Computing Krylov iterates in the time of matrix multiplication

Vincent Neiger, Clément Pernet, Gilles Villard

ISSAC, Raleigh, NC, USA.
July 19, 2024

Sorbonne Université Paris, Grenoble INP – UGA, CNRS, ÉNS de Lyon, France

Krylov matrix and Krylov basis

Definition (Krylov matrix of a single vector)

The Krylov matrix of a matrix $\mathbf{A} \in \mathbb{K}^{n \times n}$ and a vector $\mathbf{u} \in \mathbb{K}^n$, at order d, is

$$\mathbf{K}_d(\mathbf{A}, \mathbf{u}) = \begin{bmatrix} \mathbf{u} & \mathbf{A}\mathbf{u} & \dots & \mathbf{A}^{d-1}\mathbf{u} \end{bmatrix}.$$

Krylov matrix and Krylov basis

Definition (Krylov matrix of a single vector)

The Krylov matrix of a matrix $\mathbf{A} \in \mathbb{K}^{n \times n}$ and a vector $\mathbf{u} \in \mathbb{K}^{n}$, at order d, is

$$\mathbf{K}_d(\mathbf{A}, \mathbf{u}) = \begin{bmatrix} \mathbf{u} & \mathbf{A}\mathbf{u} & \dots & \mathbf{A}^{d-1}\mathbf{u} \end{bmatrix}.$$

Krylov basis: the one for the maximal d s.t. $\mathbf{K}_d(\mathbf{A}, \mathbf{u})$ has full-rank

$\rightsquigarrow$ a basis of Span$(\{\mathbf{A}^i\mathbf{u}, i \in \mathbb{N}\})$

Motivation

For a Krylov basis $\mathbf{K} = \mathbf{K}_d(\mathbf{A}, \mathbf{u})$, $\quad \mathbf{AK} = \mathbf{K} \underbrace{\begin{bmatrix} 0 & & & c_0 \\ 1 & & & c_1 \\ & \ddots & & \vdots \\ & & 1 & c_{d-1} \end{bmatrix}}_{\mathbf{C}_f}$

Motivation

For a Krylov basis $\mathbf{K} = \mathbf{K}_d(\mathbf{A}, \mathbf{u})$, $\quad \mathbf{AK} = \mathbf{K} \underbrace{\begin{bmatrix} 0 & & & c_0 \\ 1 & & & c_1 \\ & \ddots & & \vdots \\ & & 1 & c_{d-1} \end{bmatrix}}_{\mathbf{C}_f}$

Minimal and characteristic polynomials

- $f = X^d - c_{d-1}X^{d-1} - \cdots - c_0 = \text{MinPoly}(\mathbf{A})$ w.h.p.
- If $d = n$, then $\mathbf{K}^{-1}\mathbf{AK} = \mathbf{C}_f$ and $f = \text{MinPoly}(\mathbf{A}) = \text{CharPoly}(\mathbf{A})$

Generalization to multiple vectors

Definition (Krylov matrix of multiple vectors)

Krylov matrix of $\mathbf{A} \in \mathbb{K}^{n \times n}$, $\mathbf{U} = \begin{bmatrix} \mathbf{u}_1 & \ldots & \mathbf{u}_m \end{bmatrix} \in \mathbb{K}^{n \times m}$ and $\mathbf{d} = (d_1, \ldots, d_m)$:

$$\mathbf{K}_{\mathbf{d}}(\mathbf{A}, \mathbf{U}) = \left[\; \mathbf{K}_{d_1}(\mathbf{A}, \mathbf{u_1}) \;\middle|\; \mathbf{K}_{d_2}(\mathbf{A}, \mathbf{u_2}) \;\middle|\; \ldots \;\middle|\; \mathbf{K}_{d_m}(\mathbf{A}, \mathbf{u_m}) \; \right].$$

Generalization to multiple vectors

Definition (Krylov matrix of multiple vectors)

Krylov matrix of $\mathbf{A} \in \mathbb{K}^{n\times n}$, $\mathbf{U} = \begin{bmatrix} \mathbf{u}_1 & \dots & \mathbf{u}_m \end{bmatrix} \in \mathbb{K}^{n\times m}$ and $\mathbf{d} = (d_1, \dots, d_m)$:

$$\mathbf{K_d}(\mathbf{A}, \mathbf{U}) = \left[\; \mathbf{K}_{d_1}(\mathbf{A}, \mathbf{u_1}) \;\middle|\; \mathbf{K}_{d_2}(\mathbf{A}, \mathbf{u_2}) \;\middle|\; \dots \;\middle|\; \mathbf{K}_{d_m}(\mathbf{A}, \mathbf{u_m}) \; \right].$$

Krylov basis: the one for $\mathbf{d}$ lexicographically maximal s.t. $\mathbf{K_d}(\mathbf{A}, \mathbf{U})$ has full-rank

$\rightsquigarrow$ a basis of Span$(\{\mathbf{A}^i \mathbf{u}_j, i \in \mathbb{N}, j \in \{1..m\}\})$

Motivation

For a Krylov basis $\mathbf{K} = \mathbf{K_d}(\mathbf{A}, \mathbf{U})$, $\quad \mathbf{A}\mathbf{K} = \mathbf{K} \underbrace{\begin{bmatrix} \mathbf{C}_{f_1} & * & \cdots & * \\ & \mathbf{C}_{f_2} & & * \\ & & \ddots & \vdots \\ & & & \mathbf{C}_{f_m} \end{bmatrix}}_{\mathbf{H}}$

Motivation

For a Krylov basis $\mathbf{K} = \mathbf{K_d}(\mathbf{A}, \mathbf{U})$, $\quad \mathbf{AK} = \mathbf{K} \underbrace{\begin{bmatrix} \mathbf{C}_{f_1} & * & \cdots & * \\ & \mathbf{C}_{f_2} & & * \\ & & \ddots & \vdots \\ & & & \mathbf{C}_{f_m} \end{bmatrix}}_{\mathbf{H}}$

Invariant subspace decomposition

If $\sum_i d_i = n$, then $\mathbf{K}^{-1}\mathbf{AK} = \mathbf{H}$

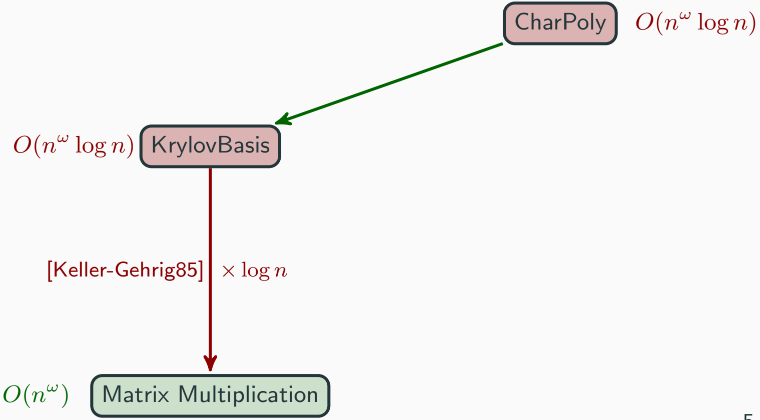
- CharPoly$(\mathbf{A}) = \prod_i f_i$
- Diag$(f_1, \ldots, f_m)$ is the Frobenius normal form of $\mathbf{A}$ w.h.p. (w.r.t. the choice of $\mathbf{U}$)
- the f_i are the invariant factors of $\mathbf{A}$ w.h.p. (w.r.t. the choice of $\mathbf{U}$)

State of the Art and open questions

- Deterministic algorithms
- Dense matrices

$$\begin{tikzcd} O(n^\omega \log n) & \boxed{\text{KrylovBasis}} \arrow[d, "\text{[Keller-Gehrig85]}"', "\times \log n"] \\ O(n^\omega) & \boxed{\text{Matrix Multiplication}} \end{tikzcd}$$

State of the Art and open questions



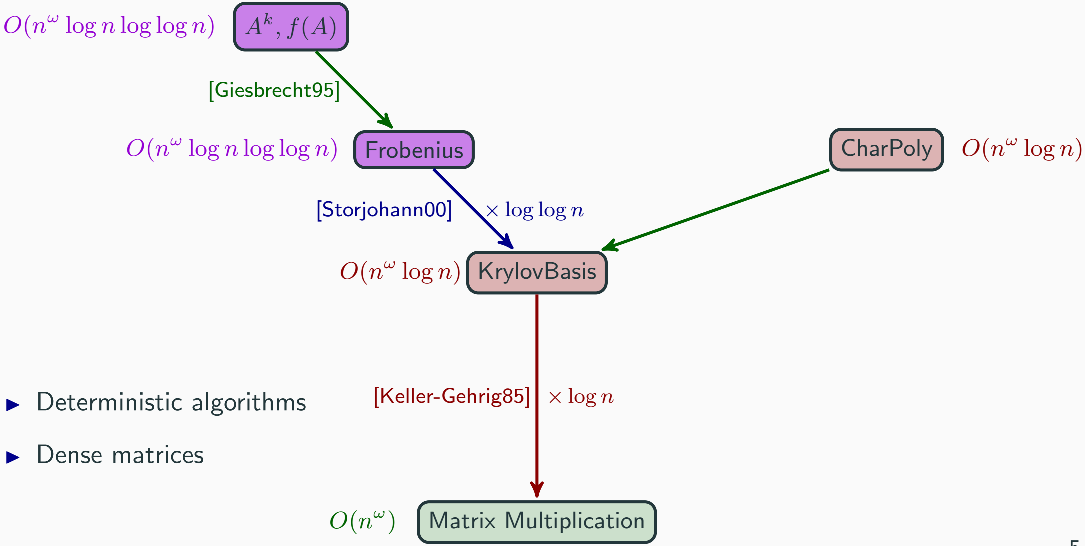
- Deterministic algorithms
- Dense matrices

State of the Art and open questions

$$\begin{tikzcd}
O(n^\omega \log n \log\log n)\ \boxed{\text{Frobenius}} \arrow[dr, "{[\text{Storjohann00}]}"', "\times \log\log n"] & & \boxed{\text{CharPoly}}\ O(n^\omega \log n) \arrow[dl] \\
& O(n^\omega \log n)\ \boxed{\text{KrylovBasis}} \arrow[d, "{[\text{Keller-Gehrig85}]}"', "\times \log n"] & \\
& O(n^\omega)\ \boxed{\text{Matrix Multiplication}} &
\end{tikzcd}$$

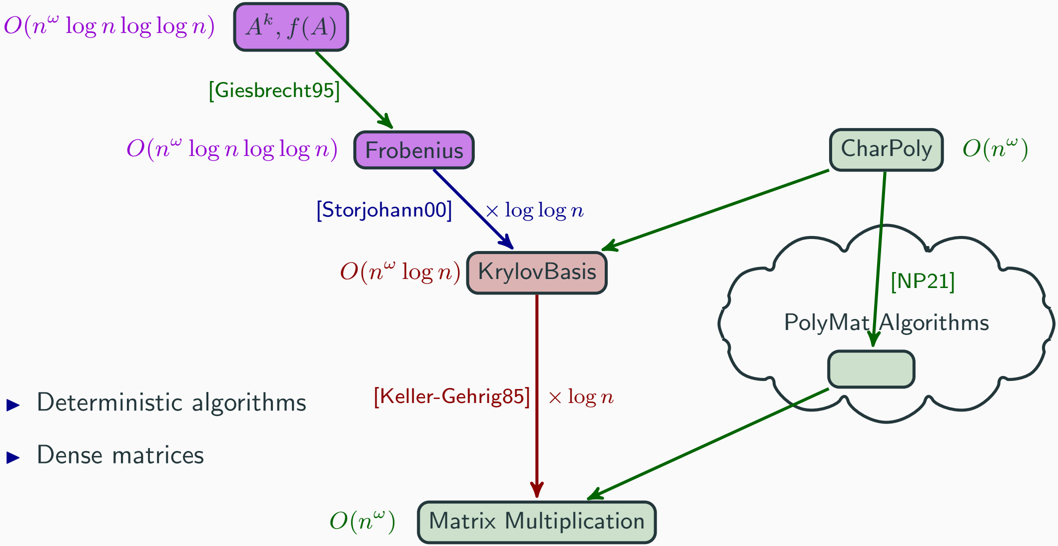
- Deterministic algorithms
- Dense matrices

State of the Art and open questions



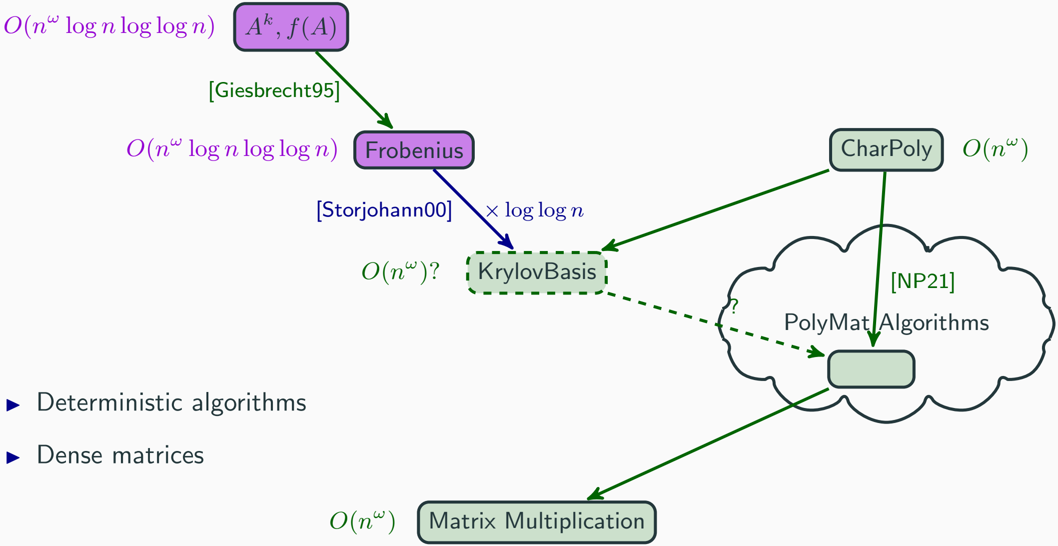
- Deterministic algorithms
- Dense matrices

State of the Art and open questions



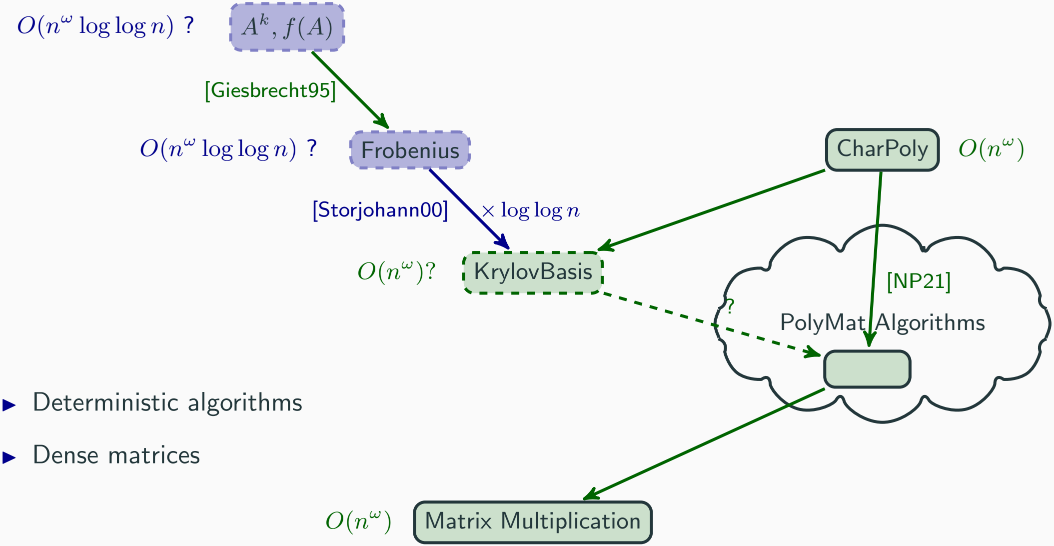
- Deterministic algorithms
- Dense matrices

State of the Art and open questions



- Deterministic algorithms
- Dense matrices

State of the Art and open questions



- Deterministic algorithms
- Dense matrices

Contributions

Theorem

A Krylov basis $\mathbf{K}_{\mathbf{d}}(\mathbf{A}, \mathbf{U})$ *of* m *vectors (i.e.* $\mathbf{U} \in \mathbb{K}^{n \times m}$*) can be computed*

1. *in* $O(n^\omega)$ *field operations* *if* $m \in O(n/log(n)^e)$ *for a constant* $e > 0$
2. *in* $O(n^\omega \log \log n)$ *field operations* *if* $m \in O(n)$

Contributions

Theorem

A Krylov basis $\mathbf{K_d}(\mathbf{A}, \mathbf{U})$ *of* m *vectors (i.e.* $\mathbf{U} \in \mathbb{K}^{n\times m}$*) can be computed*

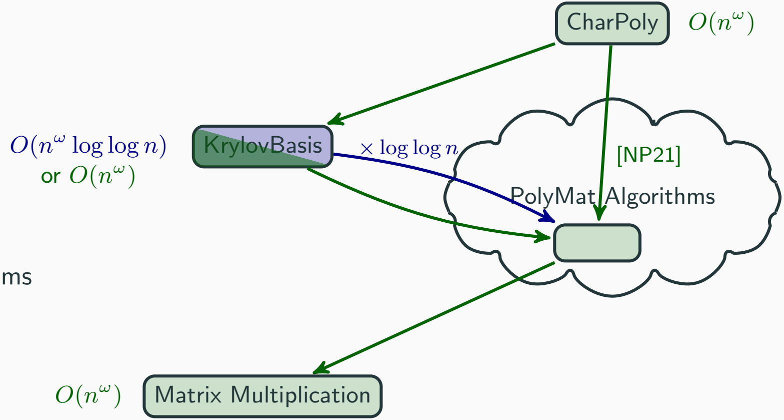
1. *in* $O(n^\omega)$ *field operations* — *if* $m \in O(n/log(n)^e)$ *for a constant* $e > 0$
2. *in* $O(n^\omega \log\log n)$ *field operations* — *if* $m \in O(n)$

Corollary (from [Storjohann'00], [Giesbrecht'95])

There is a deterministic algorithm using $O(n^\omega(\log\log n)^2)$ *field operations to compute*

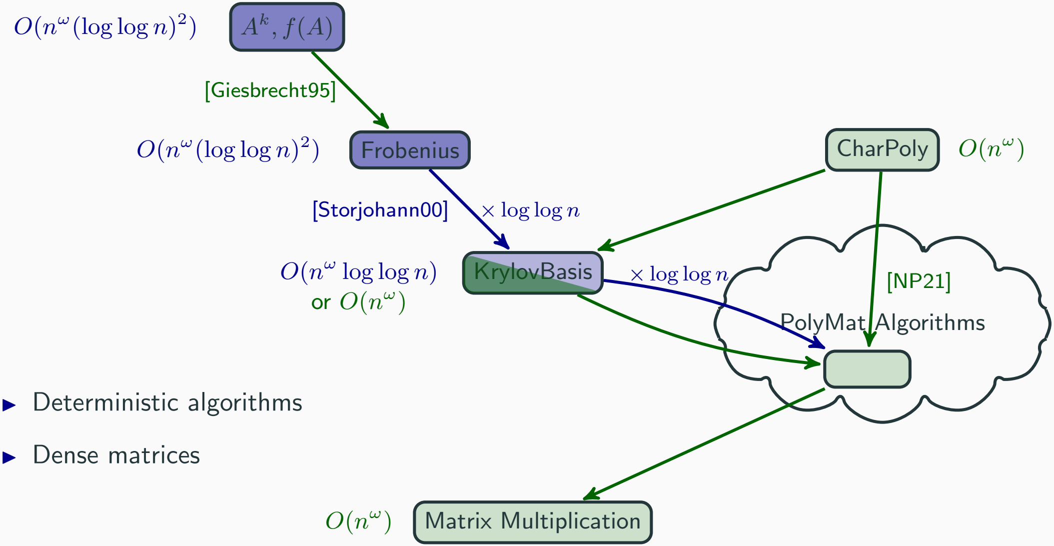
1. *the Frobenius normal form of* $\mathbf{A} \in \mathbb{K}^{n\times n}$ *with a transformation matrix*
2. $\mathbf{A}^k$ *for* $\log(k) \in O(n^{\omega-1-\varepsilon})$.
3. $f(\mathbf{A})$ *for* $f \in \mathbb{K}[X]$ *of degree* $k \in O(n^{\omega-\varepsilon})$.

Contributions



- Deterministic algorithms
- Dense matrices

Contributions



- Deterministic algorithms
- Dense matrices

Krylov matrix and basis of a single vector

Krylov basis of a single vector

Main idea

$$(\mathbf{I} - x\mathbf{A})^{-1} = \sum_{k \geq 0} x^k \mathbf{A}^k$$

1. Series expansion of the inverse of the characteristic matrix

Krylov basis of a single vector

Main idea

$$(\mathbf{I} - x\mathbf{A})^{-1}\mathbf{u} = \sum_{k\geq 0} x^k \mathbf{A}^k \mathbf{u}$$

1. Series expansion of the inverse of the characteristic matrix
2. Projection

Krylov basis of a single vector

Main idea

$$\mathbf{s}(x)t(x)^{-1} = (\mathbf{I} - x\mathbf{A})^{-1}\mathbf{u} = \sum_{k\geq 0} x^k \mathbf{A}^k \mathbf{u}$$

1. Series expansion of the inverse of the characteristic matrix
2. Projection

Krylov basis of a single vector

Main idea

$$\mathbf{s}(x)t(x)^{-1} = (\mathbf{I} - x\mathbf{A})^{-1}\mathbf{u} = \sum_{k\geq 0} x^k \mathbf{A}^k \mathbf{u}$$

$$\Updownarrow$$

$$(\mathbf{I} - x\mathbf{A})\mathbf{s}(x) = \mathbf{u}\; t(x)$$

1. Series expansion of the inverse of the characteristic matrix
2. Projection
   $\rightsquigarrow$ Linear system solving

Krylov basis of a single vector

Main idea

$$\mathbf{s}(x)t(x)^{-1} = (\mathbf{I} - x\mathbf{A})^{-1}\mathbf{u} = \sum_{k\geq 0} x^k \mathbf{A}^k \mathbf{u}$$

$$\Updownarrow$$

$$(\mathbf{I} - x\mathbf{A})\mathbf{s}(x) = \mathbf{u}\ t(x) \Leftrightarrow \begin{bmatrix} \mathbf{I} - x\mathbf{A} & -\mathbf{u} \end{bmatrix} \begin{bmatrix} \mathbf{s}(x) \\ t(x) \end{bmatrix} = 0$$

1. Series expansion of the inverse of the characteristic matrix
2. Projection
   $\rightsquigarrow$ Linear system solving
   $\rightsquigarrow$ Minimal kernel basis

Krylov matrix of a single vector

Krylov matrix algorithm (single vector case)

Input: $\mathbf{A} \in \mathbb{K}^{n\times n}, \mathbf{u} \in \mathbb{K}^{n}, d \in \{1, \ldots n\}$

Output: $\mathbf{K}_d(\mathbf{A}, \mathbf{u})$

1. $\begin{bmatrix} \mathbf{s}(x) \\ t(x) \end{bmatrix} \leftarrow \text{MinimalKernelBasis}\left(\begin{bmatrix} \mathbf{I} - x\mathbf{A} & -\mathbf{u} \end{bmatrix}\right)$
2. $\mathbf{K}_d(\mathbf{A}, \mathbf{u}) \leftarrow \texttt{SeriesExpansion}(\mathbf{s}(x)t(x)^{-1}) \mod x^d$

Krylov matrix of a single vector

Krylov matrix algorithm (single vector case)

Input: $\mathbf{A} \in \mathbb{K}^{n\times n}, \mathbf{u} \in \mathbb{K}^n, d \in \{1, \dots n\}$

Output: $\mathbf{K}_d(\mathbf{A}, \mathbf{u})$

1. $\begin{bmatrix} \mathbf{s}(x) \\ t(x) \end{bmatrix} \leftarrow \text{MinimalKernelBasis}\left(\begin{bmatrix} \mathbf{I} - x\mathbf{A} & -\mathbf{u} \end{bmatrix}\right)$
2. $\mathbf{K}_d(\mathbf{A}, \mathbf{u}) \leftarrow \texttt{SeriesExpansion}(\mathbf{s}(x)t(x)^{-1}) \mod x^d$

Algorithm [Zhou Labahn Storjohann'12]
analyzed in [Jeannerod N. Schost V.'17], [N.P.'21]
$\rightsquigarrow O(n^\omega)$

Krylov matrix of a single vector

Krylov matrix algorithm (single vector case)

Input: $\mathbf{A} \in \mathbb{K}^{n \times n}, \mathbf{u} \in \mathbb{K}^{n}, d \in \{1, \dots n\}$

Output: $\mathbf{K}_d(\mathbf{A}, \mathbf{u})$

1. $\begin{bmatrix} \mathbf{s}(x) \\ t(x) \end{bmatrix} \leftarrow$ MinimalKernelBasis $\left(\begin{bmatrix} \mathbf{I} - x\mathbf{A} & -\mathbf{u} \end{bmatrix} \right)$
2. $\mathbf{K}_d(\mathbf{A}, \mathbf{u}) \leftarrow \texttt{SeriesExpansion}(\mathbf{s}(x)t(x)^{-1}) \mod x^d$

Algorithm [Zhou Labahn Storjohann'12]
analyzed in [Jeannerod N. Schost V.'17], [N.P.'21]
$\rightsquigarrow O(n^{\omega})$

$n\times$ polynomial arithmetic in degree d : $O(n\mathbf{M}(d))$

Krylov basis of a single vector

Krylov basis: $\mathbf{K}_d(\mathbf{A}, \mathbf{u})$ with d maximal such that it has full-rank

Krylov basis algorithm (single vector case)

Input: $\mathbf{A} \in \mathbb{K}^{n \times n}, \mathbf{u} \in \mathbb{K}^n$

Output: $\mathbf{K}_d(\mathbf{A}, \mathbf{u})$

1. $\begin{bmatrix} \mathbf{s}(x) \\ t(x) \end{bmatrix} \leftarrow \text{MinimalKernelBasis}\left(\begin{bmatrix} \mathbf{I} - x\mathbf{A} & -\mathbf{u} \end{bmatrix}\right)$
2. $\mathbf{K}_n(\mathbf{A}, \mathbf{u}) \leftarrow \texttt{SeriesExpansion}(\mathbf{s}(x)t(x)^{-1}) \mod x^n$
3. Gaussian elimination to select the first d linearly independent columns

Krylov basis of a single vector

Krylov basis: $\mathbf{K}_d(\mathbf{A}, \mathbf{u})$ with d maximal such that it has full-rank

Krylov basis algorithm (single vector case)

Input: $\mathbf{A} \in \mathbb{K}^{n \times n}, \mathbf{u} \in \mathbb{K}^n$

Output: $\mathbf{K}_d(\mathbf{A}, \mathbf{u})$

1. $\begin{bmatrix} \mathbf{s}(x) \\ t(x) \end{bmatrix} \leftarrow \text{MinimalKernelBasis}\left(\begin{bmatrix} \mathbf{I} - x\mathbf{A} & -\mathbf{u} \end{bmatrix}\right)$
2. $\mathbf{K}_n(\mathbf{A}, \mathbf{u}) \leftarrow \texttt{SeriesExpansion}(\mathbf{s}(x)t(x)^{-1}) \mod x^n$
3. Gaussian elimination to select the first d linearly independent columns

Property

$t(x) = \textit{mirror}(\textit{Minpoly}(\mathbf{A}, \mathbf{u}))$

$\rightsquigarrow d = \deg(g(x))$ *where* $\begin{bmatrix} \mathbf{f}(x) \\ g(x) \end{bmatrix} \leftarrow \textit{MinimalKernelBasis}\left(\begin{bmatrix} x\mathbf{I} - \mathbf{A} & -\mathbf{u} \end{bmatrix}\right)$

Krylov basis of a single vector

Krylov basis: $\mathbf{K}_d(\mathbf{A}, \mathbf{u})$ with d maximal such that it has full-rank

Krylov basis algorithm (single vector case)

Input: $\mathbf{A} \in \mathbb{K}^{n \times n}, \mathbf{u} \in \mathbb{K}^n$

Output: $\mathbf{K}_d(\mathbf{A}, \mathbf{u})$

1. $\begin{bmatrix} \mathbf{s}(x) \\ t(x) \end{bmatrix} \leftarrow \text{MinimalKernelBasis}\left(\begin{bmatrix} \mathbf{I} - x\mathbf{A} & -\mathbf{u} \end{bmatrix}\right)$
2. $\mathbf{K}_d(\mathbf{A}, \mathbf{u}) \leftarrow \texttt{SeriesExpansion}(\mathbf{s}(x) t(x)^{-1}) \mod x^d$
3. ~~Gaussian elimination to select the first d linearly independent columns~~

Property

$t(x) = \mathit{mirror}(\mathit{Minpoly}(\mathbf{A}, \mathbf{u}))$

$\rightsquigarrow d = \deg(g(x))$ *where* $\begin{bmatrix} \mathbf{f}(x) \\ g(x) \end{bmatrix} \leftarrow \mathit{MinimalKernelBasis}\left(\begin{bmatrix} x\mathbf{I} - \mathbf{A} & -\mathbf{u} \end{bmatrix}\right)$

Krylov matrix and basis of multiple vectors

Generalization to multiple vectors

Main idea

With $\mathbf{U} \in \mathbb{K}^{n\times m}$, find $\mathbf{S} \in \mathbb{K}^{n\times m}$ and $\mathbf{T} \in \mathbb{K}^{m\times m}$ such that

$$\mathbf{S}(x)\mathbf{T}(x)^{-1} = (\mathbf{I} - x\mathbf{A})^{-1}\mathbf{U} = \sum_{k\geq 0} x^k \mathbf{A}^k \mathbf{U}$$

$$\Updownarrow$$

$$(\mathbf{I} - x\mathbf{A})\mathbf{S}(x) = \mathbf{U}\ \mathbf{T}(x) \Leftrightarrow \begin{bmatrix} \mathbf{I} - x\mathbf{A} & -\mathbf{U} \end{bmatrix} \begin{bmatrix} \mathbf{S}(x) \\ \mathbf{T}(x) \end{bmatrix} = \mathbf{0}$$

Generalization to multiple vectors

Main idea

With $\mathbf{U} \in \mathbb{K}^{n\times m}$, find $\mathbf{S} \in \mathbb{K}^{n\times m}$ and $\mathbf{T} \in \mathbb{K}^{m\times m}$ such that

$$\mathbf{S}(x)\mathbf{T}(x)^{-1} = (\mathbf{I} - x\mathbf{A})^{-1}\mathbf{U} = \sum_{k\geq 0} x^k \mathbf{A}^k \mathbf{U}$$

$$\Updownarrow$$

$$(\mathbf{I} - x\mathbf{A})\mathbf{S}(x) = \mathbf{U}\ \mathbf{T}(x) \Leftrightarrow \begin{bmatrix} \mathbf{I} - x\mathbf{A} & -\mathbf{U} \end{bmatrix} \begin{bmatrix} \mathbf{S}(x) \\ \mathbf{T}(x) \end{bmatrix} = \mathbf{0}$$

Algorithm for multiple vectors

1. $\begin{bmatrix} \mathbf{S}(x) \\ \mathbf{T}(x) \end{bmatrix} \leftarrow \text{MinimalKernelBasis}\left(\begin{bmatrix} \mathbf{I} - x\mathbf{A} & -\mathbf{U} \end{bmatrix}\right)$

2. $\mathbf{K_d}(\mathbf{A}, \mathbf{U}) \leftarrow \texttt{SeriesExpansion}(\mathbf{S}(x)\mathbf{T}(x)^{-1}) \mod x^{\mathbf{d}}$

Generalization to multiple vectors

Main idea

With $\mathbf{U} \in \mathbb{K}^{n\times m}$, find $\mathbf{S} \in \mathbb{K}^{n\times m}$ and $\mathbf{T} \in \mathbb{K}^{m\times m}$ such that

$$\mathbf{S}(x)\mathbf{T}(x)^{-1} = (\mathbf{I} - x\mathbf{A})^{-1}\mathbf{U} = \sum_{k\geq 0} x^k \mathbf{A}^k \mathbf{U}$$

$$\Updownarrow$$

$$(\mathbf{I} - x\mathbf{A})\mathbf{S}(x) = \mathbf{U}\ \mathbf{T}(x) \Leftrightarrow \begin{bmatrix} \mathbf{I} - x\mathbf{A} & -\mathbf{U} \end{bmatrix} \begin{bmatrix} \mathbf{S}(x) \\ \mathbf{T}(x) \end{bmatrix} = \mathbf{0}$$

Algorithm for multiple vectors

1. $\begin{bmatrix} \mathbf{S}(x) \\ \mathbf{T}(x) \end{bmatrix} \leftarrow$ MinimalKernelBasis $\left(\begin{bmatrix} \mathbf{I} - x\mathbf{A} & -\mathbf{U} \end{bmatrix}\right)$
2. $\mathbf{K_d}(\mathbf{A}, \mathbf{U}) \leftarrow \texttt{SeriesExpansion}(\mathbf{S}(x)\mathbf{T}(x)^{-1}) \boxed{\bmod x^{\mathbf{d}}}$ ⚠ multiple truncation orders

Generalization to multiple vectors

Series expansion of a matrix fraction at multiple orders

$$\mathbf{S}(x)\mathbf{T}(x)^{-1} \mod \begin{bmatrix} x^{d_1} & & \\ & \ddots & \\ & & x^{d_m} \end{bmatrix}$$

Generalization to multiple vectors

Series expansion of a matrix fraction at multiple orders

$$\mathbf{S}(x)\mathbf{T}(x)^{-1} \mod \begin{bmatrix} x^{d_1} & & \\ & \ddots & \\ & & x^{d_m} \end{bmatrix}$$

1. $\mathbf{Q} \leftarrow \texttt{TruncatedInverse}(\mathbf{T}, \mathbf{d})$ `// column i truncated at order` d_i
2. $\mathbf{K_d}(\mathbf{A}, \mathbf{U}) \leftarrow \texttt{TruncatedProduct}(\mathbf{S}, \mathbf{Q}, \mathbf{d})$ `// column i truncated at order` d_i

Truncated inverse

Truncated inverse $\mathbf{T}(x)^{-1} \mod x^{\mathbf{d}}$

Obstacles:

- Heterogeneity in $\mathbf{d}$

- Heterogeneity the column degree of $\mathbf{T}(x)$

Truncated inverse

Truncated inverse $\mathbf{T}(x)^{-1} \mod x^{\mathbf{d}}$

Obstacles:

- Heterogeneity in $\mathbf{d}$
  $\rightsquigarrow$ High order lifting [Storjohann'03]
- Heterogeneity the column degree of $\mathbf{T}(x)$
  $\rightsquigarrow$ Partial linearization [Gupta, Sarkar, Storjohann Valeriote'12]

Truncated inverse

Truncated inverse $\mathbf{T}(x)^{-1} \mod x^{\mathbf{d}}$

Obstacles:

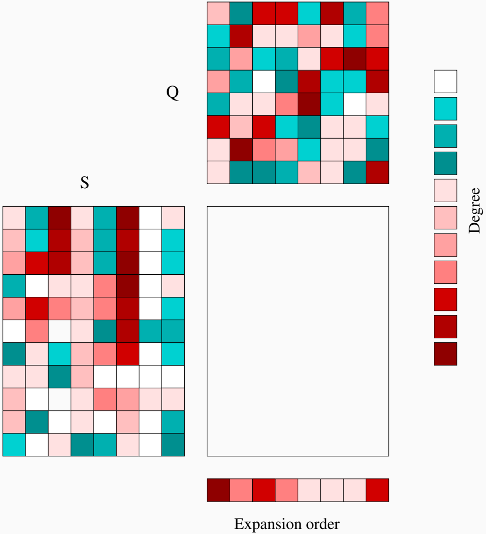
- Heterogeneity in $\mathbf{d}$
  $\rightsquigarrow$ High order lifting [Storjohann'03]

- Heterogeneity the column degree of $\mathbf{T}(x)$
  $\rightsquigarrow$ Partial linearization [Gupta, Sarkar, Storjohann Valeriote'12]

$\rightsquigarrow O(m^\omega \mathbf{M}(\frac{n}{m}) \log n \log m)$

Truncated products

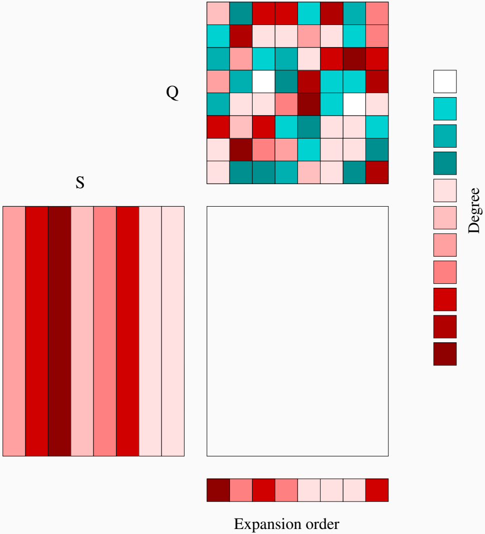
Truncated product $\mathbf{S}(x)\mathbf{Q}(x) \mod x^{\mathbf{d}}$



Truncated products

Truncated product $\mathbf{S}(x)\mathbf{Q}(x) \mod x^{\mathbf{d}}$

- Consider column degrees of $\mathbf{S}$



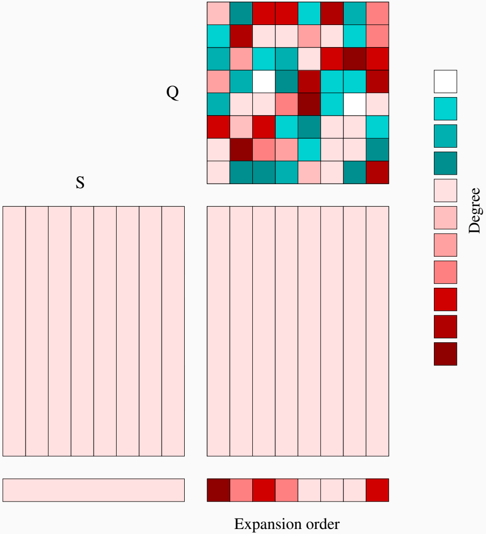
Truncated products

Truncated product $\mathbf{S}(x)\mathbf{Q}(x) \mod x^{\mathbf{d}}$

- Consider column degrees of $\mathbf{S}$
- Split degrees in geometric progression:

$$\left(\mathbf{S}^{(0)} + \sum_k \mathbf{S}^{(k)} x^{2^k\delta}\right)\mathbf{Q}(x) \mod \begin{bmatrix} x^{d_1} & & \\ & \ddots & \\ & & x^{d_m} \end{bmatrix}$$

with $\deg \mathbf{S}^{(k)} \leq 2^k\delta$



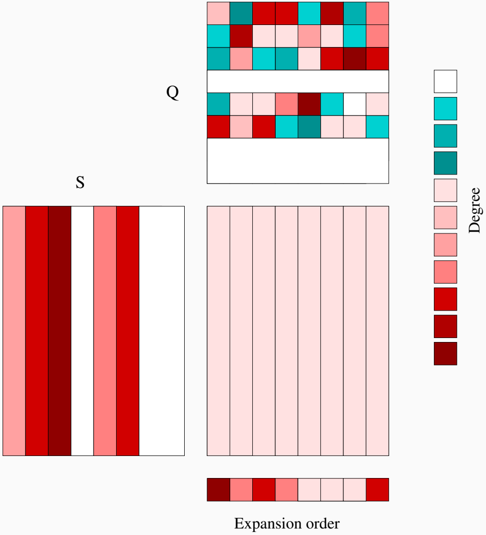
Truncated products

Truncated product $\mathbf{S}(x)\mathbf{Q}(x) \mod x^{\mathbf{d}}$

- Consider column degrees of $\mathbf{S}$
- Split degrees in geometric progression:

$$\left(\mathbf{S}^{(0)} + \sum_k \mathbf{S}^{(k)} x^{2^k\delta}\right)\mathbf{Q}(x) \mod \begin{bmatrix} x^{d_1} & & \\ & \ddots & \\ & & x^{d_m} \end{bmatrix}$$

with $\deg \mathbf{S}^{(k)} \leq 2^k\delta$



Truncated products

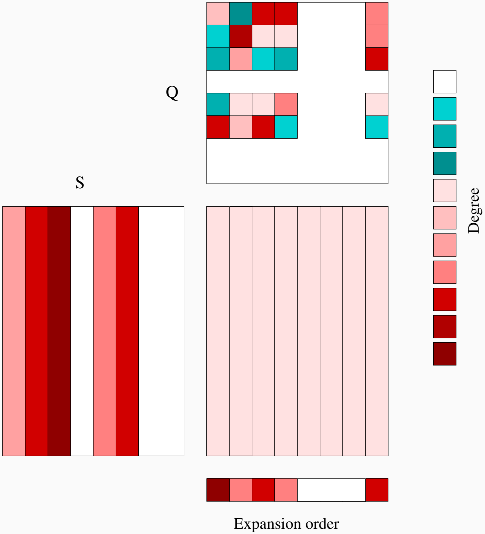
Truncated product $\mathbf{S}(x)\mathbf{Q}(x) \mod x^{\mathbf{d}}$

- Consider column degrees of $\mathbf{S}$
- Split degrees in geometric progression:

$$\left(\mathbf{S}^{(0)} + \sum_k \mathbf{S}^{(k)} x^{2^k \delta}\right) \mathbf{Q}(x) \mod \begin{bmatrix} x^{d_1} & & \\ & \ddots & \\ & & x^{d_m} \end{bmatrix}$$

with $\deg \mathbf{S}^{(k)} \leq 2^k \delta$

- degree-dimension tradeoff



Truncated products

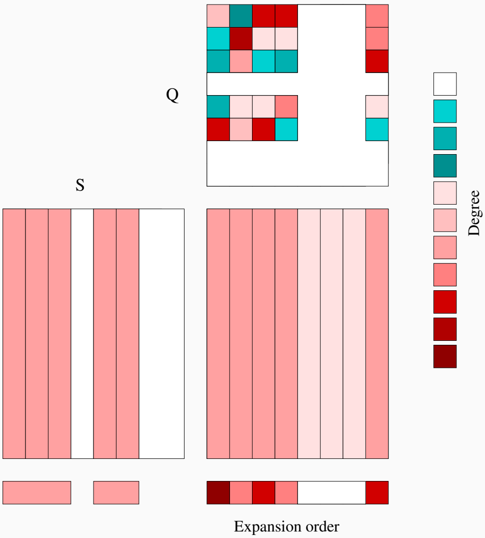
Truncated product $\mathbf{S}(x)\mathbf{Q}(x) \mod x^{\mathbf{d}}$

- Consider column degrees of $\mathbf{S}$
- Split degrees in geometric progression:

$$\left(\mathbf{S}^{(0)} + \sum_k \mathbf{S}^{(k)} x^{2^k \delta}\right) \mathbf{Q}(x) \mod \begin{bmatrix} x^{d_1} & & \\ & \ddots & \\ & & x^{d_m} \end{bmatrix}$$

with $\deg \mathbf{S}^{(k)} \leq 2^k \delta$

- degree-dimension tradeoff



Truncated products

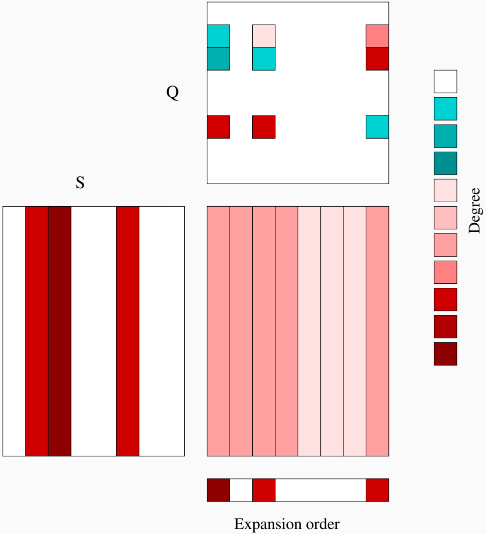
Truncated product $\mathbf{S}(x)\mathbf{Q}(x) \mod x^{\mathbf{d}}$

- Consider column degrees of $\mathbf{S}$
- Split degrees in geometric progression:

$$\left(\mathbf{S}^{(0)} + \sum_k \mathbf{S}^{(k)} x^{2^k\delta}\right) \mathbf{Q}(x) \mod \begin{bmatrix} x^{d_1} & & \\ & \ddots & \\ & & x^{d_m} \end{bmatrix}$$

with $\deg \mathbf{S}^{(k)} \leq 2^k \delta$

- degree-dimension tradeoff



Truncated products

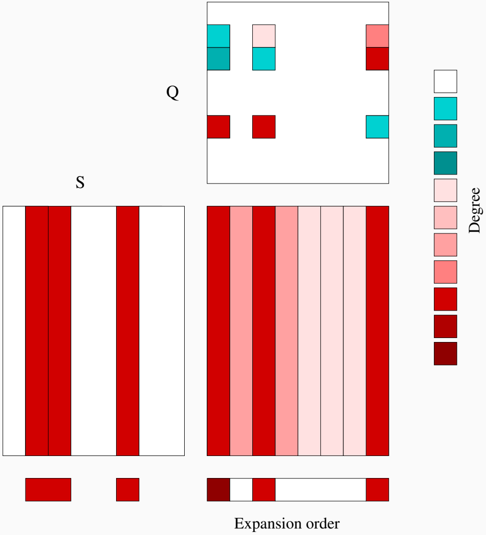
Truncated product $\mathbf{S}(x)\mathbf{Q}(x) \mod x^{\mathbf{d}}$

- Consider column degrees of $\mathbf{S}$
- Split degrees in geometric progression:

$$\left(\mathbf{S}^{(0)} + \sum_k \mathbf{S}^{(k)} x^{2^k \delta}\right) \mathbf{Q}(x) \mod \begin{bmatrix} x^{d_1} & & \\ & \ddots & \\ & & x^{d_m} \end{bmatrix}$$

with $\deg \mathbf{S}^{(k)} \leq 2^k \delta$

- degree-dimension tradeoff



Truncated products

Truncated product $\mathbf{S}(x)\mathbf{Q}(x) \mod x^{\mathbf{d}}$

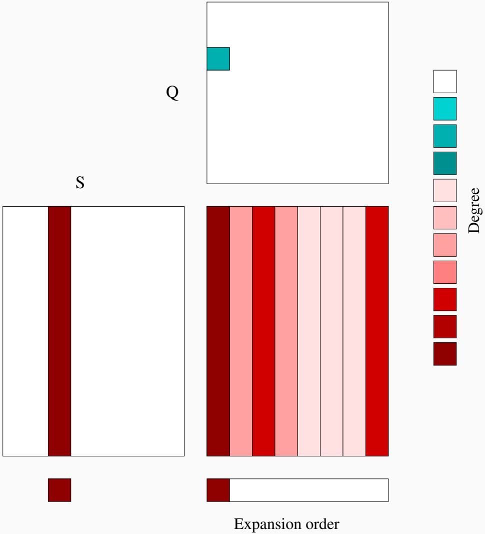
- Consider column degrees of $\mathbf{S}$
- Split degrees in geometric progression:

$$\left(\mathbf{S}^{(0)} + \sum_k \mathbf{S}^{(k)} x^{2^k\delta}\right)\mathbf{Q}(x) \mod \begin{bmatrix} x^{d_1} & & \\ & \ddots & \\ & & x^{d_m} \end{bmatrix}$$

with $\deg \mathbf{S}^{(k)} \leq 2^k\delta$

- degree-dimension tradeoff

$\rightsquigarrow O(m^{\omega-2} n \mathbf{M}(n))$



Krylov matrix and basis of multiple vectors

Krylov matrix (prescribed orders d)

Overall $O(n^\omega + m^{\omega-2}n^2(\log n)^4)$

Krylov matrix and basis of multiple vectors

Krylov matrix (prescribed orders d)

Overall $O(n^\omega + m^{\omega-2}n^2(\log n)^4)$

Krylov basis (targeting the unknown lex maximal orders d)

Find the maximal orders by

1. $\begin{bmatrix} \mathbf{F} \\ \mathbf{G} \end{bmatrix} \leftarrow \texttt{MinimalKernelBasis}\left(\begin{bmatrix} x\mathbf{I} - \mathbf{A} & -\mathbf{U} \end{bmatrix}\right)$ [Zhou Labahn Storjohann'12]
2. $\mathbf{d} \leftarrow$ Diagonal degrees of HNF$(\mathbf{G})$ [Labahn N. Zhou'17]

[Kailath'80]: HNF$(\mathbf{G})$ is a polynomial compression of the Hessenberg form $\begin{bmatrix} \mathbf{C}_{f_1} & * & \cdots & * \\ & \mathbf{C}_{f_2} & & * \\ & & \ddots & \vdots \\ & & & \mathbf{C}_{f_m} \end{bmatrix}$

Krylov matrix and basis of multiple vectors

Krylov matrix (prescribed orders d)

Overall $O(n^\omega + m^{\omega-2} n^2 (\log n)^4)$

Krylov basis (targeting the unknown lex maximal orders d)

Find the maximal orders by

1. $\begin{bmatrix} \mathbf{F} \\ \mathbf{G} \end{bmatrix} \leftarrow \texttt{MinimalKernelBasis}\left(\begin{bmatrix} x\mathbf{I} - \mathbf{A} & -\mathbf{U} \end{bmatrix}\right)$ [Zhou Labahn Storjohann'12]
2. $\mathbf{d} \leftarrow$ Diagonal degrees of HNF$(\mathbf{G})$ [Labahn N. Zhou'17]

[Kailath'80]: HNF$(\mathbf{G})$ is a polynomial compression of the Hessenberg form $\begin{bmatrix} \mathbf{C}_{f_1} & * & \cdots & * \\ & \mathbf{C}_{f_2} & & * \\ & & \ddots & \vdots \\ & & & \mathbf{C}_{f_m} \end{bmatrix}$

$\rightsquigarrow O(m^{\omega-2} n (\log n)^c + n^\omega)$

Krylov matrix and basis of multiple vectors

Krylov matrix (prescribed orders d)

Overall $O(n^\omega + m^{\omega-2} n^2 (\log n)^4)$

Krylov basis (targeting the unknown lex maximal orders d)

Find the maximal orders by

1. $\begin{bmatrix} \mathbf{F} \\ \mathbf{G} \end{bmatrix} \leftarrow \texttt{MinimalKernelBasis}\left(\begin{bmatrix} x\mathbf{I} - \mathbf{A} & -\mathbf{U} \end{bmatrix}\right)$ [Zhou Labahn Storjohann'12]
2. $\mathbf{d} \leftarrow$ Diagonal degrees of HNF$(\mathbf{G})$ [Labahn N. Zhou'17]

[Kailath'80]: HNF$(\mathbf{G})$ is a polynomial compression of the Hessenberg form $\begin{bmatrix} \mathbf{C}_{f_1} & * & \cdots & * \\ & \mathbf{C}_{f_2} & & * \\ & & \ddots & \vdots \\ & & & \mathbf{C}_{f_m} \end{bmatrix}$

$\rightsquigarrow O(m^{\omega-2} n (\log n)^c + n^\omega)$

As long as $m \in O(n/(\log n)^e)$, with $e = \max(c, 4)$ $\rightsquigarrow O(n^\omega)$

Dealing with numerous vectors: Keller-Gehrig's bootstrap

When $m = \Theta(n)$:

- Iterate all vectors a few times k
- Necessarily, only few will not be completed $\leq n/k$

Dealing with numerous vectors: Keller-Gehrig's bootstrap

When $m = \Theta(n)$:

- Iterate all vectors a few times k
  [Keller-Gehrig'85] in $O(n^\omega \log k)$
- Necessarily, only few will not be completed $\leq n/k$

Dealing with numerous vectors: Keller-Gehrig's bootstrap

When $m = \Theta(n)$:

- Iterate all vectors a few times k
  [Keller-Gehrig'85] in $O(n^\omega \log k)$
- Necessarily, only few will not be completed $\leq n/k$
- Run previous algorithm on the remaining ones

Dealing with numerous vectors: Keller-Gehrig's bootstrap

When $m = \Theta(n)$:

- Iterate all vectors a few times $k = (\log n)^e$
  [Keller-Gehrig'85] in $O(n^\omega \log k) = O(n^\omega \log \log n)$
- Necessarily, only few will not be completed $\leq n/k = n/(\log n)^e$
- Run previous algorithm on the remaining ones in $O(n^\omega)$

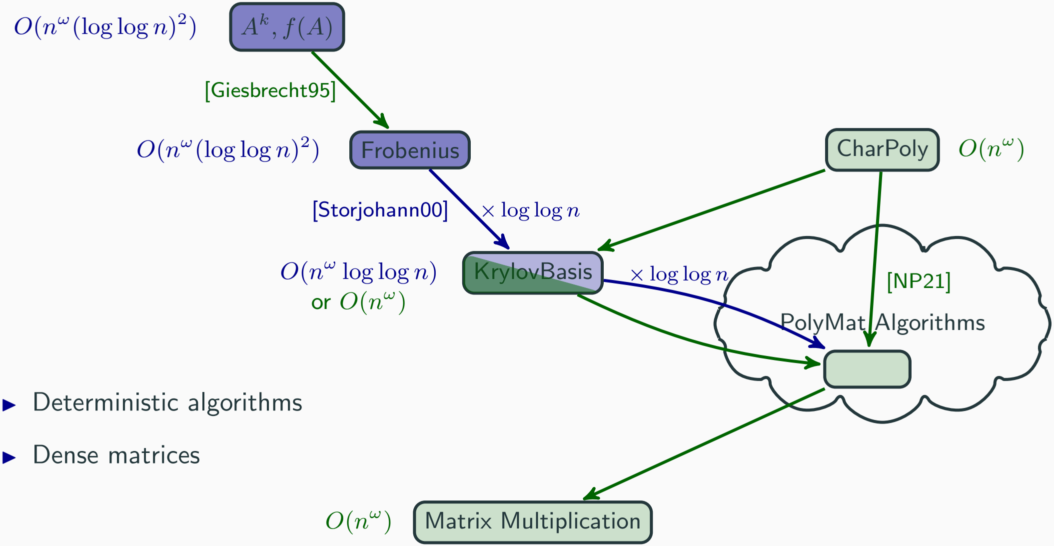
Dealing with numerous vectors: Keller-Gehrig's bootstrap

When $m = \Theta(n)$:

- Iterate all vectors a few times $k = (\log n)^e$
  [Keller-Gehrig'85] in $O(n^\omega \log k) = O(n^\omega \log\log n)$
- Necessarily, only few will not be completed $\leq n/k = n/(\log n)^e$
- Run previous algorithm on the remaining ones in $O(n^\omega)$

$\rightsquigarrow$ Overall $O(n^\omega \log\log n)$

Conclusion



- Deterministic algorithms
- Dense matrices