# Choosing a variable ordering for truth-table invariant cylindrical algebraic decomposition by incremental triangular decomposition

**James Davenport: The University of Bath**

Joint work with: Russell Bradford, Matthew England and David Wilson

ICMS: 9 August 2014

Supported by EPSRC Grant EP/J003247/1.

# Outline

# Cylindrical algebraic decomposition

A Cylindrical Algebraic Decomposition (CAD) is a partition of $\mathbb{R}^n$ into cells arranged cylindrically (meaning their projections are either equal or disjoint) such that each cell is defined by a semi-algebraic set.

Defined by Collins who gave an algorithm (projection/lifting) to produce a sign-invariant CAD for a set of polynomials, meaning each polynomial had constant sign on each cell. In some sense, makes the induced geometry of $\mathbb{R}^n$ explicit

Originally motivated for use in quantifier elimination (variable ordering partially determined by quantifiers).
Have also been applied directly on problems as diverse as algebraic simplification and (at least theoretically) robot motion planning (variable ordering free).

# Variable ordering matters

**Theory** [BD07] has an example with $O(1)$ cells in one order, and $2^{2^{n/3+O(1)}}$ in another.

**PL CAD** For $\forall x (px^2 + qx + r + x^4 \geq 0)$ [DSS04] quotes 0.54 seconds to CAD for the cheapest ordering, 83.39 for the most expensive to terminate, and 2 orderings (out of 6) didn't in 600 seconds.

**Also** for the collision problem, 2/3 of orderings failed to terminate, but the cheapest took 48 seconds.

# How to choose variable ordering?

Ideas for projection/lifting

**Brown** (depends on input $P_n$ only)

1. First eliminate the variable of least degree
2. Tiebreak by $\max_{f \in P_n} \mathrm{tdeg}$(monomial in $f$ containing $v$)
3. Tiebreak by number of occurrences

**and** repeat for second variable etc.

**sotd** Compute all $P_i$ for all orderings; choose ordering for which $\sum_{m \in p \in \bigcup P_i} \mathrm{tdeg}(m)$ is minimal

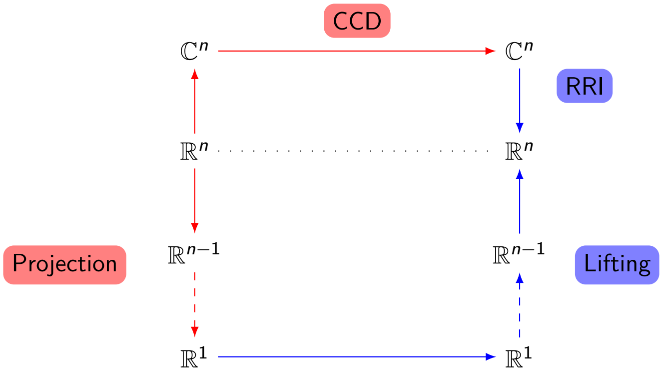
**greedy** Compute $P_{n-1}$ for all choices of first variable: choose variable for which $\sum_{m \in p \in P_{n-1}} \mathrm{tdeg}(m)$ is minimal

**ndrr** Compute $P_1$ for all orderings; choose ordering for which $\mathbb{R}^1$ is minimally divided

**ML** Use machine learning to decide which of the above to use [HEW$^+$14]

# An alternative approach [CMMXY09]

Proceed via the complex numbers,




Do a complex cylindrical decomposition via Regular Chains
Can be combined with truth table ideas [BCD+14a]

# Ordering for Regular Chains

Triangular (depends on input $P_n$ only): implemented as `SuggestVariableOrder`

1. First eliminate the variable of least degree
2. Tiebreak by $\max_{f \in P_n} \mathrm{tdeg}(\mathrm{lcoeff}(f, v))$
3. Tiebreak by $\sum_{f \in P_n} \deg_v(f)$

and repeat for second variable etc.

Brown as before

sotd as before

greedy as before

ndrr as before

Why are we suggesting sotd/greedy/ndrr based on projection, when we're not doing projection?

# Regular Chains versus Projection

- Projection produces a **global** set of polynomials
- Regular Chains does case discussion

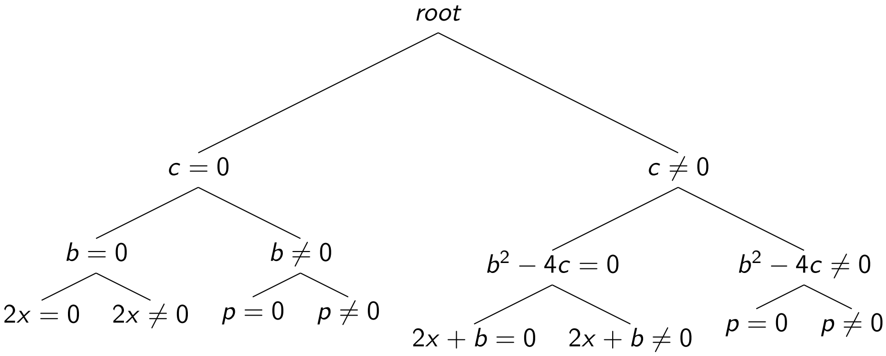



Figure: Complete complex cylindrical tree for the general monic quadratic equation, $p := x^2 + bx + c$, under variable ordering $c \prec b \prec x$.

# Results of previous heuristics

Table: Comparing the savings (as a percentage of the problem average) in cells (C) and net timings (NT) from various heuristics.

| Heuristic | 22 | | 20 | | 10 | | 00 | |
|---|---|---|---|---|---|---|---|---|
| | C | NT | C | NT | C | NT | C | NT |
| **Triangular** | 32.6 | 33.9 | 47.9 | 46.8 | 47.7 | 47.2 | 56.0 | 58.8 |
| **Brown** | 37.6 | 39.1 | 45.0 | 44.3 | 51.6 | 50.9 | 61.9 | 64.5 |
| **Sotd** | 36.7 | 23.9 | 42.8 | 39.5 | 56.3 | 53.9 | 59.9 | 61.8 |
| **Ndrr** | 40.1 | 21.2 | 35.7 | 34.4 | 54.8 | 51.3 | 54.0 | 54.3 |
| **Sotd/NDRR** | 37.0 | 24.3 | 42.5 | 39.6 | 56.0 | 53.5 | 60.4 | 62.5 |
| **NDRR/Sotd** | 41.3 | 22.6 | 38.7 | 36.0 | 57.1 | 51.7 | 58.4 | 60.2 |
| **Greedy** | 35.0 | 32.7 | 39.8 | 38.9 | 52.3 | 52.1 | 52.5 | 55.9 |

Brown is nearly always best

# Truth-Table Invariant CAD [BDE+13, BDE+14]

Assume our formula is in disjunctive normal form.
If one of the clauses is $f = 0 \wedge g > 0$, then we do not care about $g$ *except when* $f = 0$. Hence, for projection CAD, $g$ need only figure in $\mathrm{res}_x(f, g)$ and not in any other resultant/discriminant. This makes the final projection set significantly smaller
The same logic can apply to regular chains CAD [BCD+14b].
Hence we ought to measure sotd etc., not on the original projection, but on the TTI projection.
(Note that TTI has many other choices: [EBC+14].)

# Results of tailored heuristics

Table: Comparing the savings (as a percentage of the problem average): $(f_1\sigma 0 \wedge f_2\sigma 0) \vee (f_3\sigma 0 \wedge f_4\sigma 0)$: numbers are how many $\sigma$ are $=$. 12 and 11 missing from slide

| **Heur** | **22** | | **20** | | **10** | | **00** | | **All** | |
|---|---|---|---|---|---|---|---|---|---|---|
| | C | NT | C | NT | C | NT | C | NT | C | NT |
| **Tr** | 32.6 | 33.9 | 47.9 | 46.8 | 47.7 | 47.2 | 56.0 | 58.8 | 43.0 | 43.6 |
| **Br** | 37.6 | 39.1 | 45.0 | 44.3 | 51.6 | 50.9 | 61.9 | 64.5 | 46.8 | 47.5 |
| **S-TTI** | 42.7 | 40.4 | 48.4 | 48.1 | 61.2 | 60.2 | 59.9 | 61.7 | 52.2 | 50.3 |
| **N-TTI** | 48.5 | 37.1 | 47.8 | 46.9 | 59.0 | 55.3 | 54.0 | 54.3 | 50.7 | 46.0 |
| **GS-TTI** | 46.4 | 47.2 | 49.3 | 50.2 | 56.7 | 57.5 | 52.8 | 55.9 | 51.1 | 51.6 |

Generally better, except for 00 (non-TTI) case.

# Can we do even better?

Remember that Brown did pretty well initially, and is the best single heuristic for PL-CAD [HEW+14].
Consider the following set of polynomials:

- the discriminants, leading coefficients and cross-resultants of the polynomials forming the first constraint in each QFF;
- if a QFF has no EC then also the (other) discriminants, leading coefficients and cross resultants of all polynomials defining constraints there;
- if a QFF has more than one EC then also the resultant of the polynomial defining the first with that of the second.

This set does not contain all polynomials computed by `RC-TTICAD`, but those which are considered in their own right rather than modulo others.

# Regarding these as the “drivers” of TTICAD

NH: we define a new heuristic to pick an orderings in two stages: First variables are ordered according to maximum degree of the polynomials forming the input (as with Triangular and Brown). Then ties are broke by calculating the set of polynomials above for each unallocated variable and ordering according to sum of degree (in that variable).

NH+: we can use the degree of the omitted discriminants, resultants and leading coefficients as a third tie-break.

Table: Comparing the savings (as a percentage of the problem average): $(f_1\sigma 0 \wedge f_2\sigma 0) \vee (f_3\sigma 0 \wedge f_4\sigma 0)$: numbers are how many $\sigma$ are $=$. 12 and 11 missing from slide

| **Heur** | **22** | | **20** | | **10** | | **00** | | **All** | |
|---|---|---|---|---|---|---|---|---|---|---|
| | C | NT | C | NT | C | NT | C | NT | C | NT |
| **Tr** | 32.6 | 33.9 | 47.9 | 46.8 | 47.7 | 47.2 | 56.0 | 58.8 | 43.0 | 43.6 |
| **Br** | 37.6 | 39.1 | 45.0 | 44.3 | 51.6 | 50.9 | 61.9 | 64.5 | 46.8 | 47.5 |
| **S-TTI** | 42.7 | 40.4 | 48.4 | 48.1 | 61.2 | 60.2 | 59.9 | 61.7 | 52.2 | 50.3 |
| **N-TTI** | 48.5 | 37.1 | 47.8 | 46.9 | 59.0 | 55.3 | 54.0 | 54.3 | 50.7 | 46.0 |
| **GS-TTI** | 46.4 | 47.2 | 49.3 | 50.2 | 56.7 | 57.5 | 52.8 | 55.9 | 51.1 | 51.6 |
| **NH** | 45.9 | 45.5 | 48.2 | 47.6 | 56.4 | 52.4 | 67.0 | 68.5 | 51.7 | 51.3 |
| **NH$^+$** | 46.2 | 45.9 | 49.3 | 49.5 | 55.9 | 52.0 | 67.0 | 68.5 | 52.0 | 51.7 |

# Conclusions

- Variable ordering matters, with the "right" ordering being twice or more as good as "average".

And "bad" is really bad — see paper for statistics: on a multi-core, one could race several orderings.

- There is no "one size fits all" [HEW$^+$14].
- If you're doing TTICAD, your heuristics should match.

R. Bradford, C. Chen, J.H. Davenport, M. England, M. Moreno Maza, and D. Wilson.
Truth table invariant cylindrical algebraic decomposition by regular chains.
*Proc. CASC '14 (to appear). Preprint available at http://opus.bath.ac.uk/38344/*, 2014.

R.J. Bradford, C. Chen, J.H. Davenport, M. England, M. Moreno Maza, and D.J. Wilson.
Truth Table Invariant Cylindrical Algebraic Decomposition by Regular Chains.
In *Proceedings CASC 2014*, pages 44–58, 2014.

C.W. Brown and J.H. Davenport.
The Complexity of Quantifier Elimination and Cylindrical Algebraic Decomposition.
In C.W. Brown, editor, *Proceedings ISSAC 2007*, pages 54–60, 2007.

R.J. Bradford, J.H. Davenport, M. England, S. McCallum, and D.J. Wilson.
Cylindrical Algebraic Decompositions for Boolean Combinations.
In *Proceedings ISSAC 2013*, pages 125–132, 2013.

R.J. Bradford, J.H. Davenport, M. England, S. McCallum, and D.J. Wilson.
Truth Table Invariant Cylindrical Algebraic Decomposition.
http://arxiv.org/abs/1401.0645, 2014.

C. Chen, M. Moreno Maza, B. Xia, and L. Yang.
Computing Cylindrical Algebraic Decomposition via Triangular Decomposition.
In J. May, editor, *Proceedings ISSAC 2009*, pages 95–102, 2009.

A. Dolzmann, A. Seidl, and Th. Sturm.
Efficient Projection Orders for CAD.
In J. Gutierrez, editor, *Proceedings ISSAC 2004*, pages 111–118, 2004.

M. England, R. Bradford, C. Chen, J.H. Davenport, M.M. Maza, and D.J. Wilson.
Problem formulation for truth-table invariant cylindrical algebraic decomposition by incremental triangular decomposition.

In S.M. Watt *et al.*, editor, *Proceedings CICM 2014*, pages 45–60, 2014.

Z. Huang, M. England, D. Wilson, J.H. Davenport, L.C. Paulson, and J. Bridge.
Applying machine learning to the problem of choosing a heuristic to select the variable ordering for cylindrical algebraic decomposition.
In S.M.Watt *et al.*, editor, *Proceedings CICM 2014*, pages 92–107, 2014.