Software for Symbolic-Numeric Solutions of Polynomial Systems

Jan Verschelde

Department of Math, Stat & CS
University of Illinois at Chicago
Chicago, IL 60607-7045, USA

Email: `jan@math.uic.edu`

URL: `http://www.math.uic.edu/~jan`

Real Number Complexity workshop, FoCM'05, 7-9 July

joint work with Andrew Sommese and Charles Wampler;
and Anton Leykin and Ailing Zhao

A Two Line Summary of the Talk

Software – key new features in PHCpack v2.3:

(1) **intersection of positive dimensional solution sets**;

(2) **accurate computation of isolated singular solutions**.

Goal: explain symbolic-numeric aspects of the new algorithms.

This material is based upon work supported by the National Science Foundation under Grant No. 0134611 and Grant No. 0410036.
Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

Related Work: Geometric Resolutions

M. Giusti and J. Heintz: **La détermination de la dimension et des points isolées d'une variété algébrique peuvent s'effectuer en temps polynomial**. In *Computational Algebraic Geometry and Commutative Algebra, Cortona 1991*, edited by D. Eisenbud and L. Robbiano, pages 216–256, Cambridge UP, 1993.

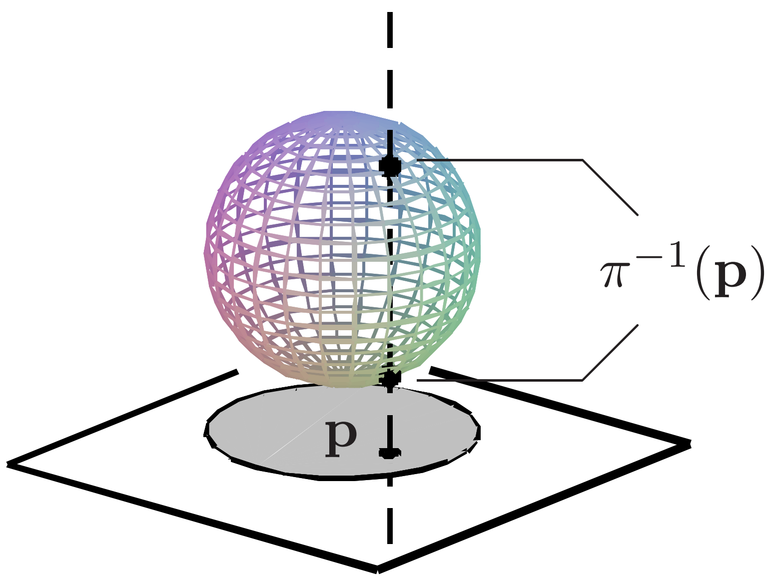
M. Giusti and J. Heintz: **Kronecker's smart, little black boxes**. In *Foundations of Computational Mathematics*, edited by R.A. DeVore, A. Iserles and E. Süli, pages 69–104, Cambridge UP, 2001.

M. Giusti, G. Lecerf, and B. Salvy: **A Gröbner free alternative for polynomial system solving**. *J. Complexity* 17(1): 154–211, 2001.

G. Lecerf: **Computing the equidimensional decomposition of an algebraic closed set by means of lifting fibers**. *J. Complexity* 19(4): 564–596, 2003.

Representing Pure Dimensional Solution Sets

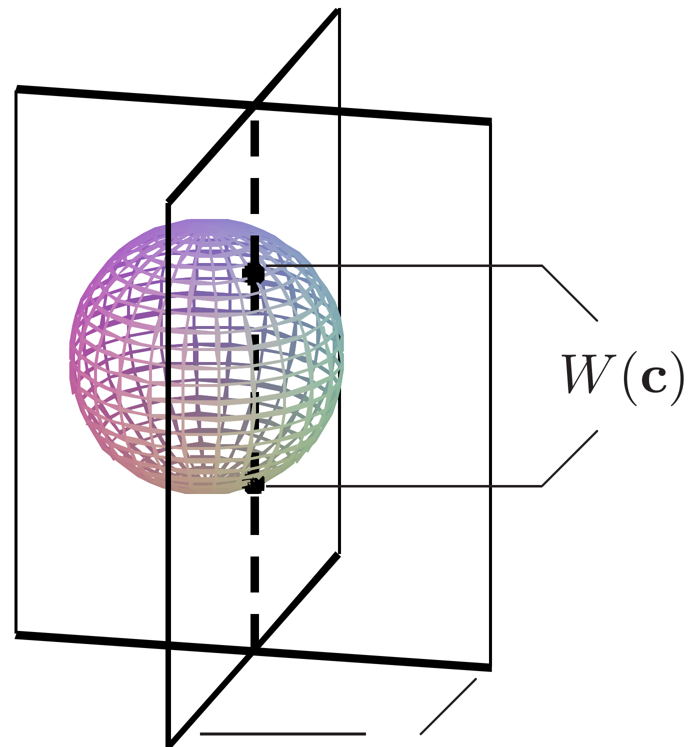
lifting fiber



sphere has 2 degrees of freedom:

choose $\mathbf{p} = (x, y)$, $z \in \pi^{-1}(\mathbf{p})$

witness set



Generic Points on a Pure Dimensional Solution Set V	
SYMBOLIC: lifting fiber $\pi^{-1}(\mathbf{p})$	**NUMERIC**: witness set $W(\mathbf{c})$
computational field k: numbers in $\mathbb{Q}$ (or in a finite extension) field operations done ***symbolically***	***numeric*** field $\mathbb{C}$: floating point complex numbers with machine arithmetic
With a ***symbolic*** coordinate change we bring V to Noether position: replace $\mathbf{x}$ by $M\mathbf{y}$, $M \in k^{n \times n}$	We slice V ***numerically*** with some randomly chosen hyperplanes: $A\mathbf{x} = \mathbf{c}$, $A \in \mathbb{C}^{r \times n}$, $\mathbf{c} \in \mathbb{C}^r$, rank$(A) = r$
choose M for coordinate change	**choose** A for slicing hyperplanes
$\dim V = r$: specialize r free variables	$\dim V = r$: cut with r hyperplanes
$\pi^{-1}(\mathbf{p}) = \{ \mathbf{y} \in \mathbb{C}^n \mid f(\mathbf{y}) = \mathbf{0}$ and $y_1 = p_1, \ldots, y_r = p_r \}$	$W(\mathbf{c}) = \{ \mathbf{x} \in \mathbb{C}^n \mid$ $f(\mathbf{x}) = \mathbf{0}$ and $A\mathbf{x} = \mathbf{c} \}$
choice of values $\mathbf{p} = (p_1, p_2, \ldots, p_r)$ for free variables $(y_1, y_2, \ldots, y_r)$ such that the fiber $\pi^{-1}(\mathbf{p})$ is finite	**choice** of r constants $\mathbf{c} = (c_1, c_2, \ldots, c_r)$ so that $\begin{cases} f(\mathbf{x}) = \mathbf{0} \\ A\mathbf{x} = \mathbf{c} \end{cases}$ has isolated solutions
for almost all $\mathbf{p} \in k^r$: $\pi^{-1}(\mathbf{p})$ consists of $\deg V$ smooth points	**for almost all** $\mathbf{c} \in \mathbb{C}^r$: $W(\mathbf{c})$ consists of $\deg V$ smooth points
where ***for almost all*** means except for a proper algebraic subset of bad choices	

Homotopy Membership Test

Does the point $\mathbf{p}$ *belong to a component* V *of* $f^{-1}(\mathbf{0})$*?*

Given: a point in space $\mathbf{p} \in \mathbb{C}^N$; a system $f(\mathbf{x}) = \mathbf{0}$;

and a witness set W, $W = (Z, L)$:

for all $\mathbf{w} \in Z : f(\mathbf{w}) = \mathbf{0}$ and $L(\mathbf{w}) = \mathbf{0}$.

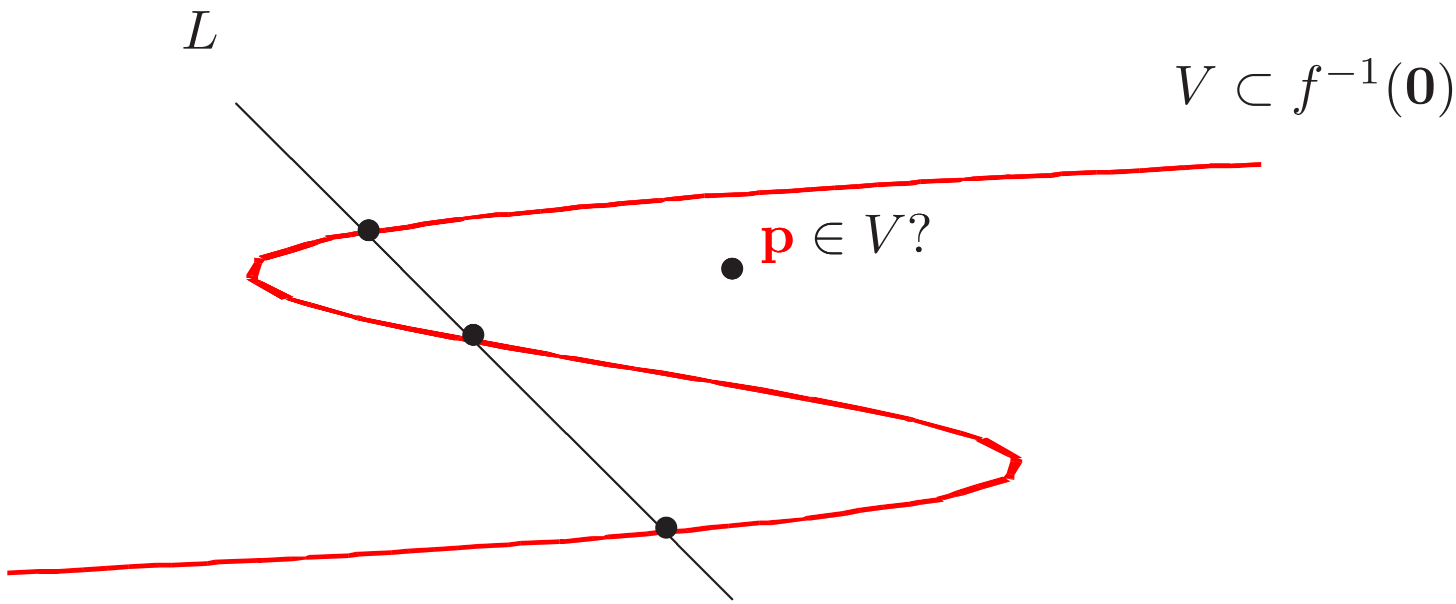
1. Let $L_{\mathbf{p}}$ be a set of hyperplanes through $\mathbf{p}$, and define

$$H(\mathbf{x}, t) = \left\{ \begin{array}{c} f(\mathbf{x}) = \mathbf{0} \\ L_{\mathbf{p}}(\mathbf{x})t + L(\mathbf{x})(1-t) = \mathbf{0} \end{array} \right.$$

2. Trace all paths starting at $\mathbf{w} \in Z$, for t from 0 to 1.

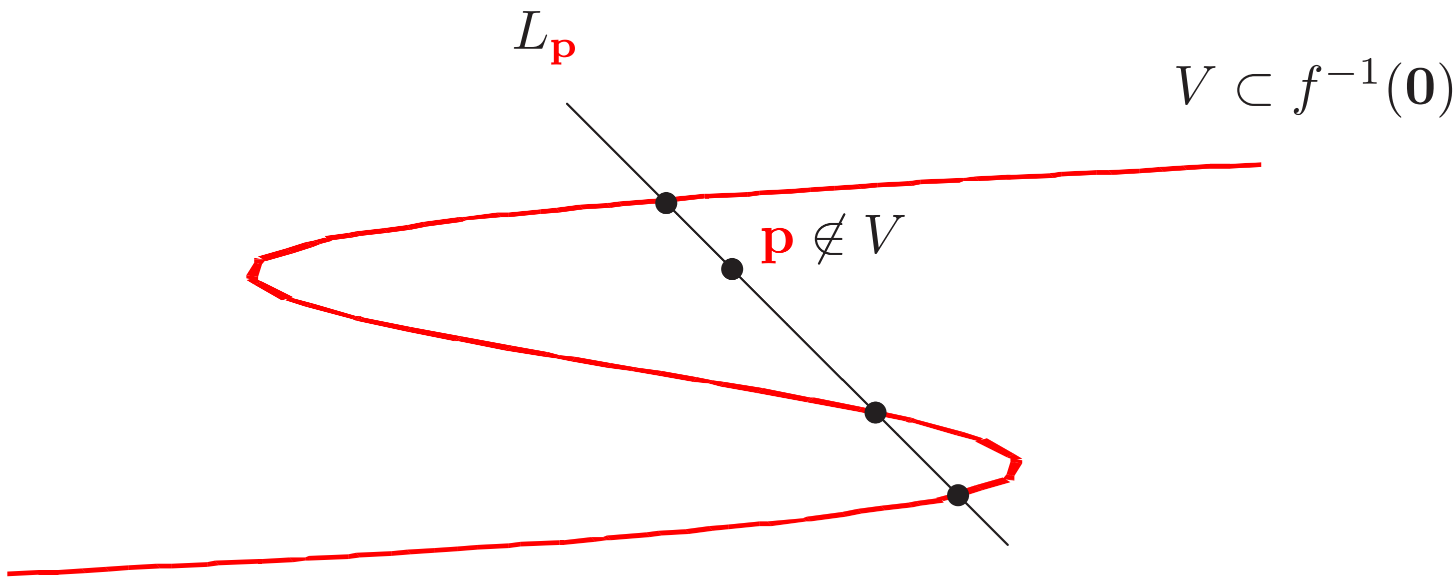
3. The test $(\mathbf{p}, 1) \in H^{-1}(\mathbf{0})$? answers the question above.

Homotopy Membership Test – an example



V is represented by a witness set $V \cap L$.

Homotopy Membership Test – an example



new witness set $V \cap L_{\mathbf{p}}$, $\mathbf{p} \notin V \cap L_{\mathbf{p}}$

$$H(\mathbf{x}, t) = \begin{cases} f(\mathbf{x}) = \mathbf{0} \\ L_{\mathbf{p}}(\mathbf{x})t + L(\mathbf{x})(1-t) = \mathbf{0} \end{cases}$$

Diagonal Homotopies: Problem Statement

Input: two irreducible components A and B, given by polynomial systems f_A and f_B (possibly identical), random hyperplanes L_A and L_B, and the solutions to

$$\underbrace{\begin{array}{c} \left\{ \begin{array}{l} f_A(\mathbf{x}) = \mathbf{0} \\ L_A(\mathbf{x}) = \mathbf{0} \end{array} \right. \\ \#L_A = \dim(A) = a \\ \{ \alpha_1, \alpha_2, \ldots, \alpha_{\deg A} \} \\ \deg A \text{ generic points} \end{array}}_{\text{a } \textit{witness set} \text{ for } A} \quad \text{and} \quad \underbrace{\begin{array}{c} \left\{ \begin{array}{l} f_B(\mathbf{x}) = \mathbf{0} \\ L_B(\mathbf{x}) = \mathbf{0} \end{array} \right. \\ \#L_B = \dim(B) = b \\ \{ \beta_1, \beta_2, \ldots, \beta_{\deg B} \} \\ \deg B \text{ generic points} \end{array}}_{\text{a } \textit{witness set} \text{ for } B}$$

Output: witness sets for all pure dimensional components of $A \cap B$.

Why new homotopies are needed

stacking two (possibly identical) systems is not sufficient!

For example: find $A \cap B$,

where A is line $x_2 = 0$, solution of $f(x_1, x_2) = x_1 x_2 = 0$,

and B is line $x_1 - x_2 = 0$, solution of $g(x_1, x_2) = x_1(x_1 - x_2) = 0$.

Problem: $A \cap B = (0, 0)$ does not occur as an irreducible

solution component of $\left\{ \begin{array}{l} f(x_1, x_2) = x_1 x_2 = 0 \\ g(x_1, x_2) = x_1(x_1 - x_2) = 0. \end{array} \right.$

Diagonal Homotopies: a very special case

Assume A and B are complete intersections, $\dim(A \cap B) = 0$.

The ***diagonal homotopy***

$$h(\mathbf{x}, \mathbf{y}, t) = \left\{ \begin{array}{c} f_A(\mathbf{x}) = \mathbf{0} \\ f_B(\mathbf{y}) = \mathbf{0} \\ (1-t)\begin{pmatrix} L_A(\mathbf{x}) \\ L_B(\mathbf{y}) \end{pmatrix} + t(\mathbf{x} - \mathbf{y}) = \mathbf{0} \end{array} \right.$$

starts at $t = 0$ at the $\deg A \times \deg B$ solutions in $A \times B \in \mathbb{C}^{n+n}$.

At $t = 1$, we find solutions at the diagonal $\mathbf{x} = \mathbf{y}$, in $A \cap B$.

Witness $V(f_1, \ldots, f_k)$ | Witness $V(f_{k+1})$

In:	W_1^k	W_2^k	$\cdots$	W_j^k	$\cdots$	W_k^k	X^{k+1}

one step in the equation-by-equation solver:

In: witness sets for the first k equations;

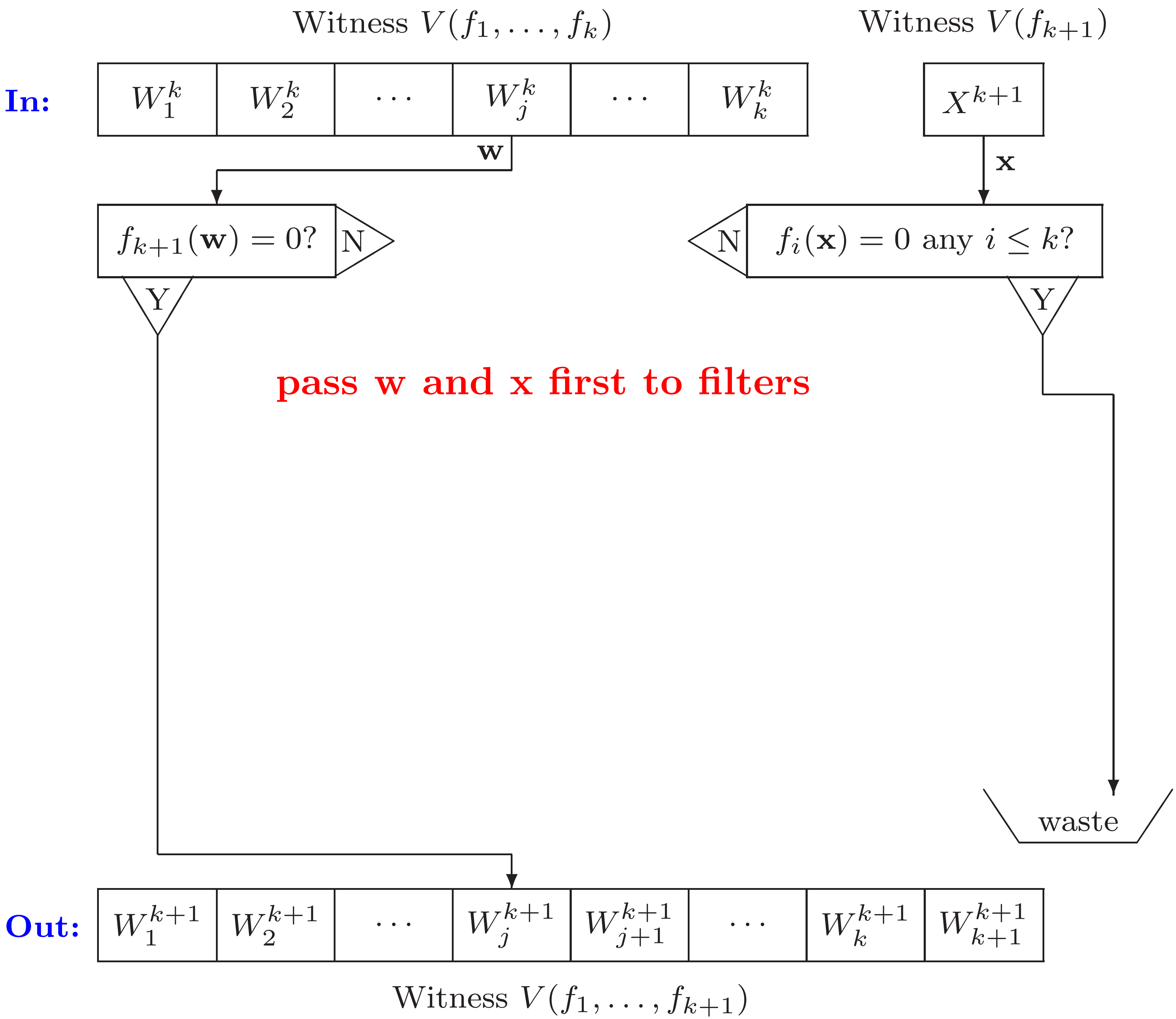
witness set for the $(k+1)$-th equation.

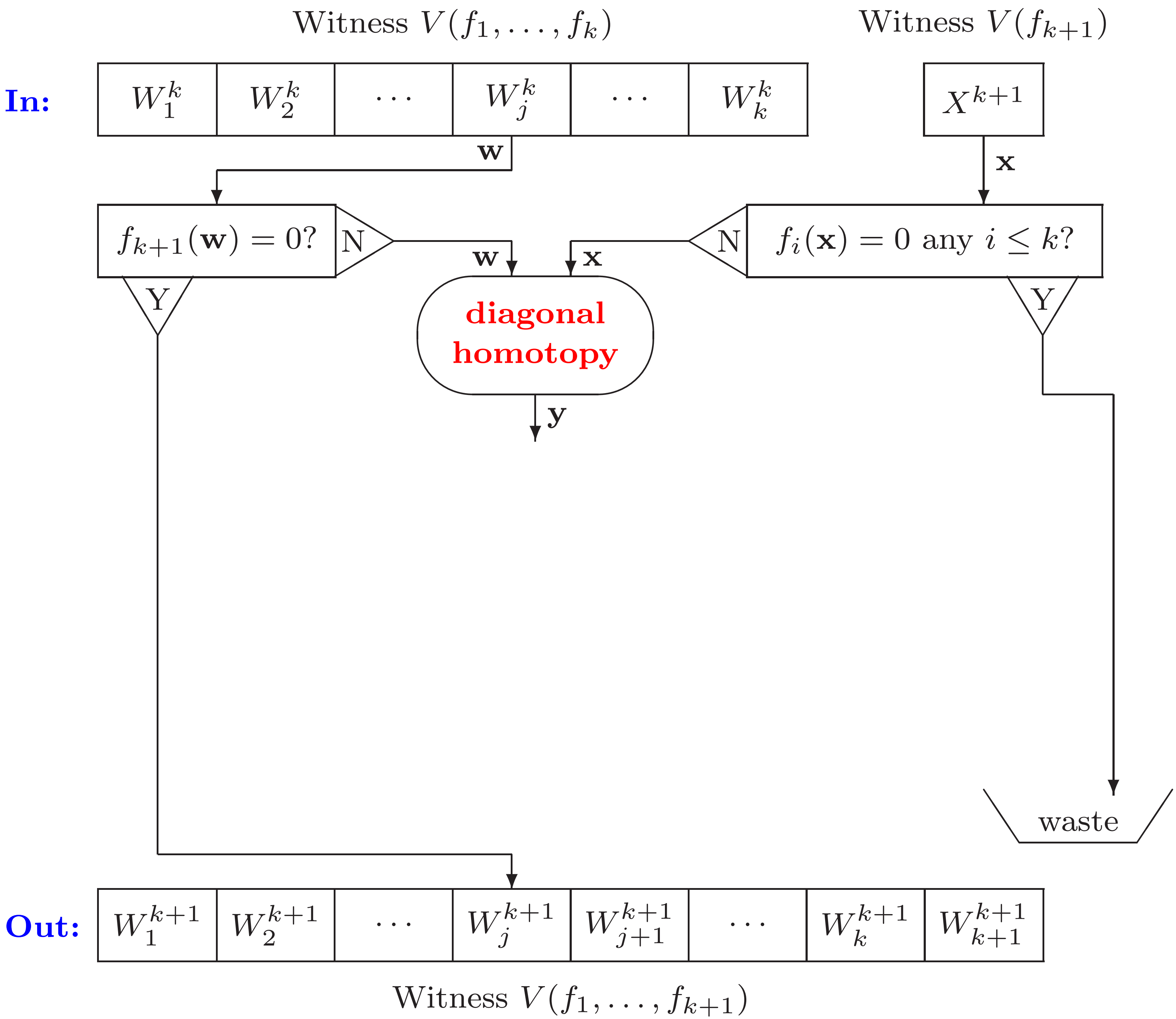
Out: witness sets for the first $k+1$ equations.

Notation: W_j^k, j = codimension of witness set

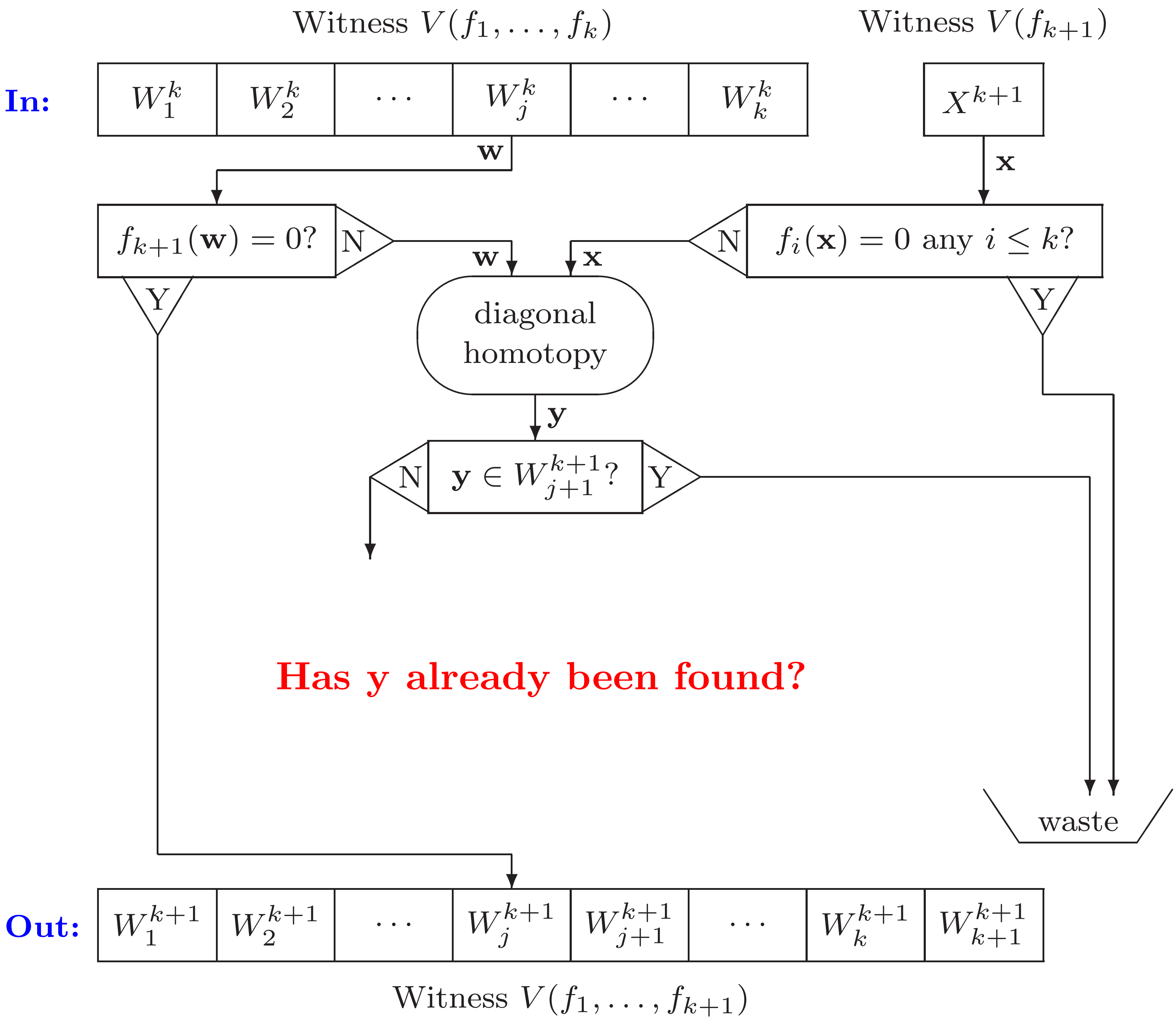
Out:	W_1^{k+1}	W_2^{k+1}	$\cdots$	W_j^{k+1}	W_{j+1}^{k+1}	$\cdots$	W_k^{k+1}	W_{k+1}^{k+1}

Witness $V(f_1, \ldots, f_{k+1})$

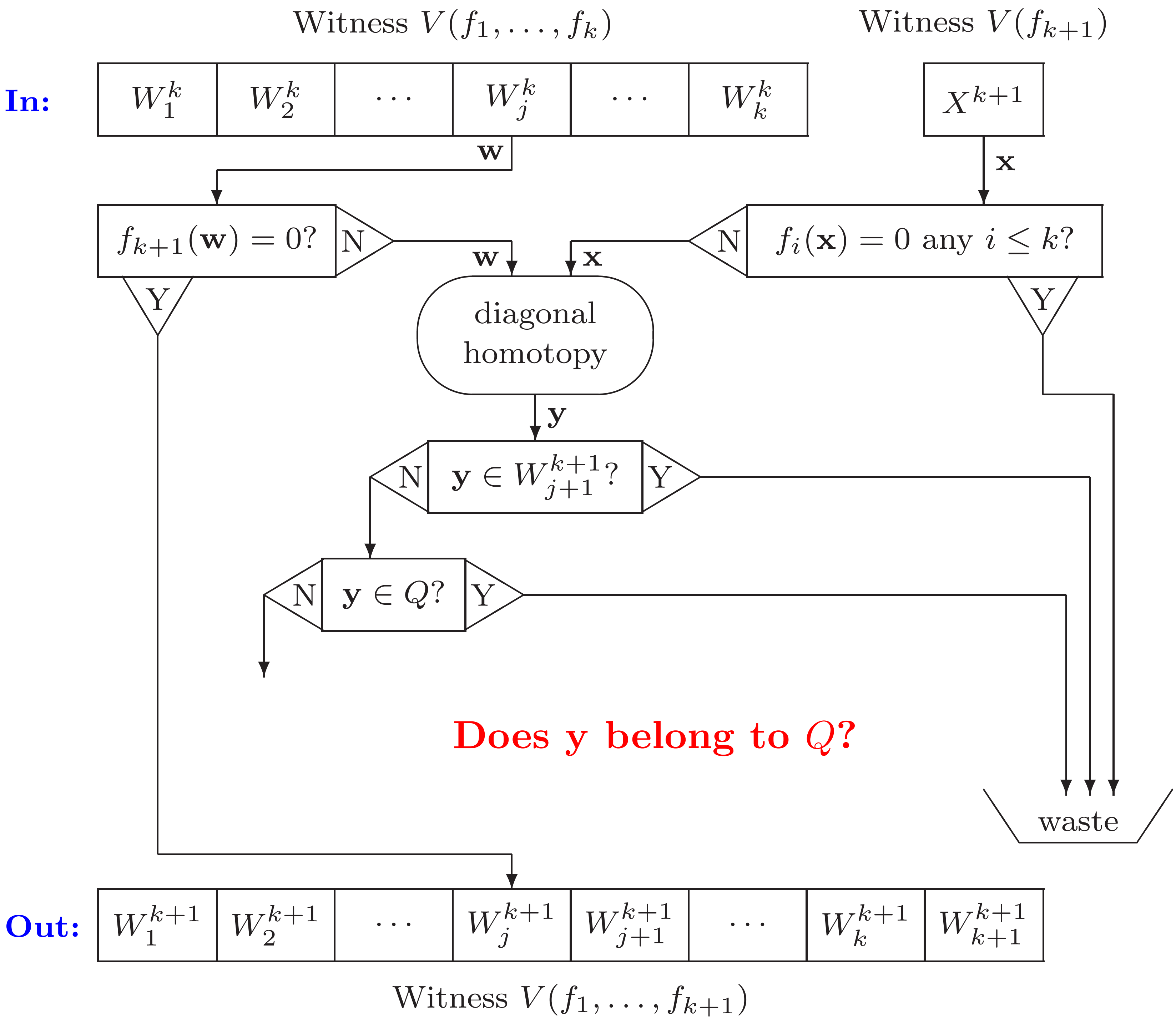

Witness $V(f_1, \dots, f_k)$
Witness $V(f_{k+1})$
In:
W_1^k
W_2^k
$\cdots$
W_j^k
$\cdots$
W_k^k
X^{k+1}
$\mathbf{w}$
$\mathbf{x}$
$f_{k+1}(\mathbf{w}) = 0?$
N
Y
N
$f_i(\mathbf{x}) = 0$ any $i \le k$?
Y
pass w and x first to filters
waste
Out:
W_1^{k+1}
W_2^{k+1}
$\cdots$
W_j^{k+1}
W_{j+1}^{k+1}
$\cdots$
W_k^{k+1}
W_{k+1}^{k+1}
Witness $V(f_1, \dots, f_{k+1})$

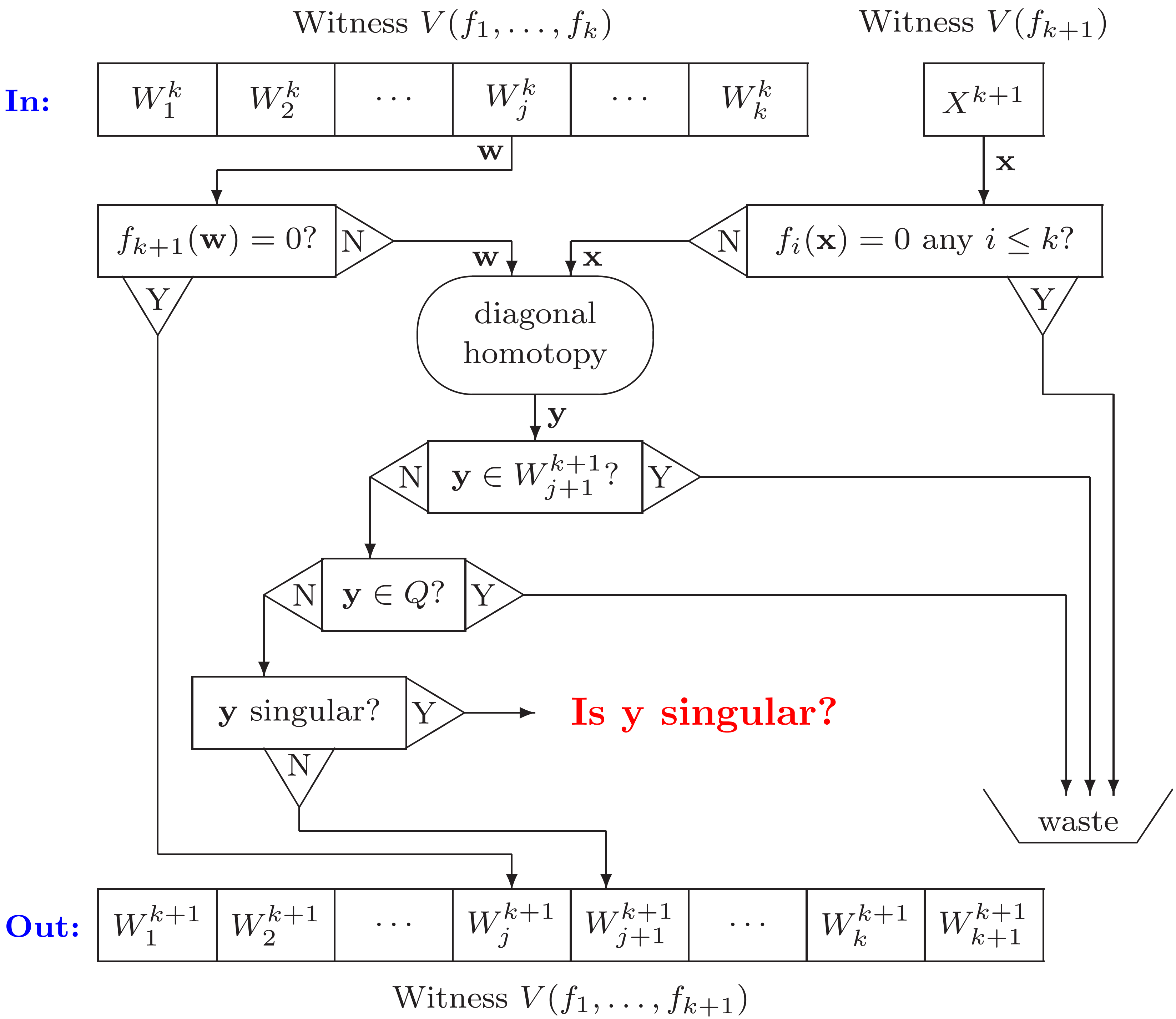


Witness $V(f_1, \ldots, f_k)$
Witness $V(f_{k+1})$
In:
W_1^k
W_2^k
$\cdots$
W_j^k
$\cdots$
W_k^k
X^{k+1}
$\mathbf{w}$
$\mathbf{x}$
$f_{k+1}(\mathbf{w}) = 0?$
N
$\mathbf{w}$
$\mathbf{x}$
N
$f_i(\mathbf{x}) = 0$ any $i \leq k$?
Y
Y
diagonal homotopy
$\mathbf{y}$
waste
Out:
W_1^{k+1}
W_2^{k+1}
$\cdots$
W_j^{k+1}
W_{j+1}^{k+1}
$\cdots$
W_k^{k+1}
W_{k+1}^{k+1}
Witness $V(f_1, \ldots, f_{k+1})$

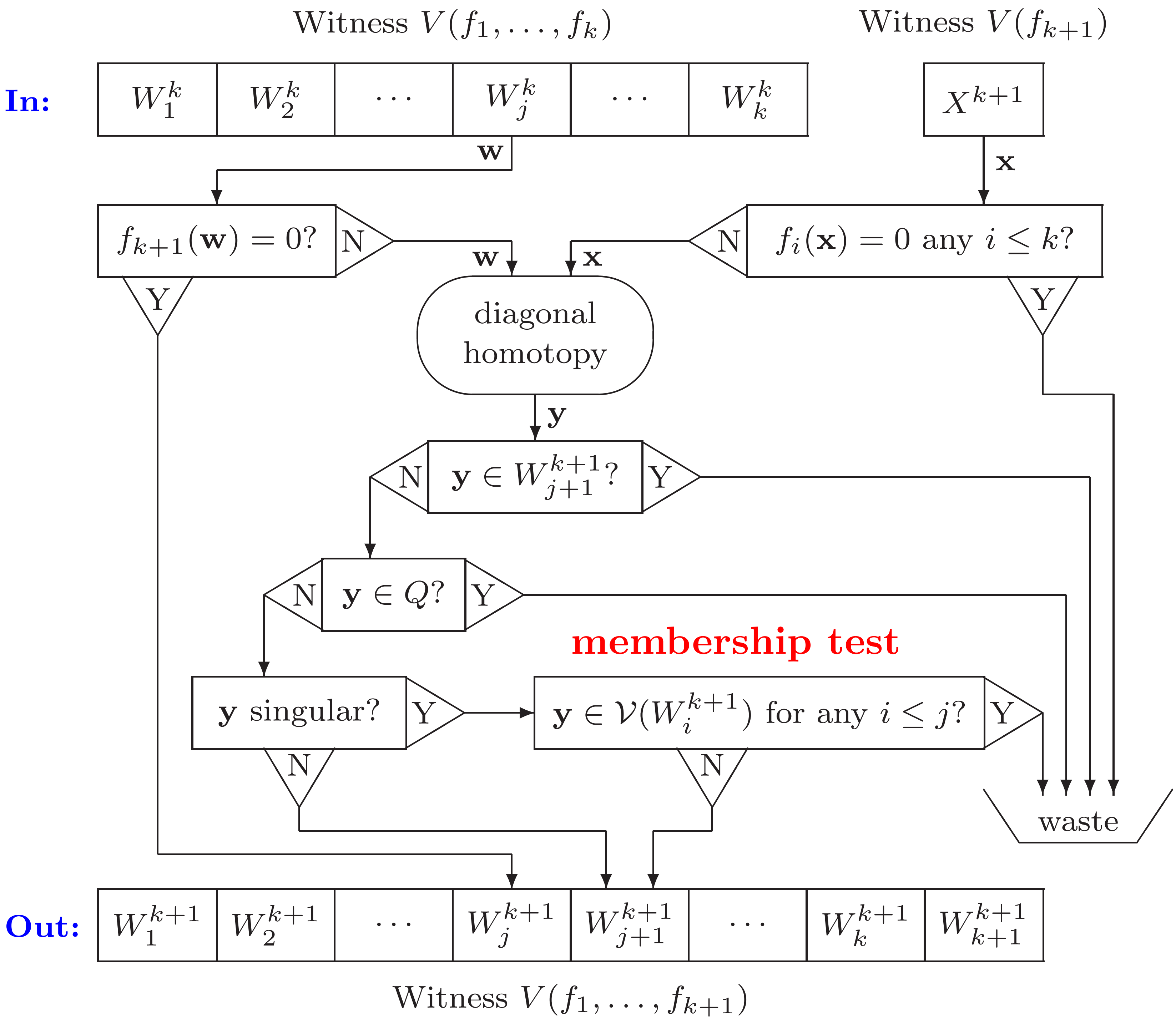


Witness $V(f_1,\ldots,f_k)$
Witness $V(f_{k+1})$
In:
W_1^k W_2^k $\cdots$ W_j^k $\cdots$ W_k^k
X^{k+1}
$\mathbf{w}$
$\mathbf{x}$
$f_{k+1}(\mathbf{w}) = 0?$
N
Y
$\mathbf{w}$
$\mathbf{x}$
N
$f_i(\mathbf{x}) = 0$ any $i \le k$?
Y
diagonal homotopy
$\mathbf{y}$
N
$\mathbf{y} \in W_{j+1}^{k+1}$?
Y
Has y already been found?
waste
Out:
W_1^{k+1} W_2^{k+1} $\cdots$ W_j^{k+1} W_{j+1}^{k+1} $\cdots$ W_k^{k+1} W_{k+1}^{k+1}
Witness $V(f_1,\ldots,f_{k+1})$


Witness $V(f_1, \ldots, f_k)$
Witness $V(f_{k+1})$
In:
W_1^k
W_2^k
$\cdots$
W_j^k
$\cdots$
W_k^k
X^{k+1}
w
x
$f_{k+1}(\mathbf{w}) = 0?$
N
Y
w
x
N
$f_i(\mathbf{x}) = 0$ any $i \leq k$?
Y
diagonal homotopy
y
N
$\mathbf{y} \in W_{j+1}^{k+1}$?
Y
N
$\mathbf{y} \in Q$?
Y
Does y belong to Q?
waste
Out:
W_1^{k+1}
W_2^{k+1}
$\cdots$
W_j^{k+1}
W_{j+1}^{k+1}
$\cdots$
W_k^{k+1}
W_{k+1}^{k+1}
Witness $V(f_1, \ldots, f_{k+1})$



Witness $V(f_1, \ldots, f_k)$
Witness $V(f_{k+1})$
In:
W_1^k
W_2^k
$\cdots$
W_j^k
$\cdots$
W_k^k
X^{k+1}
$\mathbf{w}$
$\mathbf{x}$
$f_{k+1}(\mathbf{w}) = 0?$
N
Y
$\mathbf{w}$
$\mathbf{x}$
N
$f_i(\mathbf{x}) = 0$ any $i \leq k?$
Y
diagonal homotopy
$\mathbf{y}$
N
$\mathbf{y} \in W_{j+1}^{k+1}?$
Y
N
$\mathbf{y} \in Q?$
Y
$\mathbf{y}$ singular?
Y
N
Is y singular?
waste
Out:
W_1^{k+1}
W_2^{k+1}
$\cdots$
W_j^{k+1}
W_{j+1}^{k+1}
$\cdots$
W_k^{k+1}
W_{k+1}^{k+1}
Witness $V(f_1, \ldots, f_{k+1})$


Witness $V(f_1,\ldots,f_k)$
Witness $V(f_{k+1})$
In:
W_1^k
W_2^k
$\cdots$
W_j^k
$\cdots$
W_k^k
X^{k+1}
$\mathbf{w}$
$\mathbf{x}$
$f_{k+1}(\mathbf{w})=0?$
$f_i(\mathbf{x})=0$ any $i\le k$?
diagonal homotopy
$\mathbf{y}$
$\mathbf{y}\in W_{j+1}^{k+1}$?
$\mathbf{y}\in Q$?
membership test
$\mathbf{y}$ singular?
$\mathbf{y}\in\mathcal{V}(W_i^{k+1})$ for any $i\le j$?
Y
N
waste
Out:
W_1^{k+1}
W_2^{k+1}
$\cdots$
W_j^{k+1}
W_{j+1}^{k+1}
$\cdots$
W_k^{k+1}
W_{k+1}^{k+1}
Witness $V(f_1,\ldots,f_{k+1})$

Numerical Irreducible Decomposition

$$f = \begin{bmatrix} (y-x^2)(x^2+y^2+z^2-1)(x-0.5) \\ (z-x^3)(x^2+y^2+z^2-1)(y-0.5) \\ (y-x^2)(z-x^3)(x^2+y^2+z^2-1)(z-0.5) \end{bmatrix} = \mathbf{0}.$$

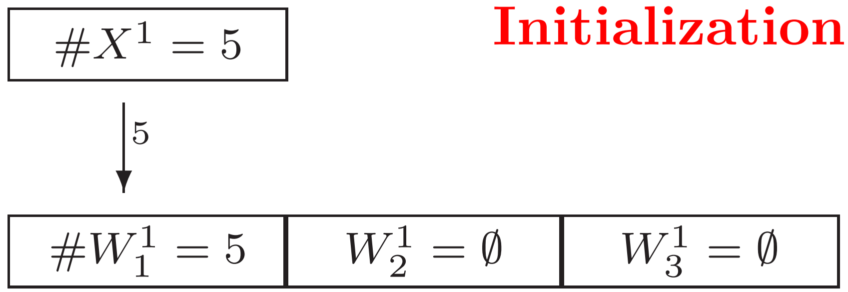
The **irreducible decomposition** of $Z = f^{-1}(\mathbf{0})$ is

$$Z = Z_2 \cup Z_1 \cup Z_0 = \{Z_{21}\} \cup \{Z_{11} \cup Z_{12} \cup Z_{13} \cup Z_{14}\} \cup \{Z_{01}\}$$

where Z_{21} is the sphere $x^2+y^2+z^2-1=0$;
Z_{11} is the line $(x=0.5, z=0.5^3)$;
Z_{12} is the line $(x=\sqrt{0.5}, y=0.5)$;
Z_{13} is the line $(x=-\sqrt{0.5}, y=0.5)$;
Z_{14} is the twisted cubic $(y-x^2=0, z-x^3=0)$;
Z_{01} is the point $(x=0.5, y=0.5, z=0.5)$.

Previous approach: 197 paths to find all candidate witness points.

Initialization with f_1

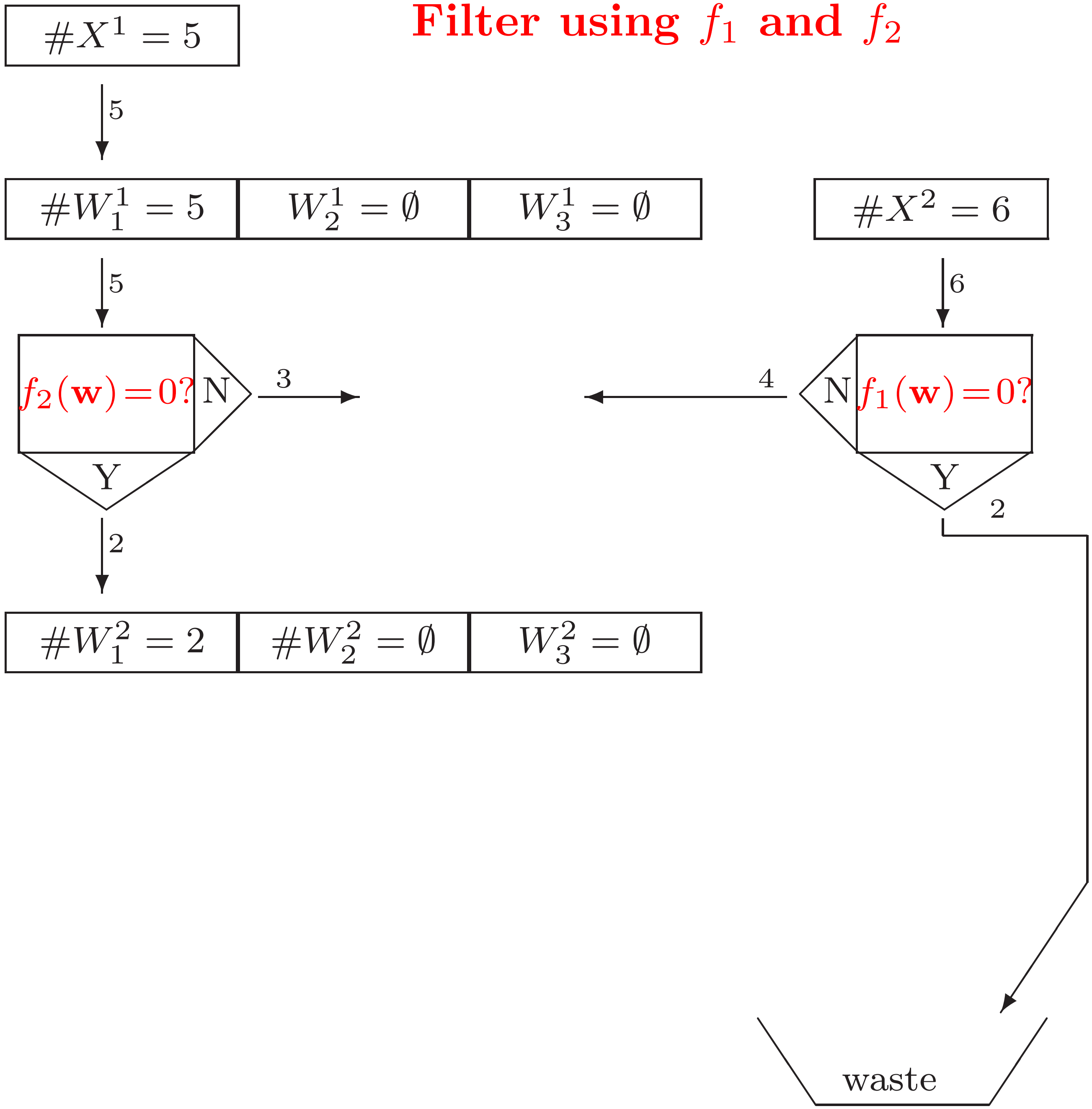


$$f_1(x,y,z) = (y - x^2)(x^2 + y^2 + z^2 - 1)(x - 0.5) = 0$$

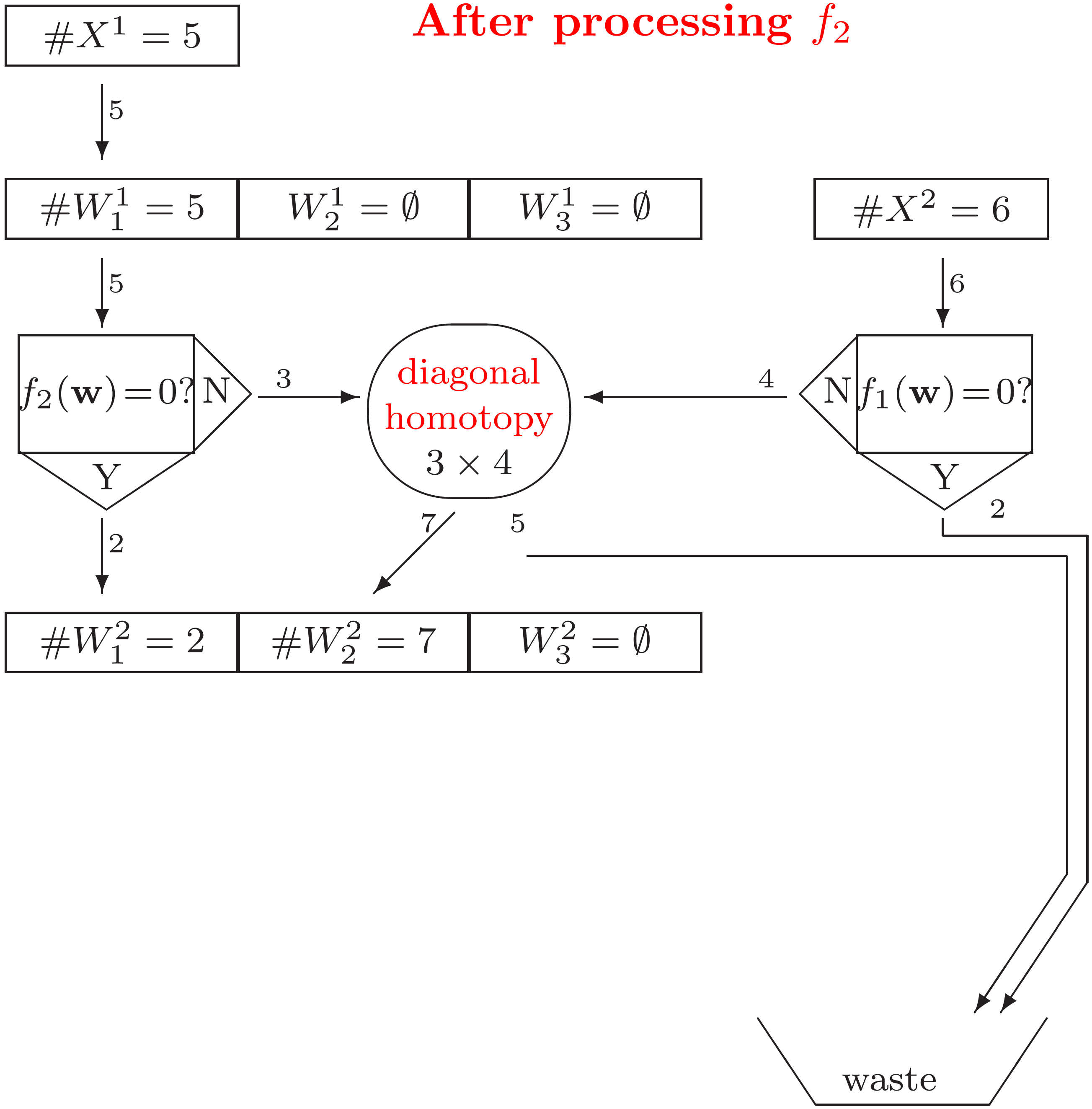
$$f_2(x,y,z) = (z - x^3)(x^2 + y^2 + z^2 - 1)(y - 0.5) = 0$$

$$f_3(x,y,z) = (y - x^2)(z - x^3)(x^2 + y^2 + z^2 - 1)(z - 0.5) = 0$$

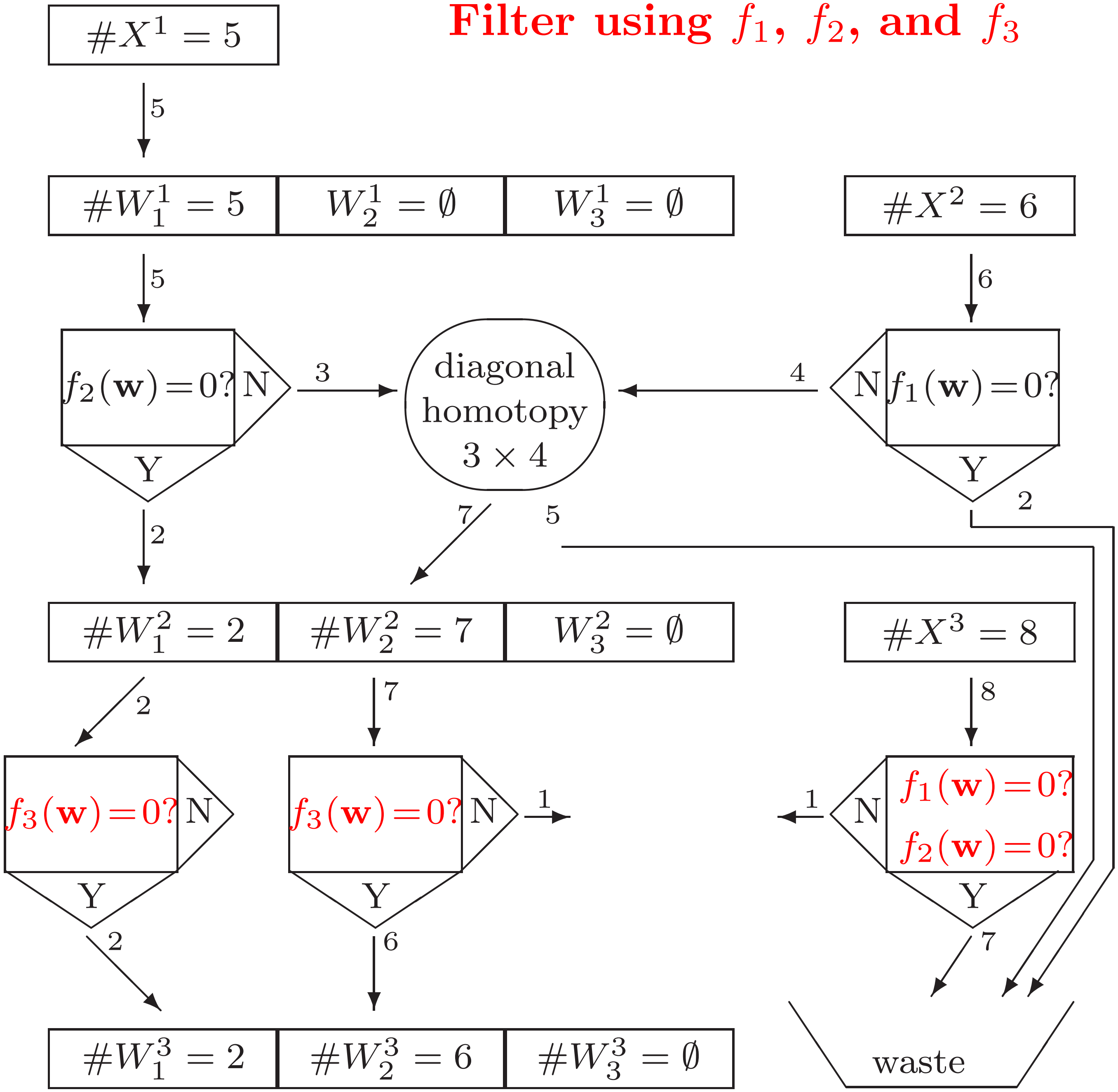
Filter using f_1 and f_2



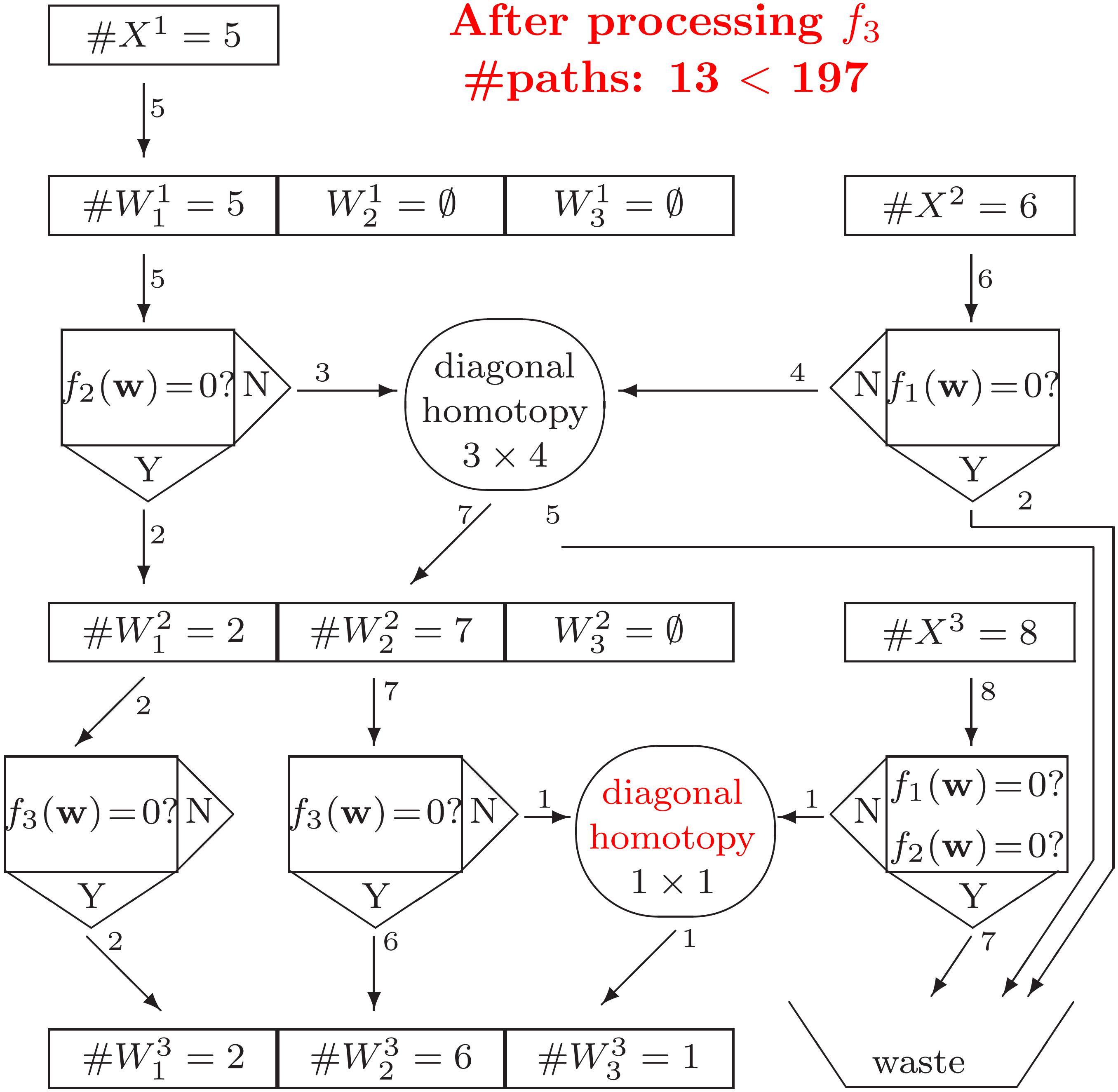
After processing f_2



Filter using f_1, f_2, and f_3



After processing f_3
#paths: 13 < 197



Adjacent Minors of a General 2-by-8 Matrix

from algebraic statistics (Diaconis, Eisenbud, Sturmfels, 1998):

$$\begin{bmatrix} x_{11} & x_{12} & x_{13} & x_{14} & x_{15} & x_{16} & x_{17} & x_{18} & x_{19} \\ x_{21} & x_{22} & x_{23} & x_{24} & x_{25} & x_{26} & x_{27} & x_{28} & x_{29} \end{bmatrix}$$

8 quadrics in 18 unknowns: 10-dimensional surface of degree 256

stage	#paths			user cpu time		
1	4	=	2×2	0.11s	=	110ms
2	8	=	4×2	0.41s	=	410ms
3	16	=	8×2	1.61s	=	1s 610ms
4	32	=	16×2	3.75s	=	3s 750ms
5	64	=	32×2	12.41s	=	12s 410ms
6	128	=	64×2	34.89s	=	34s 890ms
7	256	=	128×2	104.22s	=	1m 44s 220ms
total user cpu time				157.56s	=	2m 37s 560ms

8m 22s for direct (extrinsic) homotopy

Apple PowerBook G4 1GHz

A General 6-by-6 Eigenvalue Problem

$f(\mathbf{x}, \lambda) = \lambda \mathbf{x} - A\mathbf{x} = \mathbf{0}, \quad A \in \mathbb{C}^{6 \times 6}$, A is random matrix

6 equations in 7 unknowns: curve of degree $7 < 64 = 2^6$

stage in solver	1	2	3	4	5	total
#convergent paths	3	4	5	6	7	25
#divergent paths	1	2	3	4	5	15
#paths tracked	4	6	8	10	12	40

15 is much less than $64 - 6 = 58$ divergent paths with direct homotopy, using the plain theorem of Bézout

Singularities are keeping us in business

numerical analysis: bifurcation points and endgames

Rall (1966); Reddien (1978); Decker-Keller-Kelley (1983);
Griewank-Osborne (1981); Hoy (1989);
Deuflard-Friedler-Kunkel (1987); Kunkel (1989, 1996);
Morgan-Sommese-Wampler (1991); Li-Wang (1993, 1994);
Govaerts (2000)

computer algebra: standard bases (SINGULAR)

Mora (1982); Greuel-Pfister (1996)

numerical polynomial algebra: duality, "multiplicity structure"

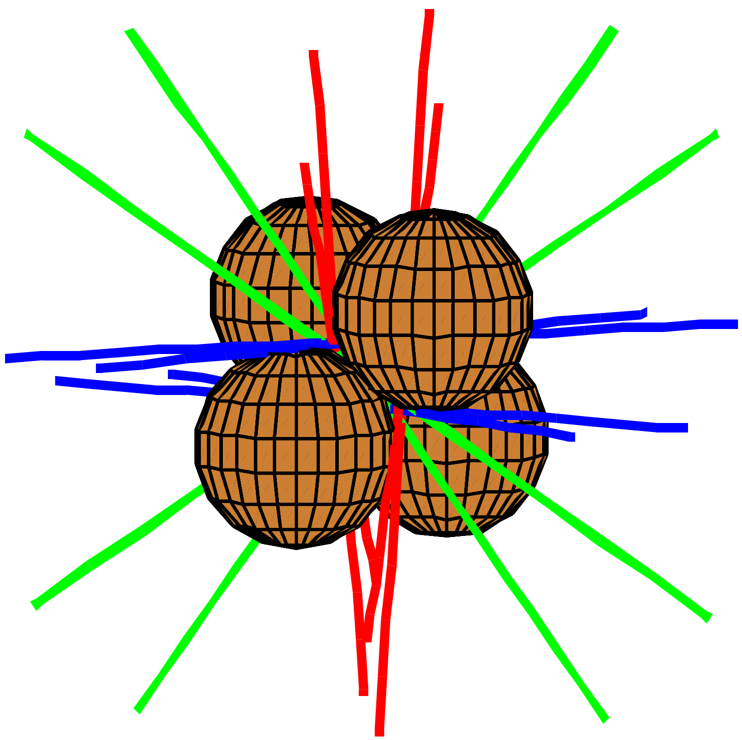
Möller-Stetter (1995); Mourrain (1997);
Stetter-Thallinger (1998); Dayton-Zeng (2005)

deflation: Ojika-Watanabe-Mitsui (1983); Lecerf (2003)

Twelve lines tangent to four spheres

Frank Sottile and Thorsten Theobald: Lines tangents to $2n - 2$ spheres in $\mathbb{R}^n$

Trans. Amer. Math. Soc. 354
pages 4815-4829, 2002.



Problem:

Given 4 spheres,
find all lines tangent
to all 4 given spheres.

Observe:

12 solutions in groups of 4.

An Input Polynomial System

```
x0**2 + x1**2 + x2**2 - 1;
x0*x3 + x1*x4 + x2*x5;
x3**2 + x4**2 + x5**2 - 0.25;
x3**2 + x4**2 - 2*x2*x4 + x2**2 + x5**2 + 2*x1*x5 + x1**2 - 0.25;
x3**2 + 1.73205080756888*x2*x3 + 0.75*x2**2 + x4**2 - x2*x4 + 0.25*x2**2
+ x5**2 - 1.73205080756888*x0*x5 + x1*x5
+ 0.75*x0**2 - 0.86602540378444*x0*x1 + 0.25*x1**2 - 0.25;
x3**2 - 1.63299316185545*x1*x3 + 0.57735026918963*x2*x3
+ 0.66666666666667*x1**2 - 0.47140452079103*x1*x2 + 0.08333333333333*x2**2
+ x4**2 + 1.63299316185545*x0*x4 - x2*x4 + 0.66666666666667*x0**2
- 0.81649658092773*x0*x2 + 0.25*x2**2
+ x5**2 - 0.57735026918963*x0*x5 + x1*x5 + 0.08333333333333*x0**2
- 0.28867513459481*x0*x1 + 0.25*x1**2 - 0.25;
```

Original formulation as polynomial system: **Cassiano Durand**.
Centers of the spheres at the vertices of a tetrahedron: **Thorsten Theobald**.
Algebraic numbers sqrt(3), sqrt(6), etc. approximated by double floats.

The system has 6 isolated solutions, each of multiplicity 4.

Deflation Operator Dfl reduces to Corank One

Consider $f(\mathbf{x}) = \mathbf{0}$, N equations in n unknowns, $N \geq n$.

Suppose Rank$(A(\mathbf{z}_0)) = R < n$ for $\mathbf{z}_0$ an isolated zero of $f(\mathbf{x}) = 0$.

Choose $\mathbf{h} \in \mathbb{C}^{R+1}$ and $B \in \mathbb{C}^{n \times (R+1)}$ at random.

Introduce $R + 1$ new multiplier variables $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \ldots, \lambda_{R+1})$.

$$\mathbf{Dfl}(f)(\mathbf{x}, \boldsymbol{\lambda}) := \left\{ \begin{array}{rcl} f(\mathbf{x}) & = & \mathbf{0} \\ A(\mathbf{x})B\boldsymbol{\lambda} & = & \mathbf{0} \\ \mathbf{h}\boldsymbol{\lambda} & = & 1 \end{array} \right. \qquad \begin{array}{c} \text{Rank}(A(\mathbf{x})) = R \\ \Downarrow \\ \text{corank}(A(\mathbf{x})B) = 1 \end{array}$$

Compared to the deflation of Ojika, Watanabe, and Mitsui:
(1) we do not compute a maximal minor of the Jacobian matrix;
(2) we only add new equations, we never replace equations.

Newton with Deflation – A Simple Example

$$f(x,y) = \begin{cases} x^2 = 0 \\ xy = 0 \\ y^2 = 0 \end{cases} \qquad A(x,y) = \begin{bmatrix} 2x & 0 \\ y & x \\ 0 & 2y \end{bmatrix} \qquad \begin{array}{c} \mathbf{z}_0 = (0,0), m = 3 \\ \mathrm{Rank}(A(\mathbf{z}_0)) = 0 \end{array}$$

A nontrivial linear combination of the columns of $A(\mathbf{z}_0)$ is zero.

$$\mathbf{Dfl}(f)(x,y,\lambda_1,\lambda_2) = \left\{ \begin{array}{rcl} f(x,y) & = & 0 \\ \begin{bmatrix} 2x & 0 \\ y & x \\ 0 & 2y \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_2 \end{bmatrix} & = & \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \\ c_1\lambda_1 + c_2\lambda_2 & = & 1, \qquad \text{random } c_1, c_2 \in \mathbb{C} \end{array} \right.$$

$\mathbf{Dfl}(f)(x,y,\lambda_1,\lambda_2) = \mathbf{0}$ has $(0,0,\lambda_1^*,\lambda_2^*)$ as **regular** zero!

Newton's Method with Deflation

Input: $f(\mathbf{x}) = \mathbf{0}$ polynomial system;
$\mathbf{x}_0$ initial approximation for $\mathbf{x}^*$;
ϵ tolerance for numerical rank.

Newton's Method with Deflation

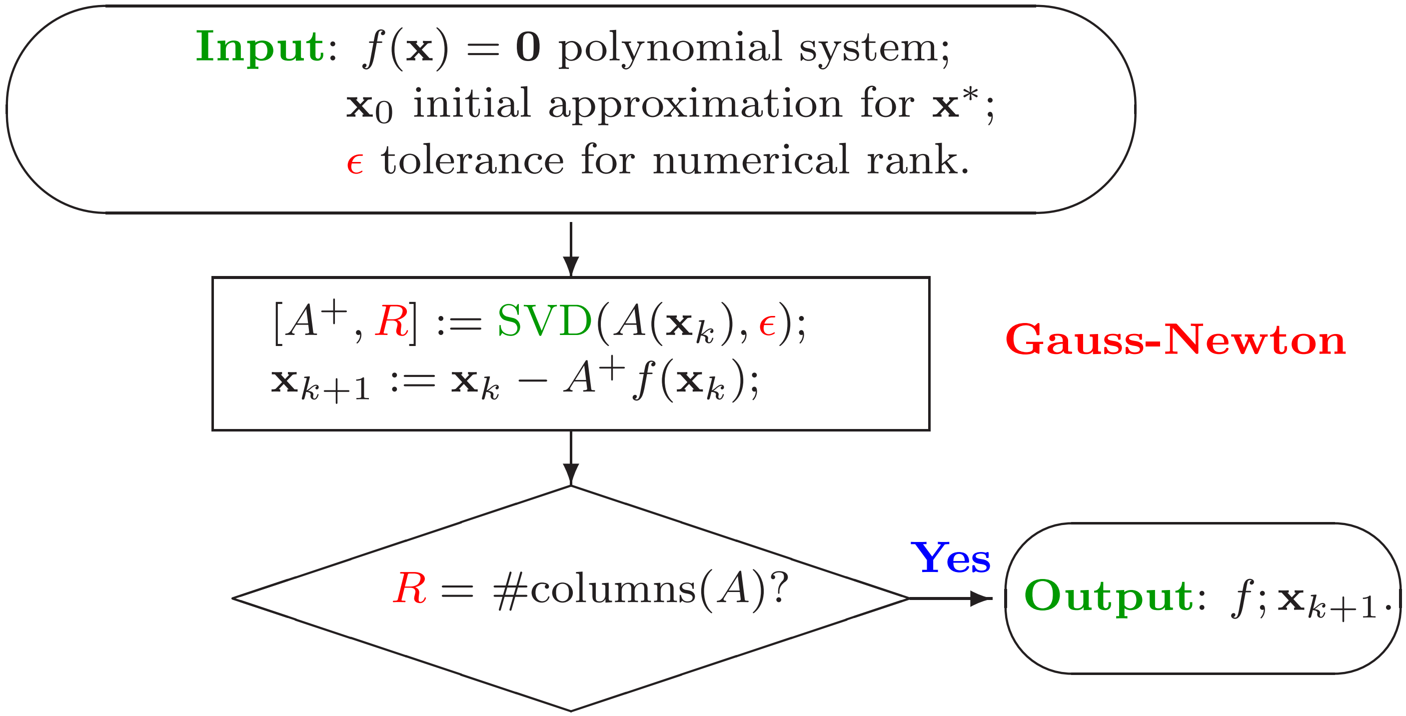
Input: $f(\mathbf{x}) = \mathbf{0}$ polynomial system;
$\mathbf{x}_0$ initial approximation for $\mathbf{x}^*$;
ϵ tolerance for numerical rank.

↓

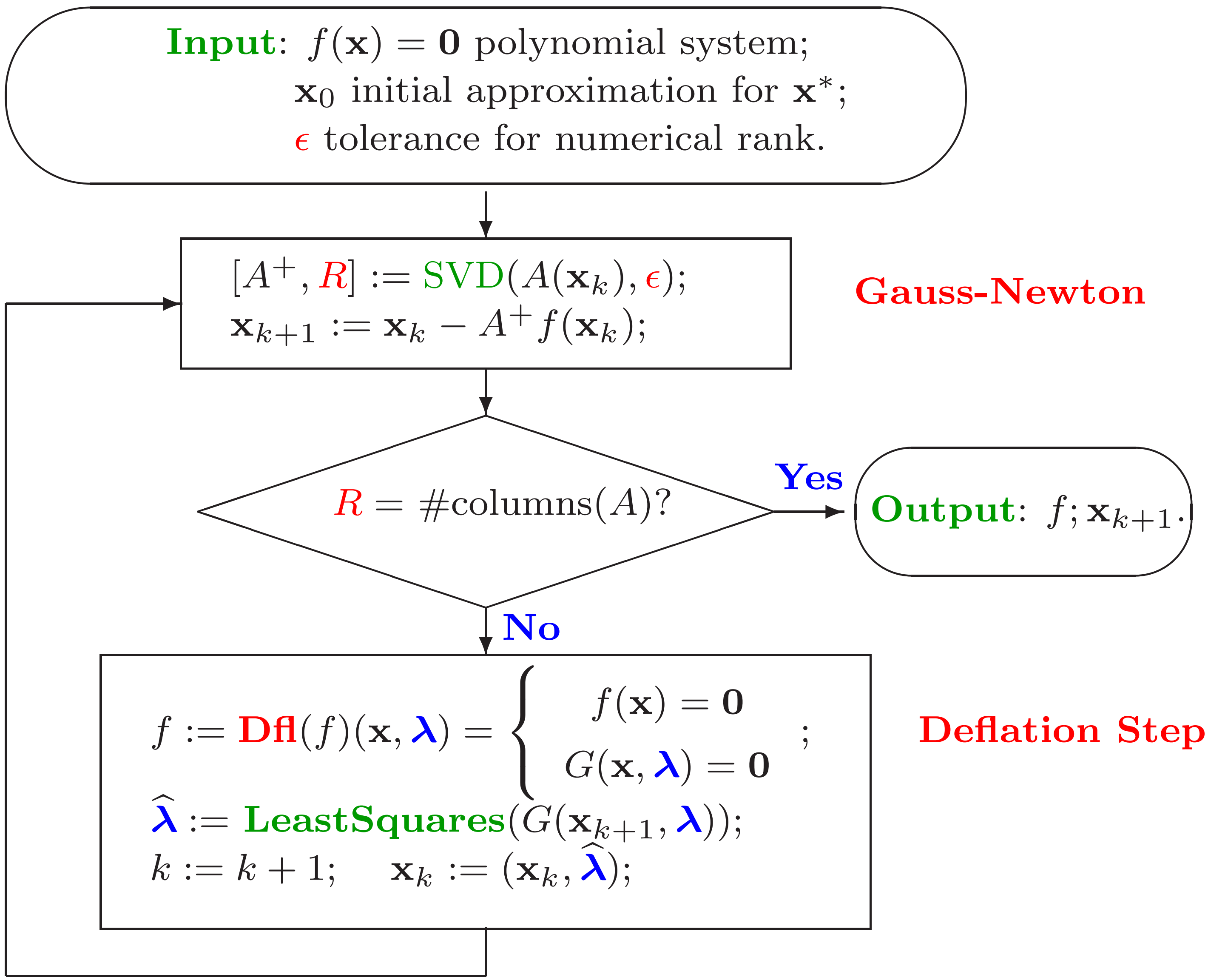
$[A^+, R] := \text{SVD}(A(\mathbf{x}_k), \epsilon);$
$\mathbf{x}_{k+1} := \mathbf{x}_k - A^+ f(\mathbf{x}_k);$

Gauss-Newton

Newton's Method with Deflation



Newton's Method with Deflation



A Bound on the Number of Deflations

Theorem (Anton Leykin, JV, Ailing Zhao):

The number of deflations needed to restore the quadratic convergence of Newton's method converging to an isolated solution is strictly less than the multiplicity.

Duality Analysis of Barry H. Dayton and Zhonggang Zeng:

(1) tighter bound on number of deflations; and

(2) special case algorithms, for corank $= 1$.

(to appear in ISSAC 2005)

Numerical Results (double float)

System	n	m	D	corank($A(\mathbf{x})$)	Inverse Condition#	#Digits
baker1	2	2	1	$1 \to 0$	1.7e-08 $\to$ 3.8e-01	$9 \to 24$
cbms1	3	11	1	$3 \to 0$	4.2e-05 $\to$ 5.0e-01	$5 \to 20$
cbms2	3	8	1	$3 \to 0$	1.2e-08 $\to$ 5.0e-01	$8 \to 18$
mth191	3	4	1	$2 \to 0$	1.3e-08 $\to$ 3.5e-02	$7 \to 13$
decker1	2	3	2	$1 \to 1 \to 0$	3.4e-10 $\to$ 2.6e-02	$6 \to 11$
decker2	2	4	3	$1 \to 1 \to 1 \to 0$	4.5e-13 $\to$ 6.9e-03	$5 \to 16$
decker3	2	2	1	$1 \to 0$	4.6e-08 $\to$ 2.5e-02	$8 \to 17$
ojika1	2	3	2	$1 \to 1 \to 0$	9.3e-12 $\to$ 4.3e-02	$5 \to 12$
ojika2	3	2	1	$1 \to 0$	3.3e-08 $\to$ 7.4e-02	$6 \to 14$
ojika3	3	2	1	$1 \to 0$	1.7e-08 $\to$ 9.2e-03	$7 \to 15$
		4	1	$2 \to 0$	6.5e-08 $\to$ 8.0e-02	$6 \to 13$
ojika4	3	3	2	$1 \to 1 \to 0$	1.9e-13 $\to$ 2.4e-04	$6 \to 11$
cyclic9	9	4	1	$2 \to 0$	5.6e-10 $\to$ 1.8e-03	$5 \to 15$

Structure of the Deflated Systems

At stage k in the deflation:

$$f_k(\mathbf{x}, \boldsymbol{\lambda}_1, \ldots, \boldsymbol{\lambda}_{k-1}, \boldsymbol{\lambda}_k) = \begin{cases} f_{k-1}(\mathbf{x}, \boldsymbol{\lambda}_1, \ldots, \boldsymbol{\lambda}_{k-1}) & = & 0 \\ A_{k-1}(\mathbf{x}, \boldsymbol{\lambda}_1, \ldots, \boldsymbol{\lambda}_{k-1}) B_k \boldsymbol{\lambda}_k & = & 0 \\ \mathbf{h}_k \boldsymbol{\lambda}_k & = & 1, \end{cases}$$

$f_0 = f$, $A_0 = A$, $R_k = \operatorname{rank}(A_{k-1}(\mathbf{z}_0))$, $R_k + 1$ multipliers in $\boldsymbol{\lambda}_k$.

random vector $\mathbf{h}_k \in \mathbb{C}^{R_k+1}$ and n_{k-1}-by-$(R_k + 1)$ matrix B_k

$$\begin{array}{rcl} \#\text{rows in } A_k & : & N_k = 2N_{k-1} + 1, \quad N_0 = N, \\ \#\text{columns in } A_k & : & n_k = n_{k-1} + R_k + 1, \quad n_0 = n. \end{array}$$

Multiplication of the polynomial matrix A_{k-1} with the random matrix B_k and vector of $R_k + 1$ multiplier variables $\boldsymbol{\lambda}_k$ is expensive!

Structure of the Jacobian Matrices

At stage k in the deflation:

$$f_k(\mathbf{x}, \boldsymbol{\lambda}_1, \ldots, \boldsymbol{\lambda}_{k-1}, \boldsymbol{\lambda}_k) = \left\{ \begin{array}{rcl} f_{k-1}(\mathbf{x}, \boldsymbol{\lambda}_1, \ldots, \boldsymbol{\lambda}_{k-1}) & = & 0 \\ A_{k-1}(\mathbf{x}, \boldsymbol{\lambda}_1, \ldots, \boldsymbol{\lambda}_{k-1}) B_k \boldsymbol{\lambda}_k & = & 0 \\ \mathbf{h}_k \boldsymbol{\lambda}_k & = & 1. \end{array} \right.$$

Jacobian matrix at the kth deflation:

$$A_k(\mathbf{x}, \boldsymbol{\lambda}_1, \ldots, \boldsymbol{\lambda}_{k-1}, \boldsymbol{\lambda}_k) = \left[\begin{array}{cc} A_{k-1} & \mathbf{0} \\ \left[\frac{\partial A_{k-1}}{\partial \mathbf{x}} \, \frac{\partial A_{k-1}}{\partial \boldsymbol{\lambda}_1} \cdots \frac{\partial A_{k-1}}{\partial \boldsymbol{\lambda}_{k-1}} \right] B_k \boldsymbol{\lambda}_k & A_{k-1} B_k \\ \mathbf{0} & \mathbf{h}_k \end{array} \right].$$

The multiplier variables $\boldsymbol{\lambda}_i$, $i = 1, 2, \ldots, k$, occur linearly,

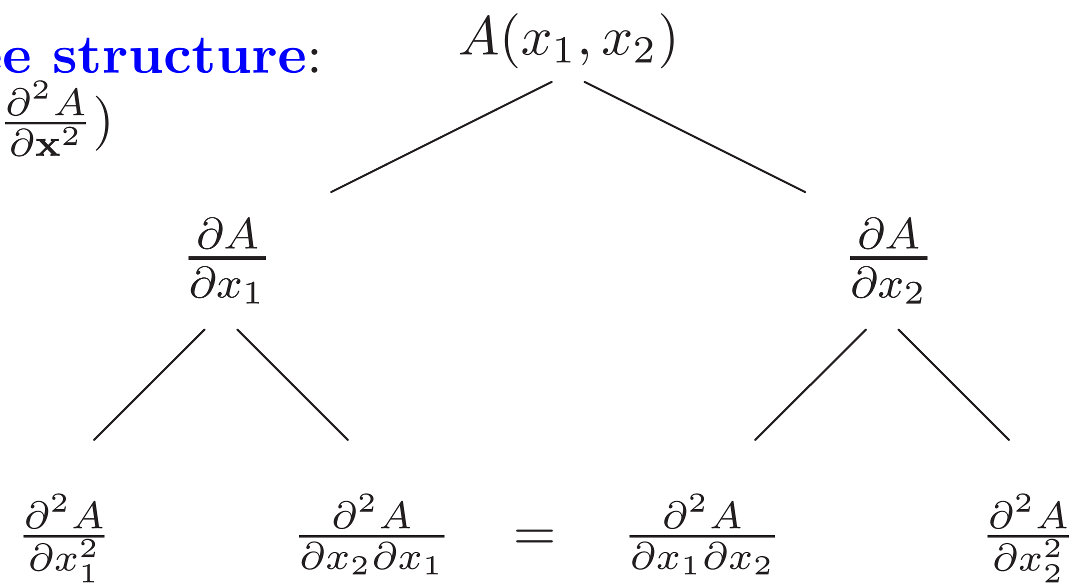
we compute derivatives only with respect to the original variables $\mathbf{x}$.

Derivatives of Jacobian matrices

Define $\frac{\partial A}{\partial \mathbf{x}} = \left[\frac{\partial A}{\partial x_1} \;\; \frac{\partial A}{\partial x_2} \;\; \cdots \;\; \frac{\partial A}{\partial x_n} \right]$, $\mathbf{x} = (x_1, x_2, \ldots, x_n)$,

where $\frac{\partial A}{\partial x_k} = \left[\frac{\partial a_{ij}}{\partial x_k} \right]$ for $A = [a_{ij}(\mathbf{x})]^{i \in \{1,2,\ldots,N\}}_{j,k \in \{1,2,\ldots,n\}}$.

Natural tree structure:
(to compute $\frac{\partial^2 A}{\partial \mathbf{x}^2}$)



Complexity of $\frac{\partial^k A}{\partial \mathbf{x}^k} \neq O(n^k)$,

but $O(\#\text{monomials of degree } k \text{ in } n \text{ variables})$.

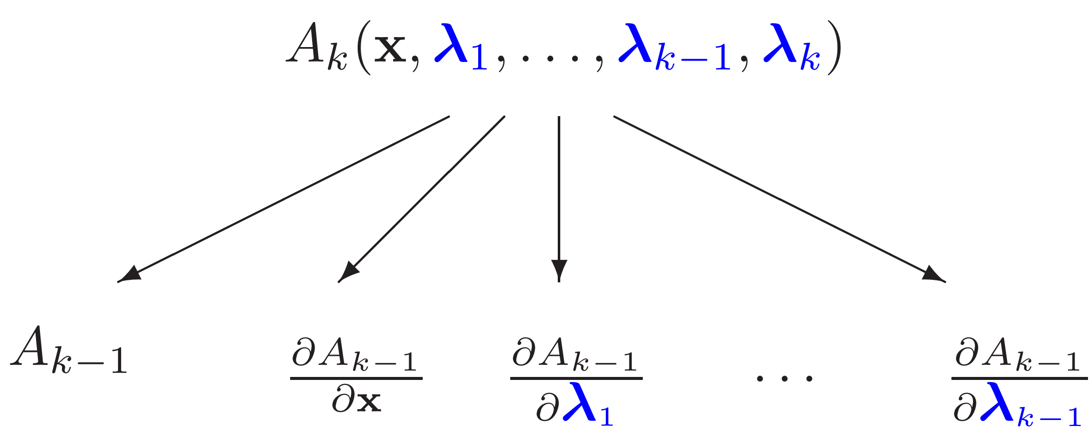
e.g.: $k = 10$, $n = 3$:
$3^{10} = 59049 \gg 66$

Column Format of Jacobian Matrices

Jacobian matrix at the kth deflation:

$$A_k(\mathbf{x}, \boldsymbol{\lambda}_1, \ldots, \boldsymbol{\lambda}_{k-1}, \boldsymbol{\lambda}_k) = \begin{bmatrix} A_{k-1} & \mathbf{0} \\ \left[\frac{\partial A_{k-1}}{\partial \mathbf{x}} \; \frac{\partial A_{k-1}}{\partial \boldsymbol{\lambda}_1} \cdots \frac{\partial A_{k-1}}{\partial \boldsymbol{\lambda}_{k-1}} \right] B_k \boldsymbol{\lambda}_k & A_{k-1} B_k \\ \mathbf{0} & \mathbf{h}_k \end{bmatrix}.$$

Tree with k children:



Unwinding the Multipliers

$$A_2(\mathbf{x}, \boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2) = \begin{bmatrix} A_1 & \mathbf{0} \\ \left[\frac{\partial A_1}{\partial \mathbf{x}} \;\; \frac{\partial A_1}{\partial \boldsymbol{\lambda}_1}\right] B_2 \boldsymbol{\lambda}_2 & A_1 B_2 \\ \mathbf{0} & \mathbf{h}_2 \end{bmatrix}$$

$$A_2$$

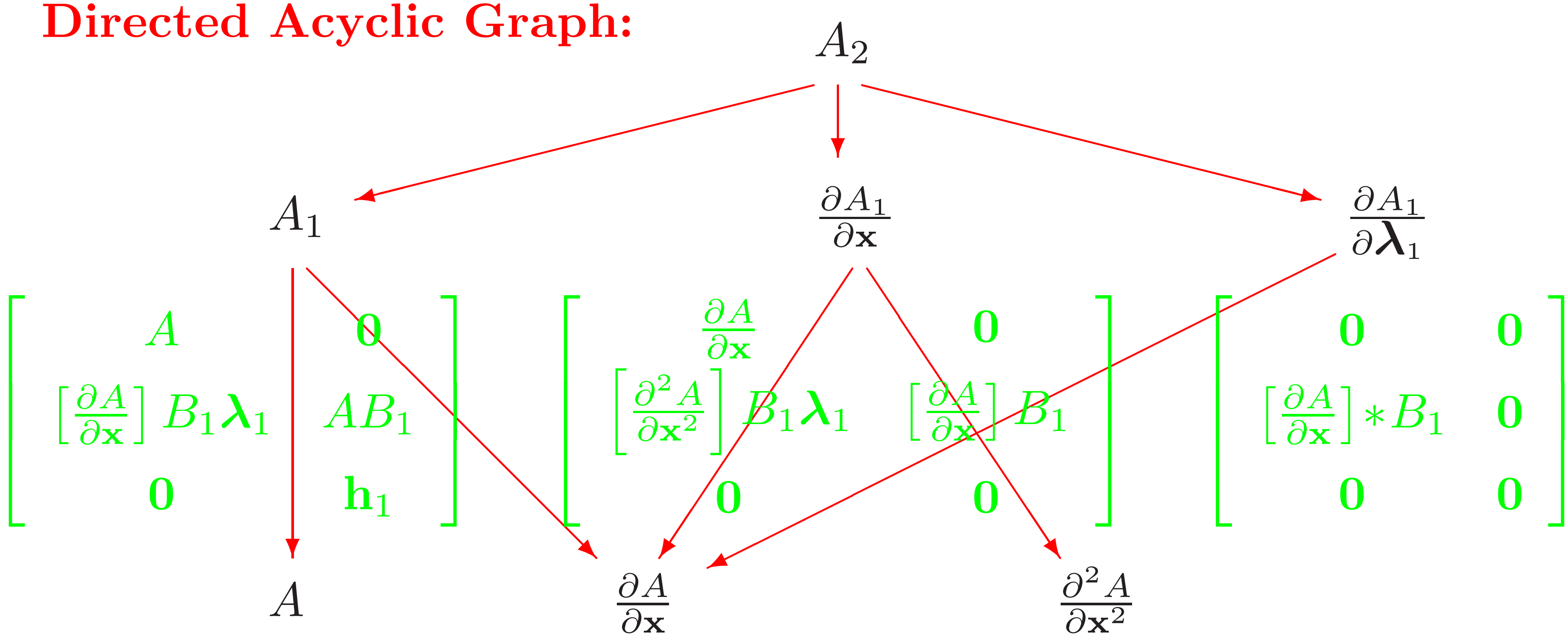
$$A_1 \qquad \frac{\partial A_1}{\partial \mathbf{x}} \qquad \frac{\partial A_1}{\partial \boldsymbol{\lambda}_1}$$

$$\begin{bmatrix} A & \mathbf{0} \\ \left[\frac{\partial A}{\partial \mathbf{x}}\right] B_1 \boldsymbol{\lambda}_1 & A B_1 \\ \mathbf{0} & \mathbf{h}_1 \end{bmatrix} \qquad \begin{bmatrix} \frac{\partial A}{\partial \mathbf{x}} & \mathbf{0} \\ \left[\frac{\partial^2 A}{\partial \mathbf{x}^2}\right] B_1 \boldsymbol{\lambda}_1 & \left[\frac{\partial A}{\partial \mathbf{x}}\right] B_1 \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \qquad \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \left[\frac{\partial A}{\partial \mathbf{x}}\right] * B_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}$$

Unwinding the Multipliers

$$A_2(\mathbf{x}, \boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2) = \begin{bmatrix} A_1 & \mathbf{0} \\ \left[\frac{\partial A_1}{\partial \mathbf{x}} \; \frac{\partial A_1}{\partial \boldsymbol{\lambda}_1}\right] B_2 \boldsymbol{\lambda}_2 & A_1 B_2 \\ \mathbf{0} & \mathbf{h}_2 \end{bmatrix}$$

Directed Acyclic Graph:



The Operator $*$

For a matrix B: $\left[\frac{\partial A}{\partial \mathbf{x}}\right] B = \left[\frac{\partial A}{\partial x_1} B \;\; \frac{\partial A}{\partial x_2} B \;\; \cdots \;\; \frac{\partial A}{\partial x_n} B\right]$.

However, $\frac{\partial}{\partial \boldsymbol{\lambda}}\left(\left[\frac{\partial A}{\partial \mathbf{x}}\right] B \boldsymbol{\lambda}\right) = \left[\underbrace{\frac{\partial A}{\partial \mathbf{x}} \mathbf{b}_1}_{\partial \lambda_1} \;\; \underbrace{\frac{\partial A}{\partial \mathbf{x}} \mathbf{b}_2}_{\partial \lambda_2} \;\; \cdots \;\; \underbrace{\frac{\partial A}{\partial \mathbf{x}} \mathbf{b}_m}_{\partial \lambda_m}\right]$,

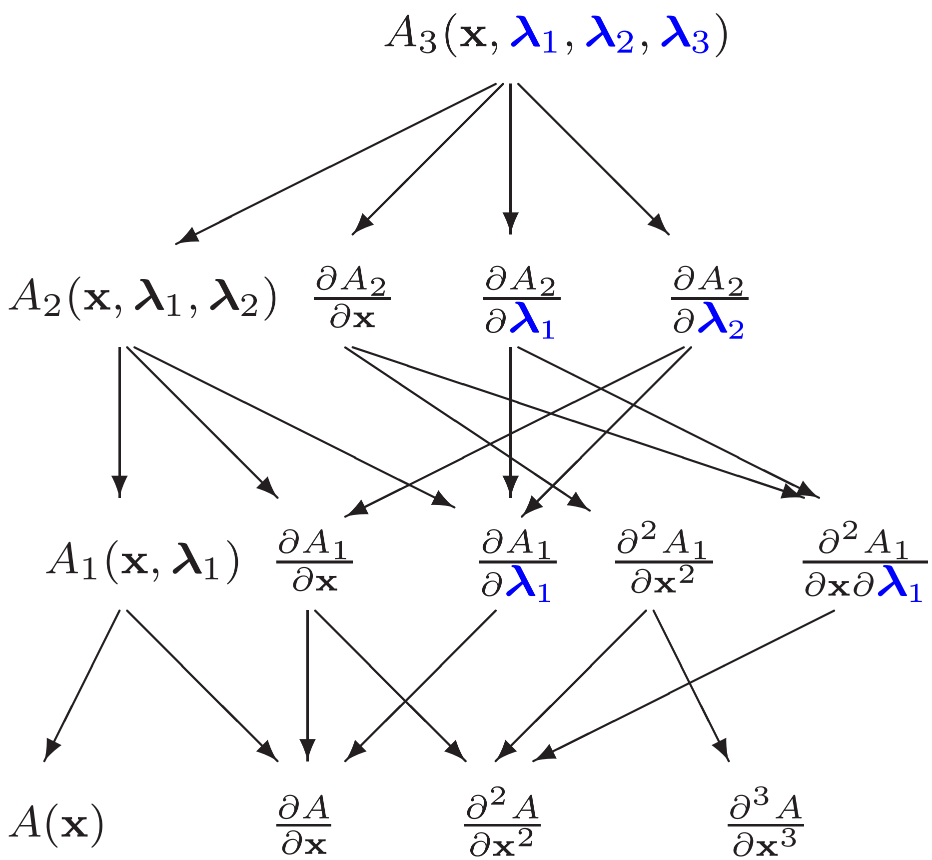
$$\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \ldots, \lambda_m), \; B = [\mathbf{b}_1 \;\; \mathbf{b}_2 \;\; \cdots \mathbf{b}_m].$$

Unlike scalar differentiation: $\frac{\partial}{\partial \boldsymbol{\lambda}}\left(\left[\frac{\partial A}{\partial \mathbf{x}}\right] B \boldsymbol{\lambda}\right) \neq \left[\frac{\partial A}{\partial \mathbf{x}}\right] B$.

With the operator $*$ we permute $\left[\frac{\partial A}{\partial \mathbf{x}}\right] B$ into $\frac{\partial}{\partial \boldsymbol{\lambda}}\left(\left[\frac{\partial A}{\partial \mathbf{x}}\right] B \boldsymbol{\lambda}\right)$.

So we have: $\frac{\partial}{\partial \boldsymbol{\lambda}}\left(\left[\frac{\partial A}{\partial \mathbf{x}}\right] B \boldsymbol{\lambda}\right) = \left[\frac{\partial A}{\partial \mathbf{x}}\right] * B$.

A Directed Acyclic Graph of Derivative Operators



Growth of Number of Nodes in DAG

The growth of the number of nodes in the directed acyclic graph, for increasing deflations k:

k	1	2	3	4	5	6	7	8	9	10
#nodes	3	7	14	26	46	79	133	221	364	596

#nodes ranges between $O(k^2)$ and $O(k^3)$

Conclusion

Deflation is effective to recondition an isolated singularity.

Software available at `http://www.math.uic.edu/~jan`.