



Hunting the multiplicity and approximating multiple roots of polynomial systems

Marc Giusti¹ Jean-Claude Yakoubsohn²

¹ Laboratoire d'Informatique
CNRS-Polytechnique
Palaiseau, France

²Institut de Mathématiques de Toulouse
Université Paul Sabatier
Toulouse, France

Computation of the roots by homotopy method

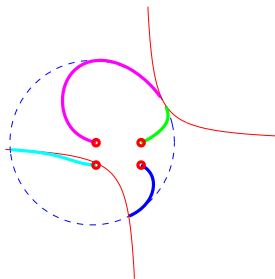
Goal : Compute all the isolated roots of a multivariate polynomial system $f(x) = 0$.

A way to do that. Define the homotopy

$$h(t, x) = (1 - t)f_0(x) + tf(x)$$

where we know all the roots of $g(x)$. We can prove that there exist generically $x(t)$ such that $h(t, x(t)) = 0$.

From a zero of $g(x)$ we can follow a curve such that $f(x(1)) = 0$.



What is the dream?

- 1- $x = (x_1, \dots, x_n) \in \mathbb{C}^n$.
- 2- a polynomial system $f = (f_1, \dots, f_m) \in \mathbb{C}^m[x]$.
- 3- w an isolated multiple root of f , i.e. $Df(w)$ has not a full rank.

The quadratic convergence of the Newton method is lost in a neighborhood of w .

To recover it : determine a regular system at w admitting the same root as the initial one, **WITHOUT ADDING NEW VARIABLES**. We will say an equivalent system.

An example

Example : $f(x, y) = (x - y^2, 2cy^3 - 2xy) = 0$ studied by Griewank and Osborne in 1983.

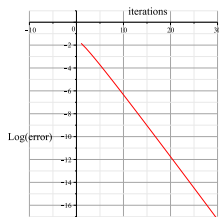


Fig. 1. Linear convergence of Newton sequence from $(0.1, -0.2)$ with $c = 5/32$.

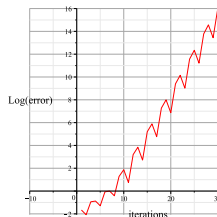


Fig. 2. Divergence of Newton sequence from $(0.1, -0.2)$ with $c = 29/32$.

In this example we remark that the gradient of $2cy^3 - 2xy$ is zero at $(0, 0)$. Hence we can replace the polynomial $2cy^3 - 2xy$ by the two following : y and $3cy^2 - x$.

We note that the new system $(x - y^2, y, 3cy^2 - x) = 0$ is regular at $(0, 0)$.

Reference

Marc Giusti, Jean-Claude Yakoubsohn

Multiplicity hunting and approximating multiple roots of
polynomials systems

Contemporary Mathematics Contemporary Mathematics 604

American Mathematical Society

Recent Advances in Real Complexity and Computation, 2013.

One single polynomial and one single variable

$$f(w) = f^0(w) = \dots = f^{(\mu-1)}(w) = f^{(\mu)}(w) = 0, \quad f^{(\mu+1)}(w) \neq 0.$$

Schröder-Newton operator: $N(f, \mu, x) = x - \mu \frac{f(x)}{f'(x)}$.

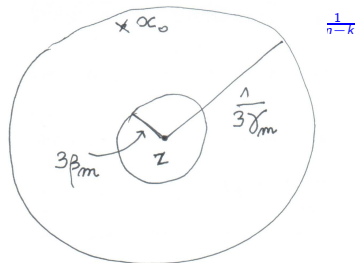
We defined quantities $\gamma_\mu, \beta_\mu, \alpha_\mu$ and proved for clusters of roots in the spirit of Shub-Smale theory a γ_μ -theorem and a α_μ -theorem.

Reference:

M. Giusti, G. Lecerf, B. Salvy, J.-C. Yakoubsohn, On location and approximation of clusters of zeros of analytic functions, Found. Comput. Math. 5 (2005), no. 3, 257–311.

One variable : existence of clusters of roots

$$\blacktriangleright \gamma_m : \sup_{k \geq m+1} \left(\frac{m! |f^{(k)}(z)|}{k! |f^{(m)}(z)|} \right)^{\frac{1}{k-m}}$$



One variable : existence of clusters of roots cont.

α_m -Theorem.

Let $\alpha_0(1) = 3 - 2\sqrt{2}$ and $\alpha_0(m) = 1/9$ if $m > 1$.

If we have $\alpha_m(f, x_0) \leq \alpha_0(m)$ then the univariate polynomial $f(x)$ has m zeros counting multiplicities in the ball $B(x_0, r)$ for any r such that $r^- \leq r \leq r^+$.

The quantities r^- and r^+ depends of the value m .

$$\text{If } m = 1 \text{ then } r^- = \frac{2\beta_m}{1 + \alpha_m + \sqrt{1 - 6\alpha_m + \alpha_m^2}} \text{ and } r^+ = \frac{1 + \alpha_m + \sqrt{1 - 6\alpha_m + \alpha_m^2}}{4\gamma_m}.$$

$$\text{In the case where } m > 1 \text{ we have } r^- = \frac{4\beta_m}{1 + 3\alpha_m + \sqrt{1 - 10\alpha_m + 9\alpha_m^2}} \text{ and } r^+ = \frac{1 + 3\alpha_m + \sqrt{1 - 10\alpha_m + 9\alpha_m^2}}{4\gamma_m}.$$

Case of regular roots : ball of quadratic convergence.

f analytic : $E \rightarrow F$, E, F two Banach spaces.

Let $f(w) = 0$ and $Df(w)$ invertible.

Let

$$\gamma(f, w) = \sup_{k \geq 2} \left(\frac{1}{k!} \|Df(w)^{-1} D^k f(w)\| \right)^{1/(k-1)} .$$

γ -**Theorem** (Shub-Smale). Let x_0 be such that

$$\|x_0 - w\| \leq \frac{3 - \sqrt{7}}{2\gamma(f, w)} .$$

Then the Newton method converges quadratically from x_0 to w .

Number of roots of a polynomial system

Theorem Let $I = \langle f_1, \dots, f_m \rangle$ and $V(I) \subset \mathbb{C}^n$ the associated variety.

- 1- The dimension of $C[x]/I$ is finite iff the dimension of $V(I)$ is zero.
- 2- In the finite dimension case one has

$$\dim C[x]/I = \#V(I)$$

where $\#V(I)$ is the number of points of $V(I)$.

Example

$$f_1(x, y) = x^2 + x^3, \quad f_2(x, y) = x^3 + y^2 \text{ and}$$

$$V(I) = \{(0, 0), (-1, 1), (-1, -1)\}.$$

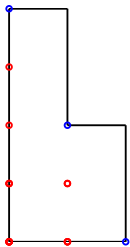
A Groebner basis of I , wrt the order induced by $x > y$ on the monomials, is:

$$g_1(x, y) = x^0 y^4 - y^2, \quad g_2(x, y) = x^1 y^2 + y^2,$$

$$g_3(x, y) = x^2 y^0 - y^2,$$

We deduce

$$\dim C[x]/I = 6$$



Multivariate case : multiplicities and local rings

Let w be a root of $f = (f_1, \dots, f_m)$, denote by I the ideal $\langle f_1, \dots, f_m \rangle$, and assume that

$V(I) = \{x \in \mathbb{C}^n, : \forall g \in I, g(x) = 0\}$ is finite.

Let $\mathbb{C}\{x - w\}$ the local ring of convergent power series in n variables with maximal ideal generated by $x_1 - w_1, \dots, x_n - w_n$.

Let $I\mathbb{C}\{x - w\}$ the ideal generated by I in $\mathbb{C}\{x - w\}$

We note $A_w = \mathbb{C}\{x - w\}/I\mathbb{C}\{x - w\}$.

Theorem.

Let $V(I) = \{w_1, w_2, \dots, w_N\}$. Then

$$1- \mathbb{C}[x]/I \sim A_{w_1} \times \dots \times A_{w_N}.$$

$$2- \dim \mathbb{C}[x]/I = \sum_{i=1}^N \dim A_{w_i}.$$

Definition. The multiplicity of w is $\dim A_w$.

Example

$f_1(x, y) = x^2 + x^3$, $f_2(x, y) = x^3 + y^2$ and

$V(I) = \{(0, 0), (-1, 1), (-1, -1)\}$.

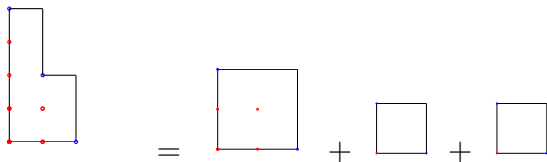
1- A standard basis of $I\mathbb{C}\{x, y\}$ is : $g_1 = x^2$, $g_2 = y^2$.

Hence $\dim A_{(0,0)} = 4$.

2- A standard basis of $I\mathbb{C}\{x + 1, y - 1\}$ (resp.

$I\mathbb{C}\{x + 1, y + 1\}$) is : $g_1 = 3x$, $g_2 = 2y$.

Hence $\dim A_{(-1,1)} = \dim A_{(-1,-1)} = 1$.



Duality and Multiplicities

We define

$$\mathcal{D}_w^j(I) = \left\{ L = \sum_{|\alpha| \leq j} L_\alpha \partial_w^\alpha \quad : \quad \forall g \in I, \quad L(g) = 0 \right\}$$

Theorem

- 1- The root w is isolated iff there exists k s.t.
 $\mathcal{D}_w^{k-1}(I) = \mathcal{D}_w^k(I)$.
- 2- In this case the dimension of $\mathcal{D}_w^k(I)$ is equal to the multiplicity of w .

Macaulay matrices and multiplicity

We consider the Macaulay matrices

$$M_k = (\partial_w^\beta((x-w)^\beta f_i(x))) \quad \text{for } |\beta| \leq k-1 \text{ and } i = 1 : m.$$

Theorem $D_w^k(I)$ is isomorphic to the kernel of M_k .

Consequently $D_w^{k-1}(I) = D_w^k(I)$ when $\ker(M_{k-1}) = \ker(M_k)$

Example

$$f_1 = x^2 + y^2 - 2, \quad f_2 = xy - 1. \quad w = (1, 1).$$

		∂_{00}	∂_{10}	∂_{01}	∂_{20}	∂_{11}	∂_{02}
S_0	f_1	0	2	2	2	0	2
	f_2	0	1	1	0	1	0
-----		-					
S_1							
	-----	-	-	-			
S_2	$(x-1)f_1$	0	0	0	4	2	0
	$(x-1)f_2$	0	0	0	2	1	0
	$(y-1)f_1$	0	0	0	0	2	4
	$(y-1)f_2$	0	0	0	0	1	2

$$\text{rank}(S_0) = 0, \quad \text{corank}(S_0) = 1$$

$$\text{rank}(S_1) = 1, \quad \text{corank}(S_1) = 2$$

$$\text{rank}(S_2) = 4, \quad \text{corank}(S_2) = 2.$$

Hence $\mu = 2$.

Mantzalaris and Mourrain approach to determine a regular system from the initial singular at w .

Step 1- Compute a basis for the dual space of the local quotient ring $A_w : \Lambda = (\Lambda_1, \dots, \Lambda_\mu)$

Step 2- Compute the system $f^\Lambda := (\Lambda(f_1), \dots, \Lambda(f_m))$.

Theorem The system f^Λ is regular at w .

Example.

$$f_1(x, y) = x^2 + y^2 - 2, \quad f_2 = xy - 1. \quad w = (1, 1).$$

$$\Lambda = (1, \partial_1 - \partial_2).$$

$$f^\Lambda = (f_1, f_2, x - y).$$

But the complexity of the computation of the Macaulay matrices is very huge!

Multiple roots and clusters of roots

Let w a multiple root of f with multiplicity μ . Let \bar{f} be a pertubated system of f . In an open neighborhood U_w where the Rouché theorem holds, i.e. if for all $x \in \partial U_w$ we have

$$\|\bar{f}(x) - f(x)\| < \|f(x)\|$$

then \bar{f} has μ roots in U_w .

$$\dim A_w^f = \sum_{\bar{w} \in \bar{f}^{-1}(0) \cap U_w} \dim A_{\bar{w}}^{\bar{f}}.$$

Recovering quadratic convergence for multiple roots of polynomial systems.

Newton method does not converge quadratically close to an isolated multiple root. To recover this quadratic convergence the idea is to determine a polynomial system which admits w as regular root, i.e, its jacobian is full rank at w .

To do that we use both numerical and symbolic computation.

Two ingredients

- 1– Deflation based on a numerical computation of derivatives.
- 2– Determination of new polynomials based on a computation of numerical ranks.

This way does not require to compute the whole structure of the local quotient algebra which is too huge.

Recovering quadratic convergence : deflating and kernelling.

We proceed by two actions :

- 1- Deflating
- 2- Kerneling

We do that without adding new variables. Only the number of equations can increase.

Exact deflating

We replace an equation $g(x) = 0$ by the n equations of its gradient $\frac{\partial g(x)}{\partial x_i} = 0$, $i = 1 : n$ if the gradient ∇g vanishes at the root w :

$$g(w) = 0, \quad \frac{\partial g(w)}{\partial x_i} = 0, \quad i = 1 : n.$$

Deflating. Example

$$f_1 = x^3/3 + y^2x + x^2 + 2yx + y^2, \quad f_2 = x^2y + x^2 + 2yx + y^2$$

The multiplicity of $(0,0)$ is 6.

We have

$$\begin{array}{cc|cc} \partial_1 & & \partial_2 & \\ x^2 + y^2 + 2x + 2y & 2yx + 2x + 2y & 2xy + 2x + 2y & x^2 + 2x + 2y \end{array}$$

$$\begin{array}{cc|cc} \partial_{11} & \partial_{12} & \partial_{21} & \partial_{22} \\ 2x + 2 & 2y + 2 & 2y + 2 & 2x + 2 \end{array} \parallel \begin{array}{cc|cc} \partial_{11} & \partial_{12} & \partial_{11} & \partial_{12} \\ 2x + 2 & 2y + 2 & 2x + 2 & 2 \end{array}$$

The new polynomial system consists of the three polynomials

$$x^2 + y^2 + 2x + 2y, \quad xy + x + y, \quad x^2 + 2x + 2y.$$

Exact kerneling

Let us consider a polynomial system $f(x) = 0$ such that $Df(w)$ is rank deficient and each row $Df_i(w) \neq 0$, $i = 1 : n$. Call $r > 0$ the rank of $Df(w)$. If

$$Df(w) = \begin{pmatrix} A(w) & B(w) \\ C(w) & D(w) \end{pmatrix} \in \mathbb{C}^{m \times n}$$

with $A(w)$ invertible matrix of size $r \times r$. The Jacobian matrix $Df(w)$ has rank r iff the Schur complement is zero

$$D(w) - C(w)A(w)^{-1}B(w) = 0$$

i.e., w is a root of $D(x) - C(x)A(x)^{-1}B(x)$.

Then we add to the initial system at most the $(m - r) \times (n - r)$ equations given by the numerators of the entries of the Schur complement:

$$D(x) - C(x)A^{-1}(x)B(x).$$

Kerneling using Schur Complement. Example

$$f_1 = x^2 + y^2 + 2x + 2y, \quad f_2 = xy + x + y, \quad f_3 = x^2 + 2x + 2y.$$

The root $(0, 0)$ has multiplicity 2.

$$J := Df(x, y, z) = \begin{pmatrix} 2x + 2 & 2y + 2 \\ y + 1 & x + 1 \\ 2x + 2 & 2y + 2 \end{pmatrix} \text{ and}$$

$$Df(0, 0, 0) = \begin{pmatrix} 2 & 2 \\ 1 & 1 \\ 2 & 2 \end{pmatrix} \text{ has rank 1.}$$

Then there are two equations to be added to f , namely $2x^2 + 4x - 2y$ and $2y$, the numerators of :

$$J(2..3, 2) - J(2..3, 1)J(1, 1)^{-1}J(1, 2).$$

Moreover the rank of the new system is 2.

Recovering quadratic convergence

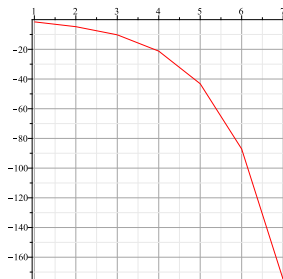
Let $GN_f(x) = x - (Df(x))^T Df(x))^{-1} Df(x))^T f(x)$ and

$$x_{k+1} = GN_f(x_k), \quad k \geq 0.$$

Consider

$$f_1 = x^2 + y^2 + 2x + 2y, \quad f_2 = xy + x + y, \quad f_3 = x^2 + 2x + 2y$$
$$f_4 = x^2 + 2x - y, \quad f_5 = y.$$

$$\begin{aligned} e_1 &= 0.22 \\ e_2 &= 0.0088 \\ e_3 &= 3 \times 10^{-5} \\ e_4 &= 6 \times 10^{-10} \\ e_5 &= 1.7 \times 10^{-19} \\ e_6 &= 1.4 \times 10^{-38} \\ e_7 &= 1 \times 10^{-76} \end{aligned}$$



Theoretical results

$f = (f_1, \dots, f_m)$ and w a multiple root of f .

Lemma If $m = n$, then the determinant of $Df(x)$ is not in the ideal generated by $\langle f \rangle$.

Theorem Let $f(w) = 0$ and $\nabla f_1(w) = 0$. Then

$$\text{multiplicity}_w(f) > \text{multiplicity}_w(\nabla f_1, f)$$

Theorem Let $f(w) = 0$, $\text{rank}(Df(w)) < n$ and $\nabla f_i(w) \neq 0$, $i = 1 : n$. Let $S = \{ \text{numerators of a Schur complement of } Df(x) \}$. Then

$$\text{multiplicity}_w(f) > \text{multiplicity}_w(f, S)$$

Complexity

- 1- At each step of deflating-kernelling the multiplicity decreases at less by one.
- 2- The number N of steps of deflating-kernelling is bounded by the multiplicity of the root.

More precisely

$$\sum_{k=1}^N m_k + (m_k - r_k)(n - r_k) \leq \mu$$

where r_k is the rank of the jacobian matrix,
 m_k the number of equations we add at each step.

From numeric to symbolic

We explain how it is possible to determine a regular system using numerical computation **without the introduction of an tolerance.**

Instead, we exhibit a numerical criterion expressing that a polynomial is small at a point.

This replaces in a certified way the notion of tolerance.

Our criterion : the α_m test.

Let $x \in \mathbb{C}^n$ and $g \in \mathbb{C}[x]$.

Definition.

Let $x_0 \in \mathbb{C}^n$. We say that $g(x)$ satisfies the α_m test at x_0 if there exists m and i such that

$$\alpha_m(\bar{g}_i, x_{0,i}) \leq \alpha_0(m)$$

where

$$g_i(\bar{z}) = g(x_{0,1}, \dots, x_{0,i-1}, z, x_{0,i+1}, \dots, x_{0,n})$$

Numerical Deflating

We will replace an equation $g(x) = 0$ by the n equations $\frac{\partial g(x)}{\partial x_i} = 0, i = 1 : n$ if $g(x)$ and the polynomials which constitute its gradient satisfy the α_m test at a point x_0 .

Numerical deflating : example.

Let us consider $f = (x^3/3 + y^2x + x^2 + 2yx + y^2, x^2y + x^2 + 2yx + y^2)$. The multiplicity of $(0, 0)$ is 6.

Let us consider $x_0 = (0.04, -0.03)$.

$x_0 = (0.04, -0.03)$	x_i	m_i	α_{m_i}
f_1	x	2	0.00069
$\partial_1 f_1 = x^2 + y^2 + 2x + 2y$	x	1	0.0052
$\partial_2 f_1 = 2xy + 2x + 2y$	x	1	0.0009
$\partial_{11} f_1 = 2x + 2$	x	1	1.04
$\partial_{12} f_1 = 2y + 2$	y	1	0.97
$\partial_{22} f_1 = 2x + 2$			

$x_0 = (0.04, -0.03)$	x_i	m_i	α_{i,m_i}
f_2	x	2	0.00181
$\partial_1 f_2 = \partial_2 f_1$			
$\partial_2 f_2 = x^2 + 2x + 2y$	x	1	0.0049
$\partial_{11} f_2 = 2y + 2$			
$\partial_{12} f_2 = 2x + 2$			
$\partial_{22} f_2 = 2$			

At this step we have $f^1 = (x^2 + y^2 + 2x + 2y, xy + x + y, x^2 + 2x + 2y)$ and all the rows of $Df^1(x_0)$ are not identically zero.

Numerical kerneling

Let us consider a polynomial system $f(x) = 0$ and $x_0 \in \mathbb{C}^n$. We suppose that each row of $Df(x_0)$ is non zero.

- ▶ **Numeric.** We determine permutation matrices P and Q such that $PDf(x_0)Q = LU$ and the modulus of the elements of the diagonal of U are in decreasing order.
- ▶ **Symbolic.** $J = PDf(x)Q$. For a given r , $2 \leq r \leq n - 1$,

$$J = \begin{pmatrix} A_{r \times r}(x) & B_{r \times (n-r)}(x) \\ C_{(m-r) \times r}(x) & D_{(m-r) \times (n-r)}(x) \end{pmatrix} \in \mathbb{C}^{m \times n}$$

we denote by $S_r(x)$ the numerators of the Schur complement associated to $A_{r \times r}(x)$.

- ▶ **Numeric.** Assume that all the elements of $S_r(x)$ satisfy the α_m test at x_0 , and that there exists an element of $S_{r-1}(x)$ which does not satisfy the α_m test at x_0 . Then we add the distinct elements of $S_r(x)$ to the system f .

Numerical kerneling : example.

Let $f = (x^2 + y^2 + 2x + 2y, xy + x + y, x^2 + 2x + 2y)$ and for $x_0 = (0.04, -0.03)$ the rows of $Df^1(x_0)$ are not identically zero. We have $Df(x_0) = \begin{bmatrix} 2.08 & 1.94 \\ 0.97 & 1.04 \\ 2.08 & 2 \end{bmatrix}$. Since

$\partial_1 f(x, y) = 2x + 2$ we have seen $\alpha_{1,1}(2x + 2, x_0) > 3 - 2\sqrt{2}$. We then consider

$A_{1 \times 1}(x, y) = 2x + 2$ $C_{2 \times 1}(x, y) = \begin{bmatrix} y + 1 \\ 2x + 2 \end{bmatrix}$, $B_{1 \times 1}(x, y) = 2y + 2$ and

$D_{2 \times 1}(x, y) = \begin{bmatrix} x + 1 \\ 2 \end{bmatrix}$. Hence the numerators of the elements of

$D_{2 \times 1}(x, y) - C_{2 \times 1}(x, y)A_{1 \times 1}(x, y)^{-1}B_{1 \times 1}(x, y)$ are the polynomials

$$f_4(x, y) = x^2 - y^2 + 2x - 2y, \quad f_5(x, y) = -2y.$$

We have :

$x_0 = (0.04, -0.03)$	x_i	m_i	α_{i,m_i}
$f_4(x, y) = x^2 - y^2 + 2x - 2y$	x	1	0.0325
$f_5(x, y) = -2y$	y	1	0.03

At this step we have obtained the system of rank 2 :

$$(x^2 + y^2 + 2x + 2y, xy + x + y, x^2 + 2x + 2y, x^2 - y^2 + 2x - 2y, 2y).$$

Certification of the numerical deflating and kerneling operations

Let $x \in \mathbb{C}^n$ and $f = (f_1, \dots, f_m) \in \mathbb{C}^m[x]$ which admits a multiple root w .

Let $F(x)$ a deflated system obtained by numerical operations of deflating and kerneling. Let

$$\gamma_1(F, w) = \sup_{k \geq 2} \left(\frac{1}{k!} \|DF(w)^\dagger\| \|D^k F(w)\| \right)^{1/(k-1)}.$$

If we have

$$\gamma_1(F, w) \|x_0 - w\| \leq \frac{3 - \sqrt{7}}{2}$$

the Gauss-Newton sequence starting from x_0 converges quadratically to w .

Point estimates of the quantities λ and ρ of the deflated system.

The deflation method consists to determine a sequence of systems

$$\Sigma_0 = f$$

$$\Sigma_1 = \text{deflated}(\Sigma_0)$$

$$\Sigma_{k+1} = \text{deflated}(\text{kerneling}(\Sigma_k)), \quad k \geq 1.$$

We note by

$$[\Sigma_0]_w = \sum_{k \geq 1} \frac{\|D^k f(w)\|}{k!} t^k.$$

We can prove that the γ quantity is related to λ_0 and ρ_0 defined by

$$[\Sigma_0]_w \leq \frac{\lambda_0 t}{1 - \rho_0 t}.$$

Point estimates of the quantities λ and ρ of the deflated system.

Theorem. Let m_k be the maximum of the order of the deflation for the system Σ_k and r_k the rank of the jacobian matrix of the system *deflated*(Σ_k). We denote by $\mu = \max \mu_k$ where μ_k is the norm of a invertible submatrix of rank r_k of $D\Sigma_k$.

Then we have

$$\lambda_k \leq (4(\sqrt{2} + 1)\mu)^k (2 + \sqrt{2})^{\sum_{j=1}^k j(m_j+1)-1} \prod_{j=0}^k (m_j + 1) \lambda_0 \rho_0^{k + \sum_{j=0}^k m_j}$$

$$\rho_k \leq (2 + \sqrt{2})^k \prod_{j=0}^k \frac{m_j + 2}{2} \rho_0.$$

where

$$[\Sigma_k]_w \leq \frac{\lambda_k t}{1 - \rho_k t}.$$

Thanks for your attention.