Iterative methods: from theory to practice (A tutorial)

Yousef Saad, ***Ruipeng Li,*** ***Yuanzhe Xi***
(Minnesota) ***(LLNL)*** ***(Emory)***

Copper Mountain Conference
on Iterative Methods

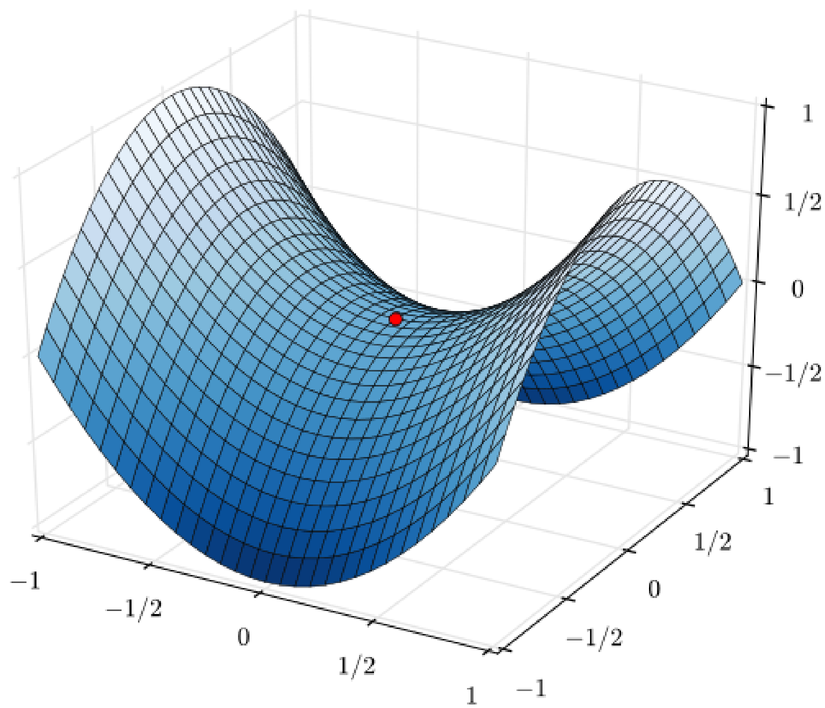
March 31, 2022

APPLICATION OF GMRES/ANDERSON IN ML & FILTERING METHODS

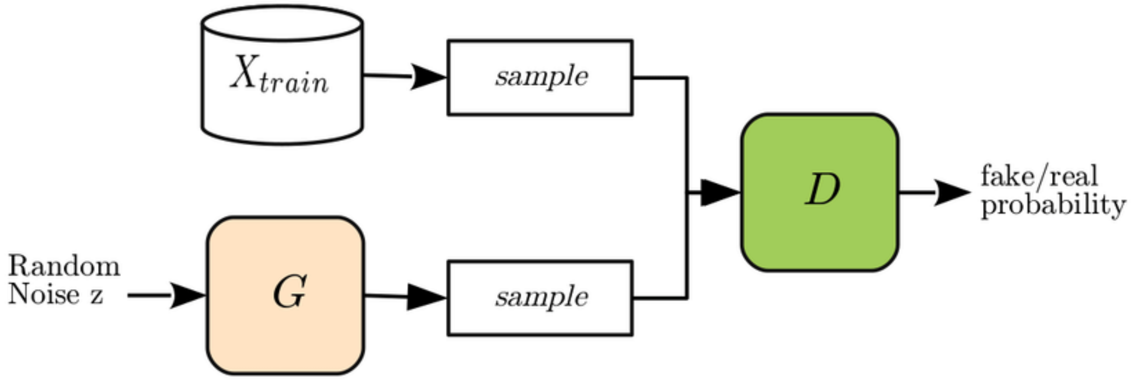
Minimax Optimization

➤ **Minimax optimization:**

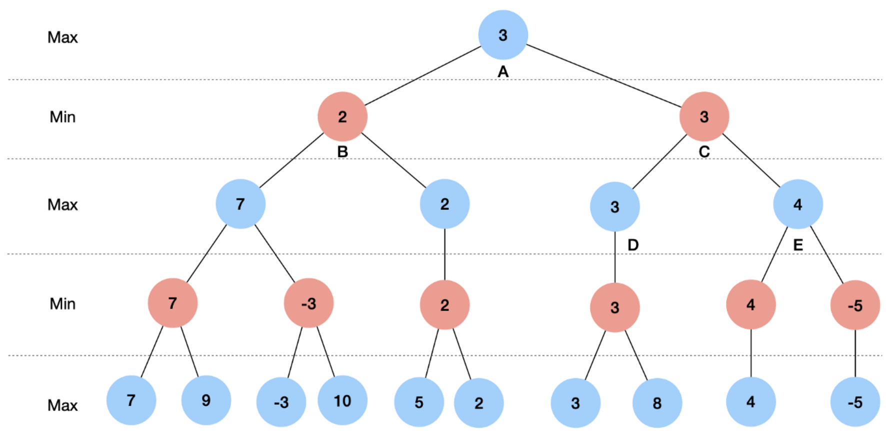
$$\arg\min_{x \in \mathcal{X}} \arg\max_{y \in \mathcal{Y}} f(x, y)$$



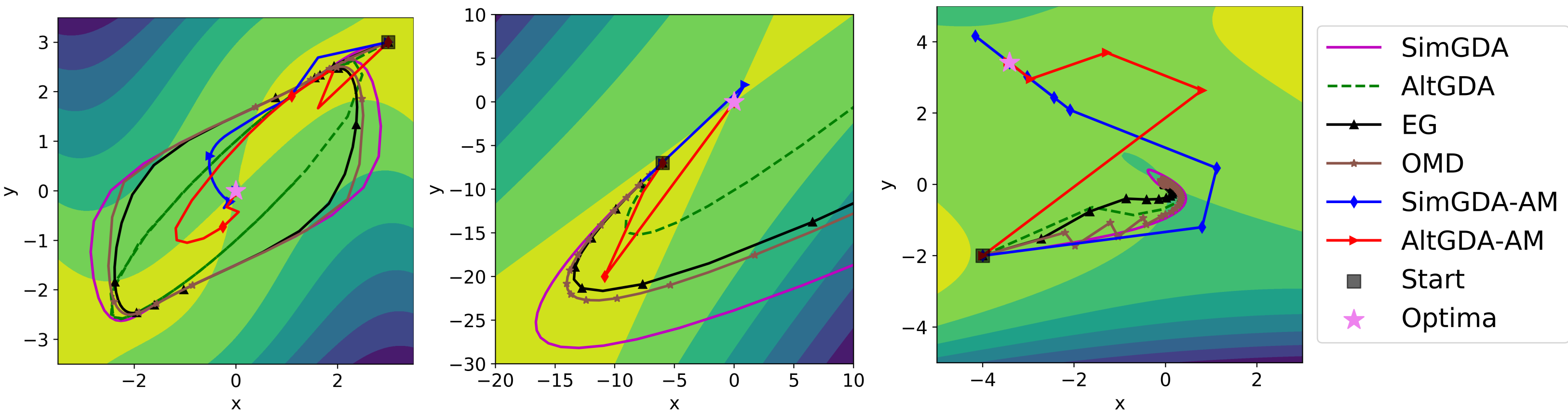
Generative Adversarial Networks (GANs)



Reinforcement Learning (RL)



Difficulty of solving minimax optimization



(a) Cycling Behavoir (b) Diverging Behavoir (c) Converging to a non-optima

Figure 1: **Left**: $f(x, y) = (4x^2 - (y - 3x + 0.05x^3)^2 - 0.1y^4)e^{-0.01(x^2+y^2)}$. **Middle**: $-3x^2 - y^2 + 4xy$. **Right**: $f(x, y) = 2x^2 + y^2 + 4xy + \frac{4}{3}y^3 - \frac{1}{4}y^4$. We can observe that baseline methods fail to converge to a local minimax, whereas the proposed Krylov subspace method always exhibits desirable behaviors.

Gradient Descent Ascent as a fixed point iteration

- **Recall**

$$\arg\min_{x\in\mathcal{X}} \arg\max_{y\in\mathcal{Y}} f(x,y)$$

- **Simultaneous GDA (SimGDA):**

$$\mathbf{x}_{t+1} = \mathbf{x}_t - \eta\nabla_{\mathbf{x}} f(\mathbf{x}_t, \mathbf{y}_t), \quad \mathbf{y}_{t+1} = \mathbf{y}_t + \eta\nabla_{\mathbf{y}} f(\mathbf{x}_t, \mathbf{y}_t)$$

- **Alternating GDA (AltGDA):**

$$\mathbf{x}_{t+1} = \mathbf{x}_t - \eta\nabla_{\mathbf{x}} f(\mathbf{x}_t, \mathbf{y}_t), \quad \mathbf{y}_{t+1} = \mathbf{y}_t + \eta\nabla_{\mathbf{y}} f(\mathbf{x}_{t+1}, \mathbf{y}_t)$$

GDA as a fixed point iteration

- Both SimGDA and AltGDA can be rewritten as a fixed point iteration $\mathbf{w}_{t+1} = G(\mathbf{w}_t)$
- SimGDA updates:

$$\mathbf{w}_{t+1} = G_{\eta}^{(\mathrm{Sim})}(\mathbf{w}_t) \triangleq \mathbf{w}_t - \eta V(\mathbf{w}_t)$$

with

$$\mathbf{w} = \begin{bmatrix} \mathbf{x} \\ \mathbf{y} \end{bmatrix}, V(\mathbf{w}) = \begin{bmatrix} \nabla_{\mathbf{x}} f(\mathbf{x}, \mathbf{y}) \\ -\nabla_{\mathbf{y}} f(\mathbf{x}, \mathbf{y}) \end{bmatrix}$$

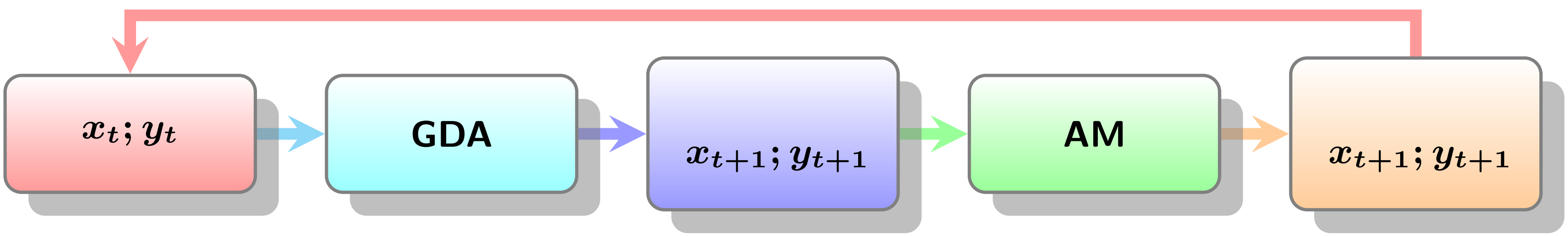
- AltGDA updates:

$$\mathbf{w}_{t+1} = G_{\eta}^{(\mathrm{Alt})}(\mathbf{w}_t)$$

Anderson mixing [Anderson, 1965]

- **Anderson mixing**

$$\mathbf{x}_{t+1} = \sum_{i=0}^{p} \beta_i \mathbf{x}_{t-p+i}, \quad \mathbf{y}_{t+1} = \sum_{i=0}^{p} \beta_i \mathbf{y}_{t-p+i}$$



- $F_t = [f_{t-p}, \ldots, f_t], f_i = G(w_i) - w_i$
- $\beta = (\beta_0, \ldots, \beta_p)^T$ **is obtained by solving**

$$\min_{\beta} \|F_t \beta\|_2, \text{ s. t. } \sum_{i=0}^{p} \beta_i = 1$$

Zero-sum bilinear games

- **Assume $\mathbf{A}$ is full rank**

$$\min_{\mathbf{x}\in R^n} \max_{\mathbf{y}\in R^n} f(\mathbf{x},\mathbf{y}) = \mathbf{x}^T\mathbf{A}\mathbf{y} + \mathbf{b}^T\mathbf{x} + \mathbf{c}^T\mathbf{y}$$

- **Nash equilibrium is given by**

$$(\mathbf{x}^*,\mathbf{y}^*) = (-\mathbf{A}^{-T}\mathbf{c}, -\mathbf{A}^{-1}\mathbf{b})$$

➤ **SimGDA** can be written as:

$$\begin{bmatrix} \mathbf{x}_{t+1} \\ \mathbf{y}_{t+1} \end{bmatrix} = \underbrace{\begin{bmatrix} \mathbf{I} & -\eta \mathbf{A} \\ \eta \mathbf{A}^T & \mathbf{I} \end{bmatrix}}_{\mathbf{G}^{(Sim)}} \underbrace{\begin{bmatrix} \mathbf{x}_t \\ \mathbf{y}_t \end{bmatrix}}_{\mathbf{w}_t^{(Sim)}} - \eta \underbrace{\begin{bmatrix} \mathbf{b} \\ \mathbf{c} \end{bmatrix}}_{\mathbf{b}^{(Sim)}}.$$

➤ **AltGDA** can be written as:

$$\begin{bmatrix} \mathbf{x}_{t+1} \\ \mathbf{y}_{t+1} \end{bmatrix} = \underbrace{\begin{bmatrix} \mathbf{I} & -\eta \mathbf{A} \\ \eta \mathbf{A}^T & \mathbf{I} - \eta^2 \mathbf{A}^T \mathbf{A} \end{bmatrix}}_{\mathbf{G}^{(Alt)}} \underbrace{\begin{bmatrix} \mathbf{x}_t \\ \mathbf{y}_t \end{bmatrix}}_{\mathbf{w}_t^{(Alt)}} - \eta \underbrace{\begin{bmatrix} \mathbf{b} \\ \mathbf{c} \end{bmatrix}}_{\mathbf{b}^{(Alt)}}.$$

➤ Convergence of GDA-AM can be studied via the convergence of GMRES [Walker, Na, SINUM, 2011]

$$(\mathbf{I} - \mathbf{G}^{(\cdot)})\mathbf{w} = \mathbf{b}^{(\cdot)}, \quad \text{with } \mathbf{w}_0 = \mathbf{w}_0^{(\cdot)}$$

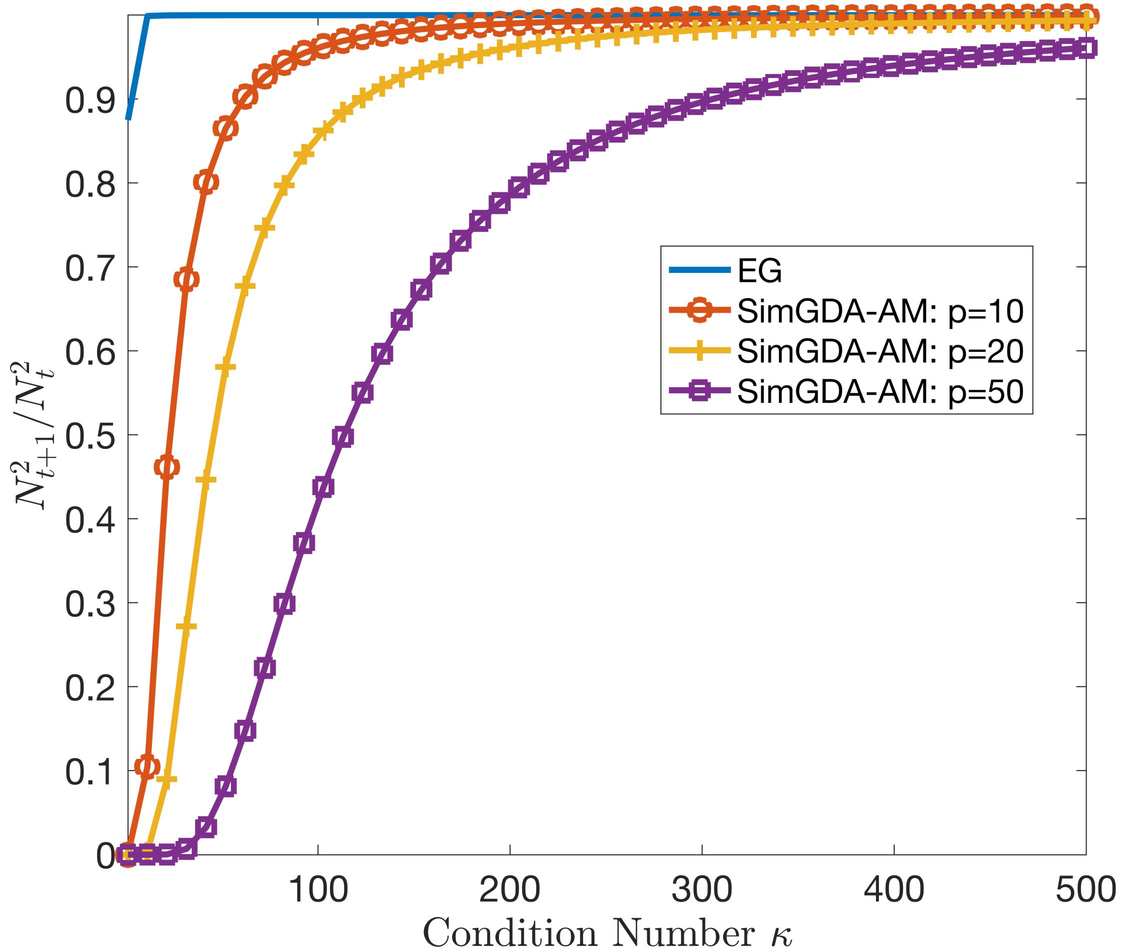
Global convergence for SimGDA-AM on bilinear problems

- p: as the restart dimension
- $N_{(k+1)p} = \|\mathbf{w}^* - \mathbf{w}_{(k+1)p}\|$
- T_p: Chebyshev polynomial of first kind of degree p

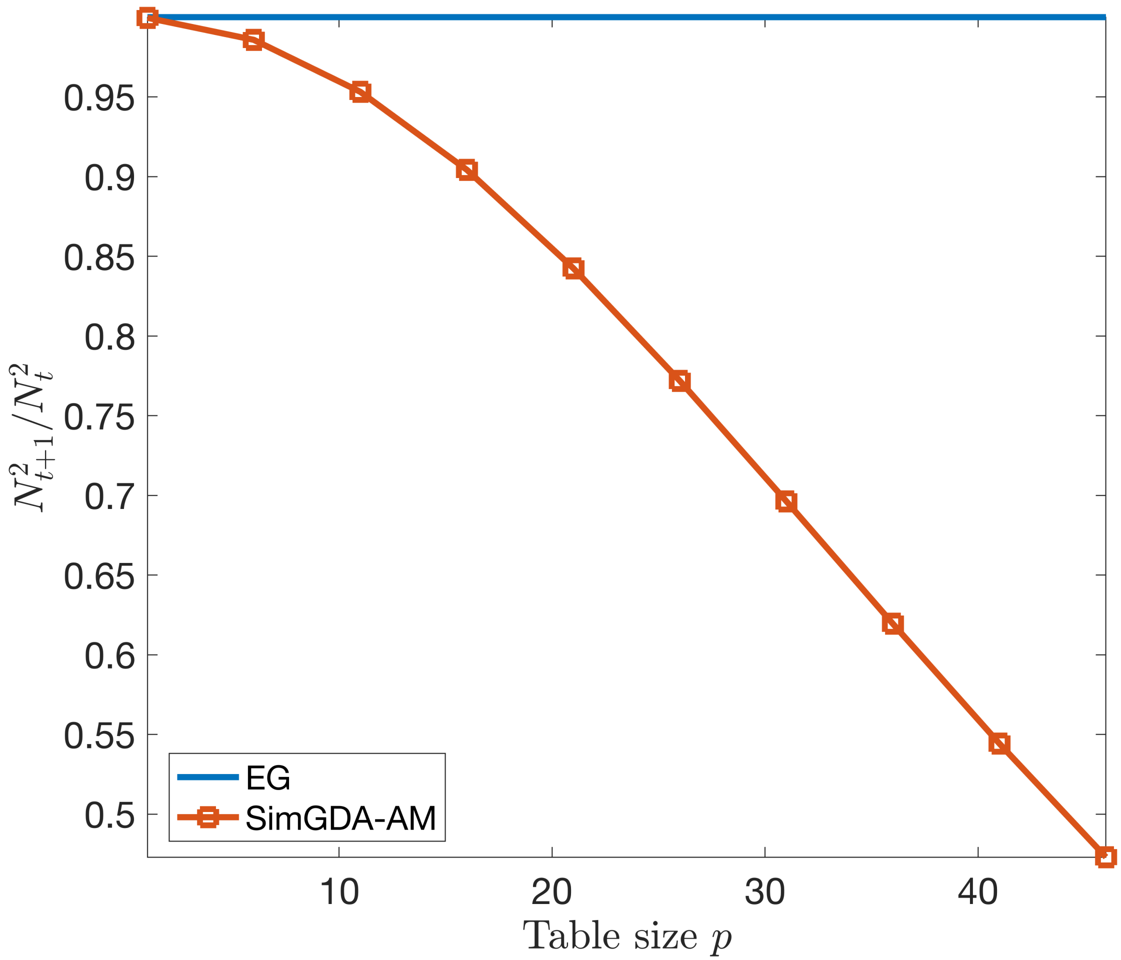
$$N^2_{(k+1)p} \leq \rho(A) N^2_{kp} \tag{1}$$

where $\rho(A) = (\frac{1}{T_p(1+\frac{2}{\kappa(\mathrm{A}^T\mathrm{A})-1})})^2$.

➤ Comparison between SimGDA-AM and EG for different condition numbers and fixed table size $p = 10, 20, 50$.



➤ Comparison between SimGDA-AM and EG for increasing table size on a matrix $\mathbf{A}$ with condition number 100.



Convergence for AltGDA-AM on bilinear problem

- p: as the restart dimension
- $N_{(k+1)p} = \|\mathrm{w}^* - \mathrm{w}_{(k+1)p}\|$

Assume A is normalized such that its largest singular value is equal to 1. Then when the learning rate η is less than 2

$$N^2_{(k+1)p} \leq \sqrt{1 + \frac{2\eta}{2-\eta}(\frac{r}{c})^p N^2_{kp}} \qquad (2)$$

where c and r are the center and radius of a disk $D(c, r)$ which includes all the eigenvalues of G. Especially, $\frac{r}{c} < 1$.

$$\min_{\mathbf{x}} \max_{\mathbf{y}} f(\mathbf{x}, \mathbf{y}) = \mathbf{x}^T A \mathbf{y} + \mathbf{b}^T \mathbf{x} + \mathbf{c}^T \mathbf{y}$$

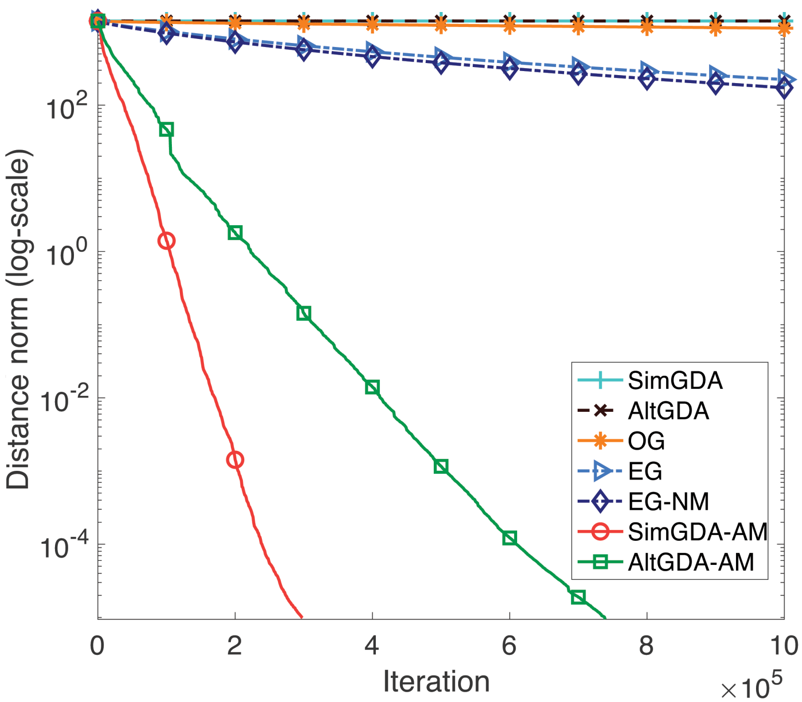


Figure 2: n=500

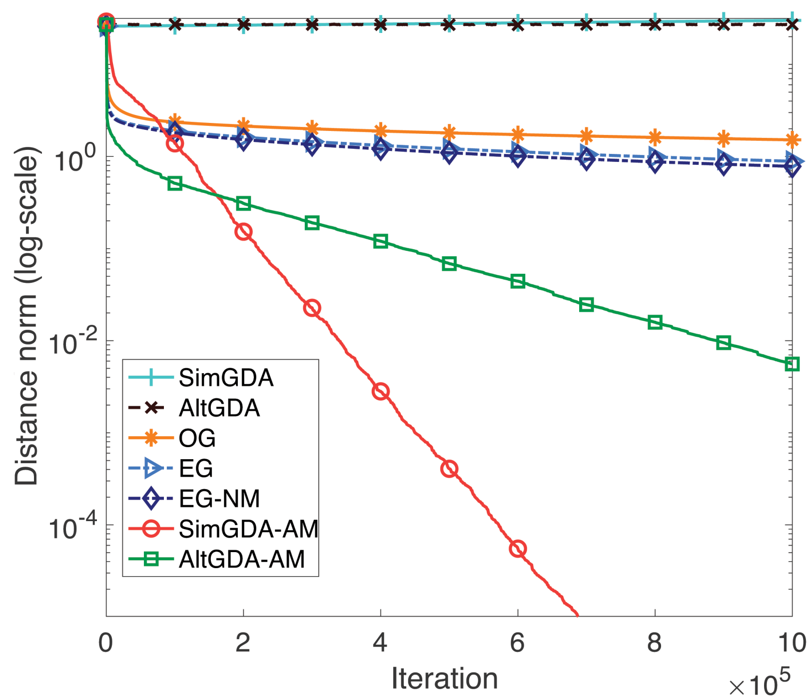


Figure 3: n=1000

$$\min_{\mathbf{x}} \max_{\mathbf{y}} f(\mathbf{x}, \mathbf{y}) = \mathbf{x}^T \mathbf{A} \mathbf{y} + \mathbf{b}^T \mathbf{x} + \mathbf{c}^T \mathbf{y}$$

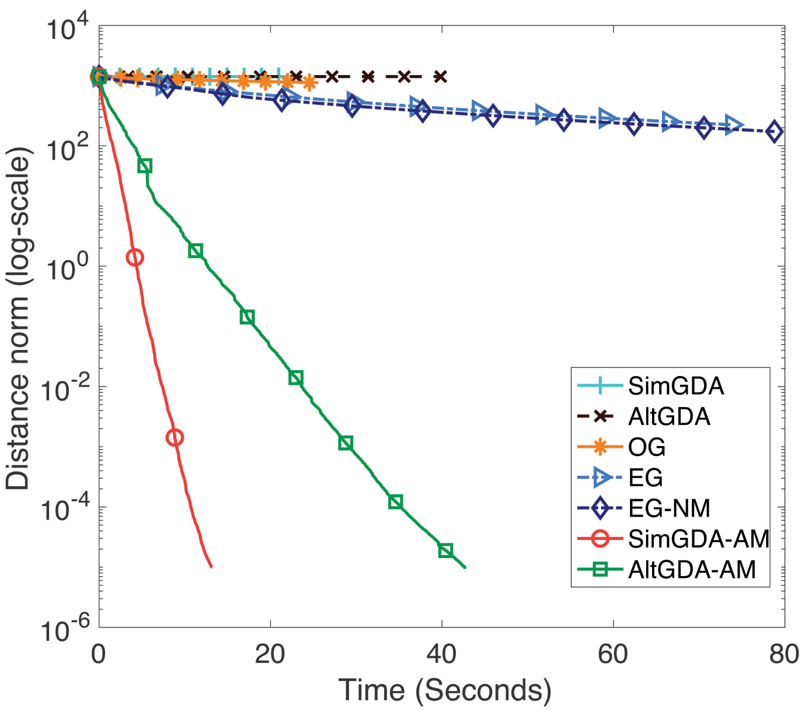


Figure 4: n=500

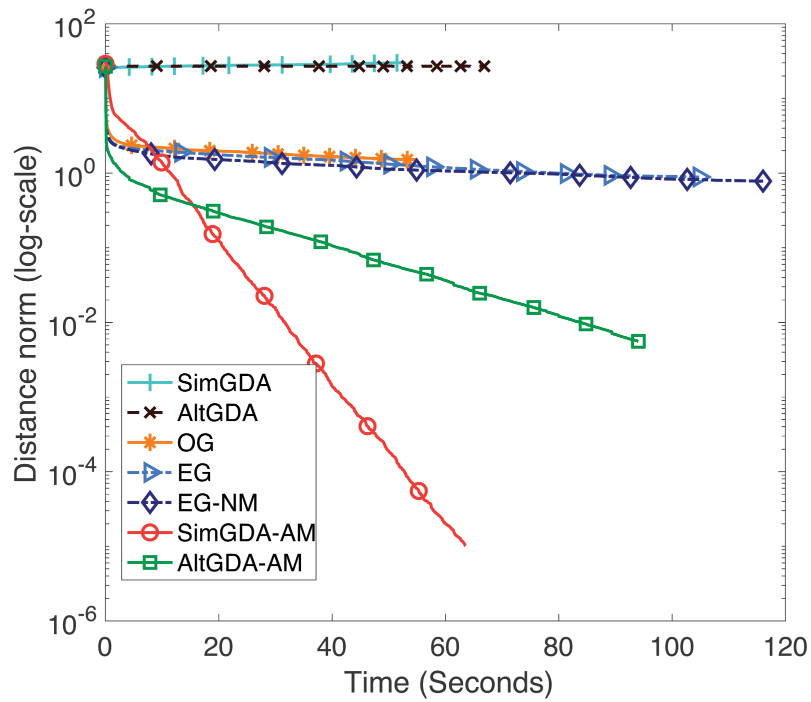
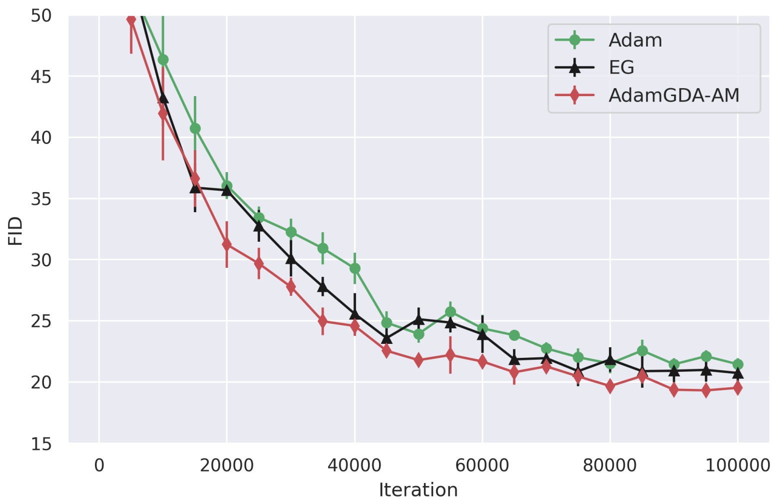


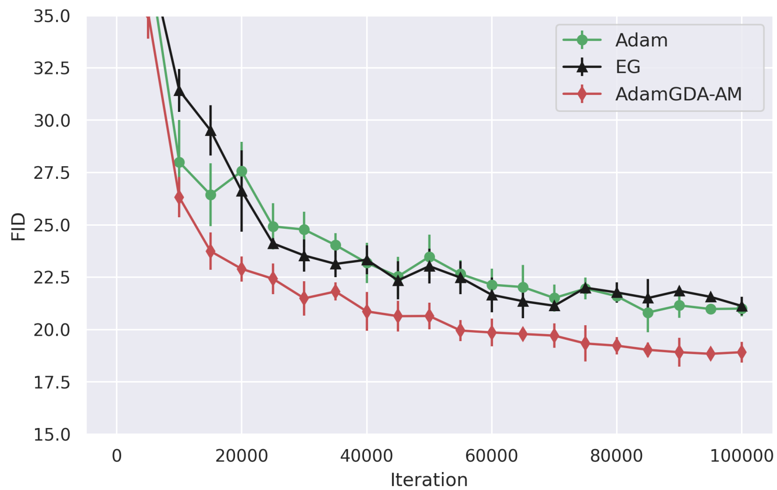
Figure 5: n=1000

Empirical performance on GANs

	WGAN-GP(ResNet)			SNGAN(ResNet)	
	CIFAR10		CelebA	CIFAR10	
Method	IS ↑	FID ↓	FID	IS	FID
Adam	7.76 ±.11	22.45 ±.65	8.43 ±.05	8.21 ±.05	20.81 ±.16
EG	7.83 ±.08	20.73 ±.22	8.15 ±.06	8.15 ±.07	21.12 ±.19
GDA-AM	8.05 ±.06	19.32 ±.16	7.82 ±.06	8.38 ±.04	18.84 ±.13



(a) WGAN-GP (ResNet)



(b) SNGAN (ResNet)

Generated Images for CIFAR10 and CelebA

(a) Generated images for CIFAR10

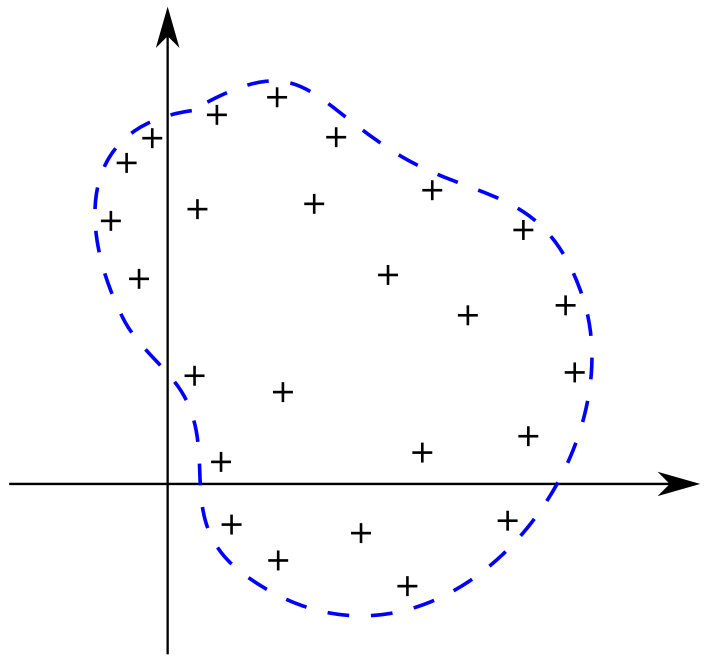
(b) Generated images for CelebA

Polynomial preconditioning

- Starting from a initial guess x_0 and initial residual $r_0 = b - Ax_0$
- The approximated solution $\tilde{x}$ at a specific iteration is $\tilde{x} = x_0 + p(A)r_0$ where p is a polynomial
- The residual is

$$\tilde{r} = b - A = (I - Ap(A))r_0 = r(A)r_0$$

where $r(z)$ is the residual polynomial $r(z) = 1 - zp(z)$



- **Goal: $\tilde{r} = r(A)r_0$ close to zero**
- **Assume A is close to be normal**
- **Small $|r(\lambda)|$ on eigenvalues $\lambda \rightarrow$ small $\|r(A)\|$**
- **By maximum modulus principle, $|r(z)|$ be small on Γ**

Optimal polynomial filter

➤ A "good" polynomial p can be solved from the minimax problem

$$\min_{r \in \mathcal{P}_m^0} \max_{z \in \Gamma} |r(z)| \quad \text{or} \quad \min_{p \in \mathcal{P}_{m-1}} \max_{z \in \Gamma} |1 - zp(z)|$$

- $\mathcal{P}_{m-1}$ is the polynomial space of degree $\leq m-1$
- $\mathcal{P}_m^0 = \{p \in \mathcal{P}_m | p(0) = 1\}$

Chebyshev approximation problem in function approximation theory.

Discretized version

➤ Let $\Gamma_n = \{z_1, z_2, \ldots, z_n\}$ be a discretization of Γ, instead we consider a discrete minimax problem

$$\min_{r \in \mathcal{P}_m^0} \max_{z \in \Gamma_n} |r(z)| \qquad \text{or} \qquad \min_{p \in \mathcal{P}_{m-1}} \max_{z \in \Gamma_n} |1 - zp(z)|$$

➤ With a basis for $\mathcal{P}_{m-1}$, the problem can be rewritten in matrix form

$$\min_{\alpha} \|e - F\alpha\|_\infty$$

➤ This is not a numerical stable approach for nonsymmetric problems and high degree polynomials!

Implicit representation of p

- Represent polynomial p by $[p(z_1), p(z_2), \ldots, p(z_n)]^T$
- Define an inner product of two polynomials p_1 and p_2 by

$$\langle p_1, p_2 \rangle = \sum_{i=1}^{n} p_1(z_i)\overline{p_2(z_i)}$$

and denote by $\| \cdot \|_w$

- Define the objective function by sum of squares

$$\min_{r \in \mathcal{P}_m^0} \|r\|_w^2 \quad \text{or} \quad \min_{r \in \mathcal{P}_m^0} \sum_{z \in \Gamma_n} |r(z)|^2$$

Arnoldi process in polynomial space

1. set $q_1 = 1/\|1\|_w$
2. for $j = 1, 2, \ldots, m$
3. compute $q := zq_j$
4. for $i = 1, 2, \ldots, j$ do
5. compute $h_{ij} = \langle q, q_i \rangle$
6. compute $q = q - h_{ij} q_i$
7. end for
8. compute $h_{j+1,j} = \|q\|_w$
9. compute $q_{j+1} = q/h_{j+1,j}$
10. end for

This process generates an orthonormal basis $\{q_1, q_2, \ldots, q_m\}$ for the polynomial space P_{m-1} under the norm $\|\cdot\|_w$.

- compute $q_1 = [1, 1, \ldots, 1]^T / \sqrt{n}$
- the polynomial $z = [z_1, z_2, \ldots, z_n]^T$
- polynomial multiplication $\Rightarrow$ entry-wise multiplication of vectors
- inner product/norm of polynomial $\Rightarrow$ standard dot product in vector space
- addition/subtraction/scalar multiplication $\Rightarrow$ corresponding operations in vector space
- orthonormal polynomial basis $\{q_1, q_2, \ldots, q_m\}$ $\Rightarrow$ orthogonal matrix $Q_m = [q_1, q_2, \ldots, q_m]$

Application of preconditioner

- Apply $p(A)$ to a vector v note that

$$M^{-1}v = p(A)v = \sum_{i=1}^{m} \alpha_i q_i(A)v := \sum_{i=1}^{m} \alpha_i v_i$$

- The v_i's can be computed recursively

$$v_{i+1} = \frac{1}{h_{i+1,i}} \left(Av_i - \sum_{j=1}^{i} h_{ji} v_j \right), \quad 1 \le i \le m-1$$

GMRES in polynomial space

- Assume $p = \sum_{j=1}^{m} \alpha_j q_j = Q_m \alpha$

$$zp = \sum_{i=1}^{m} \alpha_i (zq_i) = \sum_{i=1}^{m} \alpha_i \sum_{j=1}^{i+1} h_{ji} q_j = Q_{m+1} H_m \alpha$$

- The minimization problem can be solved as

$$\min_{p \in \mathcal{P}_{m-1}} \|1 - zp\|_w^2 \quad \Longrightarrow \quad \min_{\alpha} \|\beta e_1 - H_m \alpha\|_2^2$$

Short-term recurrence

- Replace full orthogonalization by partial one:

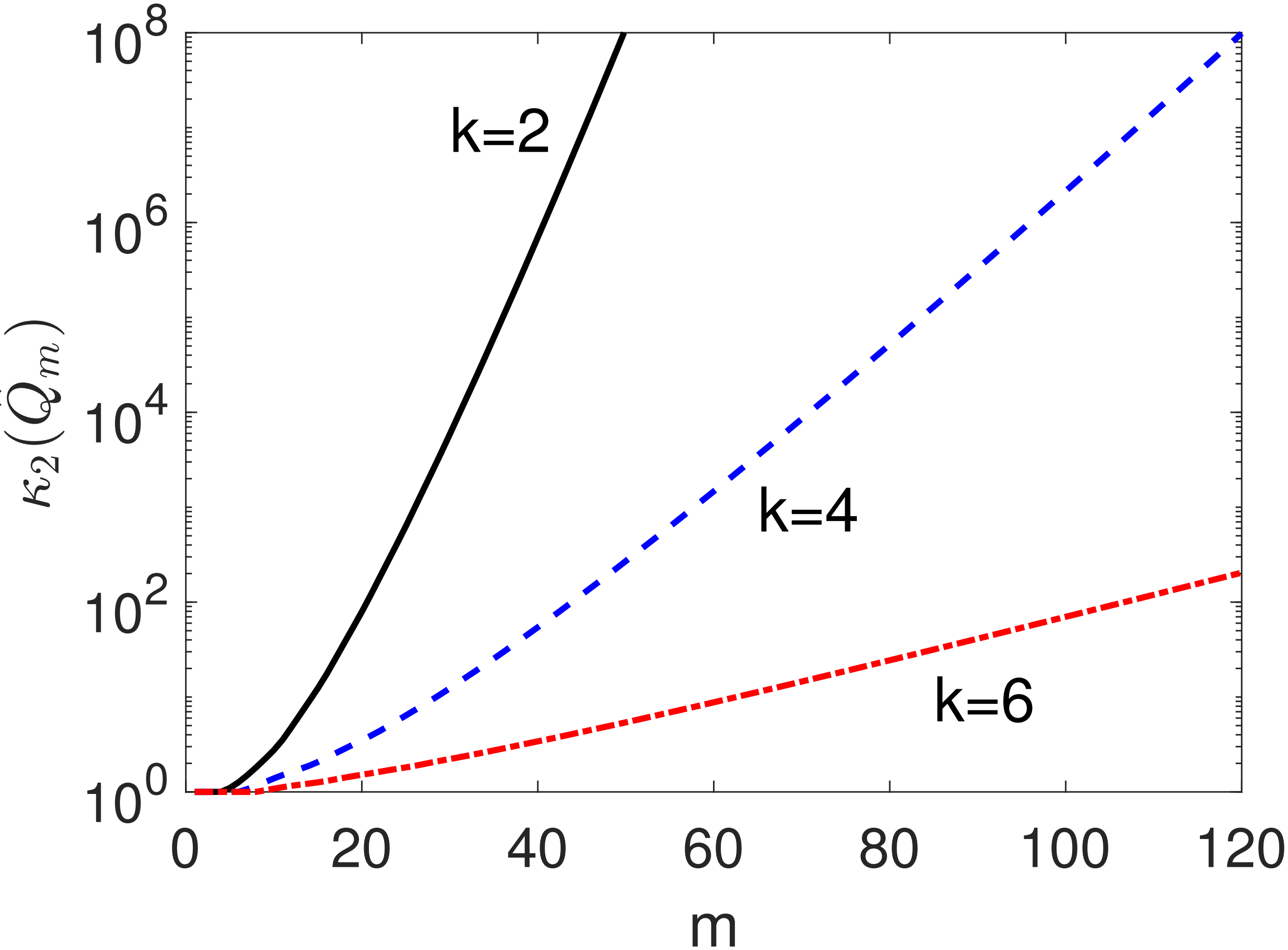
$$t_{j+1,j}\hat{q}_{j+1} = z\hat{q}_j - \sum_{i=j-k+1}^{j} t_{ij}\hat{q}_i, \quad 1 \le j \le m,$$

- Fast application of $M^{-1} = \hat{p}(A)$

$$v_{i+1} = \frac{1}{t_{i+1,i}}\left(Av_i - \sum_{j=i-k+1}^{i} t_{ji}v_j\right).$$

$\mathcal{O}(mkN)$ operations and $\mathcal{O}(kN)$ storage

➤ Conditioning of $\hat{Q}_m$ generated with k-term recurrence from Helmohltz problem



➤ The approximate boundaries of the spectrum and the approximate eigenvalues obtained from 60 steps of the Arnoldi algorithm for the $2,000 \times 2,000$ diagonal matrix.

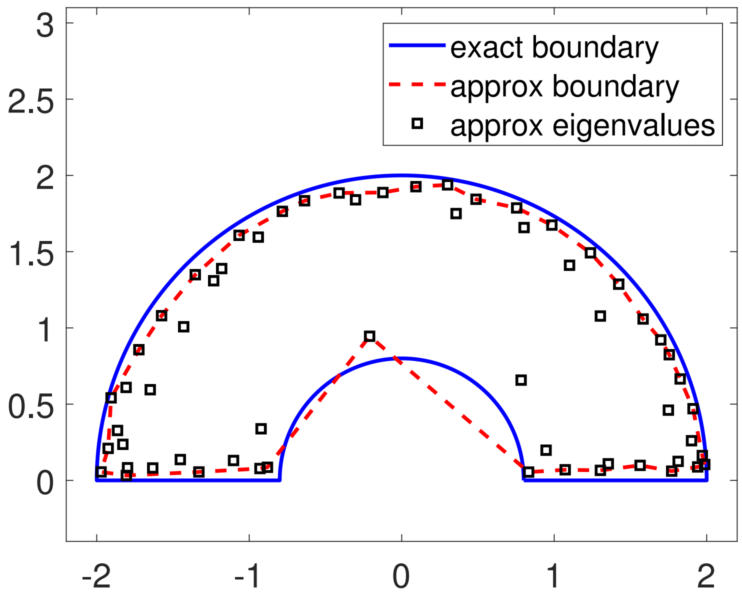


Figure 9: Approximate boundary

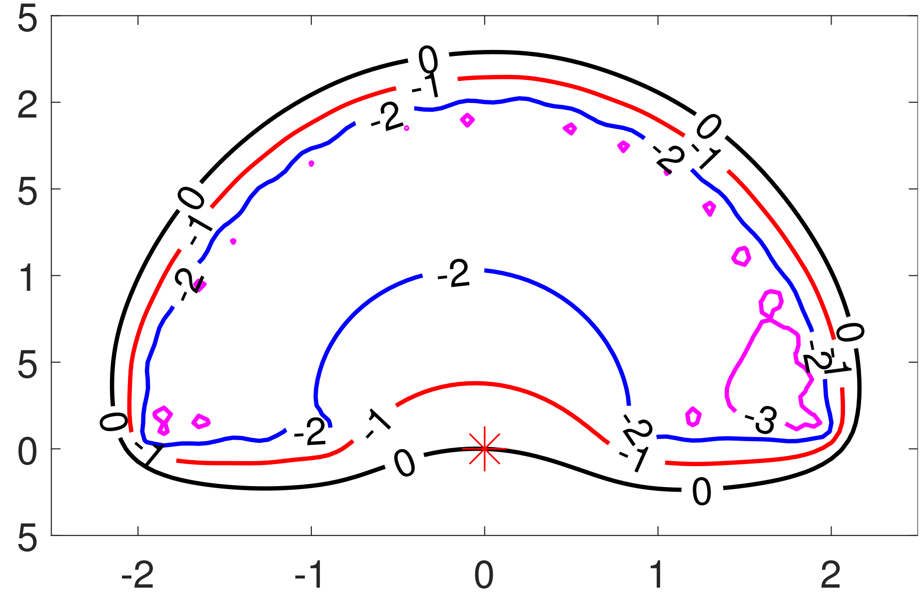


Figure 10: Contour maps of $|\mathbf{1 - zp(z)}|$ in log scale with different choice of $\mathbf{\Gamma}$ for the $\mathbf{2,000 \times 2,000}$ diagonal matrix.

➤ **Convergence results of GMRES(50) for the $2,000 \times 2,000$ diagonal matrix test with tolerance $\tau = 10^{-12}$**

		p-t	i-t	its	mv
no precond.		\	0.6525	237	237
with precond.	exact boundary	0.0045	0.5366	8	240
	approx. boundary	0.0040	0.5238	8	240

- **3D Helmholtz equation**

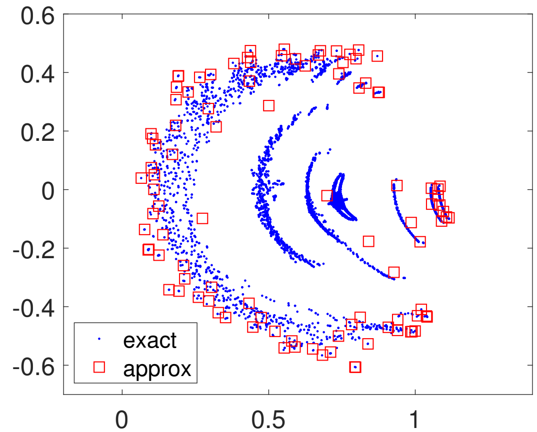
$$-\Delta u - \frac{\omega^2}{c^2(x)} u = s$$

where ω is the angular frequency and $c(x)$ is the wavespeed.

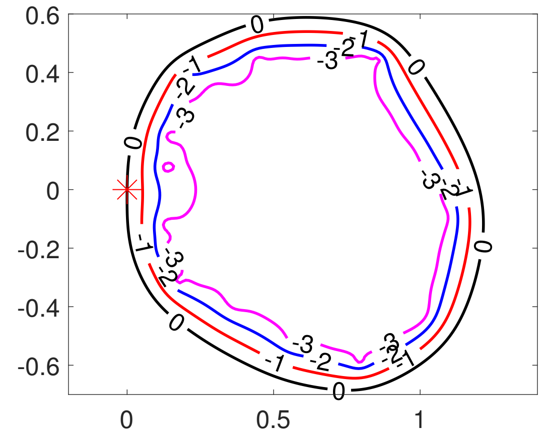
- **PML boundary conditions + 7-point stencil FD**
- **Matrix size $10^6 \times 10^6$**

Preconditioner type	p-t	i-t	its	mv
no preconditioner	\	\	F	\
ILUT	F	\	\	\
ILUT with diagonal shift $\sigma = -0.4i$	220.96	\	F	\
single polynomial of degree $600 - 1$	3.49	1484.87	16	$9,600$
compound polynomial of degree $60 \times 10 - 1$	0.05	906.41	18	$10,800$

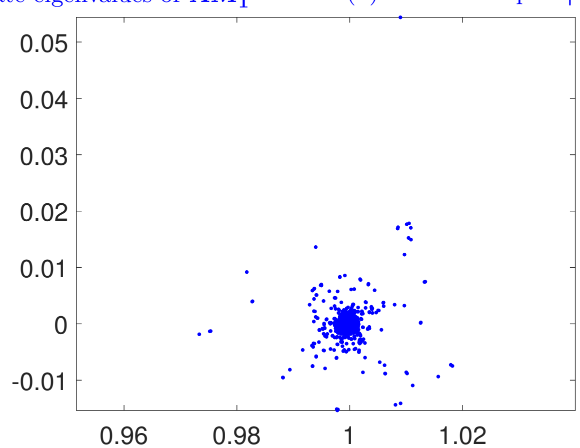
Preconditioner type	p-t	i-t	its	mv
SLR preconditioner	86.94	\	F	\
SLR with polynomial of degree $30-1$	145.85	308.03	29	870



(a) Exact and approximate eigenvalues of AM_1^{-1}



(b) Contour map of $|1-zp(z)|$ in log scale



(c) Eigenvalues of the final preconditioned matrix $A_1p(A_1)$

References

➤ **Proxy-GMRES: Preconditioning via GMRES in Polynomial Space, X. Ye, Y. Xi, Y. Saad, SIMAX, 2021 (https://github.com/xinye83/proxy-gmres)**

➤ **GDA-AM: On the Effectiveness of Solving Minimax Optimization via Anderson Mixing, H. He, S. Zhao, Y. Xi, J. Ho, Y. Saad, ICLR 2022 (https://github.com/hehuannb/GDA-AM)**