2017 SIAM Conference on Applied Algebraic Geometry

Atlanta, Georgia

Algorithm for Computing μ-Bases of Univariate Polynomials

Irina Kogan

North Carolina State University

joint work with

Hoon Hong and Zachary Hough

J. of Symbolic Comput., Vol. 8, No 3, (2017), 844 - 874

This project was supported, in part, by NSF grant CCF-1319632.

The syzygy module

- $\mathbf{a}(s) = [a_1(s), \ldots, a_n(s)] \in \mathbb{K}[s]^n$ is a univariate polynomial row vector $\boxed{\mathbf{a} \neq 0 \text{ and } n > 1}$ over a field $\mathbb{K}$.

- The syzygy module of $\mathbf{a}$ consists of column vectors in $\mathbb{K}[s]^n$, which are in the kernel of $\mathbf{a}$:

$$\mathrm{syz}(\mathbf{a}) = \{\mathbf{h} \in \mathbb{K}[s]^n \mid \mathbf{a}\,\mathbf{h} = 0\}.$$

Notation:

- n is the length of $\mathbf{a}$.

- $d = \max_{i \in \{1,\ldots,n\}} \deg a_i$ is the degree of $\mathbf{a}$.

Remark: We don't assume that $\gcd(\mathbf{a}) = 1$!

Definition of a μ-basis:

Definition: a μ-basis of $\text{syz}(\mathbf{a})$ is set of polynomial vectors $\mathbf{u}_1, \ldots, \mathbf{u}_{n-1}$, s. t.:

1. $\mathbf{u}_1, \ldots, \mathbf{u}_{n-1}$ generate $\text{syz}(\mathbf{a})$;
2. $LV(\mathbf{u}_1), \ldots, LV(\mathbf{u}_{n-1})$ are linearly independent over $\mathbb{K}$, where

   for $\mathbf{h} \in \mathbb{K}[s]^n$, such that $t = \deg \mathbf{h}$, the leading vector

$$LV(\mathbf{h}) = [\text{coeff}(h_1, t), \ldots, \text{coeff}(h_m, t)]^T \in \mathbb{K}^n.$$

Example: $\mathbf{a} = \begin{bmatrix} 1 + s^2 + s^4 & 1 + s^3 + s^4 & 1 + s^4 \end{bmatrix}$.
A μ-basis of the $\text{syz}(\mathbf{a})$ is comprised by

$$\mathbf{u}_1 = \begin{bmatrix} -s \\ 1 \\ -1 + s \end{bmatrix} \text{ and } \mathbf{u}_2 = \begin{bmatrix} 1 - 2s - 2s^2 - s^3 \\ 2 + 2s + s^2 + s^3 \\ -3 \end{bmatrix}.$$

- $\deg \mathbf{u}_1 = 1$ and $LV(\mathbf{u}_1) = [-1, 0, 1]^T$
- $\deg \mathbf{u}_2 = 3$ and $LV(\mathbf{u}_2) = [-1, 1, 0]^T$

Why are μ-bases nice?

Proposition: For a generating set $\mathbf{u}_1, \dots, \mathbf{u}_{n-1}$ of syz$(\mathbf{a})$, ordered so that deg$(\mathbf{u}_1) \leq \cdots \leq$ deg$(\mathbf{u}_{n-1})$, the following properties are equivalent:

1. [independence of the leading vectors]
   $LV(\mathbf{u}_1), \dots, LV(\mathbf{u}_{n-1})$ are independent over $\mathbb{K}$.

2. [optimality of the degrees]
   If $\mathbf{h}_1, \dots, \mathbf{h}_{n-1}$ is any generating set of syz$(\mathbf{a})$, such that
   $\deg \mathbf{h}_1 \leq \cdots \leq \deg \mathbf{h}_{n-1}$,
   then $\deg \mathbf{u}_i \leq \deg \mathbf{h}_i$ for $i = 1, \dots, n-1$.

3. [sum of the degrees]

$$\deg \mathbf{u}_1 + \cdots + \deg \mathbf{u}_{n-1} = \deg \mathbf{a} - \deg \gcd(\mathbf{a}).$$

4. [more...] [see Song and Goldman, 2009]

Remarks:

- The concept of a μ-basis was first introduced by Cox, Sederberg, Chen (1998), motivated by the search for new, more efficient methods for solving implicitization problems for rational curves, and as a further development of the method of moving lines proposed by Sederberg and Chen (1995).

- μ-basis of $\mathtt{syz}(\mathbf{a})$ is not unique, but the degrees of its elements are canonical. They where denoted by $\mu_1, \ldots, \mu_{n-1}$ in Cox, Sederberg, Chen (1998), which gave rise to the name "μ-basis".

- One can study the μ-type of $\mathbf{a}$ as in Cox and Iarrobino "Strata of rational space curves." Comput. Aided Geom. Design, 32:50–68, 2015

Algorithms to compute μ-bases

$n = 3$ algorithms:

- Cox, Sederberg and Chen (1998)
  - degrees μ_1 and μ_2 are determined prior to computing of μ-basis
  - μ-basis constructed from null vectors of two linear maps $A_1 : \mathbb{K}^{3(\mu_1+1)} \to \mathbb{K}^{\mu_1+d+1}$ and $A_2 : \mathbb{K}^{3(\mu_2+1)} \to \mathbb{K}^{\mu_2+d+1}$
  - It is not clear how to generalize to arbitrary n.
- Zheng and Sederberg (2001), Chen and Wang (2002) (Buchberger-type reduction)

arbitrary n algorithms:

- Song and Goldman (2009) (generalization of Chen and Wang to arbitrary n)
- Hong, Hough and IK (2017) (computing a "partial" reduced row-echelon form of a Sylvester-type matrix)

Main ingredients

1. Explicit isomorphism $\flat : \mathbb{K}^{n(d+1)} \to \mathbb{K}[s]_d^n$:

   Example: $n = 3$, $d = 4$

$$v = [-1, -1, 2, \quad 1, -1, 0, \quad 1, 0, 0, \quad 0, 0, 0, \quad -1, 0, 1]^T \in \mathbb{K}^{15}$$

$$\begin{aligned} v^\flat &= \begin{bmatrix} -1 \\ -1 \\ 2 \end{bmatrix} + s \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} + s^2 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + s^3 \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} + s^4 \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \\ &= \begin{bmatrix} -1 + s + s^2 - s^4 \\ -1 - s \\ 2 + s^4 \end{bmatrix}. \end{aligned}$$

Main ingredients (cont)

2. For $\mathbf{a} = \sum_{0 \le i \le d} s^i \, [c_{i1}, \ldots, c_{in}] \in \mathbb{K}[s]_d^n$ we define a Sylvester type matrix:

$$A = \begin{bmatrix} c_{01} & \cdots & c_{0n} & & & & & & & \\ \vdots & \cdots & \vdots & c_{01} & \cdots & c_{0n} & & & & \\ \vdots & \cdots & \vdots & \vdots & \cdots & \vdots & \ddots & & & \\ c_{d1} & \cdots & c_{dn} & \vdots & \cdots & \vdots & \ddots & c_{01} & \cdots & c_{0n} \\ & & & c_{d1} & \cdots & c_{dn} & \ddots & \vdots & \cdots & \vdots \\ & & & & & & \ddots & \vdots & \cdots & \vdots \\ & & & & & & & c_{d1} & \cdots & c_{dn} \end{bmatrix}$$

There are $d+1$ blocks, thus this is $(2d+1) \times n(d+1)$ matrix over $\mathbb{K}$

Example

$$\mathbf{a} = \left[\begin{array}{ccc} 2+s+s^4 & 3+s^2+s^4 & 6+2s^3+s^4 \end{array}\right]$$
$$= [2 \quad 3 \quad 6] + [1 \quad 0 \quad 0]\, s + [0 \quad 1 \quad 0]\, s^2 + [0 \quad 0 \quad 2]\, s^3 + + [1 \quad 1 \quad 1]\, s^4$$

$$A = \left[\begin{array}{ccccccccccccccc}
2 & 3 & 6 & & & & & & & & & & & & \\
1 & 0 & 0 & 2 & 3 & 6 & & & & & & & & & \\
0 & 1 & 0 & 1 & 0 & 0 & 2 & 3 & 6 & & & & & & \\
0 & 0 & 2 & 0 & 1 & 0 & 1 & 0 & 0 & 2 & 3 & 6 & & & \\
1 & 1 & 1 & 0 & 0 & 2 & 0 & 1 & 0 & 1 & 0 & 0 & 2 & 3 & 6 \\
 & & & 1 & 1 & 1 & 0 & 0 & 2 & 0 & 1 & 0 & 1 & 0 & 0 \\
 & & & & & & 1 & 1 & 1 & 0 & 0 & 2 & 0 & 1 & 0 \\
 & & & & & & & & & 1 & 1 & 1 & 0 & 0 & 2 \\
 & & & & & & & & & & & & 1 & 1 & 1
\end{array}\right]$$

is $(2d+1) \times n(d+1) = 9 \times 15$ matrix over $\mathbb{K}$

Key (but simple) observations:

$$A = \begin{bmatrix} c_{01} & \cdots & c_{0n} & & & & & & & \\ \vdots & \cdots & \vdots & c_{01} & \cdots & c_{0n} & & & & \\ \vdots & \cdots & \vdots & \vdots & \cdots & \vdots & \ddots & & & \\ c_{d1} & \cdots & c_{dn} & \vdots & \cdots & \vdots & \ddots & c_{01} & \cdots & c_{0n} \\ & & & c_{d1} & \cdots & c_{dn} & \ddots & \vdots & \cdots & \vdots \\ & & & & & & \ddots & \vdots & \cdots & \vdots \\ & & & & & & & c_{d1} & \cdots & c_{dn} \end{bmatrix}$$

- $v \in \ker A$ if and only if $v^\flat \in \mathrm{syz}_d(\mathbf{a})$.
- Let A_{*k} be a non-pivotal column: $A_{*k} = \sum_{i<k:\text{pivotal}} \alpha_i \, A_{*i}$ for some $\alpha_i \in \mathbb{K}$, then $v \in \mathbb{K}^{n(d+1)}$ such it has $-\alpha_i$ in the i-th component, 1 in the k-th component and the rest are zeros, is in the $\ker A$.
- Non-pivotal columns have n-periodic structure:

$$A_{*k} = \sum_{i<k} \alpha_i \, A_{*i} \quad \Rightarrow \quad A_{*k+n} = \sum_{i<k} \alpha_i \, A_{*i+n}$$

$$\mathbf{a} = \left[\begin{array}{ccc} 2+s+s^4 & 3+s^2+s^4 & 6+2s^3+s^4 \end{array}\right]$$

$$A = \left[\begin{array}{ccccccccccccccc}
2 & 3 & 6 & & & & & & & & & & & & \\
1 & 0 & 0 & 2 & 3 & 6 & & & & & & & & & \\
0 & 1 & 0 & 1 & 0 & 0 & 2 & 3 & 6 & & & & & & \\
0 & 0 & 2 & 0 & 1 & 0 & 1 & 0 & 0 & 2 & 3 & 6 & & & \\
1 & 1 & 1 & 0 & 0 & 2 & 0 & 1 & 0 & 1 & 0 & 0 & 2 & 3 & 6 \\
 & & & 1 & 1 & 1 & 0 & 0 & 2 & 0 & 1 & 0 & 1 & 0 & 0 \\
 & & & & & & 1 & 1 & 1 & 0 & 0 & 2 & 0 & 1 & 0 \\
 & & & & & & & & & 1 & 1 & 1 & 0 & 0 & 2 \\
 & & & & & & & & & & & & 1 & 1 & 1
\end{array}\right]$$

$p = \{1, 2, 3, 4, 5, 6, 7, 10, 13\}$, pivotal indices

$q = \{8, 9, 11, 12, 14, 15\}$, non-pivotal indices

q mod $(n = 3)$ equivalence classes: $[8, 11, 14]$ and $[9, 12, 15]$

$\tilde{q} = \{8, 9\}$, basic non-pivotal indices

μ-basis theorem (HHK 2017):

For a non-zero $\mathbf{a} \in \mathbb{K}[s]^n$

1. A has exactly $n-1$ basic non-pivotal columns.

2. The syzygies corresponding to the basic non-pivotal columns of A comprise a μ-basis of syz$(\mathbf{a})$.

$$\mathbf{a} = \left[\begin{array}{ccc} 2+s+s^4 & 3+s^2+s^4 & 6+2s^3+s^4 \end{array}\right]$$

$$A = \left[\begin{array}{ccccccccccccccc}
2 & 3 & 6 & & & & & & & & & & & & \\
1 & 0 & 0 & 2 & 3 & 6 & & & & & & & & & \\
0 & 1 & 0 & 1 & 0 & 0 & 2 & 3 & 6 & & & & & & \\
0 & 0 & 2 & 0 & 1 & 0 & 1 & 0 & 0 & 2 & 3 & 6 & & & \\
1 & 1 & 1 & 0 & 0 & 2 & 0 & 1 & 0 & 1 & 0 & 0 & 2 & 3 & 6 \\
 & & & 1 & 1 & 1 & 0 & 0 & 2 & 0 & 1 & 0 & 1 & 0 & 0 \\
 & & & & & & 1 & 1 & 1 & 0 & 0 & 2 & 0 & 1 & 0 \\
 & & & & & & & & & 1 & 1 & 1 & 0 & 0 & 2 \\
 & & & & & & & & & & & & 1 & 1 & 1
\end{array}\right]$$

$$A_{*8} = -3\,A_{*1} - 2\,A_{*2} + 2\,A_{*3} + 3\,A_{*4} - 5\,A_{*5} + 2\,A_{*6} + 1\,A_{*7},$$

$$v_8 = [3, 2, -2,\ -3, 5, -2,\ -1, 1, 0,\ 0, 0, 0,\ 0, 0, 0]^T \;\Rightarrow\; v_8^\flat = \left[\begin{array}{l} 3-3s-s^2 \\ 2+5s+s^2 \\ -2-2s \end{array}\right]$$

$$A_{*9} = -9\,A_{*1} - 8\,A_{*2} + 7\,A_{*3} + 12\,A_{*4} - 15\,A_{*5} + 5\,A_{*6} + 1\,A_{*7},$$

$$v_9 = [9, 8, -7,\ -12, 15, -5,\ -1, 0, 1,\ 0, 0, 0,\ 0, 0, 0]^T \;\Rightarrow\; v_9^\flat = \left[\begin{array}{l} 9-12s-s^2 \\ 8+15s \\ -7-5s+s^2 \end{array}\right]$$

Summary of the HHK μ-basis algorithm

Given $\mathbf{a} \in \mathbb{K}[s]^n$,

1. Construct $(2d + 1) \times n(d + 1)$ matrix A.

2. Compute "partial" reduced row-echelon E form of A, using a modified Gauss-Jordan elimination:
   (skip non-basic non-pivotal columns, stop when $n - 1$ basic non-pivotal columns are identified)

3. Read μ basis from basic non-pivotal columns.

Example

$$\mathbf{a} = \begin{bmatrix} 2 + s + s^4 & 3 + s^2 + s^4 & 6 + 2s^3 + s^4 \end{bmatrix}$$

1. $A = \begin{bmatrix} 2 & 3 & 6 & & & & & & & & & & & & \\ 1 & 0 & 0 & 2 & 3 & 6 & & & & & & & & & \\ 0 & 1 & 0 & 1 & 0 & 0 & 2 & 3 & 6 & & & & & & \\ 0 & 0 & 2 & 0 & 1 & 0 & 1 & 0 & 0 & 2 & 3 & 6 & & & \\ 1 & 1 & 1 & 0 & 0 & 2 & 0 & 1 & 0 & 1 & 0 & 0 & 2 & 3 & 6 \\ & & & 1 & 1 & 1 & 0 & 0 & 2 & 0 & 1 & 0 & 1 & 0 & 0 \\ & & & & & & 1 & 1 & 1 & 0 & 0 & 2 & 0 & 1 & 0 \\ & & & & & & & & & 1 & 1 & 1 & 0 & 0 & 2 \\ & & & & & & & & & & & & 1 & 1 & 1 \end{bmatrix}$

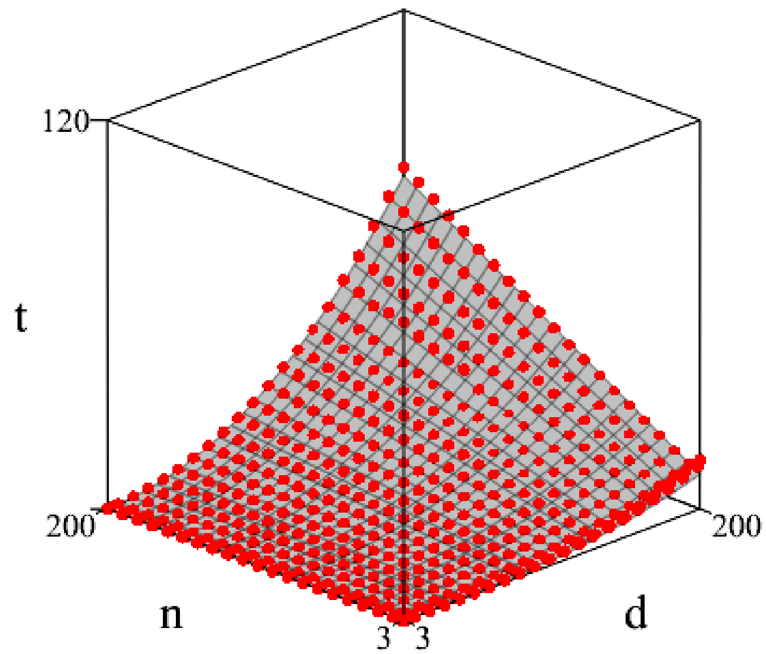
2. $E = \begin{bmatrix} 1 & & & & & & & -3 & -9 & & & & & & \\ & 1 & & & & & & -2 & -8 & & & & & & \\ & & 1 & & & & & 2 & 7 & & & & & & \\ & & & 1 & & & & 3 & 12 & 2 & 3 & 6 & & & \\ & & & & 1 & & & -5 & -15 & 1 & 0 & 0 & 2 & 3 & 6 \\ & & & & & 1 & & 2 & 5 & 0 & 1 & 0 & 1 & 0 & 0 \\ & & & & & & 1 & 1 & 1 & 0 & 0 & 2 & 0 & 1 & 0 \\ & & & & & & & & & 1 & 1 & 1 & 0 & 0 & 2 \\ & & & & & & & & & & & & 1 & 1 & 1 \end{bmatrix}$

$$M = \begin{bmatrix} 3 - 3\,s - s^2 & 9 - 12\,s - s^2 \\ 2 + 5\,s + s^2 & 8 + 15\,s \\ -2 - 2\,s & -7 - 5\,s + s^2 \end{bmatrix}$$

Comparison with Song-Goldman Algorithm

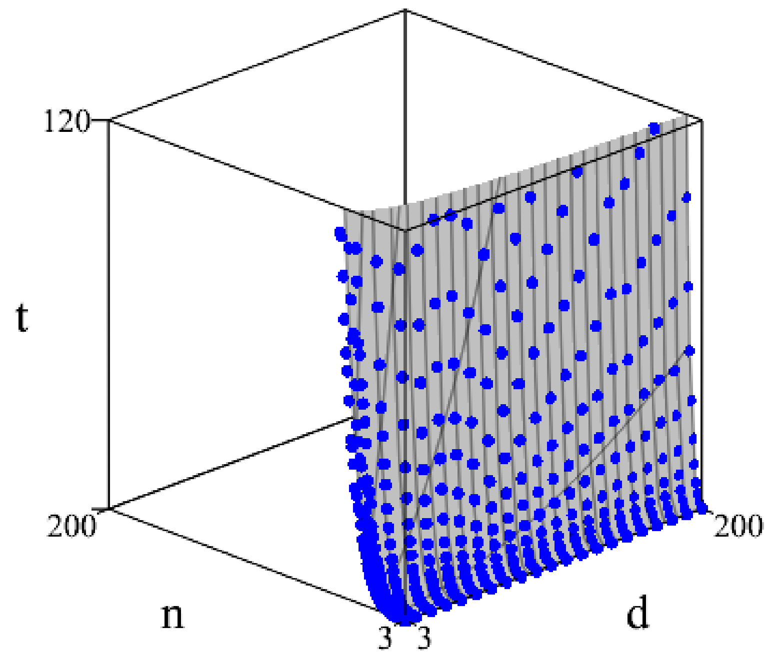
Theoretical complexity and experimental timing:

HHK algorithm:
$O(d^2n + d^3 + n^2)$

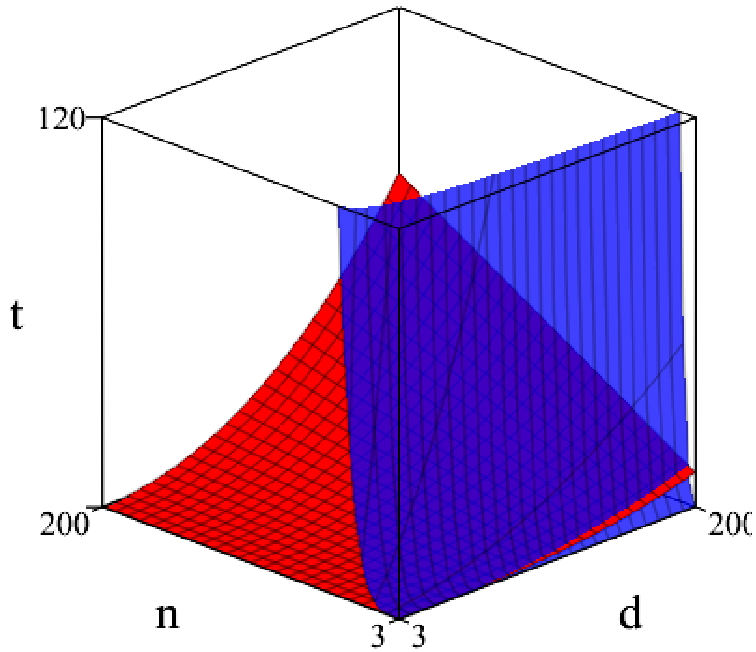


$10^{-6}\,(7.4\,d^2n + 1.2\,d^3 + 1.2\,n^2)$

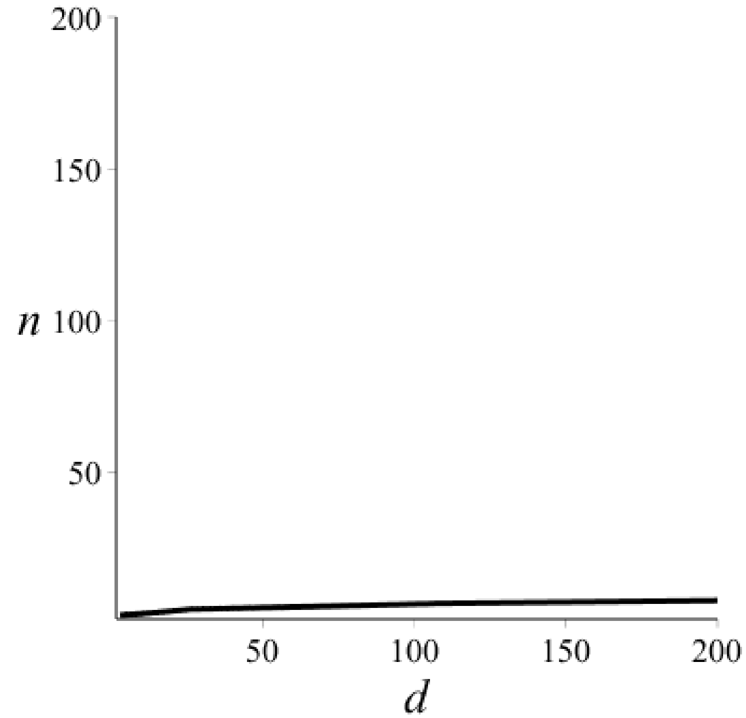
SG algorithm:
$O(dn^5 + d^2n^4)$



$10^{-7}\,(2.6\,dn^5 + 0.6\,d^2n^4)$



HHK (red) and SG (blue)



Tradeoff graph

μ-basis and $\mathsf{gcd}(\mathbf{a})$.

A μ-basis of $\mathbf{a}$ is a μ-basis of $\frac{1}{\mathsf{gcd}(\mathbf{a})}\,\mathbf{a}$

If the input vector $\mathbf{a}$ is such that $\mathsf{gcd}(\mathbf{a}) \neq 1$

- The output of the HHK algorithm is μ-basis of $\mathbf{a}$.

- The output of the SG algorithm consists of μ-basis elements multiplied by $\mathsf{gcd}(\mathbf{a})$.

μ-basis and minimal bases.

- Recall: a μ-basis $\mathbf{u}_1, \ldots, \mathbf{u}_{n-1}$ is a basis of $\ker(\mathbf{a})$, where $\mathbf{a}$ is a polynomial vector (or $1 \times n$-matrix), such that $LV(\mathbf{u}_1), \ldots, LV(\mathbf{u}_{n-1})$ are independent.

- There is a natural generalization to the problem of computing a basis $\mathbf{u}_1, \ldots, \mathbf{u}_{n-m}$ of $\ker(\mathbf{a})$, where $\mathbf{a}$ is a polynomial $m \times n$-matrix of rank m, such that $LV(\mathbf{u}_1), \ldots, LV(\mathbf{u}_{n-m})$ are independent.

- There is a body of literature on computing such bases, called minimal bases: e.g Beelen (1987), Antoniou, Vardulakis, Vologiannidis (2005), Zhou, Labahn, Storjohann (2012).

- HHK algorithm can be straightforwardly generalized for computing minimal bases. We did not yet compare this generalization with the above work.

With almost no extra cost we can modify HHK algorithm to compute:

- a minimal-degree Bézout vector

- an optimal-degree moving frame

Thank you!