients + equal variances | $\frac{(k+1)!}{2}$ | $(k+1)!$ |
| Unknown means | $k!$ | $k!$ |

| | Mixed Volume | Bezout Bound |
|---|---|---|
| Known mixing coefficients | $(2k-1)!!k!$ | $(2k)!$ |
| Known mixing coefficients + equal variances | $\frac{(k+1)!}{2}$ | $(k+1)!$ |
| Unknown means | $k!$ | $k!$ |

- Our proofs of the mixed volume in the first two cases give a start system that tracks mixed volume number of paths

| | Mixed Volume | Bezout Bound |
|---|---|---|
| Known mixing coefficients | $(2k-1)!!k!$ | $(2k)!$ |
| Known mixing coefficients + equal variances | $\frac{(k+1)!}{2}$ | $(k+1)!$ |
| Unknown means | $k!$ | $k!$ |

- Our proofs of the mixed volume in the first two cases give a start system that tracks mixed volume number of paths
- In the final case if $\lambda_i = \frac{1}{k}$ and $\sigma_i^2$ are equal, there is a unique solution up to symmetry

# Table of Contents

- Consider a parameterized polynomial system $F_b$
  - i.e. $F_b(x) = b_1x^2 + b_2x - 1$

- Consider a parameterized polynomial system $F_b$
    - i.e. $F_b(x) = b_1x^2 + b_2x - 1$

- Monodromy works by moving $b$ in a loop while tracking $x$ along that loop

- Consider a parameterized polynomial system $F_b$
  - i.e. $F_b(x) = b_1x^2 + b_2x - 1$
- Monodromy works by moving $b$ in a loop while tracking $x$ along that loop
- After one loop, usually find a new solution

# Method 2
## Monodromy

- Consider a parameterized polynomial system $F_b$
  - i.e. $F_b(x) = b_1 x^2 + b_2 x - 1$
- Monodromy works by moving $b$ in a loop while tracking $x$ along that loop
- After one loop, usually find a new solution
- This process then is repeated until some stopping criterion is fulfilled

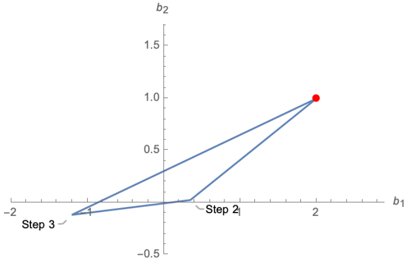
# Monodromy example

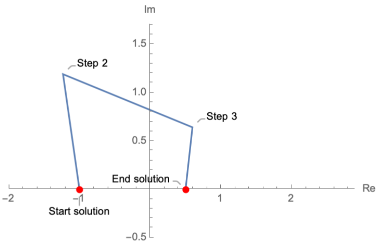
- Consider $F_b(x) = b_1 x^2 + b_2 x - 1$ where $B = \mathbb{C}^2$

# Monodromy example

- Consider $F_b(x) = b_1 x^2 + b_2 x - 1$ where $B = \mathbb{C}^2$
- For almost all $(b_1, b_2) \in \mathbb{C}^2$, $\{x \in \mathbb{C} \ : \ F_b(x) = 0\}$ has two elements

- Consider $F_b(x) = b_1x^2 + b_2x - 1$ where $B = \mathbb{C}^2$
- For almost all $(b_1, b_2) \in \mathbb{C}^2$, $\{x \in \mathbb{C} \ : \ F_b(x) = 0\}$ has two elements
- Consider $(2, 1) \in \mathbb{C}^2$ and $x = -1$, then $F_{(2,1)}(-1) = 0$



# Method 2
## Monodromy

- Monodromy methods work by taking one solution $\hat{x}$ to the system of equations $F_b$ and finding other elements via the monodromy action

# Method 2

## Monodromy

- Monodromy methods work by taking one solution $\hat{x}$ to the system of equations $F_b$ and finding other elements via the monodromy action
- This action is transitive if and only if the solution variety is irreducible

- Monodromy methods work by taking one solution $\hat{x}$ to the system of equations $F_b$ and finding other elements via the monodromy action
- This action is transitive if and only if the solution variety is irreducible
- **Major benefit:** Using monodromy we can solve up to symmetry [ALR21]

# Table of Contents

# Example 3

## Power flow equations

- Consider electric power network on $n$ nodes, $v_0, \ldots, v_{n-1}$

# Example 3

## Power flow equations

- Consider electric power network on $n$ nodes, $v_0, \ldots, v_{n-1}$
- At $v_0$ we have $x_0 = 1$, $y_0 = 0$

# Example 3

## Power flow equations

- Consider electric power network on $n$ nodes, $v_0, \ldots, v_{n-1}$
- At $v_0$ we have $x_0 = 1$, $y_0 = 0$
- At each node $v_i$, $i > 0$ we have the equation for voltage magnitude:

$$x_i^2 + y_i^2 = 1$$

# Example 3

## Power flow equations

- Consider electric power network on $n$ nodes, $v_0, \ldots, v_{n-1}$
- At $v_0$ we have $x_0 = 1$, $y_0 = 0$
- At each node $v_i$, $i > 0$ we have the equation for voltage magnitude:

$$x_i^2 + y_i^2 = 1$$

- The active power injections are described by:

$$\sum_{k=0}^{n-1} b_{ik}(x_i y_k - x_k y_i) = P_i$$

# Example 3

## Power flow equations

- Consider electric power network on $n$ nodes, $v_0, \ldots, v_{n-1}$
- At $v_0$ we have $x_0 = 1$, $y_0 = 0$
- At each node $v_i$, $i > 0$ we have the equation for voltage magnitude:

$$x_i^2 + y_i^2 = 1$$

- The active power injections are described by:

$$\sum_{k=0}^{n-1} b_{ik}(x_i y_k - x_k y_i) = P_i$$

- These are the **power flow equations**

# Method 2

## Monodromy

### Proposition (**L.**, Boston, Lesieutre 2020 [LBL20])

- The power flow equations with $P_i \neq 0$ has a solution variety that is irreducible.
- The power flow equations with zero active power injections has a solution variety that decomposes into $2^{n-1} + 1$ irreducible components.
  - $2^{n-1}$ consist of a single point of the form $(x_i, y_i) = (\pm 1, 0)$ (trivial).
  - The remaining component consists of all nontrivial solutions to the power flow equations.

## Proposition (**L.**, Boston, Lesieutre 2020 [LBL20])

- The power flow equations with $P_i \neq 0$ has a solution variety that is irreducible.
- The power flow equations with zero active power injections has a solution variety that decomposes into $2^{n-1} + 1$ irreducible components.
  - $2^{n-1}$ consist of a single point of the form $(x_i, y_i) = (\pm 1, 0)$ (trivial).
  - The remaining component consists of all nontrivial solutions to the power flow equations.

## Lemma (**L.**, Zachariah, Boston, Lesieutre [LZBL20])

- For the power flow equations with zero active power injections, solutions come in pairs.
- For the power flow equations with zero active power injections on bipartite graphs, solutions come in sets of four.
- For the power flow equations with nonzero active power injections on bipartite graphs, solutions come in pairs.

# Comparing Monodromy

| | $K_5$ | $K_6$ | $K_7$ | $K_8$ | $K_9$ | $K_{10}$ |
|---|---|---|---|---|---|---|
| time (s): monodromy | 0.05 | 0.37 | 1.97 | 16.39 | 65.33 | 357.926 |
| time (s): polyhedral | 0.37 | 2.53 | 17.10 | 112.43 | 609.49 | 2637.22 |
| time (s): total degree | 0.21 | 1.45 | 8.17 | 48.78 | 329.60 | 1510.01 |

Table: Numerical results to find all solutions for complete networks

| | $C_5$ | $C_6$ | $C_7$ | $C_8$ | $C_9$ | $C_{10}$ |
|---|---|---|---|---|---|---|
| time (s): monodromy | 0.13 | 0.158 | 1.10 | 1.46 | 2.48 | 2.60 |
| time (s): polyhedral | 2.7 | 3.03 | 5.37 | 14.8 | 56.36 | 211.24 |
| time (s): total degree | 2.11 | 3.40 | 9.76 | 31.91 | 200.41 | 862.50 |

Table: Numerical results to find all solutions for cyclic networks

# Conclusion

- Outlined methods for polynomial system solving
- Considered three applications

**Questions!?**

# References


[ALR21] Carlos Améndola, Julia Lindberg, and Jose Israel Rodriguez, *Solving parameterized polynomial systems with decomposable projections*, arXiv preprint (2021).

[Ber75] David N. Bernstein, *The number of roots of a system of equations*, Funkcional. Anal. i Priložen. **9** (1975), no. 3, 1–4. MR 0435072

[BT18] Paul Breiding and Sascha Timme, *Homotopycontinuation.jl: A package for homotopy continuation in Julia*, Mathematical Software – ICMS 2018, Springer International Publishing, 2018, pp. 458–465.

[HS95] Birkett Huber and Bernd Sturmfels, *A polyhedral method for solving sparse polynomial systems*, Math. Comp. **64** (1995), no. 212, 1541–1555. MR 1297471

[Kho78] Askold G. Khovanskii, *Newton polyhedra, and the genus of complete intersections*, Funktsional. Anal. i Prilozhen. **12** (1978), no. 1, 51–61. MR 487230

[Kou76] Anatoli G. Kouchnirenko, *Polyèdres de Newton et nombres de Milnor*, Invent. Math. **32** (1976), no. 1, 1–31. MR 419433

[KW14] Kei Kobayashi and Henry P. Wynn, *Computational algebraic methods in efficient estimation*, Geometric theory of information, Signals Commun. Technol., Springer, Cham, 2014, pp. 119–140. MR 3329739

[LAR21] Julia Lindberg, Carlos Améndola, and Jose Israel Rodriguez, *Estimating gaussian mixtures using sparse polynomial moment systems.*, arXiv preprint arXiv:2106.15675 (2021).

[LBL20] Julia Lindberg, Nigel Boston, and Bernard C. Lesieutre, *Exploiting symmetry in the power flow equations using monodromy*, ISSAC '20: Proceedings of the 45th International Symposium on Symbolic and Algebraic Computation (2020).

[LNRW21] Julia Lindberg, Nathan Nicholson, Jose Israel Rodriguez, and Zinan Wang, *The maximum likelihood degree of sparse polynomial systems*, arXiv preprint arXiv:2105.07449 (2021).

[LZBL20] Julia Lindberg, Alisha Zachariah, Nigel Boston, and Bernard C. Lesieutre, *The distribution of the number of real solutions to the power flow equations*, arXiv preprint (2020).