INRIA

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

Regularized Two Level Algorithms for Model Problems

Jichao ZHAO — Jean-Antoine DÉSIDÉRI

N° 6382

August 8, 2007

Thème NUM



ISSN 0249-6399 ISRN INRIA/RR--6382--FR+ENG



Regularized Two Level Algorithms for Model Problems

Jichao ZHAO *, Jean-Antoine DÉSIDÉRI *

Thème NUM — Systèmes numériques
Projets Opale

Rapport de recherche n° 6382 — August 8, 2007 — 29 pages

Abstract: In previous work [16], we experimented a variant of the two-level ideal algorithm for parametric shape optimization that was proposed in [8], and two of three approaches were shown superior. In this report, we try to look them in a more classic way – in terms of singular value decomposition (SVD) and regularizations, for model problems, i.e., we combine these regularization tools with these two-level ideal algorithms. Numerical results show that regularized two level algorithms are more robust.

Key-words: FMOSA, Degree elevation, Bézier parameterization, RAE2822 airfoil, Shape optimization, Nonlinear optimization, Two level idea algorithm, Full multi-grid method.

* Opale research team

Unité de recherche INRIA Sophia Antipolis
2004, route des Lucioles, BP 93, 06902 Sophia Antipolis Cedex (France)
Téléphone : +33 4 92 38 77 77 — Télécopie : +33 4 92 38 77 65

Algorithmes à deux niveaux régularisés pour les problèmes modèles

Résumé : Dans les travaux précédents [16], nous a expérimenté une variante de l'algorithme idéalá deux niveaux pour l'optimisation paramétrique de forme dans laquelle a été proposé [8], et deux de trois approches ont été montrées le supérieur. Dans ce rapport, nous essayons de les regarder d'une manière plus classique – en termes de décomposition singuliére de valeur (SVD) et régularisations, pour les problèmes modéles, c.-á-d., nous combinons ces outils de régularisation avec ces algorithmes idéaux à deux niveaux. Exposition numérique de résultats qui a régularisé des algorithmes à deux niveaux étre plus robuste.

Mots-clés : FMOSA, Altitude de degré, Paramétrisation de Bézier, RAE2822 aile, Optimisation de forme, Optimisation non-linéaire, Algorithme à deux niveaux d'idée, algorithmes multigrille.

1 Introduction

In previous work [16], we give two level correction algorithms, Y', Z' and L' methods, for a linear case – shape reconstruction problem and a nonlinear case – inverse shape test problem respectively. And we also compare these results obtained by these two level algorithms and single parameterization approach.

In this report, we try to look them in a more classic way – in terms of singular value decomposition (SVD) and regularizations, for the inverse problem.

2 Classical Regularization Methods for Inverse Problems

In this part of this report, we first introduce classical regularization methods for inverse problems, i.e., Fredholm integral equations of the first kind, which are well discussed in Hansen' work [10, 11]. In the world of developing regularization methods for inverse problems and Fredholm integral equations of the first kind, the following generic form of a first-kind Fredholm equation is often used

$$\int_0^1 K(x,s)u(s)ds = g(x)\,, \qquad 0 \le x \le 1\,, \tag{1}$$

here the function $K(x,s)$, called the kernel, is a known function of two variables, x and s, and the right side function $g(x)$ is also given.

2.1 Smoothing and Inversion

It is well understood that the integration of u with the kernel function K in the first-kind Fredholm integral equation (1) is a smoothing operation that tends to dampen high-frequency components in u, such that the right side function g is a smoother function than u. However, in the context of inverse problems, i.e. we want to compute u from the right side function g, which is expected to amplify the high-frequency components and it is thus a so-called "desmoothing process".

The above statement is well explained by the following famous example. If we choose the function u to be given by

$$u(x) = \sin(2\pi m x)\,, \qquad m = 1, 2, \cdots\,, \tag{2}$$

such that the corresponding right-hand function g is given as

$$g(x) = \int_0^1 K(x,s)u(s)ds = \int_0^1 K(x,s)\sin(2\pi m s)ds\,, \qquad m = 1, 2, \cdots\,, \tag{3}$$

then according to the Riemann-Lebesgue lemma, it should hold

$$g \to 0 \;\text{ as }\; m \to \infty\,. \tag{4}$$

From this simple example, we can see that with increasing m, the higher the frequency of $u = \sin(2\pi mx)$, the more g is damped, which is true independently of the kernel function K, consequently, the inverse process, computing u from g, will amplify the high frequencies.

2.2 Various Regularization Methods for Inverse Problems

In this part, we try to solve an example of Fredholm integral equations of the first kind (one-dimensional image restoration model "Shaw") to explain current popular approaches to regularize inverse problems. Since Hansen already made a very powerful regularization Matlab tool [10] to analysize these inverse problems, we will take advantages of it. For the test problem, the kernel K and solution g are given by

$$K(x,s) = (\cos(x) + \cos(s))^2 (\frac{\sin(\pi \sin(x) + \sin(s))}{\pi \sin(x) + \sin(s)})^2 , \qquad m = 1, 2, \cdots , \tag{5}$$

and

$$u(x) = a_1 \exp(-c_1(t - t_1)^2) + a_2 \exp(-c_2(t - t_2)^2) , \tag{6}$$

parameters $a_1 = 2$, $a_2 = 1$,$c_1 = 6$,$c_2 = 2$, $t_1 = 0.8$ and $t_2 = -0.5$.

2.2.1 The "Naive" Solution and Discrete Picard Condition

There are a variety of schemes available to discretize the integral equation (1), we can just pick up any one of them to do the work:

$$g_i = \int_0^1 K(x_i, s)u(s)ds \approx \sum_{j=1}^{N} w_j K(x_i, s_j)u(s_j) , \tag{7}$$

and we can collect them into a linear matrix equation

$$AY = b , \tag{8}$$

and it is obvious that these entries of the matrix A, the right-hand vector b, and the vector Y can be calculated. The singular value desomposition (SVD) of a matrix is often use to analysis some properties in linear algebra. For the matrix A, SVD takes the following form:

$$A = U\Sigma V^T = \sum_{i=1}^{N} u_i \sigma_i v_i^T , \tag{9}$$

both matrix U and V are orthogonal, and the middle matrix

$$\Sigma = \begin{pmatrix} \sigma_1 & 0 & 0 & \cdots & 0 \\ 0 & \sigma_2 & 0 & & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & & 0 & \sigma_{N-1} & 0 \\ 0 & \cdots & 0 & 0 & \sigma_N \end{pmatrix}_{N \times N} ,$$

and

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_N \geq 0\,. \tag{10}$$

When the matrix A is symmetric, then it satisfies $U = V$. It is also to verify the following relations with regarding to singular values and vectors:

$$Av_i = \sigma_i u_i\,, \quad \|Av_i\| = \sigma_i\,, \quad i = 1, 2, \cdots, N\,, \tag{11}$$

as well as

$$b = \sum_{i=1}^{N}(u_i^T b)u_i\,, \quad Y = \sum_{i=1}^{N}(v_i^T Y)v_i\,, \quad AY = \sum_{i=1}^{N}\sigma_i(v_i^T Y)u_i\,. \tag{12}$$

Hence, the simplest and "naive" approach to solve the linear matrix equation (8) can be written as

$$Y_{\text{naive}} = A^{-1}b = \sum_{i=1}^{N}\frac{u_i^T b}{\sigma_i}v_i\,. \tag{13}$$

However the "naive" solution has a big problem, especially when the mesh grid N is a big number, which is well explained in Hansen's Matlab demo. To best convey our points in later sections for our model problems, here we redemonstrate part results in his demo.

Define the "noise" as follows

$$b_{\text{noise}} = b + e_{\text{noise}}\,. \tag{14}$$

Take $N = 32$, Fig. 1 shows singular values σ_i, coefficients $u_i^T b$ and $\frac{u_i^T b}{\sigma_i}$ in the naive solution Y of the matrix equation, (a) has no noise, while (b) has the noise level $e_{\text{noise}} = 10^{-3}$. From the (a), we can see that when $i < 17$, values of $u_i^T b$ decrease faster than those of corresponding σ_i, in fact discrete Picard condition tells us that it should always be true as long as the integral equation has square integrable solutions, while when $i > 17$ this condition fails to hold because of rounding errors so that coefficients $\frac{u_i^T b}{\sigma_i}$ become unreasonable big, which leads to high frequency errors in the naive solution considering the multipication by $\frac{u_i^T b}{\sigma_i}v_i$. For (b), we can observe similar results, only difference is that the turning point $i = 7$, leading to the failure of the discrete Picard condition, is smaller because of the existence noise $e_{\text{noise}} = 10^{-3}$ in the right hand vector.

2.2.2 Various Regularization Methods

It is straightward to think if we throw away all the terms when the discrete Picard condition doesn't hold, i.e., $i > 17$ for the case of no noise or $i > 7$ for the noise level $e_{\text{noise}} = 10^{-3}$ in the above example, then we could cure the ill-condition problem. Yes, it is a simple approach, called truncated SVD, to stabilize the solutions, which are less sensitive to the perturbations. This process is called regularization.

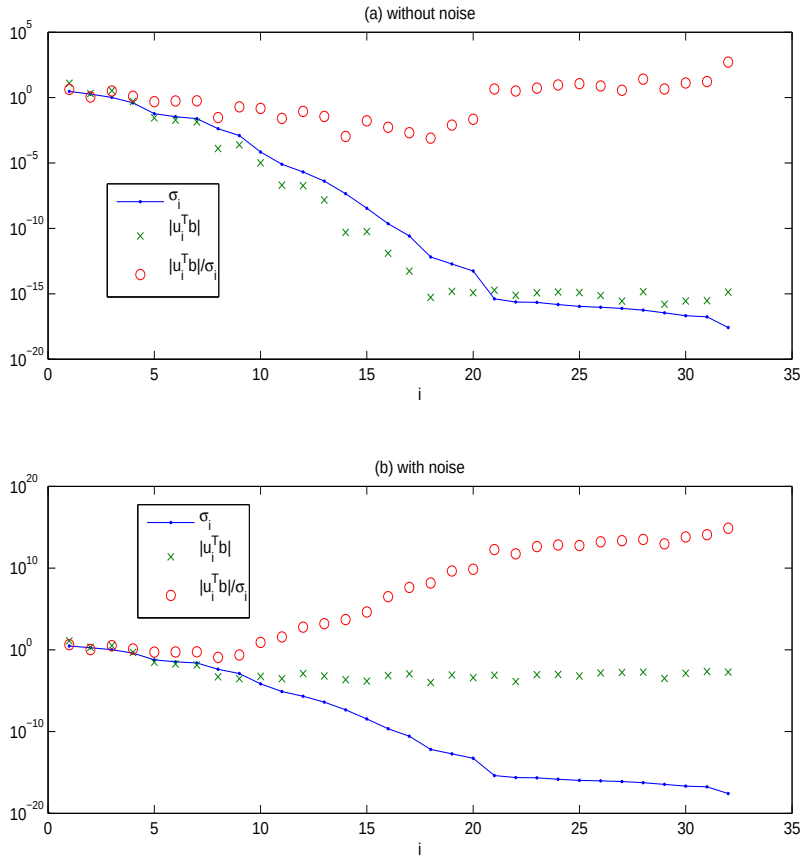


Figure 1: For $N = 32$, plot singular values σ_i, coefficients $u_i^T b$ and $\frac{u_i^T b}{\sigma_i}$ in the naive solution Y of the matrix equation. (a) has no noise, while (b) has the noise level 10^{-3}.

2.2.3 Truncated SVD

The truncated SVD (TSVD) approach is to actually compute the SVD and neglect all the undesired SVD components:

$$Y_{\text{TSVD}} = \sum_{i=1}^{k} \frac{u_i^T b}{\sigma_i} v_i \,, \tag{15}$$

where k is the maximum value which satisfies the discrete Picard condition.

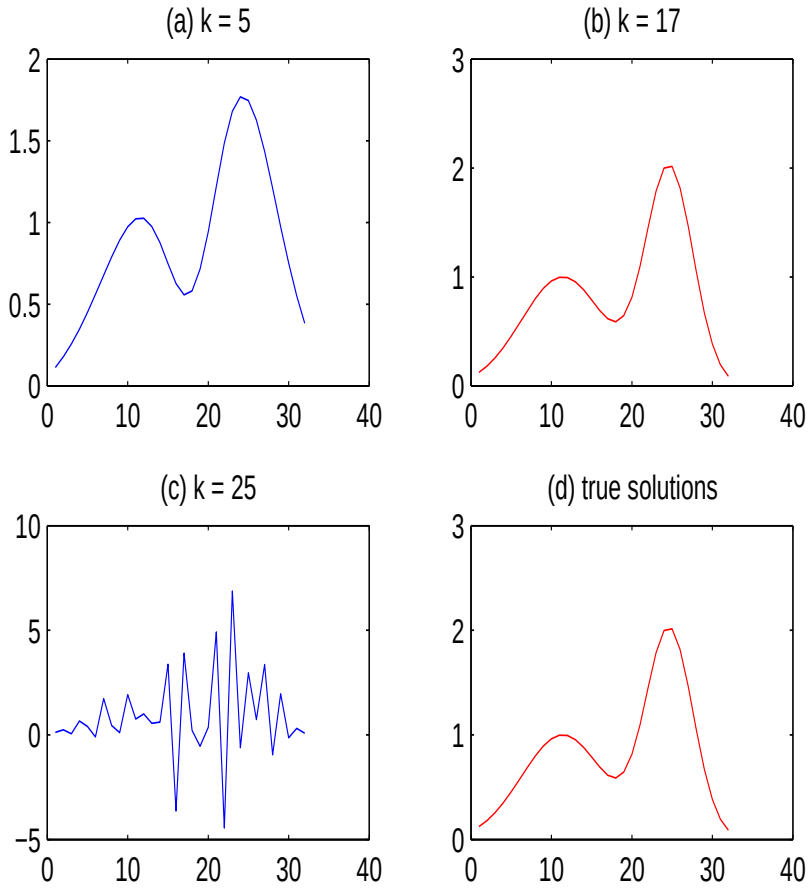


Figure 2: For $N = 32$ and without noise, plot solutions for the approach of TSVD with varying k and true solutions.

In Fig. 2, we plot solutions for the approach of TSVD with varying k and true solutions without noise, and we can see clearly that when k is smaller than the optimized value $k = 17$, approximation solutions are in some way squeezed, while k is larger than it, approximation solutions begin to oscillate.

As you can see for this TSVD method, we need to compute the singular values of the matrix, which is prohibitive for big size problems.

2.2.4 Tikhonov Regularization

The key idea of Tikhonov regularization is to accept a nonzero residual $AY - b$ and in return obtain a smaller solution norm, i.e., the problem can be formulated as follows

$$\min_Y \{\|AY - b\|_2^2 + \lambda^2 \|Y\|_2^2\} \,, \tag{16}$$

here the regularization parameter λ controls the weight of norm $\|Y\|_2$ relative to minimization of the residual norm $\|AY - b\|_2$. The deeper insight into the properties of the Tikhonov solutions Y_λ shows that

$$Y_\lambda = \sum_{i=1}^{N} \frac{\sigma_i^2}{\sigma_i^2 + \lambda^2} \frac{u_i^T b}{\sigma_i} v_i \,, \tag{17}$$

the quantities $T_i = \frac{\sigma_i^2}{\sigma_i^2+\lambda^2}$ are called the Tikhonov filter factors, which control the damping of the individual SVD components of the solution Y_λ. Specially, when $\sigma_i >> \lambda$ (i is comparable small), Tikhonov filter factors $T_i = \frac{\sigma_i^2}{\sigma_i^2+\lambda^2} \approx 1$, thus it does nothing for the low frequent components, while when $\sigma_i << \lambda$ (i is comparable large), Tikhonov filter factors $T_i = \frac{\sigma_i^2}{\sigma_i^2+\lambda^2} \approx \frac{\sigma_i^2}{\lambda^2}$, which helps to damp the coefficients of last SVD components for the high frequent components.

While in practice, we use the following normal equations equivalent to the (16):

$$(A^T A + \lambda^2 I)Y = A^T b \,, \tag{18}$$

when matrix A is symmetric and positive definite, there exists an alternative form of Tikhonov regularization:

$$(A + \lambda I)Y = b \,, \qquad \lambda \geq 0 \,, \tag{19}$$

and the regularized solution is given by

$$Y_\lambda = \sum_{i=1}^{N} \frac{w_i^T b}{\sigma_i + \lambda} w_i = \sum_{i=1}^{N} \frac{\sigma_i}{\sigma_i + \lambda} \frac{w_i^T b}{\sigma_i} w_i \,, \tag{20}$$

note that for the symmetric matrix, $w_i = u_i = v_i$.

2.2.5 Iteration Methods

For treating large-scale problems, where the direct computations involved computing SVD become prohibitive. Iteration methods, like the classic Landweber's methods and conjugate gradient (CG) algorithm, have been widely used for this problem. In these methods, iteration number k plays the role of the regularization parameter.

Another interersting idea is to use Krylov space methods. Let matrix M denote a preconditioner, thus $M^{-1}A \approx I$ and $M^{-1}(b - AY) \approx e$. The iteration can be constructed by

$$Y^{(K+1)} = Y^{(K)} + M^{-1}(b - AY^{(K)}) \,, \tag{21}$$

and it is easily that we obtain the following error relation:

$$e^{(K+1)} = (I - M^{-1}A)^K e^{(0)} \,, \tag{22}$$

3 Regularized Two level Multigrid-Type Algorithm for Shape Reconstruction Problems with Bézier Parameterization

We give two level ideal regularized correction algorithms, Y', Z' and L' methods, for a linear case – shape reconstruction problem and a nonlinear case – inverse shape test problem in the next section.

3.1 Shape Reconstruction Problem

$$\gamma : \, y(x); \text{ target: } \bar{y}(x)$$

$$J(\gamma) = \int_\gamma \frac{1}{2}(y(x) - \bar{y}(x))^2 dx \,, \tag{23}$$

or

$$J(Y) = \int_\gamma \frac{1}{2}(B_N(t)^T(Y - \bar{Y}))^2 N B_{N-1}(t)^T \Delta X^0 dt \,. \tag{24}$$

For the uniform case, we know that $x^0(t)' = NB_{N-1}(t)^T \Delta X^0 = 1$, so we only need to care about

$$J(Y) = \int_\gamma \frac{1}{2}(B_N(t)^T(Y - \bar{Y}))^2 dt \,. \tag{25}$$

On the fine level (in our experiment, we choose $N = 8$), we already know that the parametric gradient is linear (in DÉSIDÉRI and MAJD's pervious work):

$$J'(Y) = AY - b \,, \tag{26}$$

here the matrix

$$A = \int_0^1 B_N(t) B_N(t)^T dt = \{\frac{1}{2N+1} \frac{C_N^i C_N^j}{C_{2N}^{i+j}}\} \,. \tag{27}$$

and the right side known vector

$$b = A\bar{Y} \,. \tag{28}$$

It is same for all the methods on the fine level, that is, to use the following classical steepest-decent iteration:

$$Y^{j+1} = Y^j - \rho(AY^j - b) \,, \tag{29}$$

here $j = 0, 1, 2, \cdots$, Y^0 is a given initial guess, and we assume at $j = K$ we obtain the values Y^K on fine level, and go down to coarse level to make some corrections so that the approximations can converge faster than the single parameterization approach. For our analysis, we can decompose the matrix A into

$$A = W\Sigma_N W^T = \sum_{i=1}^{N} w_i \sigma_i w_i^T \,, \tag{30}$$

In the diagonal matrix Σ_N , real positive eigenvalues are arranged in **decreasing order** to keep consistence with theory of regularizations.

If we just employ iterations (29) when N is quite large, the solutions lead to oscillations. To cure the ill-conditions, we can use the alternative form of Tikhonov regularization (19) considering the matrix A is always symmetric and positive definite

$$(A + \lambda I)Y = b \,, \qquad \lambda \geq 0 \,, \tag{31}$$

and the regularized solution is given by

$$Y_\lambda = \sum_{i=1}^{N} \frac{w_i^T b}{\sigma_i + \lambda} w_i = \sum_{i=1}^{N} \frac{\sigma_i}{\sigma_i + \lambda} \frac{w_i^T b}{\sigma_i} w_i \,. \tag{32}$$

3.2 First Coarse Level Correction Method

In our experience, we choose $n = 4$ on the coarset level:

$$J(Y') = \int_\gamma \frac{1}{2} (B_N(t)^T (Y_K + E_n^N Y' - \bar{Y}))^2 dt \,. \tag{33}$$

Let $Y = E_n^N Y' + Y_K - \bar{Y}$, then we can easily obtain the following new matrix equation for coarse level corrections:

$$A_{cy} Y' = b_{cy} \,, \tag{34}$$

here the coefficient matrix

$$A_{cy} = (E_n^N)^T (A + \lambda I) E_n^N \,, \tag{35}$$

and the right side vector

$$b_{cy} = (E_n^N)^T(A + \lambda I)(-Y^K + \bar{Y}) = (E_n^N)^T(b - (A + \lambda I)Y^K)\,. \tag{36}$$

We can solve the following iteration on the coarse level for correction by initializing $Y_0' = 0$:

$$Y'^{j+1} = Y'^j - \rho(A_{cy}Y'^j - b_{cy})\,, \tag{37}$$

then we can update by $Y^K + E_n^N Y'$ on the fine level. To speed up the rate of convergence, we can use a better technique – Tchebychev iterations [9] on the fine level, i.e., it has three steps for each cycle:

$$\begin{aligned} Y^{j_1} &= Y^{j_0} - \tau_1(AY^{j_0} - b)\,, \\ Y^{j_2} &= Y^{j_1} - \tau_2(AY^{j_1} - b)\,, \\ Y^{j_3} &= Y^{j_2} - \tau_3(AY^{j_2} - b)\,, \end{aligned}$$

where τ_i $(i = 1, 2, 3)$ are given by:

$$\frac{1}{\tau_i} = \frac{b+a}{2} + \frac{b-a}{2}\, r_i \tag{38}$$

where $[a, b]$ is the targetted interval in the eigenvalue λ of A and $r_i = 0, \pm\sqrt{3}/2$ a root of the Tchebychev polynomial of degree 3.

Now let's turn to the analysis for the coarse level corrections, for simplicity, we can think $N = n$, i.e., that the elevation matrix E_n^N is identity matrix. Thus we have

$$A_{cy} = A + \lambda I\,, \tag{39}$$

according to the coarse corrections (34) of method Y', we obtain

$$Y' = \sum_{i=1}^{N} \frac{w_i^T(b_{cy} - (A + \lambda I)Y^K)}{\sigma_i + \lambda} w_i = \sum_{i=1}^{N} \frac{\sigma_i}{\sigma_i + \lambda} \frac{w_i^T(b_{cy} - (A + \lambda I)Y^K)}{\sigma_i} w_i\,, \tag{40}$$

3.3 Second Coarse Level Correction Method

For the coarse level correction of Z' method, we use the following scheme:

$$J(Z') = \int_\gamma \frac{1}{2}(B_N(t)^T(Y_k + Q_0 E_n^N Z' - \bar{Y}))^2 dt\,, \tag{41}$$

where $Q_0 = W P_N W^T$, matrix P_N is defined as

$$P_N = \begin{pmatrix} & & & 1 \\ & & 1 & \\ & \iddots & & \\ & 1 & & \\ 1 & & & \end{pmatrix}_{N \times N}\,.$$

Let $Z = (Q_0 E_n^N)Z' + Y^K - \bar{Y}$. In other words, we have the following new matrix equation:

$$A_{cz} Z' = b_{cz} \,, \tag{42}$$

here the matrix

$$A_{cz} = (E_n^N)^T Q_0 (A + \lambda I) Q_0 E_n^N = (E_n^N)^T A_1 E_n^N \,, \tag{43}$$

and the vector

$$b_{cz} = (E_n^N)^T Q_0 (b - (A + \lambda I) Y_k) \,. \tag{44}$$

Analysis for the coarse level corrections in the case of $N = n$, thus we have

$$\begin{aligned} A_{cz} &= Q_0 (A + \lambda I) Q_0 \\ &= (W P_N W^T) W \Sigma_N W^T (W P_N W^T) + \lambda I \\ &= (W P_N \Sigma_N P_N W^T) + \lambda I \\ &= (W P_N) \Sigma_N (W P_N)^T + \lambda I \,, \end{aligned}$$

from the above relation, we obtain

$$A = (W P_N)(\Sigma_N + \lambda I)(W P_N)^T = \sum_{i=1}^{N} w_{N+1-i} (\sigma_i + \lambda) w_{N+1-i}^T \,. \tag{45}$$

For the coarse corrections of method Z' and the expression of matrix P_N, we obtain

$$Z' = \sum_{i=1}^{N} \frac{w_{N+1-i}^T (b_{cz} - A_{cz} Z^K)}{\sigma_i + \lambda} w_{N+1-i} = \sum_{i=1}^{N} \frac{\sigma_i}{\sigma_i + \lambda} \frac{w_{N+1-i}^T (b_{cz} - A_{cz} Z^K)}{\sigma_i} w_{N+1-i} \,. \tag{46}$$

3.4 Third Coarse Level Correction Method

Matrix Δ is defined as

$$\Delta = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & -1 & 0 & & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & & 0 & (-1)^N & 0 \\ 0 & \cdots & 0 & 0 & (-1)^{N+1} \end{pmatrix}_{N \times N} \,,$$

L' Method uses the idea of the Z' Method to reverse pairing between eigenvalues and eigenvectors by a simple matrix Δ instead of computing the complicated matrix Q_0.

For the coarse level correction of L' method, we use the following scheme:

$$J(L') = \int_\gamma \frac{1}{2}(B_N(t)^T(Y^K + \Delta E_n^N L' - \bar{Y}))^2 dt \,. \tag{47}$$

Let $L = (\Delta E_n^N)L' + Y^K - \bar{Y}$. In other words, we have the following new matrix equation:

$$A_{cl} L' = b_{cl} \,, \tag{48}$$

here the matrix

$$A_{cl} = (E_n^N)^T \Delta (A + \lambda I) \Delta E_n^N = (E_n^N)^T A_2 E_n^N \,, \tag{49}$$

and the vector

$$b_{cl} = (E_n^N)^T \Delta (b_{cl} - (A + \lambda I) Y^K) \,. \tag{50}$$

Analysis for the coarse level corrections in the case of $N = n$, thus we have

$$\begin{aligned} A_{cl} &= \Delta(A + \lambda I)\Delta \\ &= \Delta A \Delta + \lambda I \\ &= \Delta W \Sigma_N W^T \Delta + \lambda I \\ &= (\Delta W)\Sigma_N (\Delta W)^T + \lambda I \,, \end{aligned}$$

by the defination of the matrix Δ, we have

$$A = (\Delta W)(\Sigma_N + \lambda I)(\Delta W)^T = \sum_{i=1}^{N} \tilde{w}_i (\sigma_i + \lambda) \tilde{w}_i^T \,, \tag{51}$$

where

$$\tilde{w}_i = \begin{pmatrix} w_{1,i} \\ -w_{2,i} \\ w_{3,i} \\ -w_{4,i} \\ \vdots \\ (-1)^{N-1} w_{N-1,i} \\ (-1)^N w_{N,i} \end{pmatrix}, \quad \text{note that} \quad w_i = \begin{pmatrix} w_{1,i} \\ w_{2,i} \\ w_{3,i} \\ w_{4,i} \\ \vdots \\ w_{N-1,i} \\ w_{N,i} \end{pmatrix},$$

For the coarse corrections of method L', we obtain

$$L' = \sum_{i=1}^{N} \frac{\tilde{w}_i^T (b_{cl} - A_{cl} Y^K)}{\sigma_i + \lambda} \tilde{w}_i = \sum_{i=1}^{N} \frac{\sigma_i}{\sigma_i + \lambda} \frac{\tilde{w}_i^T (b_{cl} - A_{cl} Y^K)}{\sigma_i} \tilde{w}_i \,. \tag{52}$$

3.5 Numerical Experiments

Firstly let's look at discrete Picard conditions, i.e., singular values σ_i, coefficients $u_i^T b$ and $\frac{u_i^T b}{\sigma_i}$ in the naive solution Y of the linear shape reconstruction problem with $N = 16$. In Fig. 3 (a) has no noise, while (b) has the noise level 10^{-7}, and it seems that maximum $k = 15$ for the case of no noise, and $k = 12$ for the case of (b).

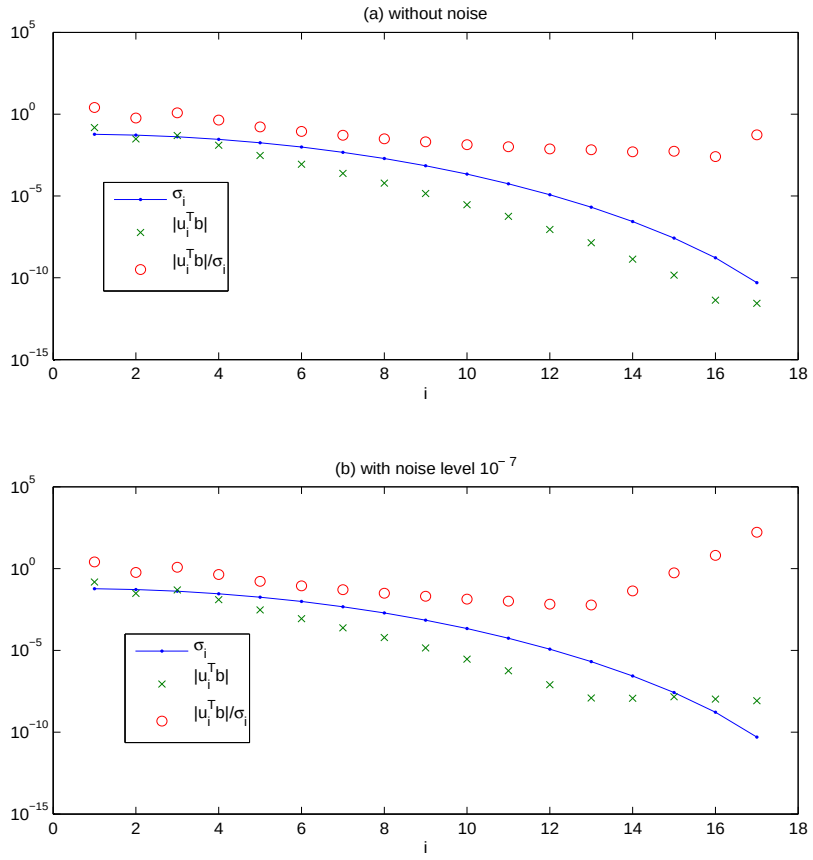


Figure 3: When $N = 16$, we plot singular values σ_i, coefficients $u_i^T b$ and $\frac{u_i^T b}{\sigma_i}$ in the naive solution Y of the linear shape reconstruction problem. (a) has no noise, while (b) has the noise level 10^{-7}.

3.5.1 Single Level Regularizations

First we use regularized iteration on a single level, in our numerical experiments, we still choose $N = 16$. And we take the case of the noise level 10^{-7}. In Fig. 5, we plot singular values σ_i, coefficients $u_i^T b$ and $\frac{u_i^T b}{\sigma_i}$ in the naive solution Y and Tikhonov regularizated solution Y_λ of the linear shape reconstruction problem. By these comparison, we can easily that see that regularizated solution Y_λ really works, while naive solution Y is very oscilated from Fig. 4 and the discrete Picard conditions shown in Fig. 5 explain the reason for it as we expect.

3.5.2 Two Level Regularizations

Both numerical experiments and theory analysis that method Y and L are very close to each other, since method Y is more simple, thus we omit discussion of method L.

We take $N = 16$ for the fine level, and $n = 6$ for the coarse level in our test. Fig. 6 and 7 show singular values σ_i, coefficients $u_i^T b'$ and $\frac{u_i^T b'}{\sigma_i}$ in Tikhonov regularizated solution Y_λ of the linear shape reconstruction problem with noise level 10^{-7} for the coarse level correction (40) of the method Y' and the coarse level correction (46) of the method Z' respectively. And we take $\lambda = 7 \cdot 10^{-5}$ for both cases, $b' = b_{cy} - (A + \lambda I)Y^K$ for the first case and $b' = b_{cz} - (A + \lambda I)Y^K$ for the latter case. Based these data, we have the following remarks and observations:

1. The discrete Picard conditions for coarse level corrections should be a little difference with classical ones. Or put in a more specific way is that coefficiences $\frac{u_i^T b'}{\sigma_i}$ in SVD should smaller that those in classic texts, and $u_i^T b'$ should have a distance below the values of σ_i instead of just a little smaller.

2. By comparing Fig. 6 and 7, we can see that method of Z' converges much faster, and it is good at reducing both high and low frequency, one shortcomings of Z' method is that it damps the middle frequency slowly.

3. From our numerical experiments, that the optimization value λ is around $7 \cdot 10^{-5}$ for the method Y', while the method Z' has more choice of the value of λ, $(7 \cdot 10^{-5} \leq \lambda \leq 5 \cdot 10^{-6})$.

4. Also from our experiments we found that the resudial $AY - b$ (about 0.01) is very close to each other among these methods $--$ regularized single parameterizetion and two level regularized algorithms when their solutions are convergent.

The difference among these methods is that the errors between approximation values and target values at mesh grids instead of residual errors. We define the maximum absolute value e as follows:

$$e = \max_{i \in \{0,1,\cdots,N\}} |Y_i - \bar{Y}_i| \,, \tag{53}$$

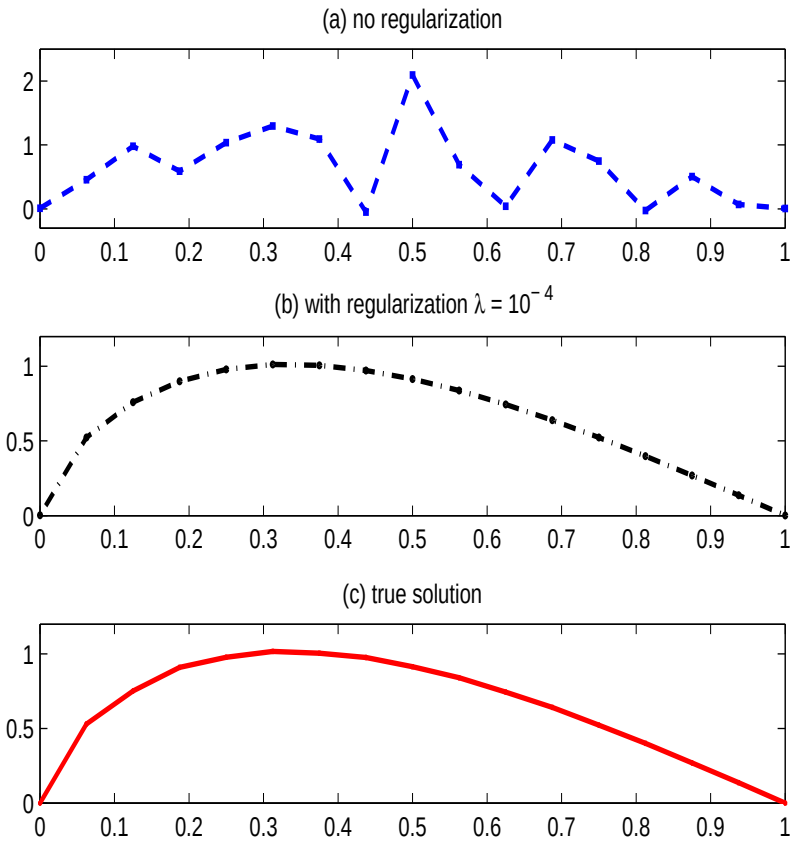


Figure 4: Numerical solutions for setting the regularization parameterization λ equate 0 and 10^{-4}, and true solutions respectively.

here Y_i are approximation values obtained by different methods, and $\bar{Y}_i$ are true values. In last experiments, we vary the number of iterations and compare these maximum absolute values among these methods. We can tell easily that method Z converges very faster.

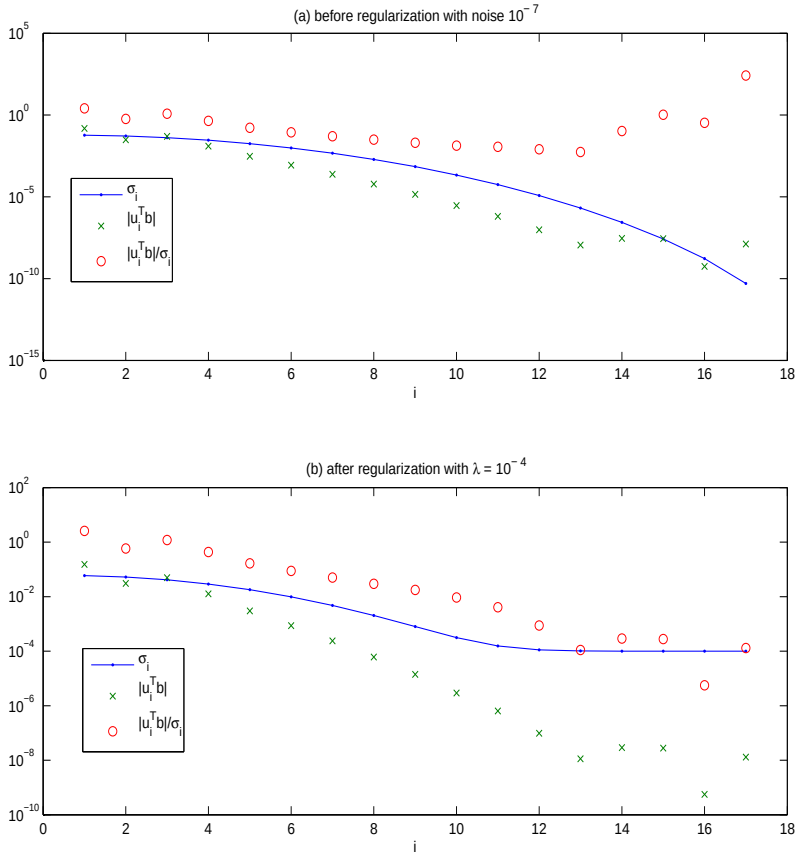


Figure 5: When $N = 16$, we plot singular values σ_i, coefficients $u_i^T b$ and $\frac{u_i^T b}{\sigma_i}$ in the naive solution Y and Tikhonov regularizated solution Y_λ of the linear shape reconstruction problem with noise level 10^{-7}. (a) is the naive approach, (b) is the approach of Tikhonov regularization.

4 Nonlinear Model Problem

We can also apply the regularized method to the nonlinear model problem as we did for the linear shape reconstruction problem. The nonlinear model problem, which we discuss here,

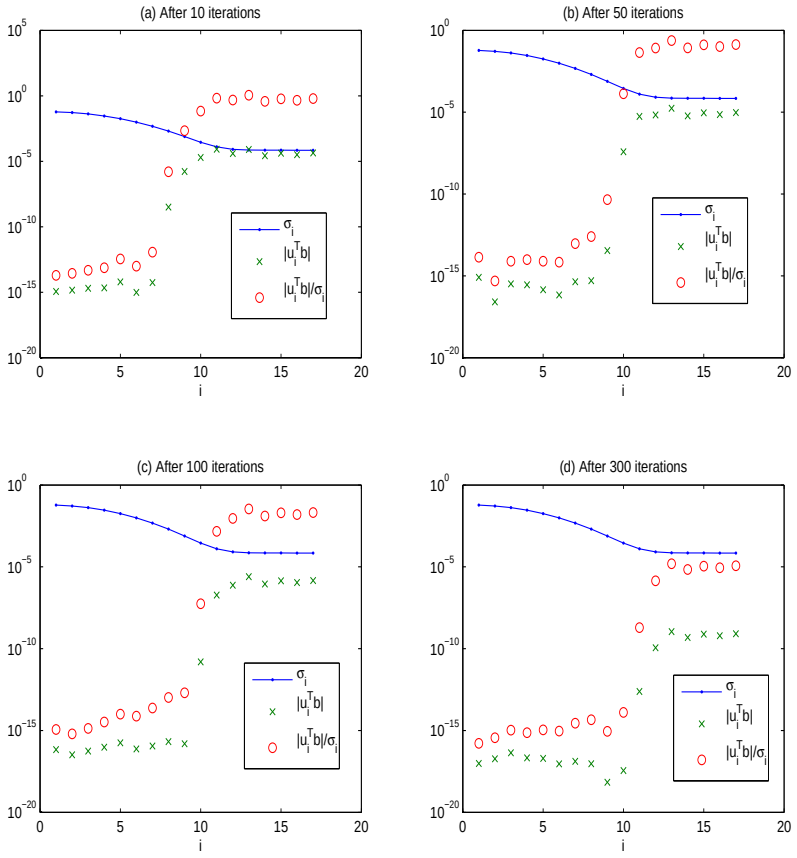


Figure 6: When $N = 16$ and $n = 6$, we plot singular values σ_i, coefficients $u_i^T b'$ and $\frac{u_i^T b'}{\sigma_i}$ in Tikhonov regularizated solution Y_λ of the linear shape reconstruction problem with noise level 10^{-7} on the coarse level correction (40) of the method Y', here $b' = b_{cy} - (A + \lambda I)Y^K$ and $\lambda = 7 \cdot 10^{-5}$.

is the following model inverse-shape test problem

$$\min_Y \mathcal{J} = \mathcal{J}(y(t)) = \frac{p^\alpha}{\mathcal{A}}\,, \tag{54}$$

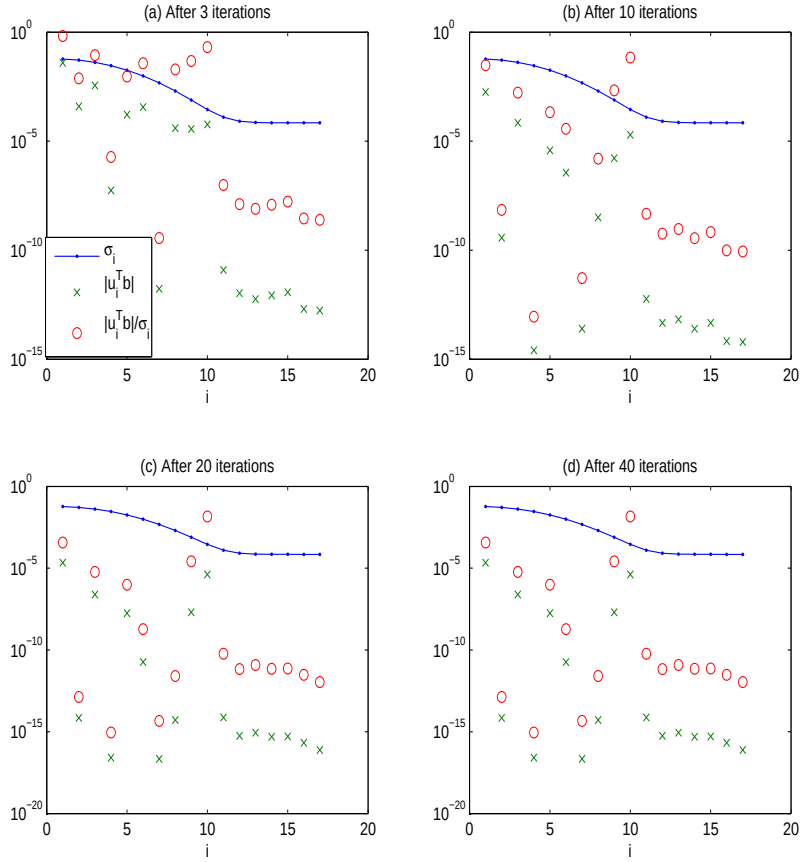


Figure 7: When $N = 16$ and $n = 6$, we plot singular values σ_i, coefficients $u_i^T b'$ and $\frac{u_i^T b'}{\sigma_i}$ in Tikhonov regularizated solution Y_λ of the linear shape reconstruction problem with noise level 10^{-7} on the coarse level correction (46) of the method Z', here $b' = b_{cz} - (A + \lambda I)Y^K$ and $\lambda = 7 \cdot 10^{-5}$.

with

$$p = \int_0^1 \sqrt{x'(t)^2 + y'(t)^2}\omega(t)\ dt\,, \qquad \mathcal{A} = \int_0^1 x'(t)y(t)\omega(t)\ dt\,, \tag{55}$$

p and $\mathcal{A}$ are, for specified $\omega(t) > 0$ and $\alpha > 1$, the pseudo-length of the arc, and the pseudo-area below the arc in which $x(t)$ is given, smooth and monotone-increasing, and both $x(t)$

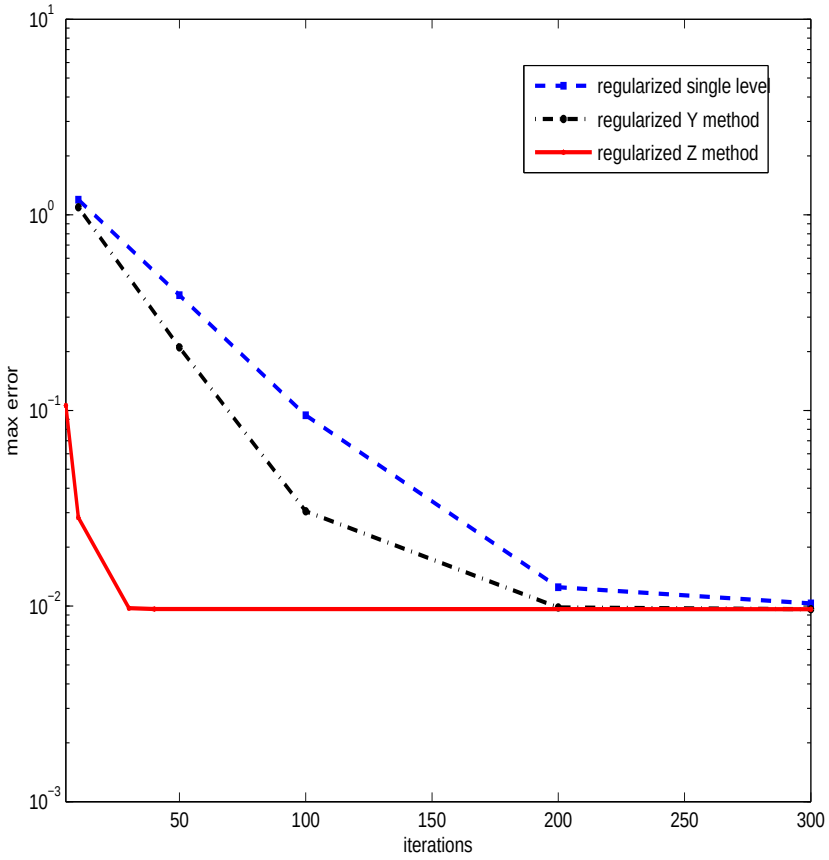


Figure 8: When $N = 16$, we compare these methods with varying the number of iterations.

and $y(t)$ are decided by control polygon

$$x(t) = \sum_{i=0}^{N} B_N^i(t) X_i \,, \qquad y(t) = \sum_{i=0}^{N} B_N^i(t) Y_i \,, \tag{56}$$

Y is a vector containing design variable Y_i for $i = 0, 1, \cdots, N$. To make our life easier, we choose $\alpha = 2$ and $\omega(t) = 1$ for $\forall\ t$, and we know the minimum value for the nonlinear problem (54) is $\mathcal{J} = 2\pi$.

As we demonstrate for linear cases, we can employ Tikhonov regularization (19) similarly for the nonlinear model problem since we can also prove that this nonlinear problem is symmetric.

$$\min_{Y} \mathcal{J}_\lambda = \mathcal{J}_\lambda(y(t)) = \frac{p^\alpha}{\mathcal{A}} + \frac{\lambda}{2}\|Y\|_2^2 \,, \tag{57}$$

here λ is the parameter of regularizations.

4.1 Iterations on Fine Level

We have known it is same for all the methods on the fine level, that is, to use the following classical steepest-decent iteration:

$$Y^{j+1} = Y^j - \rho \mathcal{J}'_\lambda(Y^j) \,, \tag{58}$$

here $j = 0, 1, 2, \cdots$, Y^0 is a given initial guess, and we assume at $j = K$ we obtain the values Y^K on fine level, and go down to coarse level to make some corrections so that the approximations can converge faster than the single parameterization approach.

To best explain our ideas, we need the Jacobian matrix of $\mathcal{J}'_\lambda(Y^j)$, denoted by $A_{\mathcal{J}'_\lambda}$, then we obtain:

$$\mathcal{J}'_\lambda(Y) = A_{\mathcal{J}'_\lambda} Y - b_{\mathcal{J}'_\lambda} = (A_{\mathcal{J}'} + \lambda I_N)Y - b_{\mathcal{J}'} \,, \tag{59}$$

note that vector $b_{\mathcal{J}'_\lambda} = b_{\mathcal{J}'}$ and $A_{\mathcal{J}'_\lambda} = A_{\mathcal{J}'} + \lambda I_N$, here Jacobian matrix $\mathcal{J}'(Y^j)$ and vector $b_{\mathcal{J}'}$ come from the original nonlinear problem (54)

$$\mathcal{J}'(Y) = A_{\mathcal{J}'} Y - b_{\mathcal{J}'} \,. \tag{60}$$

As we did for the linear shape-reconstruction problem, we can decompose matrix $A_{\mathcal{J}'_\lambda}$ into

$$A_{\mathcal{J}'} = W \Sigma_N W^T = \sum_{i=1}^{N} w_i \sigma_i w_i^T \,. \tag{61}$$

In the diagonal matrix Σ_N, real positive eigenvalues are arranged in **decreasing order**, and W holds

$$W W^T = W^T W = I \,. \tag{62}$$

Now we have done all the preparation, we are already to move to discuss various of coarse level corrections for this nonlinear model.

4.2 First Coarse Level Correction Method

For the Y' method, we set

$$Y = Y^K + E_n^N Y' \,, \tag{63}$$

here Y^K is the value obtained on the fine level, and Y' is the coarse level correction we are looking for.

By using coarse level correction (63) in the regularized nonlinear problem (57), we can easily obtain the following relations for the Y' correction by chain rules:

$$\mathcal{J}'_\lambda(Y') = {E_n^N}^T \cdot \mathcal{J}'_\lambda(Y) = A'_{\mathcal{J}'_\lambda} Y' - b'_{\mathcal{J}'} \,, \tag{64}$$

here $A'_{\mathcal{J}'_\lambda}$ stands for the coarse level Jacobian matrix, its value can be obtained by

$$A'_{\mathcal{J}'_\lambda} = {E_n^N}^T A_{\mathcal{J}'} E_n^N + \lambda (E_n^N)^T E_n^N \,, \tag{65}$$

and the constant vector $b'_{\mathcal{J}'_\lambda}$

$$b'_{\mathcal{J}'_\lambda} = {E_n^N}^T (b_{\mathcal{J}'} - (A_{\mathcal{J}'} + \lambda I_N) Y^K) \,, \tag{66}$$

$A_{\mathcal{J}'}$ is the Jacobian matrix on the fine level of the original nonlinear problem (54). Thus we can again employ classical steepest-descent iterations on the coarse level for correction by initializing $Y'_0 = 0$:

$$Y' \longleftarrow Y' - \rho \mathcal{J}'_\lambda(Y') \,, \tag{67}$$

then we can update by $Y^K + E_n^N Y'$ on the fine level. Iterations on the fine level (58) and corrections on the coarse level (67) complete a two-level correction-type idea algorithm for the nonlinear model problem. However it takes many iterations (hundreds for this model problem) to achieve complete convergence. Employing Tchebychev iterations on fine levels can converge much faster.

4.3 Second Coarse Level Correction Method

The Z' method of coarse level corrections for the nonlinear model is given as follows:

$$Y = Y^K + Q_0 E_n^N Z' \,, \tag{68}$$

here Y^K is the value obtained on the fine level, Z' is the coarse level correction we are looking for, $Q_0 = W P_n W^T$, and P_n is the same permutation matrix as we used for linear problems. The idea of the Z' Method is to reverse pairing between eigenvalues and eigenvectors by multiplying the matrix W so that larger eigenvalues pair with higher frequency on the coarse level and relaxations can remove high frequent errors efficiently [8].

By using coarse level correction (68) in the original nonlinear problem (57), we can easily obtain the following relations for the Z' correction by chain rules:

$$\mathcal{J}'_\lambda(Z') = E_n^{N^T} \cdot Q_0 \cdot \mathcal{J}'(Y) = A'_{\mathcal{J}'_\lambda} Z' - b'_{\mathcal{J}'} \,, \tag{69}$$

here $A'_{\mathcal{J}'_\lambda}$ stands for the coarse level Jacobian matrix of the Z' method (we share the same notation with Y' method for simplicity), its value can be obtained by

$$A'_{\mathcal{J}'_\lambda} = E_n^{N^T} Q_0 A_{\mathcal{J}'} Q_0 E_n^N + \lambda (E_n^N)^T E_n^N \,, \tag{70}$$

and the constant vector $b'_{\mathcal{J}'_\lambda}$

$$b'_{\mathcal{J}'_\lambda} = E_n^{N^T} Q_0 (b_{\mathcal{J}'} - (A_{\mathcal{J}'} + \lambda I_N) Y^K) \,. \tag{71}$$

Thus we can again employ classical steepest-descent iterations on the coarse level for correction by initializing $Z'_0 = 0$:

$$Z' \longleftarrow Z' - \rho \mathcal{J}'_\lambda(Z') \,, \tag{72}$$

then we can update by $Y^K + Q_0 E_n^N Z'$ on the fine level. Iterations on the fine level (58) and corrections on the coarse level (72) complete another two-level correction-type idea algorithm for the nonlinear model problem. To speed up the rate of convergence, we can use Tchebychev iterations [9] on the fine level.

4.4 Numerical Experiments

For the nonlinear problem (54), we take $N = 8$ for the fine level, and $n = 4$ for the coarse level. And in all our experiences, we take the following initial guess: Then we elevate their

Table 1: Initial Values

X^0	0	0	0.077	0.409	1.0
Y^0	0	0.01	0.01	0.01	0

degrees by the matrix E_4^8:

$$X^0 \longleftarrow E_4^8 X^0, \qquad \text{and} \quad \mathrm{Y}^0 \longleftarrow \mathrm{E}_4^8 \mathrm{Y}^0 \,. \tag{73}$$

When we analysis our numerical experiences, we define "error" as the difference between the approximation $\mathcal{J}(Y)$ obtained by different methods concerned in our work and the true solution $\mathcal{J} = 2\pi$.

First, we test the cases for methods Y and Z without regularizations. From Fig. 9, we can easily see that the control polygon obtained by the Z method is more oscillatory, but it obtains smaller object function values and converges faster.

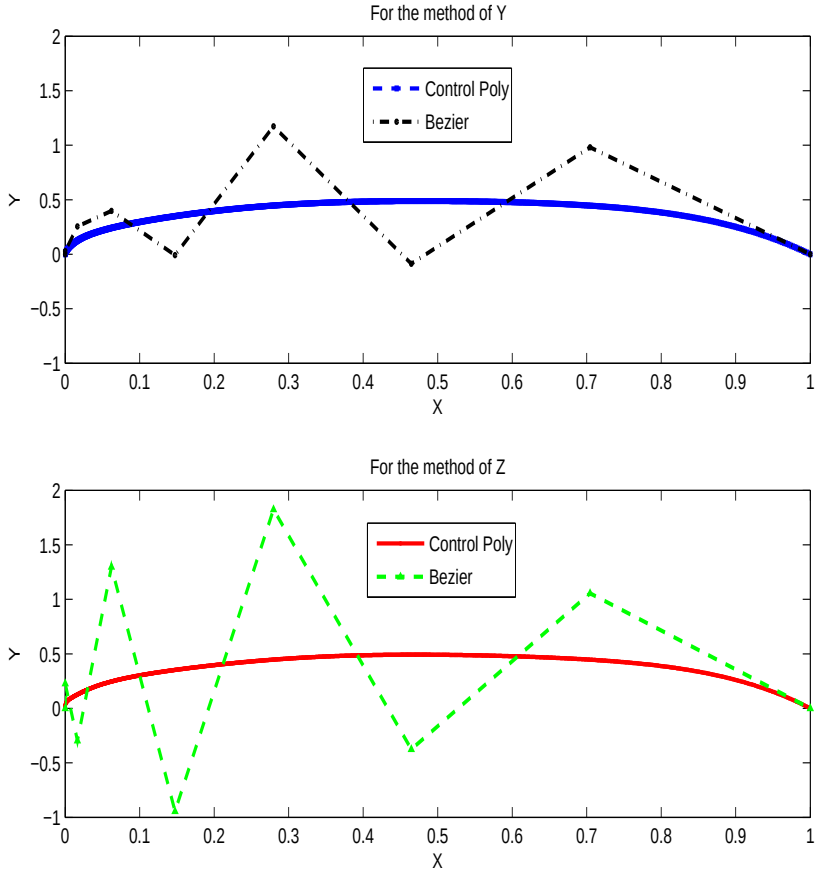


Figure 9: We plot the control polygon and Bézier profile for methods Y and Z without regularizations.

Next we use the regularization approach. We show the Bézier profile and control polygon with varying regularized parameter λ using the coarse level correction method Y and Z in Fig. 10 and 11. And from them, we can see easily that with decreasing the value of regularized parameter λ, the control polygon becomes more and more oscillation, while it may has smaller and smaller residual errors. Thus the value of regularized parameter λ is a trade-off between residual errors and oscillation of design variables. In the well developed ill-conditioned theory, we can use L-curve to decide the compromise between minimization of these two quantities, refer to Hansen [10] for details.

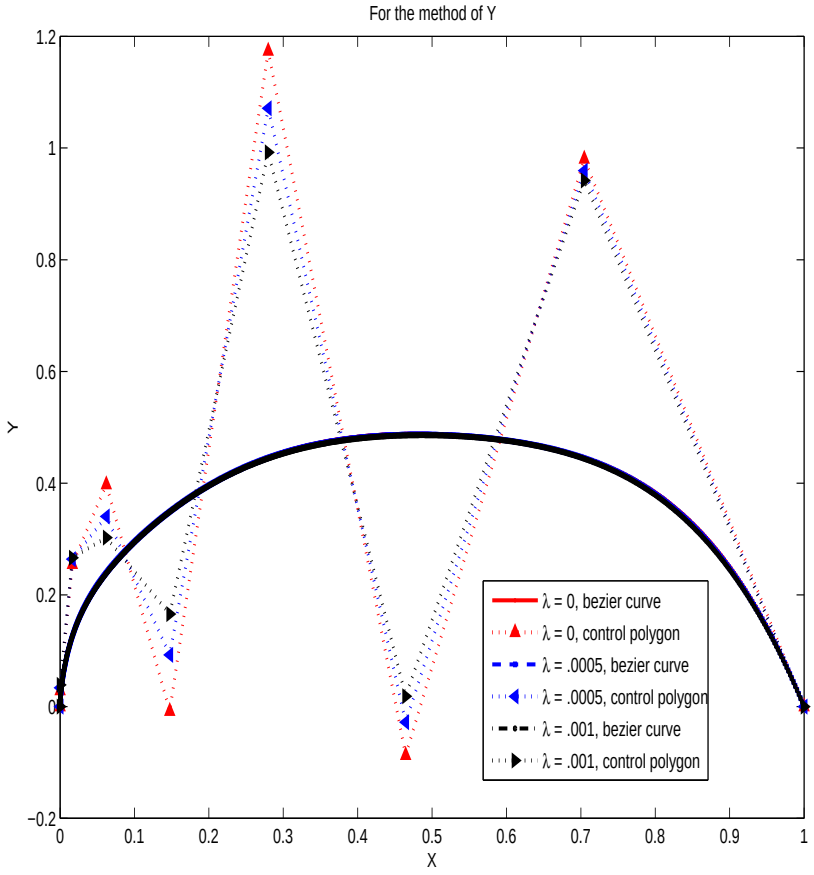


Figure 10: When $N = 8$ and $n = 4$, the Bézier profile and control polygon with varying regularized parameter λ using the coarse level correction method Y.

It may also be interesting to see what will happen for the single level method, two-level methods Y and Z with a fixing regularization parameter λ and increasing the number of iterations. In Fig. 12, we plot these errors against the number of iterations, and take $\lambda = .0002$ for the three methods. From these results, we can easily see that coarse level correction Z converges much faster, while the other two methods are almost same. Now let's focus more on the figure and try to explain in details: there are two time coarse level corrections, one happens at the iteration 22, the other one does at the iteration 68, where we can see two sharp falling of the errors for the method of Z. While for this nonlinear

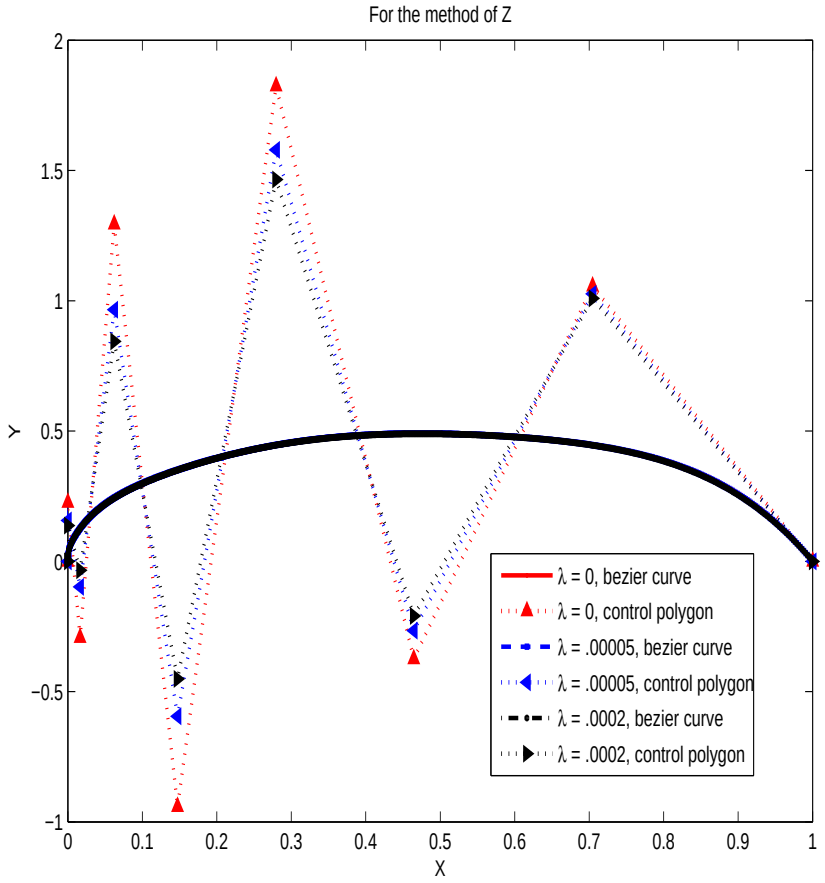


Figure 11: When $N = 8$ and $n = 4$, the Bézier profile and control polygon with varying regularized parameter λ using the coarse level correction method Z.

problem, the coarse level correction method Y doesn't help too much, as the sequence, it only obtains almost the same accuracy with the single level method with varying the number of interations.

At the end of this section, we discuss other approaches to regularize ill-condition problems. In fact for some problems, it may be better to minimize the 2-norm of a quantity that approximates a derivative of the underlying soluution, such as central difference approximations to the first and second derivatives, denoted by H_1 and H_2 respectively, where matrices

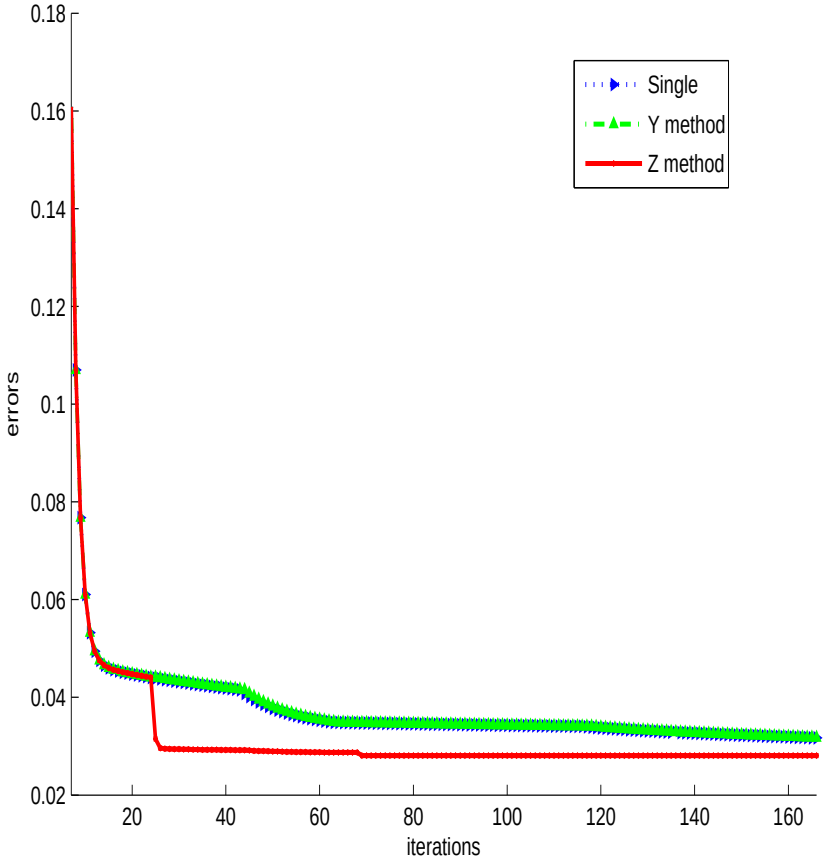


Figure 12: Compare errors among single level method, two-level methods Y and Z with increasing the number of iterations and the fixed regularization parameter $\lambda = .0002$.

H_1 and H_2 are given by

$$H_1 = \begin{pmatrix} -1 & 1 & 0 & \cdots & 0 \\ 0 & -1 & 1 & & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & & 0 & -1 & 1 \end{pmatrix}_{N \times (N+1)},$$

and

$$H_2 = \begin{pmatrix} 1 & -2 & 1 & \cdots & 0 \\ 0 & \ddots & \ddots & \ddots & \\ & 0 & 1 & -2 & 1 \end{pmatrix}_{(N-1)\times(N+1)} .$$

This leads to the following formulation of Tikhonov regularization in the general form:

$$\min_Y \mathcal{J}_\lambda = \mathcal{J}_\lambda(y(t)) = \frac{p^\alpha}{\mathcal{A}} + \frac{\lambda}{2}\|H_j Y\|_2^2 , \qquad j = 1 \text{ or } j = 2 . \tag{74}$$

For the nonlinear model problem, it seems all the three Tikhonov regularization approaches (57) and (74) obtain very similar results.

5 Conclusions

In this report, we combined the regularization approach with two level idea algorithms based on our previous work for model problems, and these results are quite promising and interesting.

References

[1] W. L. Briggs, V. E. Henson, S. F. McCormick, A multigrid tutorial, SIAM, Philadelphia, PA, 2000.

[2] J.-A. Désidéri, A. Dervieux, Hierarchical Methods for Shape Optimization in Aerodynamics – I: multilevel parametric; II: multilevel preconditioners, Introduction to Optimization and Multidisciplinary Design, Applications to Aeronautics and Turbomachinery, Von Karman Institute, March 6-10, 2006. http://www-sop.inria.fr/opale/CONFS/desideri-vki2006.pdf.

[3] J.-A. Désidéri, Multilevel Strategies for Shape Optimization in Aerodynamics, in: Numerical Analysis and Scientific Computing for PDE's and their Challenging Applications Conference, Helsinki (Finland), Center for Scientific Computing, October 6-7 2005.

[4] J.-A. Désidéri, Self-adaptive multilevel algorithms for shape optimization with application to Aerodynamics, in: MEGADESIGN Workshop, German Aerospace Center (DLR), Braunschweig (Germany), November 29-30 2005.

[5] J.-A. Désidéri, Shape optimisation algorithms in aerodynamics: parameterisation issues, in: European Research Community On Flow, Turbulence And Combustion (ERCOFTAC) Introductory Course on Design Optimisation, University of Manchester of Science and Technology (UMIST), UK, April 6-8 2005.

[6] J.-A. Désidéri, Two-Parameterization Ideal Algorithm for Shape Optimization, in: International Conference on Advances in Numerical Mathematics, Moscow (Russia), September 16-17 2005.

[7] J.-A. Désidéri, Y. Kuznetsov, P. Neittanmaki and O. Pironneau, Hierarchical Optimum-Shape Algorithms Using Embedded Bezier Parameterizations, Numerical Methods for Scientific Computing Variational Problems and Applications, Cimne, Barcelona, 2003.

[8] J.-A. Désidéri, Two-level ideal algorithm for parametric shape optimization, to appear, J. Numer. Math, 2006.

[9] J.-A. Désidéri, Modéles discrets et scheémas itératifs -application aux algorithmes multigrilles et multidomaines, HERMES, Paris, 1998.

[10] P. C. Hansen, Regularization tools: A Matlab package for analysis and solution of discrete ill-posed problems, Verson 3.1 for Matlab 6.0, 2001.

[11] P. C. Hansen, Deconvolution and regularization with Toeplitz matrices, Numerical Algorithms, Vol. (29), pp. 323-378, 2002.

[12] V. E. Henson, Multigrid methods for nonlinear problems: an overview, 2000.

[13] B. Abou El Majd, J.-A. Désidéri, T. T. Do, L. Fourment, A. Habbal, A. Janka, Multilevel Strategies and Hybrid Methods for Shape Optimization and Applications to Aerodynamics and Metal Forming, in: Evolutionary and Deterministic Methods for Design, Optimisation and Control with Applications to Industrial and Societal Problems Conference (EUROGEN 2005), Munich GermanyR. Schilling, W. Haase, J. Priaux, H. Baier (editors), ISBN: 3-00-017534-2, September 12-14 2005.

[14] B. Abou El Majd, J.-A. Désidéri, A. Habbal, A Fully Multilevel and Adaptive Algorithm for Parameterized Shape Optimization, in: European Multigrid Conference, EMG 2005, Scheveningen, The Netherlands, September 27-30 2005.

[15] C. Zillober: Software manual for SCPIP 2.3, Technical Report No. TR02-1, Informatik, Universität Bayreuth, 2002.

[16] J. Zhao, J. A. Desideri and B. Abou El Majd, Two Level Correction Algorithms for Model Problems, 2007-07, INRIA, France, Research Report 6246.



Unité de recherche INRIA Sophia Antipolis
2004, route des Lucioles - BP 93 - 06902 Sophia Antipolis Cedex (France)

Unité de recherche INRIA Futurs : Parc Club Orsay Université - ZAC des Vignes
4, rue Jacques Monod - 91893 ORSAY Cedex (France)
Unité de recherche INRIA Lorraine : LORIA, Technopôle de Nancy-Brabois - Campus scientifique
615, rue du Jardin Botanique - BP 101 - 54602 Villers-lès-Nancy Cedex (France)
Unité de recherche INRIA Rennes : IRISA, Campus universitaire de Beaulieu - 35042 Rennes Cedex (France)
Unité de recherche INRIA Rhône-Alpes : 655, avenue de l'Europe - 38334 Montbonnot Saint-Ismier (France)
Unité de recherche INRIA Rocquencourt : Domaine de Voluceau - Rocquencourt - BP 105 - 78153 Le Chesnay Cedex (France)

Éditeur
INRIA - Domaine de Voluceau - Rocquencourt, BP 105 - 78153 Le Chesnay Cedex (France)
http://www.inria.fr
ISSN 0249-6399