Enabling Efficient Analog Synthesis by Coupling Sparse Regression and Polynomial Optimization

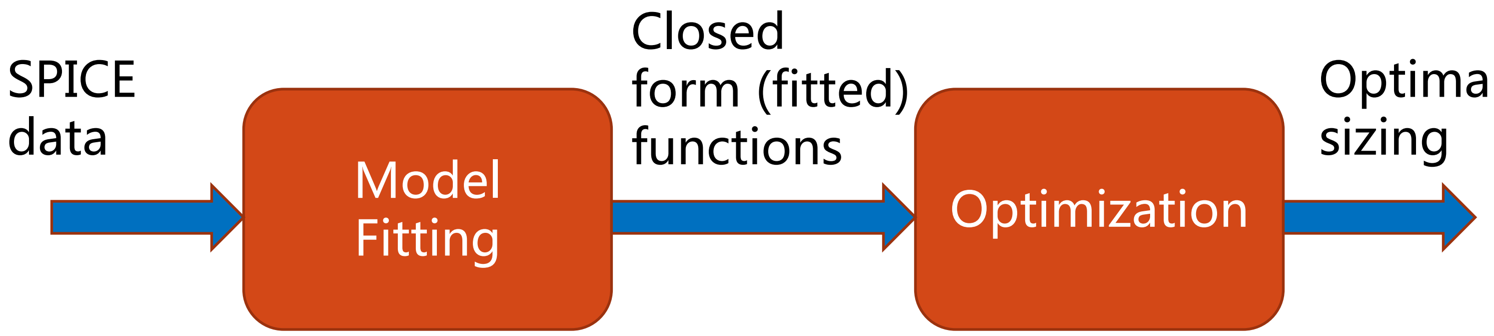
Ye Wang, Michael Orshansky, Constantine Caramanis
ECE Dept., The University of Texas at Austin

Analog Synthesis Problem

- This talk:
  - Transistor sizing for fixed topology
    - e.g. area, power, bandwidth, gain

- Optimization alternatives
  - Simulation-based [Daems2003]
    - e.g. Simulated annealing
  - Equation-based

Equation-based Approach



- 1. Choose function class. 2. Fit circuit behavior through regression. 3. Optimize

- Key issues:
  - Richer function class produces better fit
  - But can we efficiently optimize?

State-of-the-art Equation-based Methods

- Geometric programming (GP) [Boyd2001]
  - Class of functions: posynomials
  - Advantages:
    - Convex formulation
    - Efficient global optimal solutions
  - Limitations:
    - Limited function class: significant inaccuracies observed for common circuit functions

State-of-the-art Equation-based Methods

- Polynomial optimization (POP)
  - Class of functions: polynomials

  - Advantages:
    - Much richer class – allows non-linear and non-convex models [Lui2010]
    - Greatly improved model accuracy
    - Solvable *in principle* by recent advances in convex optimization, moment problems, and SDP relaxations [Lasserre2001]

SDP Formulation of POP

- POP formulation

$$\begin{aligned} \text{minimize}_x : \quad & f_0(x) \\ \text{subject to} : \quad & f_i(x) \geq 0, \quad i = 1, \ldots, p. \end{aligned}$$

- Non-linear and non-convex in x-variables
- Key: transformation to moment-variables produces linear Semidefinite Optimization Problem (SDP) – Convex

$$\begin{aligned} \text{minimize}_{\{m_\alpha\}} : \quad & \sum_\alpha p_\alpha m_\alpha \\ \text{subject to} : \quad & M_r(m) \succeq 0, \\ & M_{r-r_i}(f_i m) \succeq 0. \end{aligned}$$

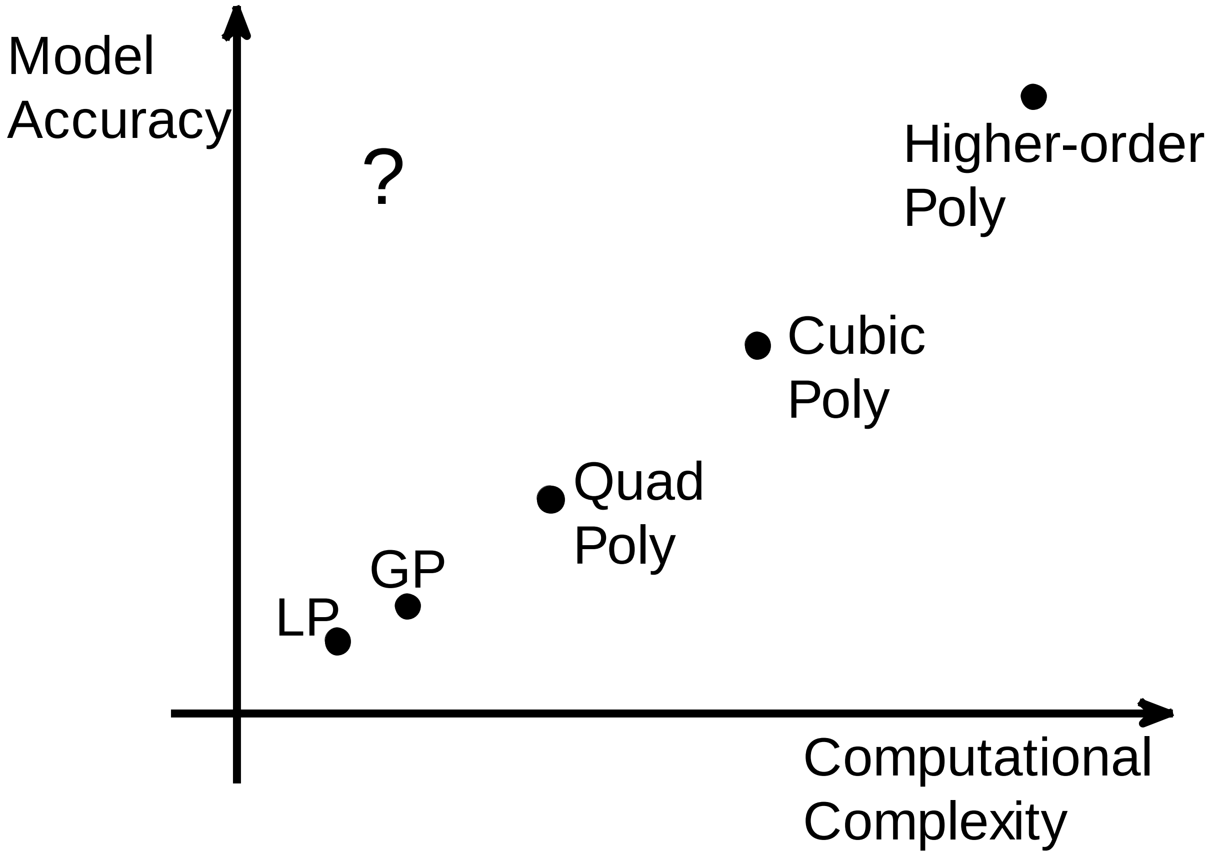
- Convex, but...

SDP formulation of POP

- Practically, restricted by SDP size: grows with number of terms and this is generally exponential in degree of the polynomial

- Thus: in practice can only use quadratic polynomials even for mid-size analog sizing problems

Accuracy-complexity Space



High Accuracy & Low Complexity

- High degree general polynomials:
  - Good fitting properties
  - But exponentially many terms hence exponentially hard to optimize

- Simply increasing polynomial order makes the problem intractable

Our Key Observation

- In fitting:
  - High degree polynomials with few non-zero terms, ALSO very good at fitting (better than low-degree and dense)
- In optimization:
  - Effort needed for solving high-degree polynomials with *certain sparsity patterns*, SIMILAR to that for low-degree polynomials

Our Contribution

- Main idea
  - Use high degree SPARSE polynomials with the particular pattern for later optimization
  - We call it coupled sparse fitting and optimization (C-SFO)

- Remaining Question: How to do the fitting step?
  - NP-hard as it is a combinatorial problem

Sparse Polynomial Regression

- The Sparsity Pattern: correlative-sparsity-pattern (CSP) matrix R [Kim2005]
  - $R(i,j) = 1$ when $x_i x_j$ appear together
    - A simple 4-variable example:

$$f(x) = x_1 + x_2^2 + x_3 + x_4^3 + x_1 x_2 + x_1^2 x_2 + x_2^2 x_4 \rightarrow \quad R = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix}$$

  - A measure for sparsity
    - Sparse CSP matrix → sparse POP → efficient to solve
- Our objective: encourage and drive CSP matrix to be sparse

Overlapping Group-Lasso Formulation

- Fitting polynomials enforcing sparse CSP matrix is still a combinatorial problem
- Lasso as the tool for sparse regression
  - Serves as a convex approximation to the original NP-hard problem
- In our case, general lasso is not enough
  - Group sparsity → Overlapping group-lasso
- ℓ_1/ℓ_2 penalty for overlapping group-lasso regression problem [Lin2006]

$$\min_{C \in \mathbb{R}^l} \frac{1}{2} \|Y - XC\|^2 + \lambda \Omega^G_{overlap}(C).$$

Experimental Results

- Use 6^{th} degree sparse polynomial
  - Better fitting results than dense quad polynomial
  - Similar level of effort to solve as dense quad polynomial
- Improved model accuracy
  - 10X compared to monomial fitting (GP)
  - 3X compared to quad poly fitting (POP)

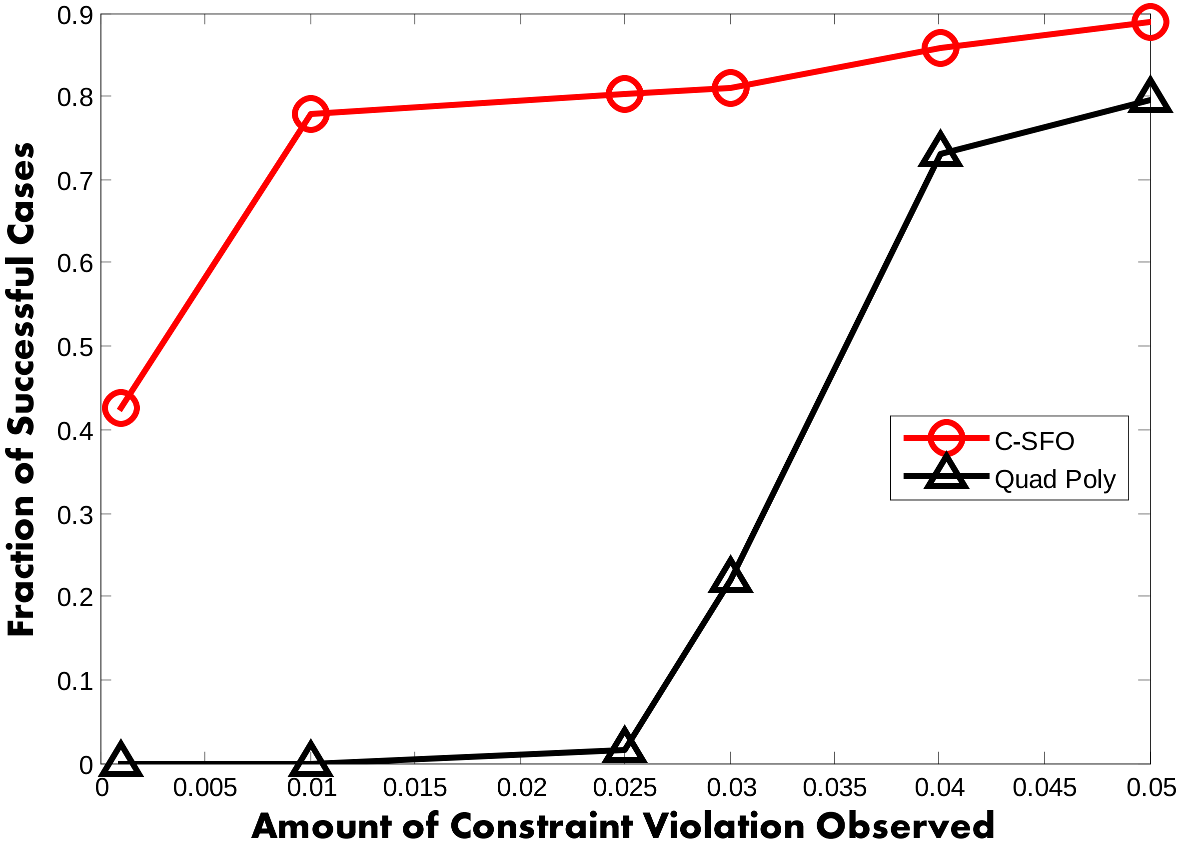
Better Optimization Results

- Optimizing a two-stage amplifier

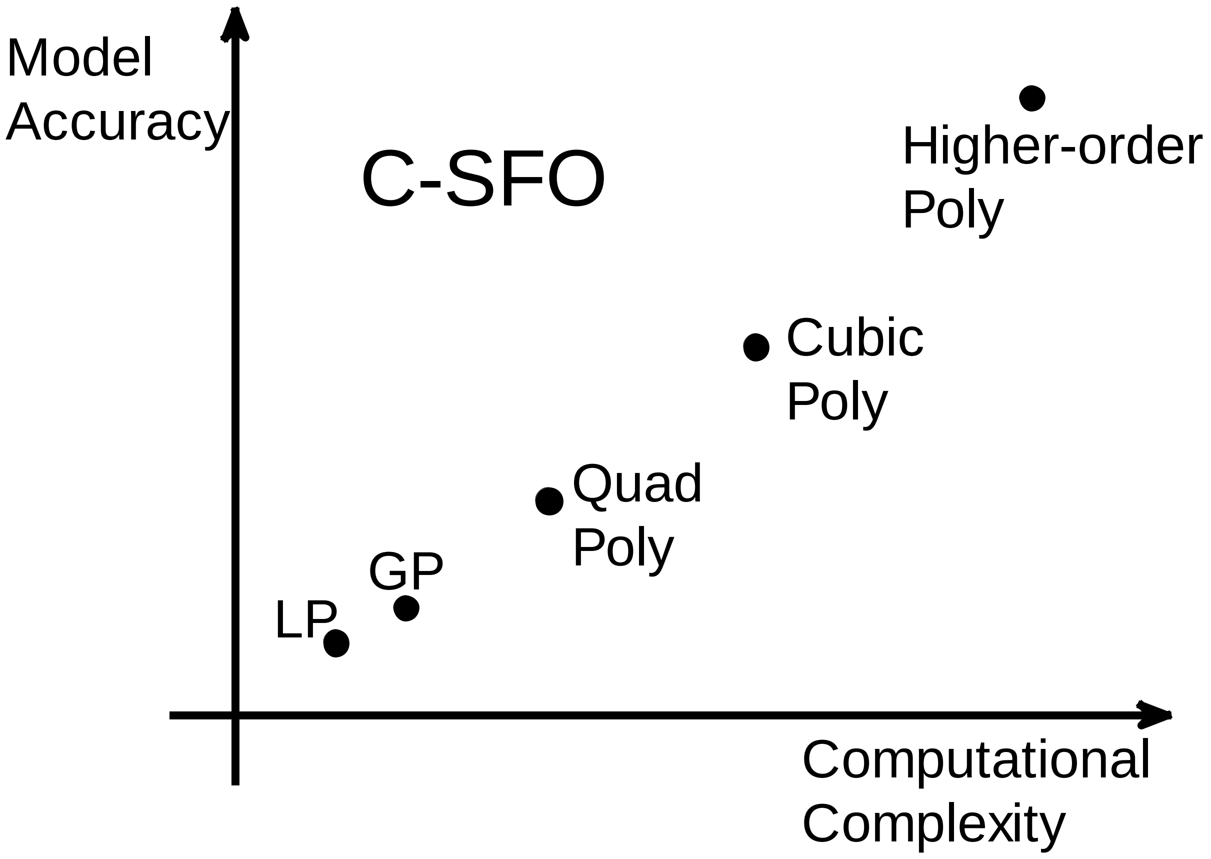
			C-SFO		Quad Poly		GP	
Case #	Metric	Spec	Model	SPICE	Model	SPICE	Model	SPICE
Case 1	Gain (10^4)	Max	1.69	1.68	1.74	1.68	1.95	0.78
	UGB (MHz)	≥10	10.2	10.5	9.98	9.96	10	9.76
	PM (°)	≥60	60.6	60.5	60	60	60	58.2
Case 2	Gain (10^4)	≥1.5	1.55	1.53	1.5	1.36	1.5	0.48
	UGB (MHz)	Max	12.6	12.6	14.8	14.8	18.4	17.8
	PM (°)	≥60	61.3	60.0	60	60.7	60	59.7
Case 3	Gain (10^4)	≥1.5	1.54	1.52	1.49	1.37	1.5	0.57
	UGB (MHz)	Max	18.9	19.0	14.9	14.8	25.6	25.6
	PM (°)	≥48	47.8	46	48	47.9	48	48.1

Better Optimization Results

- Quantitative analysis by varying design specifications



Summary


Model
Accuracy
C-SFO
Higher-order
Poly
Cubic
Poly
Quad
Poly
GP
LP
Computational
Complexity