KTH ROYAL INSTITUTE
OF TECHNOLOGY

Doctoral Thesis in Applied and Computational Mathematics

Numerical methods for parameterized linear systems

SIOBHÁN CORRENTY

Stockholm, Sweden 2024

Numerical methods for parameterized linear systems

SIOBHÁN CORRENTY

Academic Dissertation which, with due permission of the KTH Royal Institute of Technology, is submitted for public defence for the Degree of Doctor of Philosophy on Tuesday, May 7th, 2024 at 14:00 in F3, Lindstedtsvägen 26, Stockholm.

Doctoral Thesis in Applied and Computational Mathematics
KTH Royal Institute of Technology
Stockholm, Sweden 2024

TRITA-SCI-FOU 2024:08
ISBN 978-91-8040-845-5

Printed by: Universitetsservice US-AB, Sweden 2024

To my family and friends

Abstract

Solving linear systems of equations is a fundamental problem in engineering. Moreover, applications involving the solution to linear systems arise in the social sciences, business, and economics. Specifically, the research conducted in this dissertation explores solutions to linear systems where the system matrix depends nonlinearly on a parameter. The parameter can be a scalar or a vector, and a change in the parameter results in a change in the solution. Such a setting arises in the study of partial differential equations and time-delay systems, and we are interested in obtaining solutions corresponding to many values of the parameter simultaneously. The methods developed in this thesis can also be used to solve parameter estimation problems. Furthermore, software has been developed and is available online.

This thesis consists of four papers and presents both algorithms and theoretical analysis. In Paper A, a linearization based on an infinite Taylor series expansion is considered. Specifically, the linearized system is a shifted parameterized system, and the parameter is a scalar. The GMRES method is used to solve the systems corresponding to many values of the parameter, and only one Krylov subspace basis matrix is required. Convergence analysis is based on solutions to a nonlinear eigenvalue problem and the magnitude of the parameter. Notably, the algorithm is carried out in a finite number of computations.

The approach in Paper B is based on a preconditioned linearized system solved using the inexact GMRES method. In this setting, the linearization incorporates all terms in an infinite Taylor series expansion, and the preconditioner is applied approximately using iterative methods. Solutions corresponding to many values of the scalar parameter are generated from one subspace, and this is done in a finite number of linear algebra operations. Theoretical analysis, based on the error in the application of the preconditioner and the magnitude of the parameter, leads to a bound on the residual.

Paper C proposes a short recurrence Krylov subspace method for solving linear systems that depend on a scalar parameter. In particular, a Chebyshev approximation is used to construct a linearization, and the linearized system is solved in a Bi-CG setting. Additionally, shift-and-invert preconditioning leads to fast convergence of the Krylov method for many different values of the parameter. An inexact variant of the method is also derived and analyzed.

In Paper D, a reduced order model is constructed from snapshots to solve parameterized linear systems. Specifically, the parameter is a vector of dimension 2, and the sampling is performed on a sparse grid using the method proposed in Paper C. A tensor decomposition is utilized to build the model. Approaches of this kind are not always successful, and it is not known a priori if a decomposition will converge on a given set of snapshots. This work offers a novel way to generate a new set of snapshots in the same parameter space, to be used if the decomposition does not converge, with little extra computation.

Keywords: Parameterized linear systems, Krylov subspace methods, preconditioning, tensor decompositions, shifted linear systems, parameterized partial differential equations, time-delay systems, transfer functions, parameter estimation problems

Sammanfattning

Att lösa linjära ekvationssystem är ett grundläggande tekniskt problem. Dessutom uppstår tillämpningar som involverar lösningen av linjära system inom samhällsvetenskap och ekonomi. Specifikt utforskar denna avhandling lösningar till linjära system där systemmatrisen beror olinjärt på en parameter. Parametern kan vara en skalär eller en vektor, och en förändring i parametern resulterar i en förändring i lösningen. En sådant scenario uppstår vid studiet av partiella differentialekvationer och tidsfördröjningssystem, och vi är intresserade av att erhålla lösningar som motsvarar många värden på parametern samtidigt. De metoder som utvecklats i denna avhandling kan också användas för att lösa problem med parameteruppskattning. Ytterligare har programvara utvecklats och är tillgänglig online.

Denna avhandling består av fyra artiklar och presenterar både algoritmer och teoretisk analys. I artikel A behandlas en linjärisering baserad på en oändlig Taylor-serieexpansion. Specifikt är det linjäriserade systemet ett skiftat parametriserat system, och parametern är en skalär. Systemet löses med GMRES-metoden, och endast en Krylov-basmatris krävs. Konvergensanalys baseras på lösningar till ett olinjärt egenvärdesproblem och parameterns storlek. Noterbart är att algoritmen utförs i ett ändligt antal beräkningar.

Tillvägagångssättet i artikel B är baserat på ett förkonditionerat linjäriserat system löst med den inexakta GMRES-metoden. I den här kontexten innehåller linjäriseringen alla termer i en oändlig Taylor-serieexpansion, och förkonditioneringen appliceras på ett approximativt sätt med iterativa metoder. Lösningar som motsvarar många värden på den skalära parametern genereras från ett delrum, och detta görs i ett ändligt antal linjära algebraoperationer. Teoretisk analys, baserad på felet i appliceringen av förkonditioneraren och storleken på parametern, leder till en övre begränsning på residualens storlek.

Artikel C föreslår en Krylovbaserad rekursionsmetod med få termer för att lösa linjära system som är beroende av en skalär parameter. Specifikt används en Chebyshev-approximation för att konstruera en linjärisering, och det linjäriserade systemet löses med den bikonjugerade gradient-metoden. Dessutom leder förkonditionering med skifte och invertering till snabb konvergens av Krylov-metoden för många olika värden på parametern. En inexakt variant av metoden härleds och analyseras också.

I artikel D konstrueras en reducerad ordningsmodell från sampel av modellen för att lösa parametriserade linjära system. Specifikt är parametern en vektor med dimensionen 2, och samplingen utförs på ett glest rutnät med den metod som föreslås i artikel C. En tensorfaktorisering används för att bygga modellen. Tillvägagångssätt av denna typ är inte alltid framgångsrika, och det är inte känt på förhand om en tensorfaktorisering kommer att konvergera för en given uppsättning av sampel. Detta arbete presenterar ett nytt sätt att generera en ny uppsättning sampel i samma parameterrum till en låg extra kostnad. De nya lösningar kan användas om tensorfaktoriseringen misslyckas.

Nyckelord: Parameteriserade linjära system, Krylov-metoder, förkonditionering, tensordekomposition, skiftade linjära system, parametriserade partiella differentialekvationer, tidsfördröjningssystem, överföringsfunktioner, parameteruppskattningsproblem

Acknowledgements

THIS thesis represents the culmination of extensive effort, and I am grateful for the help I received to make it a reality. It is a pleasure to express my sincere appreciation to the individuals who played a role in this journey.

First, I would like to acknowledge my main supervisor, Elias Jarlebring. I value the support I received from you over the last five years. Thank you for giving me this opportunity and fostering an environment where I could become an independent researcher. I also want to express my sincere thanks to my assistant supervisor and coauthor, Kirk Soodhalter. I am grateful for all the encouragement you have given to me. Your positive attitude and kindness played an instrumental role in bringing this thesis to fruition. I would also like to mention my assistant supervisor, Johan Karlsson. Thank you for sharing your knowledge with me.

I had several opportunities to travel abroad during my PhD studies. I would like to thank Daniel Szyld for hosting my stay at Temple University and collaborating with me. Your insightful feedback not only strengthened the paper we wrote but also played a pivotal role in my growth as a researcher. Thank you to Melina Freitag for facilitating my visit to University of Potsdam. Your enthusiasm and commitment to developing high-quality work helped me immensely. I am grateful I had the opportunity to collaborate with you. An additional acknowledgement goes to Kirk Soodhalter for hosting me at Trinity College Dublin. The time spent there has left a lasting impression on me, and I am thankful for the valuable experiences and knowledge gained during my stay.

I would like to extend a thank you to Anders Szepessy for his support throughout the course of my studies. I really appreciated hearing your point of view on both scientific and non-scientific matters. Additionally, I want to express my sincere gratitude to Anna-Karin Tornberg for encouraging my development and her kindness. Your welcoming approach made me feel valued and supported. A special thank you goes to Anna Broms for her generous efforts in reading the kappa twice and offering valuable suggestions. You have always been a supportive friend. I would also like to extend my thanks to the other PhD students at the department

for creating a stimulating environment.

Attending international conferences contributed greatly to the success of this thesis. I am grateful for the members of the linear algebra community for their words of encouragement. Your belief in my abilities provided crucial support, sustaining me through the challenges I encountered on this academic journey.

I wish to express my deepest gratitude to my family for always supporting me. Mom, your kind heart and sense of *humour* keep me going, even on the hardest days. You and Dad filled my life with love and opportunities, and I know he would have gotten a kick out of this. Thank you also to my brother Dan and sister-in-law Kelly for your groundedness. I am inspired by your commitment to living a good life. I am blessed to have a large extended family. In particular, I wish to thank Melissa for picking me up at Philadelphia airport so many times. Our friendship is very dear to me. I'd also like to express my appreciation to all of my aunts, uncles, and cousins for always welcoming me in Ireland with open hearts. A special thank you to Sinéad and Malachi for being so kind, and to Peter and Luisa for making me feel at home. In addition, I wish to thank the Hinchcliff family for all they have done and continue to do for me.

I'd like to acknowledge my network of friends in Stockholm and elsewhere for filling my life with joy. I feel blessed to have you in my life. A special thank you goes to Eva and Stefan for being so generous and encouraging.

And to Matt: Thank you for standing by me throughout this project and all the others. I couldn't have gotten through this without you. Here's to many more morning coffees, long walks, jam sessions, and adventures. Every day with you is a gift.

Stockholm, May 2024
Siobhán Correnty

Contents

Notations and Abbreviations

Notation

$\mathbb{R}$	Set of real numbers
$\mathbb{R}^n$	Space of real column vectors of dimension n
$\mathbb{R}^{n\times m}$	Space of real matrices of dimension $n \times m$
$\mathbb{R}^{n\times m_1\times m_2}$	Space of real tensors of dimension $n \times m_1 \times m_2$
$\mathbb{R}_+$	Set of positive real numbers
$\mathbb{C}$	Set of complex numbers
$\mathbb{Z}$	Set of integers
$\mathbb{N}$	Set of natural numbers $1, 2, \ldots$
$\lvert\alpha\rvert$	Absolute value of a scalar $\alpha \in \mathbb{C}$
$a_{i,j}$	Entry (i, j) of a matrix A
a_j	jth column of a matrix A
$A(1:i, 1:j)$	Rows 1 to i, columns 1 to j of a matrix A
A^T	Transpose of a matrix A
A^*	Hermitian transpose of a matrix A
$\mathcal{K}_k(A, b)$	Krylov subspace of dimension k generated from the matrix $A \in \mathbb{C}^{n\times n}$ and the vector $b \in \mathbb{C}^n$
$\dim(\mathcal{K}_k(A, b))$	Dimension of $\mathcal{K}_k(A, b)$
$\mathrm{tr}(A)$	Trace of matrix A
$\mathrm{vec}(A)$	Vectorization of matrix A
$\sigma_i(A)$	ith largest singular value of A
$\sigma_{\min}(A), \sigma_{\max}(A)$	Smallest and largest singular value of A
$\lambda(A)$	Set of eigenvalues of A
$\lambda_i(A)$	ith largest eigenvalue of A
$\mathrm{diag}(D_1, \ldots, D_n)$	Block matrix with D_i on the diagonal
$\partial\Omega$	Boundary of a set Ω
I_n	Identity matrix of dimension $n \times n$

$0_{n\times m}$	Zero matrix of dimension $n \times m$
$\otimes$	Kronecker product
$\|x\|$, $\|x\|_p$	2-norm and p-norm of a vector x
$\|A\|$, $\|A\|_p$	2-norm and p-norm of a matrix A
$e_i \in \mathbb{R}^n$	ith column of I_n
$\mathcal{N}(\mu, \sigma^2)$	Gaussian distribution with mean μ and variance σ^2
$\mathbb{E}(X)$	Expectation of a random variable X
$\kappa(A)$, $\kappa_p(A)$	2-norm and p-norm condition number of A
$\text{span}\{v_1, \ldots, v_n\}$	Span of the set of vectors $v_1, \ldots, v_n$
$(u, v)_\Omega$	Integral of the dot product of $u, v \in \mathbb{C}^n$ over Ω
$\inf(S)$, $\sup(S)$	Infimum and supremum of a set S
$(\boldsymbol{K}_N - \mu \boldsymbol{M}_N)\boldsymbol{v}(\mu) = \boldsymbol{c}$	Linearization of $A(\mu)x(\mu) = b$

Abbreviations

ADI	Alternating directions implicit
AGMG	Aggregation-based algebraic multigrid
Bi-CG	Bi-conjugate gradient
Bi-CGSTAB	Bi-conjugate gradient stabilized
CP	CANDECOMP/PARAFAC
CPU	Central processing unit
FEM	Finite element method
GMRES	Generalized minimal residual
HOPGD	Higher-order proper generalized decomposition
HOSVD	Higher-order singular value decomposition
IDR	Induced dimension reduction
IRKA	Iterative rational Krylov algorithm
LTI	Linear time-invariant
MIMO	Multiple input, multiple output
MINRES	Minimal residual
NEP	Nonlinear eigenvalue problem
PDE	Partial differential equation
PEP	Polynomial eigenvalue problem
PGD	Proper generalized decomposition
SISO	Single input, single output

The hardest to learn was the least complicated

Indigo Girls

Part I: Introduction and Preliminaries

Chapter 1

Introduction

This thesis follows a compilation format and consists of two parts. In Part I, we introduce and summarize the topics covered in the thesis. Chapter 1 provides motivation for the research conducted and offers a summary of related works, aiming to place the appended papers in a broader context. Chapter 2 includes well-established findings relevant to the results presented in the articles, and Chapter 3 summarizes the appended papers and clarifies the author's contributions. Part II constitutes the main scientific contribution of this thesis and consists of a collection of four papers.

1.1 Linear systems with constant coefficients

We consider approximating the solutions to linear systems described by

$$Ax = b, \tag{1.1}$$

where $A \in \mathbb{R}^{n\times n}$ is nonsingular, $x \in \mathbb{R}^n$, and $b \in \mathbb{R}^n$. Linear systems of equations with symmetric and nonsymmetric system matrices A arise in the context of engineering and scientific problems. These problems are also present within applications of mathematics to problems in the social sciences, business, and economics [63]. For example, such systems appear when computing the approximate solution to partial differential equations (PDEs) via finite differences or the finite element method (FEM) [48], as well as within finite volume method discretizations for modeling time-dependent PDEs [36].

Linear systems also appear in the analysis and modeling of electrical circuits, where nonlinear systems are approximated by linear ones [11] and in structural analysis, when one is interested in determining the effect of loads on physical structures

and their components [25]. The possibility to solve larger linear systems allows, for example, a finer FEM-discretization and, therefore, higher accuracy and better modeling possibilities of physical systems [52].

The methods proposed in this thesis can be used to solve linear systems with complex coefficients. The formulation in (1.1) was chosen because many of the theoretical results in the appended papers are based on the assumption that the systems are real-valued. However, the algorithms developed in Paper A and some simulations shown in Paper B involve complex linear systems. In particular, this thesis develops methods for solving linear systems with nonsymmetric system matrices.

Iterative vs direct methods

There are two main classes of methods to solve the linear system described by (1.1): *direct methods* and *iterative methods*. For linear systems of the form (1.1), direct methods require $\mathcal{O}(n^3)$ work. This is too expensive for the applications in this thesis, considering that the input is usually only $\mathcal{O}(n^2)$, i.e., typically a vector paired with a sparse matrix. An example of such a method to solve (1.1) is the well-known Gaussian elimination method. This approach requires $\mathcal{O}(n)$ steps, where each step requires $\mathcal{O}(n^2)$ work. If Gaussian elimination is performed on a computer, it is typically only accurate to machine precision. Figure 1.1 is from [59]. This visualization shows that direct algorithms typically do not decrease the residual until $\mathcal{O}(n^3)$ operations have been performed. In other words, the solutions obtained through direct methods are only valuable once the entire, computationally demanding, algorithm has been executed.

On the other hand, the work per step of iterative methods depends on the structure of the matrix. Specifically, each step of the algorithm requires $x \mapsto Ax$, and some iterative algorithms also need the computation $x \mapsto A^T x$. These multiplications are the most costly part of each step, but can be carried out in an efficient way if, for example, A is very sparse. Iterative methods can, in the worst case scenario, require $\mathcal{O}(n)$ steps before terminating, where each step requires $\mathcal{O}(n^2)$ work. In certain circumstances, iterative methods achieve convergence in $\mathcal{O}(1)$ steps, and the work per step can be as low as $\mathcal{O}(n)$. This results in a method with total cost of $\mathcal{O}(n)$. Typically, however, the total cost of solving linear systems of the form (1.1) with iterative methods is $\mathcal{O}(n^2)$. Notably, iterative methods can, like direct methods, achieve accuracy to machine precision. Figure 1.1 shows that, under good conditions, iterative algorithms decrease the residual on each step [59].

The convergence of iterative methods can stagnate, and the prevention of this issue remains an area of active research; see the section on preconditioning, beginning on page 8. Papers A, B, and C of this thesis develop iterative methods for solving large, sparse linear systems, where the system matrix is nonsymmetric. Paper D makes use of the methods proposed in Paper C to obtain solutions to linear systems, then builds a reduced order model from the sampled solutions.

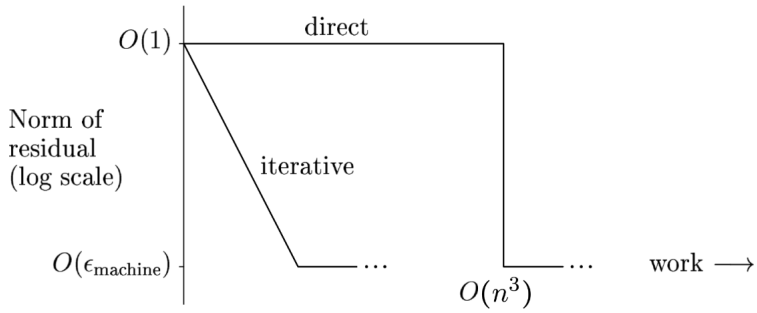

Fig. 1.1. Iterative vs direct methods [59]

Krylov subspace methods

Krylov subspace methods are among the most important *iterative* techniques available for solving large linear systems of the form (1.1). Methods of this type are based on a projection onto Krylov subspaces, where the projection can be orthogonal or oblique. The Krylov subspace of dimension $k \in \mathbb{N}$ associated with the matrix $A \in \mathbb{R}^{n\times n}$ and the vector $b \in \mathbb{R}^n$ is defined by

$$\mathcal{K}_k(A, b) \coloneqq \operatorname{span}\{b, Ab, A^2b, \ldots, A^{k-1}b\}. \tag{1.2}$$

In this way, Krylov subspaces are spanned by vectors of the form $p(A)b$, where p is a polynomial [48].

There are many different Krylov subspace methods for solving linear systems, and, depending on context, certain methods are more suitable than others. For example, some methods are preferable when the system matrix is symmetric, while for others, the convergence is characterized by the eigenvalues of A [59]. In general, Krylov subspace methods do not require direct access to the system matrix in order to approximate the solution to (1.1). Instead, one needs only a procedure to compute matrix-vector products with A and for some Krylov subspace methods with A^T. This is an attractive feature, particularly in cases where the system matrix is large.

The algorithms proposed in Papers A, B, and C to solve large, sparse, nonsymmetric linear systems are Krylov subspace methods. We describe two foundational Krylov subspace methods below. The novel algorithms proposed in this thesis were inspired by these methods and developed to inherit their robustness, accuracy, and efficiency.

GMRES for linear systems

The generalized minimal residual method (GMRES) [50] is one of the most popular iterative methods for solving linear systems with nonsymmetric system matrices [52]

and is a generalization of the minimal residual method (MINRES) for symmetric, indefinite matrices [45]. We briefly summarize the approach of GMRES as follows.

Let x_k denote the kth iterate produced by GMRES applied to the linear system (1.1), and, without loss of generality, assume $x_0 \coloneqq 0$. The approximation x_k is the vector in $\mathcal{K}_k(A,b)$ that minimizes the norm of the kth residual, i.e.,

$$x_k \coloneqq \underset{x \in \mathcal{K}_k(A,b)}{\operatorname{argmin}} \|b - Ax\|_2 .$$

The GMRES algorithm uses an Arnoldi process to generate an orthogonal basis for the Krylov subspace. After k iterations of the Arnoldi method, the relations

$$AQ_k = Q_{k+1}\underline{H}_k \tag{1.3}$$

hold, where the orthonormal columns of $Q_k \in \mathbb{R}^{n\times k}$ span $\mathcal{K}_k(A,b)$ and $\underline{H}_k \in \mathbb{R}^{k+1\times k}$ is an upper Hessenberg matrix, i.e., $(\underline{H}_k)_{i,j} = 0$, for $i > j+1$. The iterate x_k is computed as

$$x_k = Q_k z_k,$$

where $z_k = \operatorname{argmin}_{z\in\mathbb{R}^k} \|\underline{H}_k z - \beta e_1\|_2$ with $\beta \coloneqq \|b\|$ and $e_1 = \begin{bmatrix}1 & 0 & \dots & 0\end{bmatrix}^T \in \mathbb{R}^{k+1}$. Note, the full basis matrix Q_k must be stored.

In particular, the choice to orthogonalize the basis vectors stems from the observation that the vectors $A^i b$ tend to become less independent as $i \in \mathbb{N}$ increases. An orthogonalization process, typically based on a Gram–Schmidt or Householder procedure, alleviates this problem and leads to a more stable method [52]. Furthermore, the norm of the residual vectors produced by the iterates of GMRES are always nonincreasing, i.e.,

$$\|r_k\| \le \|r_{k-1}\| \le \ldots \le \|r_0\|, \tag{1.4}$$

since $\mathcal{K}_1(A,b) \subseteq \cdots \subseteq \mathcal{K}_k(A,b) \subseteq \mathcal{K}_{k+1}(A,b)$ and $r_i \in \mathcal{K}_{i+1}(A,b)$. Moreover, $\|r_n\| = 0$, ignoring rounding errors. This is because $\operatorname{span}\{q_1,\ldots,q_n\} = \mathbb{R}^n$.

The convergence of GMRES is often rapid [59]. The GMRES method applied to the linear system (1.1) is shown in Algorithm 1, and additional theoretical results concerning GMRES can be found in Chapter 2.

Bi-CG for linear systems

The bi-conjugate gradient method (Bi-CG) [17, 35] is another fundamental Krylov subspace method for solving linear systems with nonsymmetric system matrices. When used to solve systems of the form (1.1), this method requires the construction of a basis for the Krylov subspace $\mathcal{K}_k(A,b)$, as well as one for the Krylov subspace defined by

$$\mathcal{K}_k(A^T,c) \coloneqq \operatorname{span}\{c, A^T c, (A^T)^2 c, \ldots, (A^T)^{k-1} c\}, \tag{1.5}$$

Algorithm 1: GMRES

input: $A \in \mathbb{R}^{n\times n}$, $b \in \mathbb{R}^n$, $k \in \mathbb{N}$

output: Approximate solution x_k to $Ax = b$ from Krylov subspace of dimension k

1 $\beta := \|b\|$
2 $q_1 = b/\beta \in \mathbb{R}^n$
3 $Q_1 = \begin{bmatrix} q_1 \end{bmatrix} \in \mathbb{R}^{n\times 1}$
4 H_0 = empty matrix
5 **for** $i = 1, 2, \ldots, k$ **do**
6 $\quad y = Aq_i$
7 $\quad$ Orthogonalize y against $q_1, \ldots, q_i$ by a Gram–Schmidt process:
$\quad\quad h_i = Q_i^T y$
$\quad\quad y_\perp = y - Q_i h_i$
8 $\quad$ Possibly repeat Step 7
9 $\quad$ Compute $\beta_i = \|y_\perp\|$
10 $\quad$ Let $q_{i+1} = y_\perp/\beta_i$
11 $\quad$ Let $Q_{i+1} = \begin{bmatrix} Q_i, q_{i+1} \end{bmatrix} \in \mathbb{R}^{n\times(i+1)}$
12 $\quad$ Let $\underline{H}_i = \begin{bmatrix} \underline{H}_{i-1} & h_i \\ & \beta_i \end{bmatrix} \in \mathbb{R}^{(i+1)\times i}$
13 **end**
14 Compute $z_k = \operatorname{argmin}_{z\in\mathbb{R}^k} \|\underline{H}_k z - \beta e_1\|_2$
15 Return $x_k = Q_k z_k$

where $c \in \mathbb{R}^n$ and $c^T b \neq 0$. This is done by a *Lanczos biorthogonalization procedure*. Constructing the basis for (1.5) requires the action of the matrix A^T, which, in some applications, may not be available. We briefly summarize the Bi-CG method as follows.

As a result of the biorthogonalization process, the relation

$$T_k = W_k^T A V_k$$

holds. Specifically, $T_k \in \mathbb{R}^{k\times k}$ is tridiagonal and the columns of $V_k \in \mathbb{R}^{n\times k}$ are a basis for $\mathcal{K}_k(A, b)$, where $v_1 = b/\beta$ and $\beta := \|b\|$. Similarly, the columns of $W_k \in \mathbb{R}^{n\times k}$ span (1.5) and $W_k^T V_k = V_k^T W_k = I_k$. Let x_k denote the kth iterate of Bi-CG applied to (1.1). Here, we assume $x_0 := 0$ without loss of generality, and the kth iterate is defined as

$$x_k := V_k T_k^{-1}(\beta e_1),$$

where $e_1 = \begin{bmatrix} 1 & 0 & \ldots & 0 \end{bmatrix}^T \in \mathbb{R}^k$.

The nonincreasing residual norm property (1.4) does not always hold with Bi-CG. While GMRES chooses the iterate in $\mathcal{K}_k(A, b)$ that minimizes the residual norm, Bi-CG does not, despite operating within this subspace. The convergence of Bi-CG can be slow and erratic in practice [59].

Bi-CG is a *three-term recurrence method*. In particular, v_i is computed by a linear combination of v_{i-1}, v_{i-2}, and Av_{i-1}. An analogous relation holds for the basis vectors of (1.5). The full basis matrices do not need to be stored, and the Bi-CG method has lower storage and memory requirements than GMRES does. This property has particular significance when solving linear systems that take many iterations to converge.

The standard implementation of the Bi-CG algorithm applied to the linear system (1.1) is shown in Algorithm 2, and this method also returns an approximation to the so-called *dual system* given by $A^T z = c$. This implementation is based on an implicit LU factorization of the matrix T_k; see [48] for more details. Note, the Bi-CG method carries the risk of breakdown due to the underlying Lanczos biorthogonalization process [26]. In particular, this occurs when $s_i^T r_i = 0$ on line 8 of the algorithm; see Chapter 2 for a more rigorous explanation.

Algorithm 2: Bi-conjugate gradient method

input: $A \in \mathbb{R}^{n\times n}$, $b \in \mathbb{R}^n$, $c \in \mathbb{R}^n$ such that $c^T b \neq 0$, $k \in \mathbb{N}$
output: Approximate solutions x_k of $Ax = b$ and z_k of $A^T z = c$ from Krylov subspace of dimension k

1 $x_0 = 0, \quad z_0 = 0, \quad p_0 = r_0 = b, \quad q_0 = s_0 = c$
2 **for** $i = 1, 2, \ldots, k$ **do**
3 $\quad \alpha_i = (s_{i-1}^T r_{i-1})/(q_{i-1}^T A p_{i-1})$
4 $\quad x_i = x_{i-1} + \alpha_i p_{i-1}$
5 $\quad z_i = z_{i-1} + \overline{\alpha}_i q_{i-1}$
6 $\quad r_i = r_{i-1} - \alpha_i A p_{i-1}$
7 $\quad s_i = s_{i-1} - \overline{\alpha}_i A^T q_{i-1}$
8 $\quad \beta_i = (s_i^T r_i)/(s_{i-1}^T r_{i-1})$
9 $\quad p_i = r_i + \beta_i p_{i-1}$
10 $\quad q_i = s_i + \overline{\beta}_i q_{i-1}$
11 **end**
12 Return x_k, z_k

Preconditioning for Krylov subspace methods

Utilizing preconditioning in the context of Krylov subspace methods can lead to methods that are more efficient overall. The idea behind preconditioning in Krylov subspace methods is to solve a transformed system which has the same solution as the original problem but is likely to be easier to solve with an iterative method. Preconditioning is only suitable when it can be performed efficiently.

More specifically, the first step is to find a nonsingular preconditioning matrix $M \in \mathbb{R}^{n\times n}$ which is, in some way, close to A. The matrix M should also be chosen such that linear systems with this system matrix are inexpensive to solve. This

is an important requirement, since preconditioned Krylov subspace methods must perform an additional solve with the preconditioning matrix at each step [48].

There are three standard ways to apply the preconditioner when solving the linear system (1.1). One choice is to apply the preconditioner from the right, and, as a result, solve the systems

$$AM^{-1}u = b, \quad x = M^{-1}u.$$

Another option is to consider left preconditioning, i.e., approximating the transformed system

$$M^{-1}Ax = M^{-1}b.$$

Note, we do not form the matrix AM^{-1} or $M^{-1}A$, but rather rely on methods to perform a linear solve with M at each step of the Krylov algorithm. This can be done, for example, via an iterative method. A final choice is to consider

$$M_L^{-1}AM_R^{-1}u = M_L^{-1}b, \quad x = M_R^{-1}u,$$

which can be used when M is available in the factored form $M = M_L M_R$. In this case, the matrices M_L are M_R are typically triangular [48].

Preconditioning is an area of active research. Papers B, C, and D consider Krylov subspace methods with preconditioning for solving a discretized Helmholtz equation. Specifically, right preconditioning is used, though, in certain circumstances, left preconditioning could have been considered. See also, for instance, the works [13, 14, 15, 49] for other successful approaches utilizing preconditioning in the context of solving a discretized Helmholtz equation.

We summarize below two Krylov subspace methods that apply different preconditioners at each iteration. The methods developed in Papers B and C are inspired by this framework.

Flexible GMRES

The prior work [47] is based on the GMRES method with right preconditioning. Instead of applying the preconditioner M at each step, the preconditioner in this algorithm is permitted to change from iteration to iteration. Concretely, a linear system with system matrix M_i must be solved at the ith iteration, and the method is referred to as *flexible GMRES*.

Unlike iterates from the standard GMRES method, approximations obtained from flexible GMRES do not come from a Krylov subspace; see [54] for details. Additionally, the flexible GMRES procedure requires the storage of two basis matrices, of which one is the basis for a Krylov subspace. In contrast, the original GMRES algorithm, applied with or without preconditioning, requires only the storage of a single basis matrix for a Krylov subspace. Experimentally, using flexible GMRES can be worth the extra cost. See, for example, [9, 10, 24], for works based on its success.

Inexact Krylov subspace methods

In [5], [6], and [7], the authors experimented with the accuracy of the matrix-vector multiplications $x \mapsto Ax$ in Krylov subspace methods. When approximating a system of the form (1.1), it was observed that the norm of the error in the multiplication could increase as the norm of the *outer residual*, i.e., $r_i \coloneqq b - Ax_i$, decreased, and this could be done without destroying the convergence of the iterative method. Later, in [54], the authors provided rigorous theory describing how inexactly the matrix-vector products could be carried out while maintaining convergence for many fundamental Krylov subspace methods. Moreover, the authors provided computable bounds on the residual of these *inexact Krylov subspace methods*.

Papers B and C use the theory of inexact Krylov methods in order to apply a preconditioner in a more efficient way. Specifically, the preconditioners are applied using iterative solvers, and the tolerance is relaxed as the outer Krylov iteration proceeds. In this way, the approach is also similar to the framework proposed in flexible GMRES [47], and two basis matrices need to be stored. Similar approaches have been taken in, for example, the prior works [34] and [62].

1.2 Parameterized linear systems

Specifically, the research performed in this thesis considers parameterized linear systems of the form

$$A(\mu)x(\mu) = b, \qquad \mu = \begin{bmatrix}\mu_1, \ldots, \mu_\ell\end{bmatrix}^T \in \mathbb{R}^\ell, \tag{1.6}$$

for nonsingular $A(\mu) \in \mathbb{R}^{n\times n}$, $x(\mu) \in \mathbb{R}^n$, and $b \in \mathbb{R}^n$. The novel methods we propose are developed to approximate solutions $x(\mu)$ for many different values of μ. The subsequent sections provide descriptions of scientific applications that require solutions to linear systems of this form. The purpose of this is to motivate the work that was carried out and to clarify the importance of our results. An overview of how such systems are solved in the appended papers is also provided.

Parameterized shifted linear systems

Solving linear systems of the form

$$(A + \mu I)x(\mu) = b \tag{1.7}$$

in a computationally efficient way for many different values of the parameter $\mu \in \mathbb{R}$ is of scientific interest. Here the system matrix differs from $A \in \mathbb{R}^{n\times n}$ by a scalar multiple of the identity matrix, and this linear system is of the same form as (1.6) with $\ell = 1$. Specifically, the dependence on the parameter is linear, and we refer to such linear systems as *parameterized shifted linear systems*.

Linear systems of the form (1.7) arise, for instance, in the context of model order reduction [27], Tikhonov regularization for ill-posed problems [23, 32], and oscillatory hydraulic tomography [51], used to image the subsurface of the earth. Many prior works have been successfully developed for solving these linear systems; see, for example, [4, 18, 22, 56]. Parameterized shifted linear systems appear throughout this thesis.

Krylov subspace methods are attractive for solving linear systems of this variety. This is because the basis matrix for one Krylov subspace can be used to compute solutions corresponding to many different μ simultaneously, due to the *shift-invariance property of Krylov subspaces*:

$$\mathcal{K}_k(A, b) = \mathcal{K}_k(A + \mu I, b). \tag{1.8}$$

We show how one basis matrix can be used to solve (1.7) for many different μ using the GMRES method as follows.

The Arnoldi relations (1.3) can be expressed equivalently as

$$(A + \mu I)Q_k = Q_{k+1}(\underline{H}_k + \mu \underline{I}_k),$$

where $\underline{I}_k \in \mathbb{R}^{k+1\times k}$ is the identity matrix with an extra row of zeros and $Q_{k+1}\underline{I}_k = Q_k$. Furthermore, the columns of Q_k are an orthonormal basis for $\mathcal{K}_k(A + \mu I, b)$, and this basis matrix is generated independently of μ, as described in Algorithm 1. The kth iterate of the GMRES method applied to (1.7) can be computed as

$$x_k(\mu) = Q_k z_k(\mu),$$

where

$$z_k(\mu) = \underset{z\in\mathbb{R}^k}{\operatorname{argmin}} \left\|(\underline{H}_k + \mu \underline{I}_k)z - \beta e_1\right\|_2$$

and $e_1 = \begin{bmatrix}1 & 0 & \dots & 0\end{bmatrix}^T \in \mathbb{R}^{k+1}$. Thus, after the Arnoldi portion of the method has been executed, solving (1.7) requires only the solution to a small least squares problem and a matrix-vector multiplication for each μ of interest. Moreover, $x_k(\mu) \in \mathcal{K}_k(A + \mu I, b)$.

Papers A, B, and C in this thesis develop computationally effective approaches for solving shifted systems based on a similar framework. Though a GMRES procedure was described here and used in Paper A, other Krylov subspace methods can be used in an analogous way. Specifically, Paper B considers an approach based on inexact GMRES, and the algorithm proposed in Paper C is based on the Bi-CG method.

Parameterized partial differential equations

Consider a discretization of the inhomogeneous, indefinite Helmholtz equation given by

$$\left(\nabla^2 + \kappa(\mu, x)^2\right) u(x) = h(x), \quad x \in \Omega \tag{1.9}$$

with a homogeneous Dirichlet boundary condition: $u(x) = 0$, for $x \in \partial\Omega$. Here, κ is referred to as the *wave number*, and we assume the function $\kappa(\mu, x)^2 > 0$ is *separable* in μ and x, i.e.,

$$\kappa(\mu, x)^2 \coloneqq f_1(\mu)g_1(x) + \ldots + f_{nf}(\mu)g_{nf}(x),$$

where $\mu \in \mathbb{R}^\ell$, $f_i : \mathbb{R}^\ell \to \mathbb{R}$, $g_i : \mathbb{R}^d \to \mathbb{R}$, and d is the dimension of the spatial domain Ω. Furthermore, h represents sources inside the domain, and the solution $u(x)$ describes wave propagation [16]. This elliptic PDE was chosen as a representative example. We refer to μ as the *input parameter*. Observe that a change in the parameter μ leads to a change in the *response*, i.e., the solution $u(x)$. In this way, a discretization of this problem is of the form (1.6).

Parameterized PDEs similar to (1.9) arise in the study of geophysics; see, for instance, [65]. Specifically, in the context of the parameterized Helmholtz equation, the input parameter can represent the angular frequency of the wave or the direction of the incoming wave [41]. Alternatively, μ can describe the geometry of the domain or the materials through which the waves move [53].

The algorithms proposed in this thesis can be seen as tools for *uncertainty quantification*. More concretely, when there is uncertainty in the input parameter, μ can be treated as a random variable. For each *observation* of μ, we can obtain a corresponding *realization* of the solution $u(x)$, which is also a random variable. This is referred to as *forward propagation of uncertainty*. Our methods offer a computationally efficient way to approximate $u(x)$ for many values of the parameter. Having access to many such approximations allows us to compute descriptive properties, such as the mean and variance, of the solutions $u(x)$ [57].

Numerical examples

The numerical simulations in Papers A, B, C, and D consider a discretization of (1.9) by the finite element method, though other discretization methods could have been used. The resulting discretized system is of the form (1.6), where

$$A(\mu) \coloneqq A_0 + f_1(\mu)A_1 + \ldots + f_{nf}(\mu)A_{nf}$$

and b is the corresponding load vector. Specifically, $A_0, \ldots, A_{nf}$ are large and sparse, the functions f_i are nonlinear in μ, and we generate the discretized system using FEniCS [2]. The solution to the discretized problem is an approximation to the Helmholtz equation (1.9). Performing the discretization with a finer mesh leads to a more accurate model of the PDE and a larger linear system [52]. Thus, developing efficient methods for large, sparse systems is of interest.

Furthermore, we are interested in the solutions for many different μ. The obvious approach of directly interpolating a few approximations is not a suitable strategy because solutions to the parameterized Helmholtz equation are extremely sensitive to the parameters [41]. As an alternative, Papers A, B, and C develop

Krylov subspace methods to obtain approximations corresponding to many different values of a scalar μ. Paper D proposes a reduced order model to achieve the same goal with $\mu \in \mathbb{R}^2$.

More precisely, Papers A, B, and C are based on solving an equivalent linearized system of the form (1.7) with Krylov subspace methods. Though using a linearization requires the solution to a larger system, sparsity patterns keep the cost low, particularly with respect to matrix-vector multiplications. Moreover, approximations to (1.9) corresponding to many different scalar μ can be generated from one basis matrix; see the section on solving parameterized shifted linear systems with Krylov subspace methods, beginning on page 10. The algorithm developed in Paper D relies on sampling performed using a modified version of the method proposed in Paper C.

Outlook

The numerical solution to (1.9) is usually difficult to obtain with standard iterative methods, especially if the wave number κ is large. Improvements can be made by, for example, considering Krylov subspace methods with preconditioning by incomplete factorizations, domain decomposition, or multigrid methods. These approaches lead to better converge of the iterative method, though the number of iterations tends to increase with the size of the wave number. Building algorithms whose convergence is independent of the wave number remains a difficult problem, and sophisticated solvers are required [16]. In this thesis, we consider numerical simulations of (1.9) for modestly sized wave numbers without tackling the problem of increasing wave numbers. Continued research into combining our approaches with preconditioning designed for the Helmholtz equation with large wave numbers could lead to promising approaches.

Linear systems with time delays

Parameterized linear systems of the form (1.6) also arise in the context of linear time-invariant (LTI) systems. Specifically, we consider systems with time-delays. These systems are applicable in various domains, including flight control, chemical/process control, and communication networks. The delays can be in the state, the control input, or the measurements [20, 40].

In particular, we develop methods to study the system described by

$$\dot{x}(t) = A_0 x(t) + A_1 x(t-\tau) - bv(t), \tag{1.10a}$$

$$y(t) = C^T x(t). \tag{1.10b}$$

Here, the matrices A_0, $A_1 \in \mathbb{R}^{n\times n}$ have random entries, $b \in \mathbb{R}^n$ is a random vector, and $C = I \in \mathbb{R}^{n\times n}$. The delay is given by $\tau \in \mathbb{R}$, where $\tau > 0$, $v(t) \in \mathbb{R}$ describes the input, and $y(t) \in \mathbb{R}^n$ is the output. Additionally, $x(t) \in \mathbb{R}^n$ represents the state vector at time $t \in \mathbb{R}$, and $\dot{x}(t) \in \mathbb{R}^n$ describes how the state changes with respect to time. Without loss of generality, we take $\tau \coloneqq 1$.

Numerical examples

Simulations involving the *transfer function* of the system (1.10) appear in Papers A, B, and C. The transfer function of an LTI system is the ratio of the Laplace transform of the output variable to the Laplace transform of the input variable with all initial conditions set to zero. Specifically, the transfer function is parameterized in the Laplace variable and describes the input-output behavior of the system without giving any information about the system's internal structure. The transfer function is independent of the input, and, from analyzing the poles of the transfer function, the stability of the system can be determined [43]. In particular, the transfer function of the system (1.10) is the solution to the parameterized linear system (1.6), where

$$A(\mu) \coloneqq -\mu I + A_0 + A_1 e^{-\mu}; \tag{1.11}$$

see [66] for details. Here, the parameter μ is the Laplace variable, and we are interested in evaluating the transfer function for many different μ.

The algorithms developed in this thesis consider Krylov subspace methods to solve an equivalent linearized problem of the form (1.7) in an efficient way. This procedure involves solving a linear system which is larger than the original one. Sparsity patterns in the linearization are exploited in order to develop computationally efficient methods. Further, solutions corresponding to many different values of the parameter μ are generated from just one basis matrix; see the section on parameterized shifted linear, beginning on page 10.

Outlook

Improved modeling of transfer functions leads to better understanding of engineering control systems in general [44]. Particularly, in robust control of linear systems, stability and performance are often evaluated using $\mathcal{H}_\infty$-norms of transfer functions [64]. For systems with transfer function H, this norm is defined as

$$\|H\|_{\mathcal{H}_\infty} \coloneqq \sup_{\omega \in \mathbb{R}} |H(i\omega)|,$$

for stable single input, single output systems (SISO),

$$\|H\|_{\mathcal{H}_\infty} \coloneqq \sup_{\omega \in \mathbb{R}} \sigma_{\max}(H(i\omega)), \tag{1.12}$$

for stable multiple input, multiple output systems (MIMO), and $\|H\|_{\mathcal{H}_\infty} \coloneqq \infty$ for unstable systems. The quantity in (1.12) can be interpreted as the largest singular value of the transfer function H over the imaginary axis. Many methods have been proposed to approximate the $\mathcal{H}_\infty$-norm without evaluating the transfer function directly; see, for instance, the approaches developed in [8, 28, 29]. Clearly, if the dimension n is large, evaluating the transfer function for even a few different values of ω is computationally demanding.

Furthermore, the $\mathcal{H}_2$-norm is important as a robustness measure with respect to noise or external disturbances [46]. We define the $\mathcal{H}_2$-norm of a transfer function H by

$$\|H\|_{\mathcal{H}_2} \coloneqq \left(\frac{1}{2\pi} \int_{-\infty}^{\infty} |H(i\omega)|^2 \, \mathrm{d}\omega \right)^{\frac{1}{2}},$$

for stable SISO systems, and

$$\|H\|_{\mathcal{H}_2} \coloneqq \left(\frac{1}{2\pi} \int_{-\infty}^{\infty} \operatorname{tr}(H(-i\omega)^T H(i\omega)) \, \mathrm{d}\omega \right)^{\frac{1}{2}},$$

for stable MIMO systems. There are many successful methods for LTI systems which utilize the $\mathcal{H}_2$-norm. Notably, in [27], the iterative rational Krylov algorithm (IRKA) was proposed for producing the reduced order model with transfer function H_r that minimizes the quantity $\|H - H_r\|_{\mathcal{H}_2}$ in the context of SISO LTI systems without delays. Specifically, the method obtains the best stable dynamical system of dimension r for approximating the original system, where $r \ll n$.

Efficient and accurate methods for evaluating transfer functions for many different values of the parameter could lead to promising new methods for computing $\mathcal{H}_\infty$-norms and $\mathcal{H}_2$-norms. Such results could lead to improved methods for evaluating and modeling control systems.

Parameter estimation problems for PDEs

The goal of *parameter estimation* in the context of the parameterized linear system (1.6) is to determine the value of the unknown parameter μ that best fits the corresponding observed solution $x(\mu)$. The parameter estimation problem is an *inverse problem*, as we aim to infer properties of the system that produced a given solution [58]. Inverse problems are typically *ill-posed*. Specifically, there may be no solution, the solution may not be unique, or there may be a unique solution that is very sensitive to changes in the response [57].

Consider the problem of having a set of solutions to the system (1.6), and let these solutions be denoted $\{x_1, \ldots, x_N\}$, for $x_i \in \mathbb{R}^n$. Here the corresponding parameters μ_i are unknown and of interest, and the observed solutions are noisy, i.e., $x_i \coloneqq x(\mu_i) + \varepsilon_i$, where $\varepsilon_i \sim \mathcal{N}(0, \sigma^2)$ is assumed. In this context, μ_i is an *observation* of the random variable μ, and, thus, x_i is the corresponding *realization* of the random variable x. The noise in the observations could arise, for example, from measurement error.

Numerical examples

The algorithms developed in this thesis return functions $\tilde{x}(\mu) \in \mathbb{R}^n$ such that $\tilde{x}(\mu) \approx x(\mu)$, and the functions can be evaluated in a computationally efficient way for many values of the parameter. As a result, the solution of a parameter estimation problem

can be obtained by a direct application of the proposed algorithms. Concretely, we obtain an approximation to μ by considering the minimizations

$$\mu_i^* \coloneqq \operatorname*{argmin}_{\mu} \|x_i - \tilde{x}(\mu)\|_2, \quad i = 1, \ldots, N. \tag{1.13}$$

Numerical simulations involving a parameter estimation in the context of PDEs can be found in Paper D with $\mu = [\mu_1, \mu_2]^T \in \mathbb{R}^2$. In one simulation, μ describes unknown material parameters in the context of a Helmholtz equation. A second simulation describes an advection-diffusion equation, which can be used to model particle transport; see [12, 60]. Here, the unknown parameters correspond to an advection parameter as well as a diffusion coefficient. It is assumed that the solutions in these examples depend on μ in a nonlinear way.

One could, alternatively, approach this parameter estimation problem by solving (1.6) for many different values of the parameter μ using standard methods and then compare each x_i to the collection of computed solutions. This becomes infeasible when many precomputed solutions to large systems are required or the dimension of $\mu \in \mathbb{R}^\ell$ grows.

Outlook

From many observations of solutions to a parameterized linear system, we can determine a *posterior distribution* for the parameter μ. In other words, given x_i, we compute μ_i^* and construct a probability for μ using this information combined with a prior distribution of μ. Specifically, this probability is based on the realizations x_i [57].

Parameter estimation problems can, for example, help oil companies locate oil deposits to make a profit or help seismologists make predictions regarding earthquakes [57]. Other parameter estimation problems in the context of parameterized PDEs have been tackled in, for instance, [38, 42, 53]. See also [39] and references therein for a thorough summary of both theoretical and numerical methods for parameter estimation in the context of parameterized Helmholtz equations. Improved methods for parameterized linear systems could lead to better methods for tackling these types of problems.

Parameterized right-hand sides

The linear system (1.6) is the focus of the scientific work presented in this thesis. Thus, we consider only parameterized linear systems with constant right-hand sides. The prior work [55] studied the set of parameterized linear systems of dimension $n \times n$ with complex coefficients given by

$$(A + \mu_i I)x(\mu_i) = b(\mu_i), \quad \mu_i \in \mathbb{C}, \quad \text{for} \quad i = 1, \ldots, N. \tag{1.14}$$

In that setting, the right-hand side vectors $b(\mu_i)$ are unrelated, and the systems were solved simultaneously by block Krylov subspace methods. The general approach

could be adapted to our setting to, for example, solve parameterized PDEs with certain types of parameter-dependent boundary conditions. However, this would require additional research. The method presented in [55] is briefly summarized as follows.

Without loss of generality, let $x_0 \coloneqq 0 \in \mathbb{R}^n$ be the initial approximation to each system in (1.14), and let the initial residuals be given by $r_0(\mu_i) \coloneqq b(\mu_i) - Ax_0$. The shift-invariance of Krylov subspaces described by (1.8) cannot be exploited in this setting. This is because the vectors $b(\mu_i)$ are linearly independent and, thus,

$$\mathcal{K}_k(A + \mu_i I, b(\mu_i)) \neq \mathcal{K}_k(A + \mu_j I, b(\mu_j)), \quad \text{for} \quad i \neq j.$$

We focus instead on block Krylov subspace methods. Consider the Sylvester equation

$$AX^\mu + X^\mu D = B^\mu, \tag{1.15}$$

where $D \coloneqq \operatorname{diag}(\mu_1, \ldots, \mu_N) \in \mathbb{C}^{N\times N}$, $B^\mu \coloneqq \begin{bmatrix} b(\mu_1) & \ldots & b(\mu_N) \end{bmatrix} \in \mathbb{C}^{n\times N}$ are given and the matrix

$$X^\mu \coloneqq \begin{bmatrix} x(\mu_1) & \ldots & x(\mu_N) \end{bmatrix} \in \mathbb{C}^{n\times N}$$

is unknown. In this way, solutions to (1.14) and (1.15) are equivalent. Let $X_0^\mu \coloneqq 0 \in \mathbb{C}^{n\times N}$ be the initial approximation to X^μ and

$$R_0^\mu \coloneqq B^\mu - AX_0^\mu - X_0^\mu D = B^\mu$$

be the corresponding initial residual. The block Krylov subspace generated from A and B^μ is defined as

$$\mathbb{K}_k(A, B^\mu) \coloneqq \mathcal{K}_k(A, b(\mu_1)) + \mathcal{K}_k(A, b(\mu_2)) + \ldots + \mathcal{K}_k(A, b(\mu_N)),$$

and the relation

$$\mathbb{K}_k(A, B^\mu) = \mathbb{K}_k(A + \mu I, B^\mu), \quad \mu \in \mathbb{C},$$

is referred to as the *shift-invariance property of block Krylov subspaces*. It is assumed that A and B^μ are chosen such that $\dim(\mathbb{K}_k(A, B^\mu)) = kN$.

The subspace $\mathbb{K}_k(A, B^\mu)$ can be represented in terms of the orthonormal *block Arnoldi basis* given by $\{U_1, \ldots, U_k\}$. Here each $U_i \in \mathbb{C}^{n\times N}$ has orthonormal columns, and the columns of U_i are orthogonal to the columns of U_j, for $i \neq j$. We define $\boldsymbol{U}_k \coloneqq \begin{bmatrix} U_1 & \ldots & U_k \end{bmatrix} \in \mathbb{C}^{n\times kN}$ and denote the *block upper Hessenberg matrix* generated by the *block Arnoldi method* as $\underline{\boldsymbol{H}}_k \in \mathbb{C}^{(k+1)N\times kN}$. The relation

$$A\boldsymbol{U}_k = \boldsymbol{U}_{k+1}\underline{\boldsymbol{H}}_k$$

holds, as well as the relations

$$(A + \mu_i I)\boldsymbol{U}_k = \boldsymbol{U}_{k+1}\underline{\boldsymbol{H}}_k(\mu_i), \tag{1.16}$$

where

$$\underline{\boldsymbol{H}}_k(\mu_i) \coloneqq \underline{\boldsymbol{H}}_k + \mu_i \begin{bmatrix} I_{kN\times kN} \\ 0_{N\times kN} \end{bmatrix}.$$

The shifted Arnoldi relation (1.16) is used to derive Algorithm 3.1 in [55]. Specifically, approximations to the parameterized linear systems (1.14) are elements of the block Krylov subspace $\mathbb{K}_k(A, B^\mu)$, and the corresponding basis matrix $\boldsymbol{U}_k$ is used to generate the solutions. Computing each $x(\mu_i)$ requires the solution to a linear least squares problem and a matrix-vector product. This procedure is similar to the GMRES method applied to the parameterized shifted linear system (1.7), described in the section on shifted parameterized linear systems beginning on page 10.

Chapter 2

Preliminaries

This chapter contains an overview of some well-known concepts and theoretical results from the literature. Specifically, it includes fundamental results related to the Krylov subspace methods that inspired the appended papers. The purpose of this is to provide an overview and foundation for the topics included in this thesis. Since all the results in this chapter are well-established, proofs are omitted. Further details can be found in the books in the reference list at the end of this part.

In particular, the following results pertain to the linear system with constant coefficients described by (1.1) or, equivalently, the parameterized linear system (1.6) evaluated at a specific value of the parameter. We begin with some results regarding matrix factorizations, which are used both in computation and theoretical analysis throughout this dissertation.

Theorem 2.1 (LU factorization)**.** *Let $A \in \mathbb{R}^{n\times n}$ be such that*

$$\det(A(1:i,1:i)) \neq 0, \quad i = 1,\dots,n-1.$$

Then there exists an LU factorization of A, i.e., a lower triangular matrix $L \in \mathbb{R}^{n\times n}$ and an upper triangular matrix $U \in \mathbb{R}^{n\times n}$ such that $A = LU$.

The LU factorization plays an important role in various computational tasks, such as solving systems of linear equations, computing determinants, and inverting matrices. In particular, systems involving lower or upper triangular matrices can be solved efficiently by considering forward or back substitution, respectively.

Theorem 2.2 (QR factorization)**.** *If $A \in \mathbb{R}^{n\times m}$, then there exists an orthogonal matrix $Q \in \mathbb{R}^{n\times n}$ and an upper triangular matrix $R \in \mathbb{R}^{n\times m}$ such that*

$$A = QR.$$

Since the transpose of an orthogonal matrix is equal to its inverse, the QR factorization helps to simplify many computations. The following theorem is applicable in the case of overdetermined systems.

Theorem 2.3 (Reduced QR factorization)**.** *If $A \in \mathbb{R}^{n\times m}$ has full column rank and $n \geq m$, then the reduced QR factorization of A is given by $Q_1 R_1$, where*

$$A = QR = \begin{bmatrix} Q_1 & Q_2 \end{bmatrix} \begin{bmatrix} R_1 \\ 0 \end{bmatrix} = Q_1 R_1. \tag{2.1}$$

Here $R_1 \in \mathbb{R}^{m\times m}$ is an upper triangular matrix, 0 is a zero matrix of dimension $(n-m)\times m$, $Q_1 \in \mathbb{R}^{n\times m}$, $Q_2 \in \mathbb{R}^{n\times(n-m)}$, and Q_1 and Q_2 have orthogonal columns. Moreover, the factorization $A = Q_1 R_1$ where R_1 has positive diagonal entries is unique.

As the algorithms developed in this thesis are Krylov subspace methods, the following definitions are foundational.

Definition 2.4 (Krylov subspace)**.** The Krylov subspace of dimension k generated by the matrix $A \in \mathbb{R}^{n\times n}$ and the vector $b \in \mathbb{R}^n$ is given by

$$\mathcal{K}_k(A, b) \coloneqq \operatorname{span}\{b, Ab, \ldots, A^{k-1}b\}. \tag{2.2}$$

Definition 2.5 (Krylov matrix)**.** The Krylov matrix $K_k \in \mathbb{R}^{n\times k}$ associated with the Krylov subspace (2.2) is defined by

$$K_k \coloneqq \begin{bmatrix} b & Ab & A^2b & \ldots & A^{k-1}b \end{bmatrix}. \tag{2.3}$$

The parameterized shifted linear system (1.7) arises throughout this thesis. As a result, the algorithms we develop rely heavily on the following theorem.

Theorem 2.6 (Shift- and scaling-invariance)**.** *Krylov subspaces are shift- and scaling-invariant, i.e.,*

$$\mathcal{K}_k(A, b) = \mathcal{K}_k(\mu A, b) = \mathcal{K}_k(A - \mu I, b),$$

for nonzero $\mu \in \mathbb{R}$.

2.1 Regarding GMRES

The particular Krylov subspace methods developed in Papers A and B were inspired by the GMRES method and inherit many of its properties. When the standard GMRES method is applied to (1.1), an Arnoldi procedure constructs a basis matrix for (2.2) with orthogonal columns. The following theorem formalizes this.

Theorem 2.7 (Arnoldi relation)**.** *Suppose the Arnoldi algorithm applied to the matrix $A \in \mathbb{R}^{n\times n}$ with starting vector $q_1 = b/\beta \in \mathbb{R}^n$ and $\beta \coloneqq \|b\|$ does not break down after k iterations. Then, the Arnoldi relation*

$$AQ_k = Q_{k+1}\underline{H}_k \tag{2.4}$$

holds, where $Q_{k+1} = \begin{bmatrix} Q_k & q_{k+1} \end{bmatrix} \in \mathbb{R}^{n\times k+1}$, $\underline{H}_k \in \mathbb{R}^{k+1\times k}$ is an upper Hessenberg matrix, and $\underline{H}_{k-1} = \underline{H}_k(1:k, 1:k-1)$. Furthermore, the columns of Q_k are an orthonormal basis for the Krylov subspace $\mathcal{K}_k(A, b)$.

We describe the relation between the basis matrix Q_k in (2.4) and the Krylov matrix associated with the subspace (2.2) in the following theorem.

Theorem 2.8. *The matrix Q_k in* (2.4) *is the reduced QR factor of the Krylov matrix* (2.3)*, i.e.,*

$$K_k = Q_k R_k.$$

When the GMRES method is used to solve (1.1), the Krylov basis vectors are orthogonalized using a Gram–Schmidt procedure, and the full basis matrix must be stored. This can be costly when many iterations are needed to achieve convergence. We illustrate this as follows.

Corollary 2.9. *Let $q_1, \ldots, q_k$ denote the columns of Q_k. The Krylov vector q_{k+1} in* (2.4) *fulfills a $(k+1)$-term recurrence between itself and the previous Krylov vectors since*

$$\left[\begin{array}{c} \\ A \\ \\ \end{array}\right]\left[\begin{array}{c|c|c} & & \\ q_1 & \cdots & q_k \\ & & \end{array}\right] = \left[\begin{array}{c|c|c} & & \\ q_1 & \cdots & q_{k+1} \\ & & \end{array}\right]\begin{bmatrix} h_{1,1} & \cdots & \cdots & h_{1,k} \\ h_{2,1} & \cdots & \cdots & h_{2,k} \\ & h_{3,2} & \cdots & h_{3,k} \\ & & \ddots & \vdots \\ & & & h_{k+1,k} \end{bmatrix}$$

and $Aq_k = h_{1,k}q_1 + h_{2,k}q_2 + \ldots + h_{k+1,k}q_{k+1}$.

Proposition 2.10. *The cost of orthogonalizing the columns of the Krylov matrix* (2.3) *with the Gram–Schmidt procedure is $\mathcal{O}(nk^2)$.*

Definition 2.11 (Iterates of GMRES)**.** Let $x_0 \coloneqq 0$. The kth iterate of the GMRES method applied to (1.1) is given by

$$x_k \coloneqq Q_k z_k,$$

where $z_k \coloneqq \operatorname{argmin}_{z\in\mathbb{R}^k} \|\underline{H}_k z - \beta e_1\|_2$, $\beta \coloneqq \|b\|$, $e_1 = \begin{bmatrix} 1 & 0 & \ldots & 0 \end{bmatrix}^T \in \mathbb{R}^{k+1}$, and Q_k, $\underline{H}_k$ are as in (2.4).

Remark 2.12 (Solving the least squares problem)*.* The vector z_k in Definition 2.11 can be obtained in an efficient way by considering a QR decomposition of the Hessenberg matrix $\underline{H}_k$. This factorization can be reused when computing x_{k+1}.

The residual vector is used to approximate the error when the true solution is not available. The following fundamentals characterize the residual.

Definition 2.13 (Residual of iterates, GMRES)**.** The residual of the kth iterate of the GMRES method applied to the linear system (1.1) is computed efficiently as

$$\|r_k\|_2 = \|\beta e_1 - \underline{H}_k z_k\|_2 .$$

Furthermore, $r_k \in \mathcal{K}_{k+1}(A, b)$.

Definition 2.14 (Petrov-Galerkin condition, GMRES)**.** The property

$$b - Ax_k \perp A\mathcal{K}_k(A, b)$$

is referred to as the *Petrov-Galerkin condition* for the GMRES method.

Proposition 2.15 (Minimal residual, GMRES)**.** *The kth iterate of the GMRES method applied to* (1.1) *is the vector in $\mathcal{K}_k(A, b)$ which minimizes the kth residual norm*

$$\|r_k\|_2 \coloneqq \|b - Ax_k\|_2 .$$

The main convergence theorem for the GMRES method is based on describing the iterates using matrix polynomials. We begin with some basic definitions.

Definition 2.16 (Space of polynomials)**.** The space of polynomials P_k is given by

$$P_k \coloneqq \{\text{polynomials of degree} \leq k \text{ with } p(0) = 1\}. \tag{2.5}$$

Definition 2.17 (Polynomial definition of GMRES iterates)**.** The iterates of the GMRES method applied to (1.1) can be expressed as

$$x_k \coloneqq q_k(A)b$$

at iteration k, where q_k is a polynomial of degree $k - 1$.

The corresponding residual vectors can also be represented using this polynomial interpretation.

Definition 2.18 (Polynomial definition of residual, GMRES)**.** The residual of the kth iterate of GMRES can be described by

$$r_k = b - Aq_k(A)b = (I - Aq_k(A))b = p_k(A)b \tag{2.6}$$

with $p_k(z) = 1 - zq_k(z)$ and $p_k \in P_k$.

Proposition 2.19. *At iteration k, the GMRES process applied to* (1.1) *determines the coefficients of p_k such that $\|r_k\|_2$ is minimized.*

We can describe the convergence of the GMRES method applied to (1.1) using properties of the system matrix. The following is the main convergence theorem of the GMRES method.

Theorem 2.20. *Let $A \in \mathbb{R}^{n\times n}$ be nonsingular and diagonalizable with diagonalization $A = V\Lambda V^{-1}$. Then,*

$$\frac{\|r_k\|}{\|b\|} \leq \inf_{p_k\in P_k} \|p_k(A)\| \leq \kappa(V) \inf_{p_k\in P_k} \|p_k\|_{\lambda(A)},$$

where $\kappa(\cdot)$ denotes the condition number, $\lambda(A)$ is the spectrum of A, and $\|p_k\|_{\lambda(A)} \coloneqq \sup_{\lambda_i\in\lambda(A)}|p_k(\lambda_i)|$.

The bound described above includes the condition number of V. A large condition number can indicate potential linear dependence among the columns.

The following bound, based on a localization of the eigenvalues of A, is a direct consequence of Theorem 2.20.

Corollary 2.21. *Let A fulfill the same conditions as in Theorem 2.20. Furthermore, let all the eigenvalues $\lambda_i \in \lambda(A)$ be enclosed in a disc of radius $r \in \mathbb{R}$ centered at $c \in \mathbb{C}$. Then,*

$$\frac{\|r_k\|}{\|b\|} \leq \kappa(V)\left(\frac{r}{|c|}\right)^k. \tag{2.7}$$

Note, the results in Theorem 2.20 and Corollary 2.21 give a sufficient, but not necessary, condition for convergence. Moreover, the bound in (2.7) is not useful if the eigenvalues are centered too near the origin.

Breakdown can occur in a GMRES context, but is not something to be feared. The following remark formalizes this.

Remark 2.22 (Breakdown, GMRES). A so-called *lucky breakdown* occurs if $h_{k+1,k} = 0$ in Corollary 2.9. In this case, the algorithm terminates with $\|r_k\| = 0$, i.e., x_k is an exact solution.

2.2 Regarding Bi-CG

The algorithm developed in Paper C and utilized in Paper D is based on the three-term recurrence Krylov subspace method Bi-CG. Consider approximating the linear system (1.1) with this method. Unlike algorithms based on an Arnoldi process, this approach does not require the storage of a basis matrix or a separate orthogonalization process. Fundamental properties of the Lanczos biorthogonalization procedure, upon which Bi-CG is based, are described as follows.

Theorem 2.23 (Lanczos biorthogonalization)**.** *The Lanczos biorthogonalization generates matrices V_k, $W_k \in \mathbb{R}^{n\times k}$, $T_k \in \mathbb{R}^{k\times k}$, and $\underline{T}_k$, $\bar{T}_k^T \in \mathbb{R}^{(k+1)\times k}$ such that the*

relations

$$AV_k = V_k\, T_k \; + \beta_k v_{k+1} e_k^T \; = V_{k+1}\, \underline{T}_k, \tag{2.8a}$$

$$A^T W_k = W_k T_k^T + \gamma_k w_{k+1} e_k^T = W_{k+1} \bar{T}_k^T \tag{2.8b}$$

hold, where e_k is the kth column of the identity matrix of dimension $k \times k$ and $\beta_k, \gamma_k \in \mathbb{R}$ are entries of $\underline{T}_k$ and $\bar{T}_k^T$, respectively. Moreover, the columns of V_k span the subspace $\mathcal{K}_k(A, b)$, and the columns of W_k span the subspace $\mathcal{K}_k(A^T, c)$, for $c \in \mathbb{R}^n$, $c^T b \neq 0$.

In particular, this biorthogonalization procedure requires the action of the the matrix A as well as the action of the adjoint A^T, which may not be available in some applications. The following results clarify the output of this process.

Corollary 2.24 (Tridiagonal matrices, Lanczos biorthogonalization)**.** *In the relations* (2.8)*, the square matrix T_k has the form*

$$T_k \coloneqq \begin{bmatrix} \alpha_1 & \gamma_1 & & \\ \beta_1 & \ddots & \ddots & \\ & \ddots & \ddots & \gamma_{j-1} \\ & & \beta_{k-1} & \alpha_k \end{bmatrix}, \tag{2.9}$$

and the matrices $\underline{T}_k$ and $\bar{T}_k^T$ are given by

$$\underline{T}_k \coloneqq \begin{bmatrix} \alpha_1 & \gamma_1 & & \\ \beta_1 & \ddots & \ddots & \\ & \ddots & \ddots & \gamma_{k-1} \\ & & \beta_{k-1} & \alpha_k \\ & & & \beta_k \end{bmatrix}, \quad \bar{T}_k^T \coloneqq \begin{bmatrix} \alpha_1 & \beta_1 & & \\ \gamma_1 & \ddots & \ddots & \\ & \ddots & \ddots & \beta_{k-1} \\ & & \gamma_{k-1} & \alpha_k \\ & & & \gamma_k \end{bmatrix}. \tag{2.10}$$

In particular, T_k is tridiagonal and $\underline{T}_k$, $\bar{T}_k^T$ are tridiagonal, upper Hessenberg matrices.

Corollary 2.25. *The Lanczos biorthogonalization procedure gives the relation*

$$W_k^T V_k = V_k^T W_k = I_k, \tag{2.11}$$

where $I_k \in \mathbb{R}^{k \times k}$ is the identity matrix.

The iterates of the Bi-CG method applied to (1.1) are defined as follows.

Definition 2.26 (Iterates of Bi-CG)**.** Let $x_0 \coloneqq 0$. The kth iterate of the Bi-CG method applied to (1.1) is defined as

$$x_k \coloneqq V_k T_k^{-1}(\beta e_1), \tag{2.12}$$

where the matrices V_k and T_k are as in (2.8) with $v_1 = b/\beta$ and $\beta \coloneqq \|b\|$, and $e_1 = \begin{bmatrix} 1 & 0 & \dots & 0 \end{bmatrix}^T \in \mathbb{R}^k$.

Note, the standard implementation of Bi-CG does not require the storage of V_k or the solution to a tridiagonal linear system; see Algorithm 2. Instead, it is based on an implicitly formed LU decomposition of T_k, which often has better numerical behavior compared to the computation above [19].

Similar to the method GMRES, the iterates and corresponding residual vectors are elements of a Krylov subspace, formalized as follows.

Proposition 2.27 (Residual of iterates, Bi-CG)**.** *The kth iterate of Bi-CG defined in* (2.12) *is an element of the Krylov subspace* $\mathcal{K}_k(A, b)$*. Moreover, the corresponding residual vector* $r_k \coloneqq b - Ax_k$ *is included in* $\mathcal{K}_{k+1}(A, b)$*.*

The following characterizes the residual.

Definition 2.28 (Petrov-Galerkin condition, Bi-CG)**.** The property

$$b - Ax_k \perp \mathcal{K}_k(A^T, c)$$

is referred to as the *Petrov-Galerkin condition* for the Bi-CG method.

Applying the Bi-CG method to (1.1) requires a Lanczos biorthogonalization procedure. Here the Krylov subspace basis vectors v_{k+1} and w_{k+1} can be computed using a linear combination of only three vectors, shown as follows.

Corollary 2.29. *Let* $v_1, \ldots, v_k$ *and* $w_1, \ldots, w_k$ *denote the columns of* V_k *and* W_k*, respectively. The Krylov vector* v_{k+1} *in* (2.8) *fulfills a three-term recurrence between itself and the previous Krylov vectors since*

$$\left[\begin{array}{c} \\ A \\ \\ \end{array}\right]\left[\begin{array}{c|c|c} & & \\ v_1 & \cdots & v_k \\ & & \end{array}\right] = \left[\begin{array}{c|c|c} & & \\ v_1 & \cdots & v_{k+1} \\ & & \end{array}\right]\begin{bmatrix} \alpha_1 & \gamma_1 & & \\ \beta_1 & \ddots & \ddots & \\ & \ddots & \ddots & \gamma_{k-1} \\ & & \beta_{k-1} & \alpha_k \\ & & & \beta_k \end{bmatrix}$$

and $Av_k = \gamma_{k-1}v_{k-1} + \alpha_k v_k + \beta_k v_{k+1}$*. Similarly,*

$$\left[\begin{array}{c} \\ A^T \\ \\ \end{array}\right]\left[\begin{array}{c|c|c} & & \\ w_1 & \cdots & w_k \\ & & \end{array}\right] = \left[\begin{array}{c|c|c} & & \\ w_1 & \cdots & w_{k+1} \\ & & \end{array}\right]\begin{bmatrix} \alpha_1 & \beta_1 & & \\ \gamma_1 & \ddots & \ddots & \\ & \ddots & \ddots & \beta_{k-1} \\ & & \gamma_{k-1} & \alpha_k \\ & & & \gamma_k \end{bmatrix}$$

and $A^T w_k = \beta_{k-1}w_{k-1} + \alpha_k w_k + \gamma_k w_{k+1}$*.*

Breakdown can occur in the context of Bi-CG, and the situation is not always positive. We clarify this with the following remark.

Remark 2.30 (Breakdown, Bi-CG). If $v_i = 0$ or $w_i = 0$, an invariant subspace of A or A^T, respectively, has been found, and the algorithm terminates. If an invariant subspace with A has been found, then the approximate solution to (1.1) is exact. In the case that $w_i = 0$, the approximation to the dual system $A^T z = c$ is exact, though the approximation to (1.1) may not be. Either of these situations is referred to as a *lucky breakdown*. A so-called *serious breakdown* occurs if $w_i^T v_i = 0$, for $v_i \neq 0$, $w_i \neq 0$. No further progress can be made in this scenario, and the approximate solutions are not exact.

Chapter 3

Summary of Results

The main scientific contribution of this thesis is the collection of four papers included in Part II. Below are short summaries of each of the appended papers, where the author's contributions are highlighted.

3.1 Infinite GMRES

Paper A of this thesis is

E. Jarlebring and S. Correnty. Infinite GMRES for parameterized linear systems. *SIAM J. Matrix Anal. Appl.*, 43:1382–1405, 2022.

Summary: Paper A proposes a Krylov subspace method to solve parameterized linear systems of the form $A(\mu)x(\mu) = b$ for many different $\mu \in \mathbb{C}\backslash\{0\}$. In particular, the matrix $A(\mu) \in \mathbb{C}^{n\times n}$ is of the form

$$A(\mu) \coloneqq A_0 + A_1 f_1(\mu) + \ldots + A_{n_f} f_{n_f}(\mu), \quad \text{for} \quad n_f \ll n,$$

where the functions f_i are nonlinear in μ, the matrices A_i are large and sparse, and the vector $b \in \mathbb{C}^n$. The nonsingular matrix $A(\mu)$ is analytic in μ and, thus, can be described by an infinite Taylor series expansion. This work considers such an expansion centered at $\mu = 0$, i.e.,

$$A(\mu) = \sum_{i=0}^{\infty} \frac{1}{i!} A^{(i)}(0)\mu^i, \tag{3.1}$$

where we assume that many of the derivatives $A^{(i)}(0)$ are computationally available and $A(0)$ is nonsingular.

The approach of Paper A is summarized as follows. Let $A_N(\mu) \in \mathbb{C}^{n\times n}$ denote a truncated Taylor series expansion of $A(\mu)$, consisting of only the first $N+1$ terms in (3.1). A linearization of dimension $(N+1)n \times (N+1)n$ and the form

$$(\mu \boldsymbol{K}_N - \boldsymbol{I}_N)\boldsymbol{v}(\mu) = \boldsymbol{c} \tag{3.2}$$

is considered, where the matrix $\boldsymbol{K}_N$ is constructed using a linear transformation of the Taylor series coefficients of $A_N(\mu)$. Furthermore, the matrix $\boldsymbol{K}_N$ and the vector $\boldsymbol{c}$ are independent of the parameter μ, and $\boldsymbol{I}_N$ denotes the identity matrix of the appropriate dimension. Solutions to (3.2) are unique for each μ and equivalent in a certain sense to solutions of the corresponding linear systems $A_N x_N(\mu) = b$, where $A_N(\mu)$ is assumed nonsingular. Specifically, the solutions to the linearized system (3.2) are given by

$$\boldsymbol{v}(\mu) \coloneqq \begin{bmatrix} \frac{\mu^0}{0!} x_N(\mu)^T & \frac{\mu^1}{1!} x_N(\mu)^T & \dots & \frac{\mu^N}{N!} x_N(\mu)^T \end{bmatrix}^T \in \mathbb{C}^{(N+1)n},$$

proved in this work.

Applying GMRES to (3.2) requires the construction of an orthonormal basis for the Krylov subspace

$$\mathcal{K}_k(\boldsymbol{K}_N, \boldsymbol{c}) \coloneqq \operatorname{span}\left\{\boldsymbol{c}, \boldsymbol{K}_N\boldsymbol{c}, \dots, \boldsymbol{K}_N^{k-1}\boldsymbol{c}\right\}, \tag{3.3}$$

since Krylov subspaces are shift- and scaling-invariant. The associated basis matrix is built once and reused to compute approximations to the linearized systems for each μ of interest. Moreover, the basis matrix for (3.3) is constructed using the infinite Arnoldi algorithm [31]. We show that this allows us to execute the algorithm in a finite number of operations without truncation in the Taylor series expansion, i.e., $N \to \infty$. In this way, approximating the solution to (3.2) is equivalent to approximating (1.6) with $\ell = 1$ and complex coefficients. The method is referred to as *infinite GMRES*.

Paper A also provides a *tensor variant* of infinite GMRES, inspired by [30]. This version of the algorithm offers an effective way of handling the basis matrix, which is often large. Additionally, a *low-rank variant*, based on [61], is developed. The low-rank variant is more computationally efficient for problems in which the higher-order terms in the Taylor series expansion of $A(\mu)$ have a reduced rank.

A linear solve with $A(0)$ and a unique right-hand side vector is required at each iteration of the infinite Arnoldi algorithm in this context, and this is performed via a precomputed LU decomposition. The convergence analysis is based on the eigenvalues of $\boldsymbol{K}_N$ and the magnitude of the parameter μ. Numerical simulations involving a parameterized Helmholtz equation and a transfer function of a time-delay system show the suitability of these algorithms for solving large-scale problems.

Contribution: The ideas in this work were suggested by Elias Jarlebring. The convergence theory presented in Section 4 was worked out in close collaboration between the two authors. The author of this thesis generated all numerical results, and she wrote the manuscript with input from the coauthor.

3.2 Preconditioned infinite GMRES

Paper B of this thesis is

S. Correnty, E. Jarlebring, and K. M. Soodhalter. Preconditioned infinite GMRES for parameterized linear systems. *SIAM J. Sci. Comput.*, 2023. (Accepted for publication)

Summary: Paper B proposes a novel method to solve (1.6) with $\ell = 1$, where $A(\mu) \in \mathbb{R}^{n\times n}$ is nonsingular and analytic and $b \in \mathbb{R}^n$. We assume the matrix $A(\mu)$ is of the form

$$A(\mu) \coloneqq A_0 + A_1 f_1(\mu) + \ldots + A_{n_f} f_{n_f}(\mu), \quad \text{for} \quad n_f \ll n, \tag{3.4}$$

where f_i are nonlinear in $\mu \in \mathbb{R}$ and the matrices A_i are large and sparse. In particular, the proposed method returns approximate solutions to $x(\mu)$ for many different values of μ.

Let $A_N(\mu) \in \mathbb{R}^{n\times n}$ denote the Taylor series expansion of $A(\mu)$, truncated after the first $N+1$ terms. This work is based on a linearization of $A_N(\mu)$. Specifically, the linearization is of the form

$$(\boldsymbol{K}_N - \mu \boldsymbol{M}_N)\boldsymbol{v}(\mu) = \boldsymbol{c}, \tag{3.5}$$

has dimension $(N+1)n \times (N+1)n$, and includes the Taylor series coefficients of $A_N(\mu)$. Notably, the matrices $\boldsymbol{K}_N$ and $\boldsymbol{M}_N$, as well as the vector $\boldsymbol{c}$, are independent of the parameter μ. Solutions to the nonsingular system $A_N x_N(\mu) = b$ and the linearized system (3.5) are unique for each value of μ and equivalent in the sense that

$$\boldsymbol{v}(\mu) \coloneqq \begin{bmatrix} \mu^0 x_N(\mu)^T & \mu^1 x_N(\mu)^T & \ldots & \mu^N x_N(\mu)^T \end{bmatrix}^T \in \mathbb{R}^{(N+1)n},$$

which is shown in this work. Specifically, we assume that the derivatives $A^{(i)}(0)$ are computationally available.

The linear system (3.5) is solved in an inexact GMRES setting with right preconditioning. In particular, preconditioning with the matrix $\boldsymbol{K}_N$ leads to the shifted parameterized linear system

$$(\boldsymbol{I}_N - \mu \boldsymbol{M}_N \boldsymbol{K}_N^{-1})\hat{\boldsymbol{v}}(\mu) = \boldsymbol{c}, \tag{3.6}$$

where $\hat{\boldsymbol{v}}(\mu) = \boldsymbol{K}_N \boldsymbol{v}(\mu)$ and $\boldsymbol{I}_N$ is the identity matrix of the appropriate dimension. Exploiting the shift- and scaling-invariance properties of Krylov subspaces allows for the computation of solutions to (3.6) for a variety of μ simultaneously. More precisely, two basis matrices are constructed with the infinite Arnoldi algorithm [31], i.e., $N \to \infty$. Approximate solutions to (1.6) are computed from one basis matrix, and this is done in a finite number of operations.

At each infinite Arnoldi iteration, $\boldsymbol{K}_N^{-1}$ must be applied to a unique vector. This results in the need to perform many different linear solves with the nonsingular system matrix $A(0)$. Specifically, iterative methods are considered, where the accuracy is relaxed as the residual of the outer method decreases. Notably, in some

simulations, the error in the application of the preconditioner can be relatively high after just a few outer iterations, without degrading the overall convergence.

Theoretical analysis leads to a bound on the residual at each iteration. The competitiveness of the approach is illustrated through numerical simulations of solutions to a parameterized Helmholtz equation, as well as evaluations of a transfer function of a time-delay system. We refer to our algorithm as *inexact infinite GMRES*.

Contribution: Exploring the use of preconditioning to solve parameterized linear systems was suggested by Elias Jarlebring. The author of this thesis derived the main algorithm independently. She designed and produced the numerical experiments with support from the coauthors, and she wrote the manuscript with input from the coauthors.

3.3 Preconditioned Chebyshev BiCG

Paper C of this thesis is

S. Correnty, E. Jarlebring, and D. B. Szyld. Preconditioned Chebyshev BiCG method for parameterized linear systems, *Electron. Trans. Numer. Anal.*, 58:629–656, 2023.

Summary: Paper C proposes two variants of a three-term recurrence Krylov subspace method to approximate solutions to (1.6) with $\ell = 1$. In particular, solutions corresponding to many different values of the parameter $\mu \in [-a, a]$, for $a \in \mathbb{R}_+$, are returned simultaneously, and it is assumed that the matrix $A(\mu)$ is of the form described in (3.4). The approach used in Paper C is summarized as follows.

A linearization based on a Chebyshev approximation of $A(\mu)$ is constructed. In particular, we consider $P(\mu) \approx A(\mu)$ with

$$P(\mu) = P_0\tau_0(\mu) + \ldots + P_N\tau_N(\mu),$$

where $P_i \in \mathbb{R}^{n\times n}$ are coefficient matrices, $\tau_i(\mu)$ are the recursively defined Chebyshev polynomials on the interval $[-a, a]$, and N is the truncation parameter. It is assumed that N is chosen large enough that the error introduced by the Chebyshev interpolation is small. The linearization is of the form

$$(\boldsymbol{K}_N - \mu\boldsymbol{M}_N)\boldsymbol{v}(\mu) = \boldsymbol{c}$$

and has dimension $Nn \times Nn$. The matrices $\boldsymbol{K}_N$ and $\boldsymbol{M}_N$ contain the Chebyshev coefficients and $\boldsymbol{K}_N$, $\boldsymbol{M}_N$, and $\boldsymbol{c}$ are independent of the parameter μ.

This work makes use of *shift-and-invert preconditioning* with preconditioner $(\boldsymbol{K}_N - \sigma\boldsymbol{M}_N)$, for $\sigma \in (-a, a)$, selected because

$$(\boldsymbol{K}_N - \mu\boldsymbol{M}_N)(\boldsymbol{K}_N - \sigma\boldsymbol{M}_N)^{-1} \approx \boldsymbol{I}_N$$

when $\sigma \approx \mu$. Consequently, the corresponding right preconditioned system is given by

$$(\boldsymbol{K}_N - \mu\boldsymbol{M}_N)(\boldsymbol{K}_N - \sigma\boldsymbol{M}_N)^{-1}\hat{\boldsymbol{v}}(\mu) = \boldsymbol{c},$$

which can be expressed equivalently as

$$\begin{aligned} & (\boldsymbol{K}_N - \mu \boldsymbol{M}_N + \sigma \boldsymbol{M}_N - \sigma \boldsymbol{M}_N)(\boldsymbol{K}_N - \sigma \boldsymbol{M}_N)^{-1} \hat{\boldsymbol{v}}(\mu) = \boldsymbol{c} \\ \iff \quad & \left(\boldsymbol{I}_N + (-\mu + \sigma)\boldsymbol{M}_N (\boldsymbol{K}_N - \sigma \boldsymbol{M}_N)^{-1}\right) \hat{\boldsymbol{v}}(\mu) = \boldsymbol{c}. \end{aligned}$$

Here $\hat{\boldsymbol{v}}(\mu) = (\boldsymbol{K}_N - \sigma \boldsymbol{M}_N)\boldsymbol{v}(\mu)$ and $\boldsymbol{I}_N$ is the identity matrix of dimension $Nn \times Nn$. The short-term recurrence method known as *preconditioned Bi-CG* is applied to the linearized system above, and the shift- and scaling-invariance properties of Krylov subspaces are exploited. Concretely, one basis matrix for

$$\mathcal{K}_k(\boldsymbol{M}_N(\boldsymbol{K}_N - \sigma \boldsymbol{M}_N)^{-1}, \boldsymbol{c})$$

is generated and used to compute approximations to (1.6) for many different μ. This is done using the procedure described in [1, 21], where, notably, the basis matrix does not need to be stored. In this setting, the preconditioner, as well as its adjoint, are applied efficiently based on the prior works [3, 33]. We refer to this method as *preconditioned Chebyshev Bi-CG*.

A variant of preconditioned Chebyshev Bi-CG which considers an inexact application of the preconditioner $(\boldsymbol{K}_N - \sigma \boldsymbol{M}_N)$ is derived as well. Further, a bound on the residual of the inexact method at each iteration is shown. Numerical examples of the solution to a parameterized Helmholtz equation and the evaluation of a transfer function of a time-delay system show the competitiveness of the approach in solving large-scale systems for many values of the parameter.

Contribution: The idea of investigating short recurrence methods for solving parameterized linear systems was suggested by Elias Jarlebring. The author of this thesis derived the main algorithm independently, and she designed and produced all numerical experiments. The work was significantly improved from discussions with and feedback from Daniel B. Szyld. The author of this thesis wrote the manuscript with input from the coauthors.

3.4 Chebyshev HOPGD

Paper D of this thesis is

S. Correnty, M. A. Freitag, and K. M. Soodhalter. Chebyshev HOPGD with sparse grid sampling for parameterized linear systems, Preprint arXiv:2309.14178, 2024. (Submitted)

Summary: The method proposed in Paper D constructs a reduced order model to approximate parameterized linear systems of the form (1.6) with $\ell = 2$ and a nonlinear dependence on the parameters. In this way, the model can be evaluated to approximate the solution $x(\mu)$ for many different values of μ, where

$$\mu = \begin{bmatrix} \mu_1 & \mu_2 \end{bmatrix}^T \in \mathbb{R}^2 \quad \text{and} \quad \mu_1 \in [a_1, b_1], \quad \mu_2 \in [a_2, b_2].$$

The main idea of this work is summarized as follows.

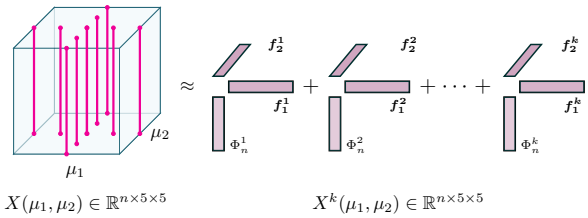

FIG. 3.1. Tensor matrix $X(\mu_1,\mu_2)$ with $N = 9$ snapshots (dots connected by vertical lines) and approximation $X^k(\mu_1,\mu_2)$.

The method requires precomputed solutions $x(\mu_1,\mu_2)$, referred to as *snapshots*, corresponding to a set of nodes $\boldsymbol{\mu}$, where

$$\boldsymbol{\mu} \coloneqq \{(\mu_{1_1},\mu_{2_1}),\ldots,(\mu_{1_N},\mu_{2_N})\}.$$

The nodes are chosen in a structured way on a sparse grid, and the snapshots $x(\mu_{1_i},\mu_{2_i})$ are generated using a modified version of the method proposed in Paper C. This process is referred to as *sampling*, and the snapshots are assembled in a sparse tensor matrix $X(\mu_1,\mu_2) \in \mathbb{R}^{n\times m_1\times m_2}$, for $m_1,m_2 \in \mathbb{N}$. Specifically, $X(\mu_{1_i},\mu_{2_i}) = x(\mu_{1_i},\mu_{2_i})$, for $i = 1,\ldots,N$, and zero otherwise.

The reduced order model is constructed via a decomposition of $X(\mu_1,\mu_2)$ as proposed in [37], and the resulting approximation $X^k(\mu_1,\mu_2)$ is a sum of k rank-one tensors; see Figure 3.1. An evaluation of this model, denoted by $x^k(\mu_1,\mu_2) \in \mathbb{R}^n$, is given by

$$x^k(\mu_1,\mu_2) \coloneqq \sum_{j=1}^{k} \Phi_n^j F_1^j(\mu_1) F_2^j(\mu_2) \tag{3.8a}$$

$$= x^{k-1}(\mu_1,\mu_2) + \Phi_n^k F_1^k(\mu_1) F_2^k(\mu_2), \tag{3.8b}$$

where $\Phi_n^j \in \mathbb{R}^n$, F_1^j and F_2^j are scalar functions, and $(\mu_1,\mu_2) \in \boldsymbol{\mu}$. The k terms in the summation above, referred to as *modes*, are obtained by an *alternating directions, greedy algorithm*, where the cost of each step scales linearly with the number of unknowns n. Modes are added to the approximation until

$$\frac{\left\|x(\mu_1,\mu_2) - x^k(\mu_1,\mu_2)\right\|}{\left\|x(\mu_1,\mu_2)\right\|} < \varepsilon \quad \text{for all} \quad (\mu_1,\mu_2) \in \boldsymbol{\mu} \tag{3.9}$$

and a predetermined tolerance $\varepsilon > 0$. Note, $x^k(\mu_1,\mu_2)$ is separable in the two parameters, and this model can only be evaluated for $(\mu_1,\mu_2) \in \boldsymbol{\mu}$.

In general, we cannot guarantee that (3.9) will be satisfied for a given set of snapshots, even when ε is chosen modestly. If the approximation fails to converge adequately on the tensor corresponding to

$$\{x(\mu_{1_1},\mu_{2_1}),\ldots,x(\mu_{1_N},\mu_{2_N})\},$$

the approach proposed in Paper D offers a way to generate a different set of snapshots in the same parameter space. Notably, this is done with little extra computation. As generating the snapshots is the dominating part of the algorithm, and it is not known a priori if a given tensor decomposition will succeed, this is an advantageous feature.

Once the model described by (3.8) has been constructed, the one variable functions F_1^j, F_2^j can be interpolated to approximate $x(\mu_1, \mu_2)$, for $(\mu_1, \mu_2) \notin \boldsymbol{\mu}$. Evaluating the resulting model requires, essentially, a linear combination of vectors of dimension n. Thus, the proposed method is one way to overcome the so-called *curse of dimensionality* associated with solving parameterized linear systems of the form (1.6).

Numerical examples arising from both a parameterized Helmholtz equation and a parameterized advection-diffusion equation are included. Simulations show the effectiveness of this approach in approximating the solution to parameterized PDEs and solving parameter estimation problems. Moreover, this method can be generalized for solving (1.6) with $\ell = 3, 4, \ldots$. We refer to our algorithm as *Chebyshev HOPGD with sparse grid sampling.*

Contribution: The author of this thesis came up with the idea for this work. She derived the algorithm and designed and produced the numerical experiments with support from Kirk M. Soodhalter. The simulations section was extended as a result of discussions with Melina A. Freitag. The author of this thesis wrote the manuscript with input from the coauthors.

References

[1] M. I. Ahmad, D. B. Szyld, and M. B. van Gijzen. Preconditioned multishift BiCG for $\mathcal{H}_2$-optimal model reduction. *SIAM J. Matrix Anal. Appl.*, 38(2):401–424, 2017.

[2] M. S. Alnaes, J. Blechta, J. Hake, A. Johansson, B. Kehlet, A. Logg, C. Richardson, J. Ring, M. E. Rognes, and G. N. Wells. The FEniCS project version 1.5. *Arch. Numer. Softw.*, 3, 2015.

[3] A. Amiraslani, R. M. Corless, and P. Lancaster. Linearization of matrix polynomials expressed in polynomial bases. *IMA J. Numer. Anal.*, 29(1):141–157, 2009.

[4] M. Baumann and M. B. van Gijzen. Nested Krylov methods for shifted linear systems. *SIAM J. Sci. Comput.*, 37, 2015.

[5] A. Bouras and V. Frayssé. A relaxation strategy for inexact matrix-vector products for Krylov methods. Technical Report TR/PA/00/15, CERFACS, Toulouse, France, 2000.

[6] A. Bouras and V. Frayssé. A relaxation strategy for the Arnoldi method in eigenproblems. Technical report TR/PA/00/17, CERFACS, Toulouse, France, 2000.

[7] A. Bouras, V. Fraysse, and L. Giraud. A relaxation strategy for inner-outer linear solvers in domain decomposition methods. Technical Report TR/PA/00/16, CERFACS, Toulouse, France, 2000.

[8] S. Boyd, V. Balakrishnan, and P. Kabamba. A bisection method for computing the H_∞ norm of a transfer matrix and related problems. *Math. Control. Signals, Syst.*, 2:207–219, 1989.

[9] H. Calandra, S. Gratton, R. Lago, X. Vasseur, and L. M. Carvalho. A modified block flexible GMRES method with deflation at each iteration for the solution of non-Hermitian linear systems with multiple right-hand sides. *SIAM J. Sci. Comput.*, 35:345–367, 2012.

[10] H. Calandra, S. Gratton, J. Langou, X. Pinel, and X. Vasseur. Flexible variants of block restarted GMRES methods with application to geophysics. *SIAM J. Sci. Comput.*, 34:714–736, 2012.

[11] W.-K. Chen. *The Electrical Engineering Handbook*. Elsevier, New York, 2005.

[12] B. A. Egan and J. R. Mahoney. Numerical modeling of advection and diffusion of urban area source pollutants. *J. Appl. Meteorol. (1962-1982)*, 11(2):312–322, 1972.

[13] H. C. Elman, O. G. Ernst, and D. P. O'Leary. A multigrid method enhanced by Krylov subspace iteration for discrete Helmholtz equations. *SIAM J. Sci. Comput.*, 23:1291–1315, 2001.

[14] Y. A. Erlangga. Advances in iterative methods and preconditioners for the Helmholtz equation. *Archives Comput. Methods Engin.*, 15:37–66, 2008.

[15] Y. A. Erlangga, C. Vuik, and C. W. Oosterlee. On a class of preconditioners for solving the Helmholtz equation. *Appl. Numer. Math.*, 50:409–425, 2004.

[16] O. G. Ernst and M. J. Gander. Why it is difficult to solve Helmholtz problems with classical iterative methods. In I. Graham, T. Hou, O. Lakkis, and R. Scheichl, editors, *Numerical Analysis of Multiscale Problems.*, volume 83 of *Lecture Notes in Computational Science and Engineering*, pages 325–363. Berlin, Heidelberg, Springer, 2011.

[17] R. Fletcher. Conjugate gradient methods for indefinite systems. In G. A. Watson, editor, *Numerical Analysis*, volume 506 of *Lecture Notes in Math.*, pages 73–89. Springer Berlin, 1976.

[18] R. W. Freund. Solution of shifted linear systems by quasi-minimal residual iterations. In L. Reichel, A. Ruttan, and R. S. Varga, editors, *Numerical Linear Algebra: Proceedings of the Conference in Numerical Linear Algebra and Scientific Computation*, pages 101–122. Berlin, New York, de Gruyter, 1993.

[19] R. W. Freund and N. M. Nachtigal. An implementation of the QMR method based on coupled two-term recurrences. *SIAM J. Sci. Comput.*, 15(2):313–337, 1994.

[20] E. Fridman. *Introduction to Time-Delay Systems: Analysis and Control*. Springer International Publishing Switzerland, 2014.

[21] A. Frommer. BiCGStab(ℓ) for families of shifted linear systems. *Computing*, 70:87–109, 2003.

[22] A. Frommer and U. Glässner. Restarted GMRES for shifted linear systems. *SIAM J. Sci. Comput.*, 19(1):15–26, 1998.

[23] A. Frommer and P. Maass. Fast CG-based methods for Tikhonov–Phillips regularization. *SIAM J. Sci. Comput.*, 20(5):1831–1850, 1999.

[24] S. Gazzola and M. Sabaté Landman. Flexible GMRES for total variation regularization. *BIT*, 59:721–746, 2019.

[25] A. Ghali, A. Neville, and T. Brown. *Structural Analysis: A Unified Classical and Matrix Approach*. CRC Press, Boca Raton, 7th edition, 2017.

[26] G. H. Golub and C. F. Van Loan. *Matrix Computations*. The Johns Hopkins University Press, 3rd edition, 1996.

[27] S. Gugercin, A. C. Antoulas, and C. Beattie. $\mathcal{H}_2$ model reduction for large-scale linear dynamical systems. *SIAM J. Matrix Anal. Appl.*, 30:609–638, 2008.

[28] N. Guglielmi, M. Gürbüzbalaban, and M. L. Overton. Fast approximation of the H_∞ norm via optimization over spectral value sets. *SIAM J. Matrix Anal. Appl.*, 34(2):709–737, 2013.

[29] S. Gumussoy and W. Michiels. Computing $\mathcal{H}_\infty$ norms of time-delay systems. In *Proceedings of the 48th IEEE Conference on Decision and Control*, pages 263–268, 2009.

[30] E. Jarlebring, G. Mele, and O. Runborg. The waveguide eigenvalue problem and the tensor infinite Arnoldi method. *SIAM J. Sci. Comput.*, 39(3), 2017.

[31] E. Jarlebring, W. Michiels, and K. Meerbergen. A linear eigenvalue algorithm for the nonlinear eigenvalue problem. *Numer. Math.*, 122(1):169–195, 2012.

[32] M. E. Kilmer and D. P. O'Leary. Choosing regularization parameters in iterative methods for ill-posed problems. *SIAM J. Matrix Anal. Appl.*, 22(4):1204–1221, 2001.

[33] D. Kressner and J. E. Roman. Memory-efficient Arnoldi algorithms for linearizations of matrix polynomials in Chebyshev basis. *Numer. Linear Algebra Appl.*, 21:569–588, 2014.

[34] P. Kürschner and M. A. Freitag. Inexact methods for the low rank solution to large scale Lyapunov equations. *BIT*, 60:1221–1259, 2020.

[35] C. Lanczos. Solution of linear equations by minimized iterations. *J. Res. Natl. Bur. Stand.*, 49:33–53, 1952.

[36] R. J. LeVeque. *Finite Volume Methods for Hyperbolic Problems.* Cambridge Texts in Applied Mathematics. Cambridge University Press, 2002.

[37] Y. Lu, N. Blal, and A. Gravouil. Adaptive sparse grid based HOPGD: Toward a nonintrusive strategy for constructing a space-time welding *computational vademecum. Internat. J. Numer. Methods Engrg.*, 114:1438–1461, 2018.

[38] Y. Lu, N. Blal, and A. Gravouil. Datadriven HOPGD based computational vademecum for welding parameter identification. *Comput. Mech.*, 64:47–62, 2019.

[39] L. Marin, L. Elliott, P. J. Heggs, D. B. Ingham, D. Lesnic, and X. Wen. Parameter identification in Helmholtz-type equations with a variable coefficient using a regularized DRBEM. *Inverse Probl. Sci. Eng.*, 14(8):837–858, 2006.

[40] W. Michiels and S.-I. Niculescu. *Stability, Control, and Computation for Time-delay Systems: An Eigenvalue-based Approach.* SIAM, Philadelphia, 2nd edition, 2014.

[41] D. Modesto, S. Zlotnik, and A. Huerta. Proper generalized decomposition for parameterized Helmholtz problems in heterogeneous and unbounded domains: Application to harbor agitation. *Comput. Methods Appl. Mech. Engrg.*, 295:127–149, 2015.

[42] E. Nadal, F. Chinesta, P. Díez, F. J. Fuenmayor, and F. D. Denia. Real time parameter identification and solution reconstruction from experimental data using the proper generalized decomposition. *Comput. Methods Appl. Mech. Engrg.*, 296:113–128, 2015.

[43] I. J. Nagrath and M. Gopal. *Control Systems Engineering.* New Academic Science, 6th edition, 2021.

[44] S. Okiy, C. C. Nwobi-Okoye, and A. C. Igboanugo. Transfer function modelling: A literature survey. *Res. J. Appl. Sci. Eng. Technol.*, 11:1265–1279, 2015.

[45] C. C. Paige and M. A. Saunders. Solution of sparse indefinite systems of linear equations. *SIAM J. Numer. Anal.*, 12(4):617–629, 1975.

[46] J. Peeters and W. Michiels. Computing the $\mathcal{H}_2$-norm of large-scale time-delay systems. *IFAC Proc. Vol., 46 (2013), pp. 114–119.*

[47] Y. Saad. A flexible inner-outer preconditioned GMRES algorithm. *SIAM J. Sci. Comput.*, 14(2):461–469, 1993.

[48] Y. Saad. *Iterative Methods for Sparse Linear Systems.* SIAM, Philadelphia, 2nd edition, 2003.

[49] Y. Saad and D. Osei-Kuffuor. Preconditioning Helmholtz linear systems. *Appl. Numer. Math.*, 60:420–431, 2010.

[50] Y. Saad and M. H. Schultz. GMRES: A generalized minimal residual algorithm for solving nonsymmetric linear systems. *SIAM J. Sci. Stat. Comput.*, 7:856–869, 1986.

[51] A. K. Saibaba, T. Bakhos, and P. K. Kitanidis. A flexible Krylov solver for shifted systems with application to oscillatory hydraulic tomography. *SIAM J. Sci. Comput.*, 35:3001–3023, 2013.

[52] T. Sauer. *Numerical Analysis*. Pearson, 2nd edition, 2012.

[53] M. Signorini, S. Zlotnik, and P. Díez. Proper Generalized Decomposition solution of the parameterized Helmholtz problem: application to inverse geophysical problems. *Internat. J. Numer. Methods Engrg.*, 109:8:1085–1102, 2017.

[54] V. Simoncini and D. Szyld. Theory of inexact Krylov subspace methods and applications to scientific computing. *SIAM J. Sci. Comput.*, 25, May 2002.

[55] K. M. Soodhalter. Block Krylov subspace recycling for shifted systems with unrelated right-hand sides. *SIAM J. Sci. Comput.*, 38(1):A304–A324, 2016.

[56] K. M. Soodhalter, D. B. Szyld, and F. Xue. Krylov subspace recycling for sequences of shifted linear systems. *Appl. Numer. Math.*, 81:105–118, 2014.

[57] T. J. Sullivan. *Introduction to Uncertainty Quantification*, volume 63. Springer International Publishing Switzerland, 2015.

[58] A. Tarantola. *Inverse Problem Theory and Methods for Model Parameter Estimation*. SIAM, Philadelphia, 2005.

[59] L. N. Trefethen and D. Bau. *Numerical Linear Algebra*. SIAM, 1997.

[60] S. Ulfah, S. A. Awalludin, and Wahidin. Advection-diffusion model for the simulation of air pollution distribution from a point source emission. *J. Phys.: Conf. Ser.*, 948(1):012067, 2018.

[61] R. Van Beeumen, E. Jarlebring, and W. Michiels. A rank-exploiting infinite Arnoldi algorithm for nonlinear eigenvalue problems. *Numer. Linear Algebra Appl.*, 23(4):607–628, 2016.

[62] J. van den Eshof, G. L. G. Sleijpen, and M. B. van Gijzen. Iterative linear system solvers with approximate matrix-vector products. In A. Bori~i, A. Frommer, J. Bálint, A. Kennedy, and B. Pendleton, editors, *QCD and Numerical Analysis III*, volume 47 of *Lecture Notes in Computational Science and Engineering*, pages 133–142. Springer Berlin Heidelberg, 2005.

[63] L. Zheng and X. Zhang. *Modeling and Analysis of Modern Fluid Problems*. Mathematics in Science and Engineering. Elsevier Science, 2017.

[64] K. Zhou, J. C. Doyle, and K. Glover. *Robust and optimal control*. Prentice Hall, 1995.

[65] S. Zlotnik, P. Díez, D. Modesto, and A. Huerta. Proper generalized decomposition of a geometrically parameterized heat problem with geophysical applications. *Internat. J. Numer. Methods Engrg.*, 103(10):737–758, 2015.

[66] K. J. Åström and R. M. Murray. *Feedback Systems: An Introduction for Scientists and Engineers*. Princeton University Press, 2008.

They might be bigger,
but we're faster and never scared

Taylor Swift

Part II: Scientific Papers