Bases of relations in one or several variables: fast algorithms and applications

Vincent Neiger

ENS de Lyon – C.-P. Jeannerod, G. Villard
U. of Waterloo – É. Schost

PhD defense

November 30, 2016
Relations

polynomials $\in \mathbb{K}[X] = \mathbb{K}[X_1, \ldots, X_r]$  
submodule of $\mathbb{K}[X]^n$

\[
\begin{bmatrix}
  p_1 & \cdots & p_m \\
  \vdots & & \vdots \\
  f_1 & & f_m
\end{bmatrix} = 0 \mod M
\]

a relation (or syzygy)

elements of $\mathbb{K}[X]^n/M$ (finite dimension $D$)

$\leadsto$ relations form a submodule of $\mathbb{K}[X]^m$
Hermite-Padé approximation

Over $\mathbb{K} = \mathbb{Z}/7\mathbb{Z}$,

$$\begin{bmatrix} p_1 & p_2 & p_3 & p_4 \end{bmatrix} \begin{bmatrix} 5X^3 + 4X^2 + 6X + 4 \\ 2X^3 + X^2 + X + 3 \\ 2X + 1 \\ 4X^3 + X^2 + 4X \end{bmatrix} = 0 \mod X^4$$

trivial relation $\leadsto p = \begin{bmatrix} X^4 & 0 & 0 & 0 \end{bmatrix}$

relation of small degree $\leadsto p = \begin{bmatrix} X + 5 & 1 & 5 & 1 \end{bmatrix}$

basis of relations $\leadsto P = \begin{bmatrix} X + 2 & 0 & 6 & 0 \\ X^2 & X^2 & 0 & 0 \\ X + 2 & 3X + 2 & X & 0 \\ X + 5 & 1 & 5 & 1 \end{bmatrix}$
Bivariate interpolation with degree constraints

\( \mathcal{M} = \) polynomials vanishing at \( \{(24,80),(31,73),(15,73),(32,35),(83,66),(27,46),(20,91),(59,64)\} \)

\[ \begin{align*}
M &= (X - 24) \cdots (X - 59) \\
L &= \text{Lagrange interpolant}
\end{align*} \quad \rightarrow \quad \mathcal{M} = \langle M(X), Y - L(X) \rangle
\]

Degree constraints: \( p(X, Y) = p_0(X) + p_1(X) Y + p_2(X) Y^2 \)

\[ \begin{align*}
\deg &\leq 4 \\
\deg &\leq 2 \\
\deg &\leq 0
\end{align*} \]

Equation: \( p(X, Y) \equiv 0 \mod \mathcal{M} \quad \Leftrightarrow \quad [p_0 \ p_1 \ p_2] \begin{bmatrix} 1 \\ L \\ L^2 \end{bmatrix} = 0 \mod M(X) \)

\[ p(X, Y) = (2X^4 + 56X^3 + 42X^2 + 48X + 15) + (72X^2 + 12X + 30)Y + Y^2 \]
Basis of relations

Problem: given $\mathcal{M}$ and $\mathbf{f}$,

- compute a basis of the module of relations $\mathcal{R}$
- with nice properties: unique, minimal degrees, computing mod $\mathcal{R}$, ...

**univariate**

shift $s = (s_1, \ldots, s_m) \in \mathbb{Z}^m$

s-Popov basis

Hermite:

$$\begin{bmatrix} X^2 + 3X + 2 \\ 5X + 6 \\ 4X \\ 3 \\ 1 \end{bmatrix}$$

Popov:

$$\begin{bmatrix} X & 6 & 2 \\ 6 & X + 6 & 4 \\ 2 & 5 & X + 5 \end{bmatrix}$$

**multivariate**

monomial order $\prec$ on $\mathbb{K}[X]^m$

$\prec$-Gröbner basis

```
Y
```

```
Z
```

```
X
```
Basis of relations

\[ pf = 0 \mod M \]
knowing multiplication matrices

**Hermite-Padé approximation**

\[ pf = 0 \mod X^D \]

**Multivariate interpolation**

\[ \leadsto \text{list-decoding algorithms} \]
\[ p(x_i, y_i) = 0 \text{ for } 1 \leq i \leq D \]
Basis of relations

\[ \text{pf} = 0 \mod M \]

knowing multiplication matrices

Change of monomial order

\[ \leadsto \text{polynomial system solving} \]

\[ \prec_1 \text{-GB of } M \leadsto \prec_2 \text{-GB of } M \]

Hermite-Padé approximation

\[ \text{pf} = 0 \mod X^D \]

Multivariate interpolation

\[ \leadsto \text{list-decoding algorithms} \]

\[ p(x_i, y_i) = 0 \text{ for } 1 \leq i \leq D \]

Normal forms of matrices

\[ M \in \mathbb{K}[X]^{m \times m} \xrightarrow{\text{unimodular}} P \]
Basis of relations

\[ \text{pf} = 0 \mod \mathcal{M} \]

knowing multiplication matrices

Hermite-Padé approximation

\[ \text{pf} = 0 \mod X^D \]

Multivariate interpolation

\[ p(x_i, y_i) = 0 \text{ for } 1 \leq i \leq D \]

Change of monomial order

\[ \leadsto \text{polynomial system solving} \]

\[ \prec_1\text{-GB of } \mathcal{M} \rightarrow \prec_2\text{-GB of } \mathcal{M} \]

Normal forms of matrices

\[ \mathbb{K}[X]^{m \times m} \xrightarrow{\text{unimodular}} \mathcal{P} \]
V = \mathbb{K}[X_1, \ldots, X_r]^n/M is a \mathbb{K}\text{-vector space of dimension } D

**Linear algebra viewpoint:**

- matrix \( E = \begin{bmatrix} e_1 \\ \vdots \\ e_m \end{bmatrix} \in \mathbb{K}^{m \times D} \) (equation \( \begin{bmatrix} f_1 \\ \vdots \\ f_m \end{bmatrix} \in \mathbb{V}^{m \times 1} \))
- matrix \( M_i \in \mathbb{K}^{D \times D}, \ 1 \leq i \leq r \) (multiplying by \( X_i \) in \( \mathbb{V} \))

\[
[p_1 \ \cdots \ \ p_m] \begin{bmatrix} f_1 \\ \vdots \\ f_m \end{bmatrix} = \sum_{j,i} \alpha_{j,i} X_1^{j_1} \cdots X_r^{j_r} f_i
\]

relation = \( \mathbb{K}\text{-linear} \) relation between \( \{ e_i M_1^{j_1} \cdots M_r^{j_r} \}_{\mathbb{K}^{1 \times D}} \)
Bases of relations via linear algebra

basis of relations = subset of nullspace of multi-Krylov matrix

≺

\text{order:}

\begin{equation}
\begin{bmatrix}
E \\
EM_1 \\
\vdots \\
EM_D
\end{bmatrix}
\begin{bmatrix}
E \\
EM_1 \\
\vdots \\
EM_D
\end{bmatrix}
\begin{bmatrix}
M_2 \\
M_2^D
\end{bmatrix}
\end{equation}

\omega:

D \times D \text{ matrix multiplication in } \mathcal{O}(D\omega) \text{ operations}

• [Keller-Gehrig, 1985]: charpoly\((M)\) in \mathcal{O}(D\omega \log(D)) \text{ (one variable, } E=\text{Id}, \text{ output = Hermite)}

• [FGLM, 1993]: general case in \mathcal{O}(rD^3)

• [Beckermann&Labahn, 2000]: \mathcal{O}(mD^2) \text{ for structured } M \text{ (one variable)}

• [Faug`ere et al., 2014]: for \text{\textless\textless\textless lex and Shape position, } \mathcal{O}(D\omega \log(D) + rM(D) \log(D))

General case with fast matrix multiplication?
Bases of relations via linear algebra

basis of relations = subset of nullspace of multi-Krylov matrix

≺\text{top}_{\text{lex}}\text{ order: }\omega: D \times D \text{ matrix multiplication in } \mathcal{O}(D^\omega) \text{ operations}

- [Keller-Gehrig, 1985]: \text{charpoly}(\mathbf{M}) \text{ in } \mathcal{O}(D^\omega \log(D)) \text{ (one variable, } \mathbf{E} = \text{Id}, \text{ output } = \text{Hermite})
- [FGLM, 1993]: general case in \mathcal{O}(rD^3)
- [Beckermann&Labahn, 2000]: \mathcal{O}(mD^2) \text{ for structured } \mathbf{M} \text{ (one variable)}
- [Faugère et al., 2014]: for ≺_{\text{lex}} \text{ and Shape position, } \mathcal{O}(D^\omega \log(D) + rM(D) \log(D))

? General case with fast matrix multiplication?
General algorithm incorporating fast linear algebra

Algorithm:
1. compute monomial basis = first independent rows
2. find $\prec$-Gröbner basis by nullspace computation

Difficulty: incorporate fast multiplication in 1. for any $\prec$

- $X_1, \ldots, X_r \rightsquigarrow$ gather operations involving $M_i$
- $X_i, X_i^2, X_i^4, \ldots \rightsquigarrow$ gather operations involving $M_i^{2j}$ as if $\prec_{lex}^{top}$
- insert new rows according to the order $\prec$

Cost bound: $O(rD^\omega \log(D))$ field operations

Size of dense representations: \[
\begin{array}{c|c}
\text{input} & \text{output} \\
\hline
rD^2 & \leq rD^2 \\
\end{array}
\]
Change of monomial order

**Problem:** $\prec_1$-GB of $\mathcal{M} \rightarrow \prec_2$-GB of $\mathcal{M}$

$= \prec_2$-GB of relations: $p_1 = 0 \mod \mathcal{M}$

**Approach:** [FGLM, 1993]

1. compute $M_1, \ldots, M_r$ from $\prec_1$-GB
2. compute the $\prec_2$-GB of relations

**Result (case of ideals):**

step 1. in $O(rD^\omega \log(D))$

assuming the $\prec_1$-initial ideal is Borel-fixed

$\leadsto$ extends [Faugère et al., 2014]
Change of monomial order

\[ \rightsquigarrow \text{polynomial system solving} \]

\[ \prec_1 \text{-GB of } \mathcal{M} \longrightarrow \prec_2 \text{-GB of } \mathcal{M} \]

Basis of relations

\[ pf = 0 \mod \mathcal{M} \]

knowing multiplication matrices

Hermite-Padé approximation

\[ pf = 0 \mod X^D \]

Multivariate interpolation

\[ \rightsquigarrow \text{list-decoding algorithms} \]

\[ p(x_i, y_i) = 0 \text{ for } 1 \leq i \leq D \]

Normal forms of matrices

\[ \mathbf{M} \in \mathbb{K}[X]^{m \times m} \xrightarrow{\text{unimodular}} \mathbf{P} \]
Approximant bases: divide-and-conquer via multiplication

\[ m \times n \text{ matrix of degree } < d = D/n \]

\[ \begin{bmatrix} X^d \\ \vdots \\ X^d \end{bmatrix} \]

module \( M = X^d K[X]^n \)

\[ p_f = 0 \mod M \]

\[ P^{(1)} := s\text{-reduced for } f \text{ and } d/2 \]
\[ g \text{ and } t := \text{update } f \text{ and } s \]
\[ P^{(2)} := t\text{-reduced for } g \text{ and } d/2 \]
\[ \text{return } P^{(2)} P^{(1)} \]

\( P^{(1)} \) and \( P^{(2)} \) matrices \( m \times m \) of degree \( d/2 \)

Cost bound: \( O(m^{\omega} M(d) \log(d)) \)

\( \sim \) very efficient in \textbf{balanced case} \((n \approx m)\)

size of input: \( mnd \), with \( n \in O(m) \)
Approximant bases: degree control assuming small shifts

very efficient algorithm in **balanced case**

\[ \rightarrow \] difficulty for improvements: controlling the output degrees

**Example:**

\[
\begin{bmatrix}
  f_1 \\
  f_2 \\
  f_3 \\
  f_4
\end{bmatrix} = \begin{bmatrix}
  1 \\
  1 + X \\
  X + X^2 \\
  X^2 + X^3
\end{bmatrix} \quad \rightarrow \quad \begin{bmatrix}
  X + 1 & -1 \\
  X - 1 & X + 1 & -1 \\
  X + 1 & X - 1 & X + 1 & -1 \\
  X^{125} & X^{125} & -X^{125} & X^{125}
\end{bmatrix}
\]

**Assume** \( s = (s_1, \ldots, s_m) \) **almost uniform**

\[ \Rightarrow \text{average row degree } \mathcal{O}(nd/m) \]

\[ \Rightarrow \text{output size } = \text{input size } = \mathcal{O}(mnd) \]

if \( n \) is small, \( \mathcal{O}(m^\omega d) \) not satisfactory

**Under this assumption:** \( \mathcal{O}(m^{\omega-1} nd) \) [Zhou-Labahn, 2012]

using Storjohann’s transformations [Storjohann, 2006]

to rely on **balanced case**
Approximant bases: degree control via normalized basis

What about arbitrary shifts? (e.g. Hermite?)

Example: \( s = (0, 0, 0, 0, d, d, d, d) \), same \( f_1, f_2, f_3, f_4 \) / random \( f_5, f_6, f_7, f_8 \)

Degrees in \( s\)-reduced basis:

\[
\begin{bmatrix}
1 & 0 \\
1 & 1 & 0 \\
1 & 1 & 1 & 0 \\
125 & 125 & 125 & 125 \\
124 & 124 & 124 & 124 & 0 \\
124 & 124 & 124 & 124 & 0 \\
124 & 124 & 124 & 124 & 0 \\
124 & 124 & 124 & 124 & 0 \\
\end{bmatrix}
\]

size \( m^2d \)

Degrees in \( s\)-Popov basis:

\[
\begin{bmatrix}
1 & 0 \\
0 & 1 & 0 \\
0 & 0 & 1 & 0 \\
0 & 0 & 0 & 125 \\
0 & 0 & 0 & 124 & 0 \\
0 & 0 & 0 & 124 & 0 \\
0 & 0 & 0 & 124 & 0 \\
0 & 0 & 0 & 124 & 0 \\
\end{bmatrix}
\]

size \( mnd \)

size of normalized basis: \( O(mnd) \) independently of the shift

s-Popov basis: aim & means
Approximant bases: degree control via normalized basis

Degree control $\leadsto$ compute the $s$-Popov basis $P$

- $s$-Popov not compatible with multiplication
- size of product beyond target cost

$\leadsto$ change how to combine $P^{(1)}$ and $P^{(2)}$

diagonal degrees of $P$: $\delta = \delta^{(1)} + \delta^{(2)}$

knowing $\delta$, reduce to balanced case
Approximant basis: general fast algorithm

Diagonal degrees $\delta \Rightarrow \{\begin{align*}
s\text{-Popov basis} &= -\delta\text{-Popov basis} \\
-\delta &\text{ almost uniform}
\end{align*}\}

1. $B := -\delta\text{-reduced basis}$ (via [Storjohann, 2006] + balanced case)
2. $P :=$ normalize $B$ into $-\delta\text{-Popov basis}$ (constant transformation)

Result: $O(m^\omega M(D/m) \log(D)^2) \subseteq O^\sim(m^{\omega-1}D)$

- arbitrary shift $s$
- arbitrary orders $pf = 0 \mod \begin{bmatrix} X^{d_1} \\ \vdots \\ X^{d_n} \end{bmatrix}$
- returning $s\text{-Popov basis}$

$D := d_1 + \cdots + d_n$
Change of monomial order
\[ \sim \text{polynomial system solving} \]
\[ \prec_1 \text{-GB of } M \rightarrow \prec_2 \text{-GB of } M \]

Basis of relations
\[ pf = 0 \text{ mod } M \]
knowing multiplication matrices

Hermite-Padé approximation
\[ pf = 0 \text{ mod } X^D \]

Multivariate interpolation
\[ \sim \text{list-decoding algorithms} \]
\[ p(x_i, y_i) = 0 \text{ for } 1 \leq i \leq D \]

Normal forms of matrices
\[ M \in \mathbb{K}[X]^{m \times m} \xrightarrow{\text{unimodular}} P \]
List-decoding Reed-Solomon codes

Reliable delivery of data over an unreliable communication channel

\[ w = w_0 + \cdots + w_k X^k \xrightarrow{\text{encoding}} (w(x_1), \ldots, w(x_D)) \xrightarrow{\text{noise}} (y_1, \ldots, y_D) \]

Few errors during transmission: \( w(x_i) = y_i \) for many \( i \)'s

Retrieve \( w \) via bivariate interpolation + root finding [Guruswami-Sudan, 1999]

\[ p(x_i, y_i) = 0 \]
\[ \text{small degree } p(X, Y) \]
\[ \implies p(X, w(X)) = 0 \]
From bivariate interpolation to univariate relation

Constrained **bivariate** interpolation: \( p(x_i, y_i) = 0 \) for all \( i \)

- **Y-constraint**: \( \deg_Y < m \) \( \Rightarrow \) **univariate** relation

\[
\begin{bmatrix}
p_0 & p_1 & \cdots & p_{m-1}
\end{bmatrix}
\begin{bmatrix}
1 \\
L \\
\vdots \\
L^{m-1}
\end{bmatrix}
= 0 \mod (X - x_1) \cdots (X - x_D)
\]

- **X-constraint**: satisfied via s-Popov

→ **Generalization of approximants**:

relations modulo \( \mathcal{M} = \begin{bmatrix}
m_1 \\
\vdots \\
m_n
\end{bmatrix} \)

- \( m_1, \ldots, m_n \) split over \( \mathbb{K} \)
- roots and multiplicities are known
Interpolation: generalizing approximation techniques

Generalizations:
- update $g := P^{(1)}f \mod (m_1, \ldots, m_n) \rightsquigarrow$ fast via CRT
- divide and conquer via multiplication
- divide and conquer via “find & use degrees”

Efficiency: generalization of the **balanced case**?

Fact: degree of output at most $\text{lcm}(m_1, \ldots, m_n)$

\[
\begin{align*}
\mod X^D & \quad \rightarrow \quad \mod \left[ \begin{array}{c}
X^d \\
\cdots \\
X^d
\end{array} \right] \\
\mod(X - x_1) \cdots (X - x_D) & \quad \rightarrow \quad \mod \left[ \begin{array}{c}
\prod_i(X - x_i^{(1)}) \\
\cdots \\
\prod_i(X - x_i^{(n)})
\end{array} \right]
\end{align*}
\]
Interpolation: controlling the degrees

No “balanced case”, yet
known diagonal degrees ⇒ almost uniform shift ⇒ small output degrees

- shift modified in recursive calls
- shift may become far from uniform
  ⇝ intermediate bases may have large degrees

change shift processing to keep it uniform:
  - all recursive calls with uniform shift
  - correction via change of shift

efficiency: fast kernel basis [Zhou et al., 2012]
  ⇝ fast algorithm for s almost uniform

\[ P^{(1)} := \text{0-reduced} \]
\[ g \text{ and } t := \cdots \]
\[ P^{(2)} := \text{0-reduced} \]
\[ \text{Shift}(P^{(2)}, t) \]
\[ \text{return } P^{(2)}P^{(1)} \]
Bases of interpolants: results

- arbitrary shift $s$
- arbitrary diagonal with known linear factors
- returning s-Popov basis

\[
\begin{bmatrix}
    m_1 \\
    \vdots \\
    m_n
\end{bmatrix}
\]

\[
pf = 0 \mod \begin{bmatrix} m_1 & \cdots & m_n \end{bmatrix}
\]

\[
D := \deg(m_1) + \cdots + \deg(m_n)
\]

Cost bound: \( O(m^{\omega-1}M(D) \log(D)^3) \subseteq O^\sim(m^{\omega-1}D) \)

Improves upon previous algorithms:
- based on fast basis reduction [Cohn-Heninger, 2011+2012]

\(\sim\) list- and soft-decoding of Reed-Solomon codes
\(\sim\) robust Private Information Retrieval [Devet-Goldberg-Heninger, 2012]
Change of monomial order
\[ \leadsto \text{polynomial system solving} \]
\[ \prec_{1}\text{-GB of } M \rightarrow \prec_{2}\text{-GB of } M \]

Basis of relations
\[ pf = 0 \bmod M \]
knowing multiplication matrices

Hermite-Padé approximation
\[ pf = 0 \bmod X^D \]

Multivariate interpolation
\[ \leadsto \text{list-decoding algorithms} \]
\[ p(x_i, y_i) = 0 \text{ for } 1 \leq i \leq D \]

Normal forms of matrices
\[ M \in \mathbb{K}[X]^{m \times m} \xrightarrow{\text{unimodular}} P \]
Normal forms of polynomial matrices

Problem: any basis \( \mathbf{M} \) of \( \mathcal{M} \) \( \rightarrow \) s-Popov basis of \( \mathcal{M} \)

\[ = \text{s-Popov basis of relations } \mathbf{p} \cdot \mathbf{Id} = 0 \mod \mathbf{M} \]

Input: \( m \times m \) matrix

of degree \( \leq d \) \( \rightarrow \) size \( m^2 d \)

of generic det. degree \( \leq D_{\text{gen}} \) \( \rightarrow \) size \( mD_{\text{gen}} \)

Fast algorithm for arbitrary shifts?

Previous work:

- Popov: \( \mathcal{O}^\sim(m^\omega d) \) \[\text{[Giorgi et al., 2003]} \quad \text{[Gupta et al., 2011+2012]}\]
- Hermite: \( \mathcal{O}^\sim(m^\omega d) \) Las Vegas \[\text{[Gupta-Storjohann 2011]}\]
- Hermite: \( \mathcal{O}^\sim(m^{\omega-1} D_{\text{gen}}) \) \[\text{[with G. Labahn & W. Zhou]}\]
Reduction of basis matrix

Popov form

shifted Popov form

Hermite form

Reconstruction from equations

\[
\begin{bmatrix}
p_1 & \cdots & p_m
\end{bmatrix}
\begin{bmatrix}
f_{11} & \cdots & f_{1n} \\
\vdots & & \vdots \\
f_{m1} & \cdots & f_{mn}
\end{bmatrix}
= 0 \mod
\begin{bmatrix}
m_1 \\
m_2 \\
\vdots \\
m_n
\end{bmatrix}
\]

High-order lifting [Storjohann, 2003]

\[\deg(P) \leq d\]

\(P\) triangular
Normal forms via bases of relations

Compute the Smith form \( UMV = \text{diag}(m_1, \ldots, m_n) \)

\[
p \cdot \text{Id} = 0 \mod M \iff pV = 0 \mod \begin{bmatrix} m_1 & \cdots & m_n \end{bmatrix}
\]

cost: \( O^\sim(m^{\omega^{-1}}D_{\text{gen}}) \) Las Vegas \hspace{1cm} [Storjohann, 2003] [Gupta et al., 2012]

\( \sim \) it remains to compute the \( s \)-Popov basis of relations

generalization of approximants/interpolants:
\( M = \text{arbitrary diagonal matrix} \) (unknown roots, if any)

generalizing divide and conquer “find & use degrees”
\( \Rightarrow \) remains one equation (\( f \) column vector)

\( p f = 0 \mod m \)

\( ? \) arbitrary \( m \): how to divide and conquer?
Basis of relations for an arbitrary diagonal matrix

\[ pf = 0 \mod m \iff \begin{bmatrix} p & q \end{bmatrix} \begin{bmatrix} f \\ m \end{bmatrix} = 0 \text{ for some quotient } q \]

⇒ difficulty: fast s-Popov kernel basis of a column vector

Base case: s almost uniform ⇒ via approximant basis in \( O^\sim(m^{\omega-1} D) \)

New divide and conquer approach based on finding a “splitting index”

\[ s \rightarrow (s^{(1)}, s^{(2)}) \text{ with half the amplitude} \]

Result: \( O^\sim(m^{\omega-1} D) \)

- arbitrary shift s
- arbitrary diagonal
- returning s-Popov basis

\[ pf = 0 \mod \begin{bmatrix} m_1 \\ \vdots \\ m_n \end{bmatrix} \]

assumption \( n \in O(m) \)
Summary and perspectives

Results:

- Gröbner basis of relations
- Change of monomial order
- $s$-Popov basis of relations
- Normal form of $\mathbb{K}[X]$ matrix

\[
\begin{align*}
\text{cost bound} & \quad \text{i/o size} \\
O^\sim(rD^\omega) & \quad rD^2 \\
O^\sim(m^{\omega-1}D) & \quad mD
\end{align*}
\]

Perspectives and open questions:

- Implementation in Linbox (C++)
- Fast $s$-Popov kernel basis
- Deterministic Smith form
- Deterministic characteristic polynomial in $O(m^\omega)$
- Unconditional fast change of monomial order
- Exploiting double structure in bivariate interpolation