Convex Relaxation Methods in Optimisation and Robust Control

Emad Hamadi

Doctor of Philosophy

City, University of London

School of Mathematics, Computer Science and Engineering

October 2020

Contents

List of Figures

List of Tables

Acknowledgements

First and foremost, I want to express my most sincere gratitude and recognition to my supervisor Professor George Halikias for his continuous support, and invaluable guidance and encouragement during my PhD study. His immense knowledge and plentiful experience have encouraged me in all the time of my academic research and daily life.

Thank you George!

I would also like to pay my appreciation and my respects to my second supervisor Professor Nicholas Karcanias who regrettably passed away in April of 2020. He was a dedicated professor in the engineering department and truly a great and beloved person. I will always be thankful to him for his consistent support during my study in City University.

I gratefully acknowledge the funding received towards my PhD from the City, University of London PhD Scholarship and I thank them for providing various learning, enriching, networking, and professional and personal development experiences.

Additionally, I would like to to offer my special thanks to Dr Efstathios Milonidis for all the help I got from him. I would also like to express my gratitude to the School of Mathematics, Computer Science and Engineering and all its member's staff, especially the Postgraduate Research Office members for all their support and for this great opportunity I have been given to conduct my PhD research.

Finally, words cannot express my gratitude to Vafa Hamadi for her love, patience and emotional support. Without her tremendous understanding and encouragement, it would be impossible for me to complete my study.

I cannot forget to thank my family, friends and fellow students for all their unconditional support in last few years.

I dedicate this thesis to my beautiful daughters, **Rihanna** and **Hanna**.

Abstract

This thesis deals with the analysis of the real μ problem as a powerful tool for measuring the stability margins of a system subject to parametric uncertainty. Several algorithms of varying complexity are proposed for calculating upper bounds of the structured singular value of a matrix M subject to real parametric uncertainty. Our approach is based on the projection of the uncertainty set in the most critical direction. This is implicit in the set of optimal (minimum-norm) unstructured singularising perturbations and is defined by the pair of singular vectors corresponding to the largest singular value of M. Two relaxations are considered to simplify the problem. A randomised algorithm is proposed which relies on the partial enumeration of the Zonotope's vertices for high dimensional problems when the complete enumeration may not be practical. Applying this bound to our problem produces a probabilistic lower bound on the structured distance to singularity. The main results of the thesis are extended to the distance to singularity problems with "correlated" or nonlinear descriptions of uncertainty. A similar randomised algorithm is proposed for breaching the gap between the Quadratic Integer Programming (QIP) and its convex relaxation which is closely related to the structured singular value problem. It is shown that the duality gap can be reduced, provided a Reduced Rank Quadratic Integer Problem (RRQIP) can be solved. Alternatively, a sequence of increasingly tighter bounds can be obtained by solving a sequence of QIP's of progressively increasing complexity (and rank). Here, we present a randomisation algorithm for breaching the duality gap when the full enumeration is not computationally feasible. The Greatest Common Divisor(GCD) problem to calculate the nearest common root of a polynomial set under perturbations in their coefficients is also considered in this thesis. We propose a relaxation technique directly to the Sylvester structure before converting to the diagonal matrix which is the standard setting for the structured singular value estimation. This could give an upper bound tighter than the largest singular value without solving the equivalent μ problem which is potentially large-scale. Several numerical examples illustrate the main results of this work.

Chapter 1

Introduction

In the field of control systems, an accurate model is required to analyse and design a feedback controller. The quality and accuracy of a model, however, depends on how closely it describes the real system. Nevertheless, adding unnecessary characteristics can result in over-complicated models which could dramatically increase complexity and the volume of computations, limiting the sampling rate of the controller. On the other hand, provided model uncertainties are small and provided they are not amplified by feedback they could result in small deviations between the actual and the simulated responses of a system. However, the presence of uncertainty is always one of the significant difficulties in control systems analysis and design. A widely used method to control a system is feedback control which is an interconnection of the plant and the controller in a feedback loop. Therefore, a designed controller must stabilise the closed-loop system despite the presence of the uncertainty. At the same time, the controller should fulfil the design objectives in an efficient way. A good controller should be able to minimise the effect of uncertainties. In robust control, the largest size of uncertainty that the system can tolerate is critical. This is normally referred to as the stability radius of the design. The larger the uncertainty that can be tolerated, the more robust is the control system. As an example, PID (proportional integral derivative) controllers are simple but effective method in robust control systems. PID controllers use a control loop feedback mechanism to control process variables and are the most popular controller in the industry.

In general, deriving a dynamic model of a system is not straightforward since in the real world systems are non-linear and complex. In most cases, however, a relatively simple linear time-invariant model which contains modelling error and uncertainties can adequately describe the behaviour of a system for control purposes at least near equilibrium. A control

system is then said to be robust if its response is insensitive to model perturbations. Figure 1.1 shows a general feedback framework which can address a general model of uncertainties. In this diagram, M is a transfer function representing the nominal system, and Δ is a transfer function representing the uncertainty.

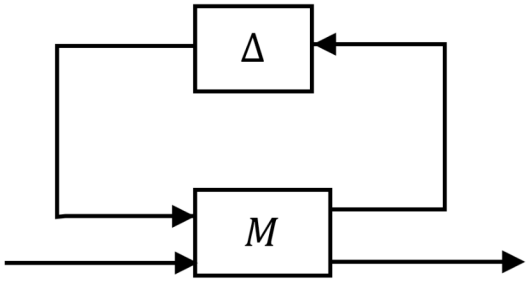

Figure 1.1: A general control system framework

The structure we impose on Δ affects the stability radius. Ignoring any structure on Δ is likely to result in a conservative design. On the other hand, adding a more refined structure of uncertainty makes the design of the controller difficult and costly. In robust control, we usually distinguish between Robust analysis and Robust Synthesis. The analysis problem consists of finding the maximum perturbation that the feedback system can tolerate before losing stability. The synthesis problem is to design a controller that minimises the effects of the perturbations or keeps them under an acceptable level. Maximising the robustness of a system in the presence of unstructured uncertainties has been addressed in the literature(see [3, 4]). Typically, unstructured uncertainties result in small stability radius compared to structured uncertainty. Stability analysis when Δ is a structured matrix (rather than a full matrix) can be addressed via Structured Singular Value (μ) techniques. In which, the perturbations are restricted to a certain class of structured matrices, and the stability radius is defined as the smallest norm in the perturbation, Δ, that destabilises the system in Figure 1.1 (denoted as $\mu_{\Delta}(M)^{-1}$). Unfortunately, in general, $\mu_{\Delta}(M)$ cannot be computed in polynomial time. In other words, this problem is numerically intractable and can be shown to be NP-hard in terms of polynomial complexity[5]. One possible way to overcome this issue is to relax the problem and find reliable convex upper and lower bounds on $\mu_{\Delta}(M)$, which is sufficient for many practical applications. At present, the standard procedure used to minimise the upper bound of $\mu_{\Delta}(M)$ is the so-called D-iteration[6] resulting in a convex upper bound on μ_{Δ}. Nonetheless, in theory, the gap between $\mu_{\Delta}(M)$ and its convex upper bound can be arbitrarily large [7]. This gives rise to another research problem, i.e. to iden-

tify conditions under which the gap between structured singular value and its upper bound can be breached [8, 9]. Although, at the present time, there is no general method available for this purpose, some algorithms have been developed for specific uncertainty classes. The fundamental disadvantage of these algorithms is their high computational cost which may dramatically increase as the size of the system grows. In this thesis the approximate GCD problem is formulated and studied as an application of the μ problem. The study of GCD of a set of polynomials has been widely studied in recent years. Due to extensive applications in Control Theory like algebraic control methods, distance to controllability or observability, determinantal assignment problems, Robust Control, the stability of dynamic systems subject to structured perturbations, Linear Systems, Numerical Analysis and other Engineering fields. Another topic which is also the focus of this research is Quadratic Integer Programming(QIP). QIP is a classical problem which is widely used in various real-world applications, including control and communications [10]. QIP is also an NP-hard problem which shares interesting similarities with μ problem. It is shown in this work that this motivates us in this research to transfer intuition and solution techniques from one problem domain to the other.

In the following sections, some of the above-mentioned topics and terms are explained in more details.

1.1 Complexity in optimisation

Computational complexity theory has been a central area of theoretical computer science since its early development in the mid-1960s[11].The subsequent rapid development in the next three decades has not only established it as a rich exciting theory, but also shown strong influence on many other related areas in computer science, mathematics, and operations research. Computational complexity refers to the the amount of communications required to solve a type of problem by systematic application of an algorithm. Since the size of the particular input to a problem will affect the amount of computation, measures of complexity needs to be carried out to identify the effectiveness of the designed algorithms. Hence categorising the algorithms or problems are a good introduction to identify the computational complexity.

When discussing complexity, “time” and “space” is often used as the basis for comparison. The “time” here is not the actual time, but rather an indication of the number of computation required to achieve a solution. Due to the fact that the computer hardware is

constantly improving, the “time” required to solve a given problem for the same algorithm will be consequently decreasing. However, the complexity of a problem, which is in fact the number of steps required to terminate the algorithm, will be always identical regardless of the processing power improvement. That is the reason why the Complexity is quantified independent of hardware. It , nevertheless, depends on the methods and algorithms used to approach the problem. A meaningful way of quantifying complexity is to introduce a formal model of computation $\mathcal{M}$ which specified decision algorithms can be implemented and to identify various parameters of this model as formal representations of computational resources. The two most basic computational resources which are studied in complexity theory are running time and tape (or memory) space [12].

1.1.1 Complexity Classes

The general notion of a complexity class for a model of computation $\mathcal{M}$ decides a language $\mathcal{X}$ is defined by TIME $t(n)$ and SPACE $s(n)$ as functions of type $(N) \rightarrow (N)$.These are related to a *Turning Machine*,$\mathcal{T}$ model which is theoretically a set of languages, $T \in \mathcal{T}$ in time $t(n)$ and space $s(n)$. i.e.,

$$TIME(t(n)) = \{X \subseteq \{0,1\} : \exists T \in \mathcal{T} \forall n(t(n) \leq t(n)) \text{ and } T \text{ decides } X\};$$
$$SPACE(s(n)) = \{X \subseteq \{0,1\} : \exists T \in \mathcal{T} \forall n(s(n) \leq s(n)) \text{ and } T \text{ decides } X\};$$

In general the complexity classes are defined as below,

P Class

The set of problems P contains problems with whose solution-time upper bound scales as a polynomial function of the input size. P problem is referred to the complexity class in which the problems can be efficiently solved by a deterministic algorithm (known as solvable or tractable problems). Other time classes are quadratic time, linear time or exponential time. Examples of polynomial time algorithms: the "quicksort" algorithm and basic arithmetic operations such as addition, subtraction, multiplication, division, and comparison.

NP Class

Non-deterministic Polynomial-time, NP complexity, contains problems whose solutions can be verified within a polynomially scaled upper bound. That does not mean they can be absolutely solved in polynomial time, but given a potential solution, its verity can be confirmed

or denied in polynomial time.

NP-Hard class

A problem is NP-hard (nondeterministic polynomial time-hard) if it it can be obtained from a NP-complete problem that is polynomial time Turing-reducible. It can be said to be "at least as hard as the hardest problems in NP."

Examples of NP-hard problems: Subset sum problem, traveling salesman problem, halting problem.

NP-Complete class

A problem is NP-complete (nondeterministic polynomial time-complete) if it belongs to both NP as well as NP-hard. NP-complete problems can be obtained by transforming every other problem in NP in polynomial time. NP-complete problems are of note because there is an apparent correlation between the quick verification of solutions and quick solving of the problems

Common approaches to solving NP-complete problems are heuristic algorithms and approximation algorithms.

Examples of NP-Complete problems: graph isomorphism problem, Boolean satisfiability problem, knapsack problem, Hamiltonian path problem, travelling salesman problem, subgraph isomorphism problem, and more. Figure 1.2 shows Euler diagram for P, NP, NP-complete and NP-hard set of problems.

1.2 Convex and Non-convex Optimisation problems

Having described the complexity class of a problem in previous section, we understood that NP- Hard problems, in general, cannot be solved in polynomial time. A common practice to approach these problems, therefore, is to relax them to approximate *convex* bounds which are solvable. An optimisation problem is said to be convex if all the constraints are convex functions and the feasible region (intersection of convex constraints) is a convex region. In this case, there is only one globally optimal solution which can be calculated efficiently. In contrast, any optimisation problem with a non-convex constraint or objective function is called a non-convex problem. Non-convex problems can have multiple feasible regions and therefore, multiple local optimal points in each region. Determining that a non-convex

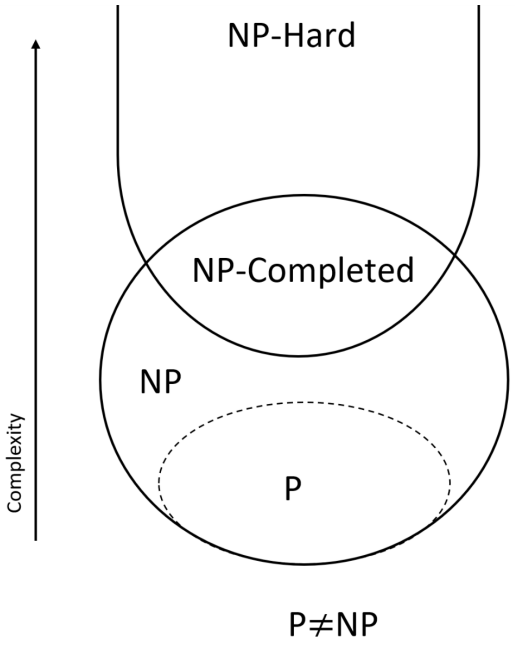

Figure 1.2: P, NP, NP-complete and NP-hard set of problems

problem is infeasible, the objective function is unbounded or finding the optimum across all feasible regions can take exponential time.

Convex optimisation problems are considerably more general than linear programming problems but have the desirable characteristics of linear programming problems. They can be solved quickly and reliably, even for large-scale problems with thousands of variables.

1.3 Structured Singular Value

In previous sections we described NP-Hard class as an important set of problems which computational algorithms fail to solve them. We have also described convex sets as solvable optimisation problem to which NP-hard problem could be approximated. Here, we briefly define the structured singular value as a powerful tool for measuring stability margins of a control system subjected to parametric uncertainty. This problem is NP-hard and normal solutions to approach this problem is to relax it to a convex bound. All models used for control design are uncertain to some extent. If uncertainty is small, it can only induce minor deviations in the system behaviour. System complexity, however, enhances the uncertainties on some occasions in ways that make the entire system behaviour more difficult to predict. Even more importantly, uncertainty can jeopardise the stability margin of the system. Accordingly, the presence of uncertainty in real systems is one of the most challenging issues of the control design methodology. One of the main objectives in control

design is to reduce the effect of uncertainty on the stability of the system. Conversely, we would like to determine the size of the largest perturbation that a system can tolerate before it loses stability. The effect of uncertainty and modelling errors have been studied by many researchers (e.g. see [13]).

A control system is considered to be robust if its performance and stability are insensitive to modelling errors. In fact, robust synthesis methods aim to design the controller so that robust performance in the presence of uncertainties is achievable. In order to analyse the effect of uncertainty on the behaviour of a system, we need to provide a model of perturbations in the system dynamics. The complexity of the error model, however, can make the controller unnecessarily complicated and costly. Conversely, unstructured error model can lead to conservative feedback designs. In control theory, it is typical to classify uncertainty as structured or unstructured. In unstructured uncertainty, perturbations are allowed to be any stable $n \times n$ transfer function, and then the analysis is carried out by considering the worst-case scenario. Thus the problem reduces to finding the unstructured perturbation ,Δ, with the smallest $\mathcal{H}_\infty$ norm (say r) that destabilizes the feedback loop [14] and hence, any Δ which $\|\Delta\|_\infty < r$ will not destabilize the system. For a system with transfer function M this value of r is called the *Stability radius* and is denoted by $\gamma_\Delta(M)$. Thus if the norm of Δ is sufficiently small, then the closed-loop system will be stable. The small gain theorem can be used to identify how large $\|\Delta\|_\infty$ can be before it destabilises the closed-loop system.

As discussed earlier, unstructured uncertainty is, in general, a special class of perturbations and may result in a conservative design. In most cases, only a few entries of the transfer function or state-space model contain perturbations and thus, only a few entries in Δ are subjected to uncertainty. In that case, the remaining entries in Δ take zero value. We particularly consider block-diagonal structures in this work which can be described by repeated full blocks and scalar blocks:

$$\mathbf{\Delta} = \{diag\left[\delta_1 I_{r_1}, ..., \delta_S I_{r_s}, \Delta_1, ..., \Delta_F\right]\} \tag{1.1}$$

where $\delta_i \in \mathbb{C}$ (or $\mathbb{R}$), $\Delta_j \in \mathcal{C}^{m_j \in m_j}$ and I_{r_i} is the $r_i \times r_i$ identity matrix.

The problem of maximising the robust stability radius of systems subject to structured uncertainty has been widely investigated in the literature, and has led to the development of μ-analysis and synthesis techniques [15, 16, 3, 14, 17, 18, 9]. The Structured Singular value (μ) is a generalisation of the singular value of a matrix. Similarly to QIP problem, the general solution to this problem is also NP-hard [19, 5] which means that for any given

algorithm to compute μ, there will be "worst-case" problems for which the algorithm is unable to find the optimal solution in polynomial time. In practice, the problem can be relaxed, and upper bounds are often used, which may give sufficiently robust stability and performance margins for certain simple uncertainty structures. This class of problems, however, is too small for most practical applications [9]. In theory, the gap between μ solution and its convex upper bound can be arbitrarily large [7]. Therefore, breaching the gap between μ and its upper bound can be highly beneficial. Authors of [9, 20, 21, 8] conducted valuable research which provides realistic estimates of the stability radius and fast algorithms. The research area is still active for the derivation of tighter bounds and improved algorithms for its calculation.

In this research several algorithms of varying computational complexity is proposed for calculating upper bounds of the structured singular value (equivalently lower bounds of the structured distance to singularity) of a matrix M subject to real parametric uncertainty. Our approach is based on the projection of the uncertainty set in the most critical direction. This is implicit in the set of optimal (minimum-norm) unstructured singularising perturbations and is defined by the pair of singular vectors corresponding to the largest singular value of M. Two relaxations are considered to simplify the problem. The first leads to the maximisation of a convex quadratic function under box constraints. The second relaxation, which is computationally more demanding but produces tighter bounds, results in a geometric problem in the complex plane involving the intersection of a Zonotope with a parametric family of Apollonius circles. For problems of low or medium complexity, this can be solved via the complete enumeration of the Zonotope's vertices, e.g. by applying the computationally efficient "reverse enumeration" algorithm. For high dimensional problems, this approach may not be practical and a randomised algorithm is proposed as an alternative which relies on the partial enumeration of the Zonotope's vertices. The convex hull of these partially-enumerated vertices approximates the Zonotope within a probabilistic Housedorff distance bound. Applying this bound to our problem produces a probabilistic lower bound on the structured distance to singularity (equivalently a probabilistic upper bound on the structured singular value). The main results of this part of thesis are extended to the distance to singularity problems with "correlated" or nonlinear descriptions of uncertainty. The results are illustrated with several numerical examples.

1.4 Quadratic Integer Programming

Quadratic Integer Programming (QIP) problem is a classical optimisation problem with many applications. For example Model Predictive Control (MPC), an advanced method of control with application in many engineering fields like process and power systems is based on optimising a performance index over a finite time-horizon on the current time slot. One of the main ideas behind MPC is its ability to formulate and solve a QIP problem where the constrains on control parameters can be formulated as linear inequalities. Apart from its unique applications, QIP share common characteristics with μ-problem which motivates their simultaneous study in this thesis. In terms of complexity, they both are NP-hard problems. Both have convex upper bounds and have an equivalent dual problem. And finally, in both problem there are conditions to breach the duality gap. The algorithms of solving these problems are quite similar. Thus, developing an algorithm in one problem could be adapted in the other problem. There are some other common characteristics that also motivate studying both problems in parallel.

As mentioned above, QIP, however, like μ-problem, cannot be solved in polynomial time, i.e. it is an NP-hard problem. A way to overcome this is to convert QIP to an equivalent convex problem or relax it to a convex problem and optimise ideally tight bounds. Many solvable cases of QIP have been presented in the literature. For instance, if matrix, Q, in the QIP problem $(\max x^T Q x : x \in \{-1, 1\})$ is of rank one the solution can be found by inspection [22]. This is also true when Q has non-positive off-diagonal entries [23], or when the system graph is associated with the Max-Cut problem [24]. All these special cases have their own complexity and have been investigated by numerous researchers. For example, imposing restriction on equivalent system's graph $G(Q)$ results in solvable classes of the problem. Such special cases include the pseudo-Boolean program [25], cases where the graph $G(Q)$ is series-parallel [26] or when the graph $G(Q)$ is a binary tree [26]. Various other special solvable cases of the QIP problem have been defined in the literature (see Chapter 10 of [27] for a survey of the methods that solve some special cases associated with the QIP problem). In all of these cases, the solution of the QIP problem is obtained by methods appropriate to each special case. The general solution to the problem, however, remains open. The low-Rank Quadratic Optimisation problem is another well-known special class of QIP problems in the literature which can be solved via special algorithms of polynomial complexity [22], [28], [29] and [30]. This class of problems is still an active area of research with many applications, e.g. in the area of renewable energy. Another application is MPC

optimal Control. In this case, the optimisation problem needs to be solved strictly within the sampling interval of the control system. Therefore, since the QIP problem is solved on-line, it is critically important that the real-time algorithm can efficiently track the process and does not impose delays to the control loop. Although powerful hardware in recent computers enables a faster sampling rate, control systems also become more complex with hundreds of control loops. For example in large scale networked systems typical objective matrices may have thousands of entries. Note that often the exact solution to the problem is not required and tight bounds are acceptable.

One of the contributions of this research is to introduce a fast track method for finding a convex upper bound with reduced computational cost. A condition to improve the accuracy of the solution of QIP problem is also derived by breaching the gap between optimal solution and its convex relaxation bound. This also helps to establish a more time-efficient algorithm by reducing the computation at each iteration. In the thesis, a randomised algorithm is proposed for calculating an upper bound of the QIP problem. First, a convex relaxation of the problem was defined, the solution of which produces an upper bound to the original problem. The duality gap of the problem, i.e. the distance between the convex upper bound and the optimal solution, can be reduced which is equivalent to the full enumeration of the vertexes of an specific convex polygon called Zonotope. Although this can be achieved by a polynomial-time algorithm, the computation may still be intractable for problems of high-dimensionality. In the present research, a randomisation algorithm is presented for breaching the duality gap when the full enumeration is computationally not feasible. It is shown that even with incomplete enumeration improved probabilistic bounds may be obtained. The bound may be further improved by combining the proposed randomised approach with the solution of a sequence of deterministic QIP problems of increasing rank. This approach is useful in practice since it can exploit fully the available computational resources to obtain the tightest bound possible with a pre-specified probability.

1.5 Greatest Common Divisor(GCD)

Finally, as a practical application of μ problem, we can name the Greatest Common Divisor (GCD) problem. One of the common methods to approach GCD problem is converting it to an equivalent μ problem in order to enable implementing various available algorithms in μ field. This will, however, increase the dimensions and consequently the computational expenses of the problem dramatically. Therefore, it can effectively illustrates the impor-

tance of low cost algorithms to calculate the solution. The study of the GCD of a set of polynomials has several applications in Control Theory (e.g. algebraic control methods, determinantal assignment problems, distance to controllability or observability), Robust Control (stability of dynamic systems subject to structured perturbations), Linear Systems, Numerical Analysis and other Engineering fields and hence has received considerable interest in recent years. Computation of the GCD is a non-generic problem. Hence, the concept of "approximately coprimeness" can be defined as a distance from the nearest common divisor in an appropriate sense. Similar definition of "almost zeros" was first introduced in [31]. This definition has been reformulated to the notion of "approximate GCD" which is, in fact, based on the relaxation of the conditions defining the exact GCD, see [32], [33], [34], [35], [36], [37] and references therein. [38] proposed one of the most recent methods for calculating the distance of a set of co-prime polynomials to the set of polynomials sharing a common root. The technique was based on singular values to define and solve approximate GCD problems by converting the corresponding Sylvester matrix in GCD problem to a diagonal matrix compatible to μ problem.

A disadvantage of this method is that the size of equivalent objective matrix in μ problem will increase considerably. This will significantly increase the computational expenses of the corresponding μ problem. One way of overcoming this issue, which has been noted in most of references mentioned earlier, is to ignore the structure of the problem and use the singular value as an approximate measure of singularity of the corresponding Sylvester resultant matrix(which implies loss of coprimeness). This, however, gives a loose bound which is far from the optimal solution in many cases.

In this research, we propose using a relaxation approach to perturbations with Sylvester matrix structure. This gives an upper bound tighter than the largest singular value while avoiding calculating the structured singular value of a high-dimensional problem. The advantages of this approach are illustrated via a numerical example.

1.6 Thesis Aims and Objectives

The research starts with working on QIP problem and studying the typical algorithms used to find feasible solutions. It is then expanded into the area of convex relaxations and explores the limitation of current techniques in these areas. The second main topic in this research involves the efficient computation of the Structured Singular Value(SSV) or μ problem and its application in robust control.

The efficient computation of the the μ problem can be extended to efficient algorithms in the area of robust control synthesis using H_∞ methods, for which ad-hoc algorithms are currently applied (D-K iteration). In this work, it is also aimed to exploit the similarity between these two problems in order to develop an efficient algorithm for a large class of convex relaxation problems.

Another topic in this research is the problem of calculating the nearest common root of a set of polynomials that are under perturbations in their coefficients which is considered as an application of the μ problem. It is aimed to generalise a convex approach to find an a cost efficient upper bound on GCD which is tighter than the largest singular value.

The main contribution of this research can be summarised as follows:

- Propose a cost-efficient method for breaching the gap between the QIP problem and its convex relaxation.
- Propose a probabilistic method of calculating a convex upper bound on μ based on the intersection of the associated convex polygon (so-called Zonotope) with a parametric family of Apollonius circles.
- Extend the main results of the proposed method to the distance to singularity problems with "correlated" or nonlinear descriptions of uncertainty.

There are also some minor contributions which can be itemise as follows:

- Improved the convex bound by combining the proposed randomised approach with the solution of a sequence of deterministic QIP problems of increasing rank
- Propose an ellipse of minimum area which contains all vertices of the Zonotope to obtain a lower bound on the distance to singularity; and propose a method to calculate an enclosing ellipse directly from the Zonotope's generating matrix without enumerating the vertices
- Establish an easy-computable bound which gives a tighter bound but the computation is immediate.
- Propose a relaxation methods to find a cost efficient upper bound on GCD which is tighter than the largest singular value.

The outcome of this work is two Manuscripts:

- "Structured singular value of matrices with real parametric uncertainty: Deterministic and Probabilistic algorithms" submitted to *International Journal Of Control* Journal and is under the review.

The other one is titled

- "A randomised algorithm for breaching the gap between the quadratic integer programming problem and its semidefinite relaxation"ready for submission.

The structure of this report is as Follows:

Chapter 2

This chapter starts with an introduction to QIP problems. A brief background and literature review is given in this chapter. The concept of convex relaxation methods, SDR, RRQIP, Zonotope is also reviewed. A novel randomised algorithm is proposed. This is the first contribution in this work which is itemised as "Propose a cost-efficient method for breaching the gap between the QIP problem and its convex relaxation" above. The main idea of this algorithm is discussed in detail. The advantages and disadvantages of this method are also evaluated. In addition, a probabilistic sufficient condition under which the gap between QIP problem and its convex upper bound can be breached is derived. In this chapter the behaviour of the Zonotope enumeration algorithm is also studied to estimate the probability that the optimal solution occurs is vertices with specific flatness characteristics.A minor contribution of "Improved the convex bound by combining the proposed randomised approach with the solution of a sequence of deterministic QIP problems of increasing rank" will be introduced in this chapter. A numerical experiment finally included to show the applicability of the proposed method.

Chapter 3

In this chapter, some preliminaries needed for chapter 4 is introduced . μ analysis is briefly described. The special conditions where a 1×1 block of the matrix takes various structures is also discussed in this chapter. This section also gives several definitions related to the structured singular value and the structured distance to singularity for real, parametric, diagonal uncertainty structures. This is mainly based on two main references, [6] and [9]. Detailed proofs of the main materials used in subsequent chapters are included, where appropriate. Some of the proofs have been transferred to Appendix to improve the readability of the Chapter.

Chapter 4

In this chapter two deterministic and probabilistic approaches are introduced to calculate upper bound on μ. First the μ problem is defined for real set of uncertainty and then the distance to a singularity is formulated as a quadratic optimisation problem. A deterministic upper bound on μ via solving an optimisation problem is then derived. Using the randomised method introduced in chapter 2, a probabilistic upper bound on the structured singular value of a matrix is derived. Numerical experiment are included in each section to show the validity of the proposed methods.

The preliminary results of this Chapter apply to the general case for which the multiplicity of the larger singular value of M is an arbitrary integer. The (generic) case $m = 1$ is analysed in detail. This allows for a concrete set of algorithms to be formulated and solved based on an interesting geometric interpretation of the problem. This is the second contribution of this thesis listed as "Propose a probabilistic method of calculating a convex upper bound on μ based on the intersection of the associated convex polygon (so-called Zonotope) with a parametric family of Apollonius circles". Extensions of the proposed methodology to correlated parameters and nonlinear uncertainty models are also presented here, which create the third main contribution of this work titled as "Extend the main results of the proposed method to the distance to singularity problems with "correlated" or nonlinear descriptions of uncertainty". Several algorithmic implementation aspects are discussed in this chapter. Some of the minor contributions like ellipse of minimum area or easy-computable bound, are formed here. A randomisation algorithm for the partial enumeration of the Zonotope's vertices is also describe along with an estimation of the Hausdorff distance between the Zonotope and the convex hull of the polytope corresponding, in order to the reduced set of vertices obtained from the randomisation algorithm. This can be used to obtain a probabilistic lowed bound on the structured distance to singularity. A numerical example is presented at the end of this Chapter to illustrate the performance of all the algorithms described in this section.

Chapter 5 In this Chapter, Greatest Common Divisor (GCD) of two polynomials as an application of the μ problem is considered. The problem is first defined and its application in control system will be reviewed. The link between GCD and μ problem is then defined and it is shown how the problem could be formulated to the solution of an equivalent μ problem. A relaxation method will be then introduced to find a cost efficient upper bound on GCD

which is tighter than the commonly used largest singular value. This is what we itemised it as a minor contribution titled "Propose a relaxation methods to find a cost efficient upper bound on GCD which is tighter than the largest singular value" above. The performance of proposed method will be illustrated by a number of numerical example.

Chapter 6

This is concluding chapter which summarises the results of the research and the main conclusions in the context of the aims and objectives of the study outlined in Chapter 1. Finally, the possible extensions of the work in future research is outlined.

1.7 Notation

The notation of the thesis is standard and is summarised here for convenience. $\mathbb{R}$ denotes the set of real numbers. For integer n, $\mathbb{R}^n$ denotes the space of n-dimensional (column) vectors whose entries are in $\mathbb{R}$, $\mathcal{X}$ denotes the set of n-dimensional vectors whose entries are either 1 or -1, i.e. $\mathcal{X} = \{-1, 1\}^n$. For integers n and m, $\mathbb{R}^{n\times m}$ denotes the space of all $n \times m$ matrices whose entries are in $\mathbb{R}$. For $A \in \mathbb{R}^{n\times m}$, A' denotes the transpose of A, trace(A) is the sum of the diagonal elements of A, $\mathcal{N}(A)$ denotes the null space of A and $\mathcal{R}(A)$ is the range of A. If $A \in \mathbb{R}^n$ is symmetric ($A = A'$), $\underline{\lambda}(A)$ denotes the smallest eigenvalue of A and we write $A \succeq 0$ if $\underline{\lambda}(A) \geq 0$ and $A \succ 0$ if $\underline{\lambda}(A) > 0$. The m-dimensional identity matrix is denoted by I_m and the $m \times n$ null matrix is denoted by $0_{m,n}$ (0_m if $m = n$) with the subscripts omitted if they can be inferred from the context. The null set is denoted by $\emptyset$. If $\mathcal{F}$ is a space, $\dim\ \mathcal{F}$ denotes the dimension of $\mathcal{F}$. For a square matrix A, diag(A) denotes A with all its off-diagonal elements set to zero. $A \in \mathbb{R}^{m\times n}$ is called orthogonal if $AA' = I_m$ ($m \leq n$) or $A'A = I_n$ ($n \leq m$). For $A = A' \in \mathbb{R}^n$ the spectral decomposition is the identity: $A = U\Lambda U'$ where $U \in \mathbb{R}^{n\times n}$ is orthogonal and $\Lambda \in \mathbb{R}^{n\times n}$ is a diagonal matrix of the eigenvalues of A. If $\mathcal{P}$ is a polytope then vert$(\mathcal{P})$ is the set of its vertices. If $\mathcal{F}$ is a finite set then $|\mathcal{F}|$ is its cardinality, i.e. the number if its elements. Finally, if $\mathcal{P} \subseteq \mathbb{R}^n$ then conv$(\mathcal{P})$ is its convex hull. Other notation is introduced as needed.

Chapter 2

A Randomised Algorithm for Enumerating Zonotope Vertices in QIP

2.1 Introduction

In the previous chapter we introduced QIP as an NP-hard problem with many practical applications. We have also mentioned computing good bounds as a practical approach. A "good" solution is considered to be one which is computationally efficient. Some QIP research guarantees a small gap to the optimal value of the problem. Other research evaluates methods for obtaining the exact solution in special cases. However, obtaining good approximate solutions for QIP is not easy in general.

Many solvable cases of QIP have been identified in the literature. For instance, when the QIP problem is of rank one [22], when the quadratic matrix defining the problem has non-positive off-diagonal elements [23], or when the system graph is associated with the Max-Cut problem [24]. Algorithms for solving all these special cases have their own time and space complexity. The Low-Rank Quadratic Optimisation problem is another well-known special class of QIP problems which can be solved via special algorithms of polynomial complexity [29], [39] and [30]. This class of problems is still an active area of research with many applications. Often in practice, the exact solution to the problem is not required, and tight bounds may be acceptable. In this Chapter, several methods for breaching the gap between the optimal solution of the QIP problem and its Semidefifinite Relaxation (SDR) is reviewed. It is known [30] that the gap may be reduced by solving an auxiliary

QIP problem of low rank, a task that can be achieved via the complete enumeration of the vertices of a a convex polytope(so-called Zonotope) corresponding to the problem. A full enumeration algorithm can be used to solve the reduced rank QIP, which, in turn, can be used to reduce the duality gap of the solution of the standard QIP problem from the solution of its SD relaxation. In this chapter, the concepts of SDR, Zonotope, duality gap and low-rank problems is also defined.

These results will be further extended using the randomised version of this algorithm. The advantage of the proposed method is its ability to terminate before all vertices of the Zonotope associated with the problem are identified. This reduces the computation load, especially for large scale problems. Our approach is based on the results in [40] which derives an estimate of the Hausdorff distance between the Zonotope and its approximation, defined as the convex hull of the partially enumerated vertices of the Zonotope (obtained by random sampling). This algorithm is used to solve the Reduced Rank QIP(RRQIP) in a probabilistic setting. In other words, a probabilistic condition is derived for which the gap between the QIP solution and its convex upper bound is breachable. The algorithm is utilised to evaluate those vertices at which the maximiser of QIP problem occurs with a high probability. In fact, the algorithm relies on the fact that "sharp" vertices of the Zonotope which are further away from the origin are selected with higher probability compared to "flat" vertices by the randomised algorithm. A similar approach has been used in Chapter 4 to derive a probabilistic bound of the (real) structured distance to the singularity of a matrix obtained from the intersection of a Zonotope with a parametric family of Apollonius circles. For the problem considered in this chapter, an estimate of the Zonotope's vertex furthest away from the origin can be used to derive a probabilistic upper bound on the solution of the QIP which is tighter than the SDR bound. Besides, we generalise the results of [41] which derive a decreasing sequence of upper bounds to the solution of the QIP problem by solving a sequence of auxiliary QIP problems of increasing rank (and hence also complexity) which are equivalent to the enumeration of the vertices of a sequence of Zonotopes defined in progressively higher dimensional spaces. By identifying the optimal transition between deterministic (full) and randomised (partial) vertex enumeration, it is possible to obtain the tightest bound compatible with the available computational resources.

A disadvantage of this method, nevertheless, is its probabilistic nature. i.e. the solution provided by the algorithm is always valid only with a certain probability. However, this probability could be increased at the cost of increasing the number of iterations. In fact, the trade-off between the accuracy and speed of the algorithm can be adjusted by the user.

Another disadvantage of this method is the possibility that the reduced polytope omits the optimiser and hence the upper bound breaches the solution of QIP. To overcome this feature, a method is introduced to extend the gap between the QIP problem and its upper bound. In fact, the gap is extended by a minimal amount which is just sufficient to cover all those vertices which potentially have not been selected by the algorithm. Nevertheless, if the maximiser is obtained at a sharp vertex, the upper bound proposed by this method may be conservative. We study the probability of recovering particular sharp vertices and relate it to the angle of the vertex. A MATLAB program code is developed to confirm the validity of this method.

The structure on this chapter is as follows: First the QIP problem is formally defined along with its Semi Definite (SD) relaxation and a review of the main results of [41] and [30] for breaching the duality gap in section 2.2. The results of section 2.2 then will be formulated in a geometric setting by showing that the improved upper bounds of section 2.2 can be obtained by enumerating the vertices of a Zonotope corresponding to a reduced-rank QIP. A suitable method for this task is the "reverse enumeration" algorithm [42]. Then probabilistic bounds on the solution of the QIP problem are proposed using a randomised algorithm for partial enumeration of the vertices of a Zonotope. The results of the proposed methods in this section are illustrated via numerical examples.

It has to be mentioned here that although the main contribution in this thesis was in the μ-problem field, it is decided to present the QIP chapter prior to the μ chapter. This is because it is believed that the description of the Zonotope concept and Randomisation algorithm in QIP field is more understandable due to a better visualisation. Hence the reader will be familiarised with the Zonotope concept before reading the μ chapter.

2.2 Unconstrained $\{-1, 1\}$ Quadratic Integer Programming

The unconstrained Quadratic Integer Programming (QIP) problem in $\{-1, 1\}$ variables is defined as:

$$(\text{QIP}) \quad \gamma := \max_{\forall x \in \mathcal{X}} x' Q x \tag{2.1}$$

where $\mathcal{X} = \{-1, 1\}^n$ and $Q = Q' \in \mathbb{R}^{n \times n}$ is given. Computationally, this is a classical NP hard problem [43].

The form of the QIP defined in (2.1) can be transformed to the zero-one QIP problem via the linear transformation $y = (x + e)/2$ where $e \in \mathbb{R}^n$ is the vector of ones and zeroes to problems involving a linear term using a simple homogenisation procedure [44]. It is

well known that other optimisation problems, such as the Quadratic Assignment Problem [45], the Maximum-Cut Problem [46], can also be transformed to the standard QIP problem considered here.

Since (2.1) is an NP-hard problem, it cannot be solved in polynomial time, and therefore it is normally relaxed to an upper bound. Tighter upper bounds will result in a better approximation. Among many methods, Semi Definite Relaxation (SDR) of the QIP is well-known in literature, [47], [48], [49], [44], [50] and [22].

In [50], the author considers three different relaxations of the QIP problem and shows that all three yield the same bound. The following relaxation is widely used in literature. Let $D \in \mathbb{R}^{n\times n}$ be any diagonal matrix for which $D - Q$ is positive semidefinite and let x be any binary vector where $x \in \mathcal{X}$. Thus the following identities can be verified:

$$D - Q \geq 0 \Leftrightarrow D \geq Q \Rightarrow x'Dx \geq x'Qx \tag{2.2}$$

For all x, 2.2 can be rewritten as

$$x'Qx = -\big(\text{trace}(D) - x'Dx\big) - x'(D-Q)x + \text{trace}(D) \tag{2.3}$$

Since $x \in \mathcal{X}$ we can write

$$x'Dx = \text{trace}(D) \tag{2.4}$$

Substituting this into (2.2) gives:

$$x'Qx \leq \text{trace}(D) \tag{2.5}$$

Inequality (2.5) is true for all diagonal D such that $D - Q \geq 0$ and for all $x \in \{-1, 1\}$. Therefore the tightest bound is obtained as:

$$\max_{x\in\{-1,1\}} x'Qx \leq \min_{\substack{D-Q\geq 0 \\ D \text{ is diagonal}}} \text{trace}(D) \tag{2.6}$$

The above relaxation reduces QIP to the problem of finding an upper bound by solving the problem in (2.7) which is within the class of convex optimization problems:

$$(SDR) \qquad \overline{\gamma} = \min_{\substack{D-Q\geq 0 \\ \text{D is diagonal}}} \text{trace}(D) \tag{2.7}$$

There are a number of efficient algorithms to solve SDR problem. For example, the interior-point algorithm introduced by Yang in[51]. In [24], Goeman et al. introduced a randomised approximation algorithm to solve the maximum cut problem as a special case of QIP. In [52], the author found an approximate solution for Lagrangian relaxation SDR. Zhang et al. also introduced a more accurate approximation for the same problem [53].

It can be shown that the optimal solution of the SDR problem is unique [30]. The following result summarises three necessary conditions for the optimal solution of the SDR problem [30]. These are used in the sequel to investigate the relaxation gap of $\bar{\gamma} - \gamma$.

Lemma 2.2.1. *[30] Let D be the unique minimiser for the SDR problem so that trace$(D) = \bar{\gamma}$ and $D - Q \succeq 0$. Then:*

1. $\dim \mathcal{N}(D - Q) \geq 1$, *(equivalently, $\underline{\lambda}(D - Q) = 0$) so that*

$$D - Q = \begin{bmatrix} V & V_+ \end{bmatrix} \begin{bmatrix} 0_r & 0 \\ 0 & \Lambda_+ \end{bmatrix} \begin{bmatrix} V' \\ V'_+ \end{bmatrix} \tag{2.8}$$

   for some orthogonal $\begin{bmatrix} V & V_+ \end{bmatrix} \in \mathbb{R}^{n \times n}$, $r = \dim \mathcal{N}(D-Q) \geq 1$ and $\Lambda_+ \succ 0$.

2. *There does not exist diagonal Z such that trace$(Z) = 0$ and $V'ZV \preceq 0$.*

3. *Every row of V has (Euclidean) norm at least $1/\sqrt{n}$. In particular, none of the rows of V is zero.*

Necessary and sufficient conditions for the absence of a gap between the QIP problem and the SDR problem, i.e. for $\gamma = \bar{\gamma}$ are established as follows:

Lemma 2.2.2. *[30] Let D be the (unique) minimiser for the SDR problem and let $D - Q$ have a spectral decomposition (2.8). Then the following statements are equivalent: (i) $\gamma = \bar{\gamma}$, (ii) $\mathcal{N}(D - Q) \cap \mathcal{X} \neq \emptyset$, and (iii) $Vy \in \mathcal{X}$ for some $y \in \mathbb{R}^r$.*

Lemma 2.2.2 suggests a simple test given in the following Lemma for the absence of the relaxation gap.

Lemma 2.2.3. *[30] Let all variables be as in Lemma 2.2.2. By rearranging the rows of V if necessary; and let $V = \begin{bmatrix} V'_{11} & V'_{21} \end{bmatrix}'$ with $V_{11} \in \mathbb{R}^{r \times r}$ non-singular. Then $\gamma = \bar{\gamma}$ if and only if $V_{21}V_{11}^{-1}z \in \{-1, 1\}^{n-r}$ for some $z \in \{-1, 1\}^r$. Equivalently a necessary and sufficient condition for $\gamma = \bar{\gamma}$ is that $\max_{x \in \mathcal{X}} \; x'VV'x = n$.*

Note that the second condition given in Corollary 2.2.3 implies that the duality gap is zero if $r = 1$. This follows from Lemma 2.2.1 part 3 which shows that in this case each element of (vector) V has modulus $\frac{1}{\sqrt{n}}$ and hence the maximum of $\|V'x\|^2$ over all $x \in \mathcal{X}$ must be equal to n. An improved bound of the QIP problem that breaches the convex SDR upper bound is established by the following Lemma:

Lemma 2.2.4. *[30] Let all variables be as in Lemma 2.2.2 and suppose that γ_r solves the reduced-rank quadratic integer problem (RRQIP)*

$$(\text{RRQIP}) \qquad \gamma_r := \frac{1}{n} \max_{x \in \mathcal{X}} x'VV'x \tag{2.9}$$

Then

$$\gamma \leq \bar{\gamma} - n(1 - \gamma_r)\underline{\lambda}(\Lambda_+) \leq \bar{\gamma} \tag{2.10}$$

Proof. Since $\left[\begin{array}{cc} V & V_+ \end{array}\right] \in \mathbb{R}^{n \times n}$ is orthogonal, then

$$x'x = x' \left[\begin{array}{cc} V & V_+ \end{array}\right] \left[\begin{array}{c} V' \\ V'_+ \end{array}\right] x = x'VV'x + x'V_+V'_+x = n$$

for any $x \in \mathcal{X}$. Therefore:

$$x'V_+V'_+x = n - x'VV'x \tag{2.11}$$

Due to the fact that trace$(D) = x'Dx$ for any $x \in \mathcal{X}$ and also considering the definition of SDR in (2.7),we can reformulate (2.1) to

$$\gamma := \max_{x \in \{-1,1\}} x'Qx = \max_{x \in \{-1,1\}} x'(\frac{\overline{\gamma}}{n}I - (D - Q))x$$

which is equivalent to

$$\gamma := \max_{x \in \{-1,1\}} x'(\frac{\overline{\gamma}}{n}I - \left[\begin{array}{cc} V & V_+ \end{array}\right] \left[\begin{array}{cc} 0_r & 0 \\ 0 & \Lambda_+ \end{array}\right] \left[\begin{array}{c} V \\ V_+ \end{array}\right])x$$

hence,

$$\gamma := \max_{x \in \{-1,1\}} x' \left[\begin{array}{cc} V & V_+ \end{array}\right] (\frac{\overline{\gamma}}{n}I - \left[\begin{array}{cc} 0_r & 0 \\ 0 & \Lambda_+ \end{array}\right]) \left[\begin{array}{c} V \\ V_+ \end{array}\right] x$$

which gives

$$\gamma := \max_{x\in\{-1,1\}} x' \left[\begin{array}{cc} V & V_+ \end{array}\right] \left(\left[\begin{array}{cc} \frac{\overline{\gamma}}{n} I & 0 \\ 0 & \frac{\overline{\gamma}}{n} - \Lambda_+ \end{array}\right] \right) \left[\begin{array}{c} V \\ V_+ \end{array}\right] x$$

Since $x'Qx \leq \overline{\lambda}(Q)x'x$, we have

$$\begin{aligned} \gamma &\leq \max_{x\in\{-1,1\}} \frac{\overline{\gamma}}{n} x'VV'x + \left(\frac{\overline{\gamma}}{n} - \underline{\lambda}(\Lambda_+)\right) x'V_+V_+'x \\ &= n\left(\frac{\overline{\gamma}}{n} - \underline{\lambda}(\Lambda_+)\right) + \underline{\lambda}(\Lambda_+) \max_{x\in\{-1,1\}} x'VV'x \\ &= \overline{\gamma} - n\underline{\lambda}(\Lambda_+)(1-\gamma_r) \end{aligned} \tag{2.12}$$

□

Note that since $V'V = I_r$, then $\gamma_r \leq 1$. Note also that $\gamma_r = 1$ implies from Lemma 2.2.3 that $\gamma = \bar{\gamma}$ and the duality gap is zero. In general, the relaxation upper bound $\bar{\gamma}$ can be "breached" provided a simple solution to the RRQIP in (2.9) can be found for which γ_r is less than one (or at least if an upper bound less than one can be found on γ_r). Although problem (2.9) is similar to the original QIP, the difference is that the matrix VV' in the cost function in (2.9) has a potentially low rank. It is shown in the next section that low-rank quadratic integer programming problems are significantly easier to solve than full-rank problems. Note also that the lower the value of γ_r and the higher the value of $\underline{\lambda}(\Lambda_+)$, the tighter the bound on γ.

Reference [41] defines a sequence of QIP problems of increasing rank resulting in a decreasing sequence of upper bounds on γ. The initial bound can be selected either as the largest eigenvalue of Q (which is a well upper bound on γ) or the SDR bound via a slight reformulation of the problem. The solution can be obtained either iteratively or via a "one-shot" approach. Again, the solution of each QIP problem resulting from either approaches can be obtained by the enumeration of the vertices of a Zonotope. For the iterative algorithm these are defined in progressively higher dimensional spaces and are of increasing complexity. In this form, the algorithm strikes a balance between accuracy (small gap) and computational complexity and can be used to obtain the tightest bound compatible with the available computational resources.

Lemma 2.2.5. *[41] Let $A = A' \in \mathbb{R}^n$ be given and assume that A has at least two distinct eigenvalues (otherwise computational of γ is straightforward). Let $\mathcal{X}_1 = \{-\frac{1}{\sqrt{n}}, \frac{1}{\sqrt{n}}\}^n$,*

$f(x) = x'Qx$ *and* $\gamma_1 = \max\{f(x) : x \in \mathcal{X}_1\}$. *Let* A *have* r *distinct eigenvalues* $\lambda_1 > \cdots > \lambda_r$ *with multiplicities* $m_1, \ldots, m_r$, *respectively, where* $2 \leq r \leq n$ *and* $\sum_{i=1}^{r} m_i = n$, *so that* A *has an ordered Schur form*

$$A = U\Lambda U' = \begin{bmatrix} U_1 & \cdots & U_r \end{bmatrix} \begin{bmatrix} \lambda_1 I_{m_1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_r I_{m_r} \end{bmatrix} \begin{bmatrix} U_1' \\ \vdots \\ U_r' \end{bmatrix} \tag{2.13}$$

where $U = \begin{bmatrix} U_1 & \cdots & U_r \end{bmatrix} \in \mathbb{R}^{n \times n}$ *is orthogonal,* $\Lambda = diag(\lambda_1 I_{m_1}, \ldots, \lambda_r I_{m_r})$ *is the diagonal matrix of the eigenvalues of* A, *with* $U_i \in \mathbb{R}^{n \times m_i}$, $i = 1, \ldots, r$. *Set* $f_0(x) = 0$, $\phi_0 = 0$, $\gamma_0 = 0$, $\lambda_{r+1} = 0$ *and for* $i = 1, \ldots, r$, *define*

$$f_i(x) = \sum_{j=1}^{i} x' U_j \frac{\lambda_j - \lambda_{i+1}}{\gamma_i - \lambda_{i+1}} U_j' x, \quad \phi_i = \max_{x \in \mathcal{X}_n} f_i(x) \tag{2.14}$$

and

$$\gamma_i = \phi_{i-1}\gamma_{i-1} + (1 - \phi_{i-1})\lambda_i. \tag{2.15}$$

Then for $i = 1, \ldots, r$: $0 \leq \phi_i \leq 1$, $\lambda_i \leq \gamma_i$, $\gamma \leq \gamma_i$ *and* $\gamma = \gamma_i$ *if and only if* $\phi_i = 1$ *or equivalently if and only if there exists* $x \in \mathcal{X}$ *such that* $\phi_{i-1} = f_{i-1}(x)$ *and* $\sum_{j=1}^{i} x' U_j U_j' x = 1$. *Hence,* $\gamma = \gamma_r \leq \cdots \leq \gamma_2 \leq \gamma_1 = \lambda_1$. *Finally, suppose that* $\phi_j < 1$ *for* $j = 1, \ldots, i-1$ *and* $\phi_i = 1$. *Then:* $\gamma = \gamma_r = \cdots = \gamma_i < \gamma_{i-1} < \cdots < \gamma_1$.

Proof. See [41]. □

The "one-shot" version of Lemma 2.2.5 is stated next:

Lemma 2.2.6. *Let* $A = A' \in \mathbb{R}^{n \times n}$ *be as given in the previous Lemma. Choose the largest* i *such that* $1 \leq i < r$ *and such that the maximisation in (2.16) below is feasible. Define*

$$\begin{aligned} V_1 &= \begin{bmatrix} U_1 & \cdots & U_i \end{bmatrix}, & \Lambda_1 &= diag(\lambda_1 I_{m_1}, \cdots, \lambda_i I_{m_i}) \\ V_2 &= U_{i+1}, & \Lambda_2 &= \lambda_{i+1} I_{m_{i+1}} \\ V_3 &= \begin{bmatrix} U_{i+2} & \cdots & U_r \end{bmatrix}, & \Lambda_3 &= diag(\lambda_{i+2} I_{m_{i+2}}, \cdots, \lambda_r I_{m_r}) \end{aligned}$$

so that A has a Schur form given by

$$A = U\Lambda U' = \begin{bmatrix} V_1 & V_2 & V_3 \end{bmatrix} \begin{bmatrix} \Lambda_1 & 0 & 0 \\ 0 & \Lambda_2 & 0 \\ 0 & 0 & \Lambda_3 \end{bmatrix} \begin{bmatrix} V_1' \\ V_2' \\ V_3' \end{bmatrix}$$

Let $S := \Lambda_1 - \lambda_{i+1} I_m$ so that $S > 0$ is diagonal and define $V = V_1 S^{1/2} \in \mathbb{R}^{n \times m}$ where $m = \sum_{j=1}^{i} m_j$. Let $\bar{\phi}_i$ be the maximum of the reduced rank quadratic integer program

$$\bar{\phi}_i = \max_{x \in \mathcal{X}} x' V V' x \tag{2.16}$$

Then

$$\gamma := \max_{x \in \mathcal{X}} x' A x \leq \lambda_{i+1} + \bar{\phi}_i =: \gamma_i. \tag{2.17}$$

Furthermore, $\gamma = \gamma_i$ if and only if there exists a maximiser $x_i \in \mathcal{X}$ for the RRQIP defined in equation (2.16) such that $V_3' x_i = 0$.

Note that (2.16) has the same form as the original QIP problem but now $VV' \geq 0$ and rank(VV')= r (potentially r can be much lower than n)

If $\gamma_r = 1$, the gap between QIP and SDR program is zero. However if $\gamma_r < 1$, then the distance between the optimal solution and its convex upper bound (duality gap) can be reduced.

2.3 Fixed rank QIP and Zonotopes

Allemand and Fukuda et al in 2001 [22] have shown that if matrix Q in the QIP problem is of fixed rank, positive semidefinite and all its eigenvalues are known, the QIP can be solved in a polynomial time. In the literature, this is known as the Fixed Rank Convex(FRC) QIP problem[1] or Reduced Rank QIP(RRQIP).

It is shown in [22] that the solution of RRQIP can be reduced to the enumeration of the extreme points of a Zonotope. This is the image of a hypercube under the linear transformation and is a convex polytope of a special type. The claim can be verified by the following consideration:

$$\max_{x \in \mathcal{X}} x' V V' x = \max_{x \in \mathcal{X}} \|V' x\|^2 = \max_{x \in \mathcal{X}} \sum_{i=1}^{d} \langle x, \nu_i \rangle^2 \tag{2.18}$$

where $\nu_i \in \mathbb{R}^n$ is the ith row of $V' \in \mathbb{R}^{n \times m}$ and $\langle x, \nu_i \rangle$ is the inner product of x and ν_i.

Note that the second equality follows from the convexity of $\sum_{i=1}^{d} \langle x, \nu_i \rangle^2$. Now consider the following linear mapping

$$\mathcal{Z} = \{V'x : x \in [-1, 1]^n\} \tag{2.19}$$

The image of the hypercube $[-1, 1]^n$ under map (2.19) defines the Zonotope $(\mathcal{Z})$. Every extreme point of x will be mapped to an extreme point of $\mathcal{Z}$ under transformation of $Z \rightarrow V'x$. Substituting (2.19) into (2.18), shows that RRQIP becomes:

$$\max_{x \in \{-1,1\}^n} x'VV'x = \max_{z \in Z} \sum_{i=1}^{m} z_i^2 \tag{2.20}$$

where $z_i \in \mathbb{R}^d$ is the ith row of $z \in \mathbb{R}^{n \times m}$. Let S be the set of extreme points of $\mathcal{Z}$. Then (2.20) will reduce to

$$\textbf{(RRQIP)} \qquad \max_{x \in \{-1,1\}^n} x'VV'x = \max_{z \in \mathcal{Z}} \sum_{i=1}^{m} z_i^2 = \max_{z \in \mathcal{S}} z'z \tag{2.21}$$

Thus, RRQIP is reduced to the enumeration of extreme points of Zonotope $\mathcal{Z}$. From discrete geometry theory, it is known that the number of extreme point of Zonotope $\mathcal{Z}$ in $\mathbb{R}^n$ is $O(n^{(m-1)})$ and therefore the point of $\mathcal{S}$ can be calculated in $O(n^{(m-1)})$ for $m \geq 3$ and $O(n^m)$ for $m \geq 2$. More specifically, if $\mathcal{Z}$ is in general position, the number of its vertices is given by [22].

$$|\text{vert}(\mathcal{Z})| = 2 \sum_{i=0}^{m-1} \binom{n-1}{i} \tag{2.22}$$

Among many attempts to develop optimal algorithms for enumerating the extreme point of Zonotope, the most well-known one is based on a "Reverse Search" and was developed by Avis and Fukuda [42]. A major disadvantage of this method is its memory requirement. It has to store all the extreme points and also all faces and their incidences [54]. Ferrez at [28], developed an algorithm based on reverse search which reduces the time complexity to $O(mnLP(n, m)vert(Z))$ where $LP(a, b)$ is the complexity of solving a linear problem with a variables and b inequalities. In [40], Stinson at al. proposed a randomised algorithm in which the Zonotope is approximated by a convex hull of a partially enumerated set of vertices. In next section we have a general review over some of enumeration algorithms but before that a geometric interpretation of Zonotope is described.

2.4 Geometric Interpretation of Zonotope

Zonotopes are convex, centrally symmetric polytopes. They can be thought of as linear projections of a high-dimensional hypercube. If $\Omega \in \mathbb{R}^{m\times n}$ with $m < n$ then the set $\mathcal{Z} = \{\Omega x : x \in [-1, 1]^n\}$ defines a Zonotope in $\mathbb{R}^m$. An equivalent way to define Zonotopes is via Minkowski sums of n line segments in $\mathbb{R}^m$. Let $\{\omega_i\}_{i=1}^m$, $\omega_i \in \mathbb{R}^m$, $i = 1, 2, \ldots, n$ be the columns of Ω. Then

$$\mathcal{Z} = \mathcal{A}_1 + \mathcal{A}_2 + \ldots \mathcal{A}_p = \left\{ \sum_{i=1}^{n} \omega_i : \omega_i \in \mathcal{A}_i, i = 1, 2, \ldots, n \right\}$$

where $\mathcal{A}_i = \{\lambda\omega_i : -1 \leq \lambda \leq 1\}$, $i = 1, 2, \ldots, n$. The vectors ω_i are called the generators of $\mathcal{Z}$. To emphasise the dependence of $\mathcal{Z}$ on its generators we will often write $\mathcal{Z} = \mathcal{Z}(\omega_1, \omega_2, \ldots, \omega_n)$ or $\mathcal{Z} = \mathcal{Z}(\Omega)$.

The solution of the RRQIP problem reduces to the enumeration of the extreme points of the Zonotope [22] $\mathcal{Z} = \{V'x : x \in \mathcal{X}\}$ since

$$n\gamma_r = \max_{x\in\mathcal{X}} x'VV'x = \max_{x\in\mathcal{X}} x'VV'x = \max_{z\in\mathcal{Z}} z'z \tag{2.23}$$

and the last maximisation is achieved at an extreme point of $\mathcal{Z}$ since $\mathcal{Z}$ is convex.

Slightly different versions of the RRQIP problem in (2.9) have also been considered in [21, 8, 20] in connection with the so-called real and complex structured singular value problems in robust control applications. The problem of enumerating the extreme points of the Zonotope $\mathcal{Z}$ for low-rank matrix V is well known, see for example [55], [56], [54], [57], [58].

2.5 Enumeration Algorithms

Searching or enumeration of vertices and faces of a graph that describe a system is a fundamental problem in computational geometry. There is a wide number of known search techniques for enumerations. The main objective in vertex enumeration is to find a polynomial-time algorithm or a linear-time algorithm [59]. One of the first methods widely used in graph-based problems was proposed by Read et al. [60] and is known as Backtracking. This algorithm examines an element to decide whether to include it into the current solution or not and afterwards, continues to the next element recursively. Dyer in [61], was the first to propose a linear-time algorithm which uses a depth-first search and a balanced tree data

structure [59]. Some researches focus only on a particular type of polytopes [62]. Among many other efficient enumeration algorithms (see [59], [60], [61], [28], [29], [63] and [39]), Avis and Fukuda introduced an efficient algorithm known as reverse search [42] for enumeration of all vertices and cells in a hyperplane arrangement in $\mathbb{R}^m$. The Avis-Fukuda algorithm is widely used in various listing problems in combinatorics and geometry. In this work, we use reverse research for the vertices of the Zonotope associated with the RRQIP as the final stage of the algorithm used to breach the duality gap of the original QIP problem. If reverse research algorithm successfully designed, its time complexity is proportional to the size of output times the size of the input. The space complexity is also a polynomial in the size of the input. Hence, both time and space complexity are severely affected by the size of the system. In this work, we try to use approximate methods to reduce the number of inputs and consequently reduce the complexity of enumeration algorithms.

In the Zonotope associated with QIP, there is only one local optimal vertex X^*. In order to explain the idea behind the reverse search, let $G = (\mathcal{V}, E)$ be a graph associated to an objective matrix where $\mathcal{V}$ is vertex set, and E is edge set. Consider the tree T spanning all vertices of G with the only sink X^*. Therefore, if we track this graph from X^* using any conventional methods like depth-first search [61], all vertices can be enumerated. Unlike the backtracking method [60] which simply performs the vertex research algorithm itself and stores the information of those vertices which included in the solution at each step, reverse search algorithm traces each edge against its orientation which, in fact, corresponds to reversing the local search without storing any information. An interesting application of reverse search is its ability to enumerate the cells, triangulation's, connected-induced subgraphs bases and spanning trees. For our work, in particular, cell enumeration is crucial as it can be used to find adjacent vertices and hence calculate the angle of each vertex. In fact, there is a relation between a vertex of a Zonotope and its dual arrangement (Theorem 3.4(Duality) in [1]).

This is illustrated graphically in Figure 2.1.

The dual associated with a Zonotope $\mathcal{Z}(\mathcal{V})$ generated by the columns of V, is a central arrangement $\mathcal{A}(\mathcal{V})$ of n hyperplanes in $\mathbb{R}^d$, having a v^j as its normal vector:

$$\mathcal{A}(\mathcal{V}) = \{h_j^0 : j = 1, 2, ..., n\} \tag{2.24}$$

where $h_j^0 = \{y \in \mathbb{R}^d : \langle v^j, y \rangle = 0\}$ for $j = 1, 2, ..., n$. Similar definition for the positive and negative side of each hyperplane is defined as: $h_j^+ = \{y \in \mathbb{R}^d : \langle v^j, y \rangle > 0\}$ and

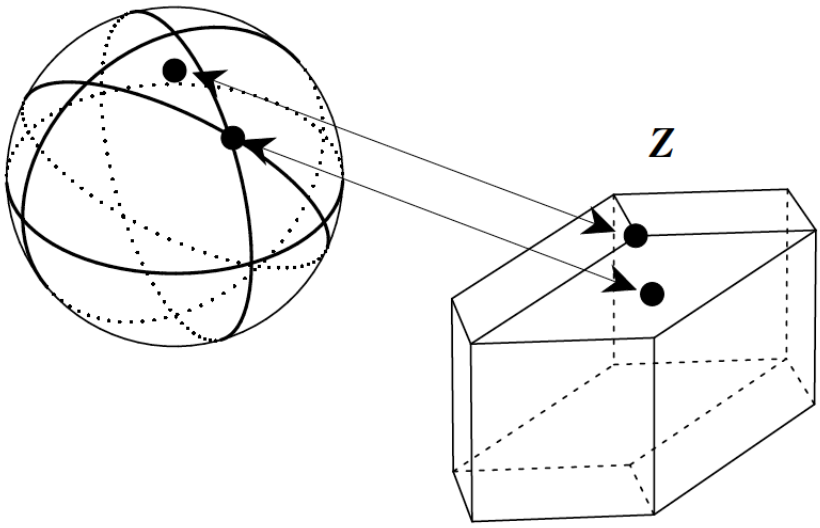

Figure 2.1: Duality of a Zonotope and its associated arrangement [1]

$h_j^- = \{y \in \mathbb{R}^d : \langle v^j, y \rangle < 0\}$.

Therefore, the location of any vector $c \in \mathbb{R}^d$ can be defined as $\sigma(c)$ where:

$$\sigma(c)_j = \begin{cases} + & \text{if} \quad c \in h_j^+ \\ 0 & \text{if} \quad c \in h_j^0 \\ - & \text{if} \quad c \in h_j^- \end{cases} \tag{2.25}$$

The benefit of this identification, is that each vertex of Zonotope $\mathcal{Z}(V)$ can be identified by a unique sign vector of its dual cell. For example if Figure 2.2 illustrates a cut subsection of 5 hyperplanes in $\mathbb{R}^3$ associated with Zonotope Figure 2.1. Each cell is represented by a sign vector of length 4 as shown in Figure 2.2

Another interesting consequence of the sign vector is the simplicity with which adjacent cells can be found. For example if corresponding sign vector of two cells differ in only one sign, these two faces are adjacent. Note that two vertices of $\mathcal{Z}$ are adjacent if and only if associated cells are adjacent. As mentioned before, this characteristic is useful to calculate the angle of vertices. There are some other improvements to the reverse search method[64] in the literature which makes this algorithm one of the most efficient methods for enumerating the vertices of Zonotope.

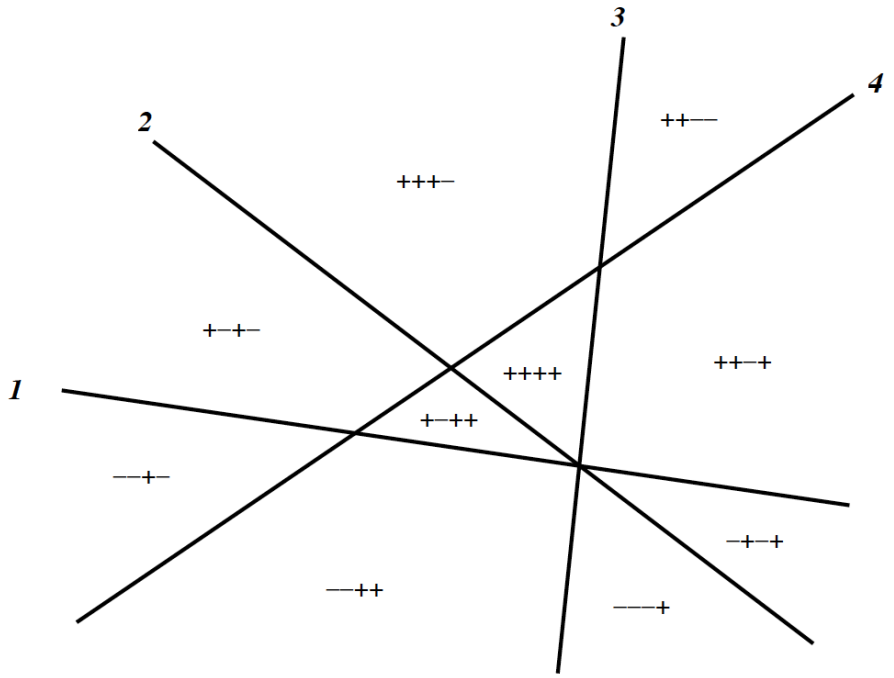

Figure 2.2: A cut subsection of a 3-dimensional central arrangement [1]

2.6 A randomised algorithm for enumerating Zonotope vertices

Breaching the SDR bound relies on the complete enumeration of the vertices of a Zonotope. This can be effectively achieved using several deterministic algorithms, e.g. the reverse-enumeration algorithm outlined in the previous section. In high dimensional problems, however, even efficient deterministic algorithms are impractical. In this case, one can rely on randomised algorithms which approximate the Zonotope by generating only a subset of its vertices. A method to obtain an improved (probabilistic) bound on γ can be obtained by using the results of [40]. As it has been mentioned in the previous section, the time complexity of enumeration algorithms can be problematic as the size of a system grows. For straightforward algorithms like Quickhull [65], the complexity scales exponentially in m, the number of vertices. An alternative approach is the reverse search [42], which has been introduced in section 2.5. Another alternative is to approximate the Zonotope via Goffin's algorithm [66]. Nevertheless, the numerical implementation of this method appears problematic [40]. Randomised methods are used in some research publications to approximate solution to the QIP problem (see [24],[67] and [68]). Among these methods, Stinson et al. [40] introduced an interesting randomised method recently. The algorithm uses the fact that a Zonotope vertex is a linear combination of Zonotope generators. The research in [40] shows if the randomised algorithm terminates before all vertices are recovered, the

convex hull of the enumerated vertices approximates the Zonotope with certain accuracy and probability. In this chapter, it is shown how this approach can be adapted to qualify the probability that the duality gap in the QIP is breachable, even if it is not possible to enumerate all vertices of the Zonotope in the solution of RRQIP.

Consider a Zonotope $\mathcal{Z}(\Omega)$ with generator a matrix $\Omega \in \mathbb{R}^{q\times p}$ where $q \leq p$. Suppose that the columns of Ω satisfy the following two conditions: (i) No column of Ω is the zero vector, and (ii) no two columns of Ω are scalar multiples of each other. Then $\mathcal{Z}(\Omega) = \{\Omega x : x \in [-1\ 1]^p\}$ is in general position and the number of its vertices is given by

$$|\text{vert}(\mathcal{Z})| = 2\sum_{i=0}^{q-1} \binom{p-1}{i} \tag{2.26}$$

Then, under the above assumption, for $x \in \mathbb{R}^n$ such that $\Omega' x$ has all nonzero elements, the point v defined by the mapping $v = m(\delta) := A\text{sign}(A'\delta)$ is a vertex of $\mathcal{Z}(\omega_1, \ldots, \omega_p)$. From central symmetry $v \in \text{vert}(\mathcal{Z}) \Rightarrow -v \in \text{vert}(\mathcal{Z})$. Moreover, if

$$H = \bigcup_{i=1}^{q} \left\{\delta \in \mathbb{R}^p : \hat{\omega}_i'\delta = 0\right\} \tag{2.27}$$

where $\hat{\omega}_i'$ are the rows of Ω, $i = 1, 2, \ldots, q$, then the mapping $m : \mathbb{R}^q \setminus H \rightarrow \text{vert}(\mathcal{Z})$ is well defined and onto [40]. A randomisation algorithm can now be used to enumerate the vertices of $\mathcal{Z}$. This is summarised in Figure 2.3. The algorithm updates a list of vertices (initialised as the empty list) by drawing independent samples $\delta \in \mathbb{R}^p$ from a p-dimensional standard Gaussian distribution, computing $v_+ = \Omega\text{sign}(\Omega'\delta)$ and $v_- = -v_+$ and adding them to the list (unless they are already listed). The algorithm can proceed until all vertices have been enumerated, or terminate after a fixed number of iterations. In the later case only a subset of vertices will be (in general) enumerated. The convex hull of these vertices will be a subset of $\mathcal{Z}$.

It is worth mentioning that Zonotopes have another geometric representation, known as the Minkowski sum [29] (also known as dilation). The Minkowski sum of two sets A and B is defined as follows:

$$A + B = \{a + b : \quad a \in A, \quad b \in B\} \tag{2.28}$$

A Zonotope $(\mathcal{Z})$ can then be defined as

$$\mathcal{Z} = A_1 + ... + A_n \tag{2.29}$$

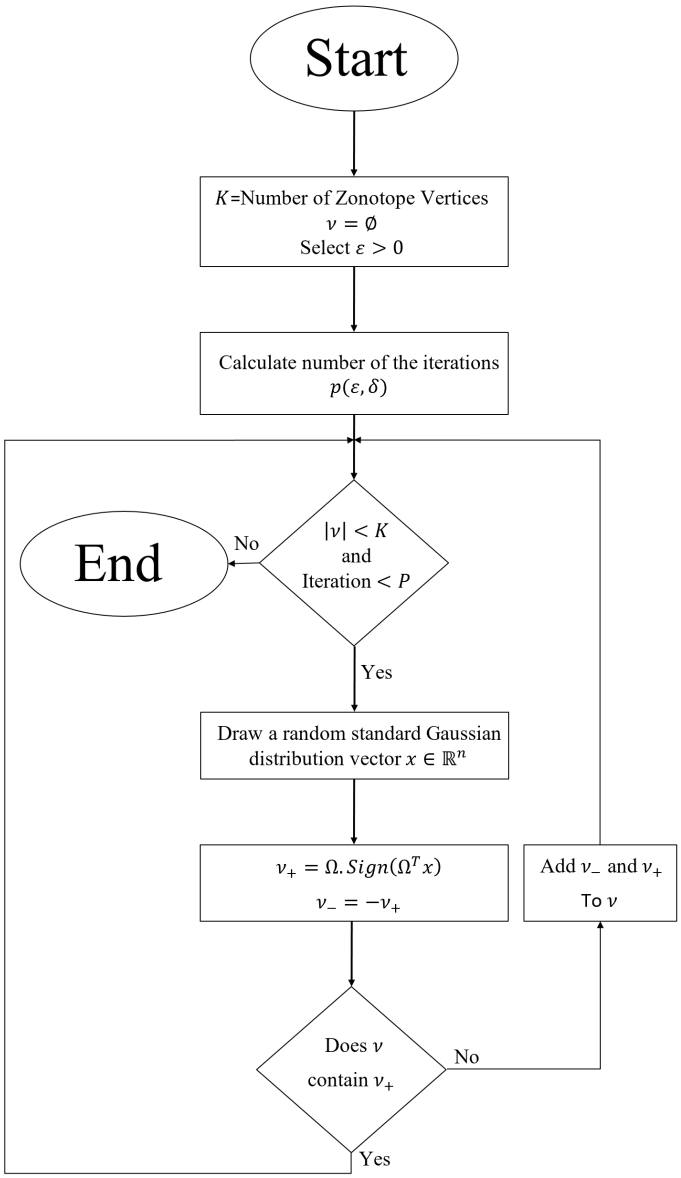

Figure 2.3: Randomised Algorithm flow-chart

where

$$A_i = \{\alpha\overline{\alpha}_i \quad | \quad \alpha \in [-1, 1]\} \tag{2.30}$$

Without lose of generality, we assume that matrix $V = [\overline{\nu}_1 \dots \overline{\nu}_n] \in \mathbb{R}^{n\times m}$ in (2.19) is in general position, i.e. V neither has a column of zero, nor are there any two columns which are a scalar multiple of each other.

The above discussion can be summarised in the following theorem:

Theorem 2.6.1 (Theorem2, Corollary 1 and Corollary 2 [40]). *Let* $\Omega = [\ \omega_1 \quad \dots \quad \omega_m\ ] \in \mathbb{R}^{n\times m}$ *be the Zonotope generator and let* $x \in \mathbb{R}^n$ *be such that none of the component of*

$A'x$ is zero, define V as

$$V = m(x) := \Omega sign\left(\Omega' x\right) \tag{2.31}$$

where sign(X) returns a vector in $\{-1, 1\}$ whose elements correspond to the sign of each component of X. Then both V and $-V$ are vertices of $Z(A)$.

Moreover, let $H \subset \mathbb{R}^m$ as defined in (2.27)*, then mapping $m : \mathbb{R}^m \backslash H \longrightarrow Ver\left(Z(A)\right)$ is well defined and $P_x[H] = 0$, where P_x is probability measure in $\mathbb{R}^m$.*

A direct consequence of Theorem 2.6.1 is that each $x \in \mathbb{R}^n$ maps under relation (2.31) to some Zonotope vertex with probability one. However, the probability that a randomly chosen x maps to a specific vertex is related to the geometric properties of the vertex. More specifically, the probability that x maps to a sharp vertex is larger than the probability that x maps to a flat vertex. In fact, [40] argue that the contribution of flat vertices to forming a Zonotope is negligible. The idea is that the flat vertices are less likely to be selected in randomised algorithm which eventually results to an enumeration over a reduced number of vertices. It is obvious that examining every vertices in order to identify non-influential vertices is not computationally efficient. In fact, if you could examine every vertex you would have a complete list and therefore distinguishing between flat an sharp vertices would be irrelevant. However, there is an interesting relation between the contributions of a vertex of the Zonotope and the probability that a random vector x maps to this vertex as defined before. To investigate this relationship we need to introduce the following two concepts:

- Normal Cone of a vertex $V(N_Z(V))$,
- Hausdorff distance between a vertex V and the convex hull of $Ver(Z(A))\backslash\{V\}$.

The normal cone of a vertex is the set of all $x \in (R)^m$ which map to the vertex under (2.31). This set forms a region denoted by $N_Z(V)$ formally defined as follows:

$$N_Z(V) = \{x \in (R)^m \quad | \quad \langle z - V, x\rangle \leq 0 \quad for \quad z \in Z\} \tag{2.32}$$

Figure 2.4 shows an example of the normal cone for a sharp and a flat vertex for a two-dimensional zonotope:

Referring to Theorem 2.6.1 the normal cone $N_Z(V)$ is the inverse of the map defined in (2.31), i.e.

$$\text{interior} \quad N_z(V) = m^{-1}(V) \tag{2.33}$$

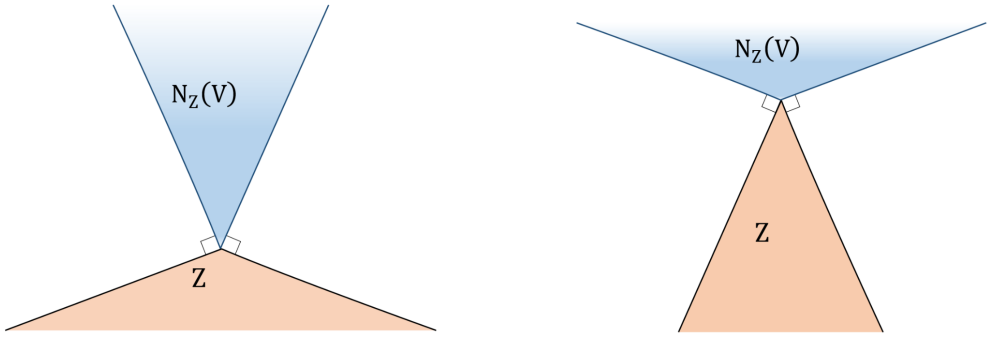

Figure 2.4: An example of normal cone for sharp and flat vertices in a 2-dimensional plane

One of the main characteristic of Normal cone in a Zonotope is that the integration of the normal cone corresponding to all vertices forms the complete space $\mathbb{R}^m$. This is a consequence of Theorem 2.6.1 and the fact that $P_x[H] = 0$ or equivalently

$$P_x\left[\{x \in \mathbb{R}^m | m(x) \in ver\left(Z(A)\right)\}\right] \tag{2.34}$$

where $m(x)$ defined in (2.31). This consequence motivates finding the probability of a vector x being mapped to a specific vertex $P_x(V_i)$. It is obvious that sharp vertices have a higher probability than flat vertices. Figure 2.5 illustrates how the normal cones of vertices of a two-dimensional Zonotope forms the entire space.

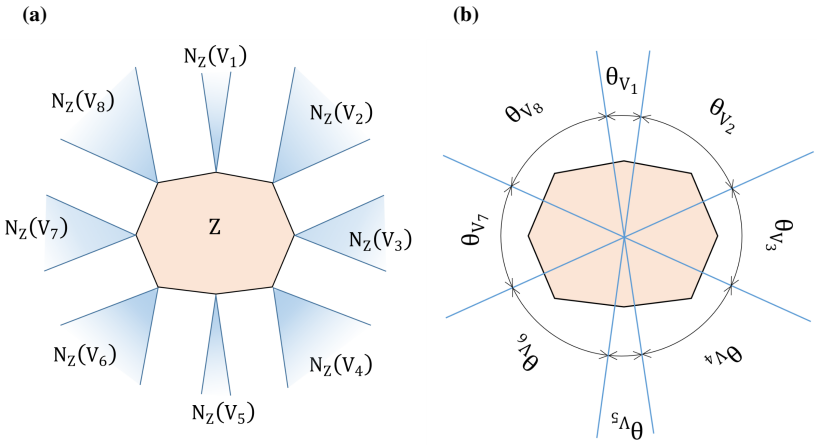

Figure 2.5: (a)Example of Normal Cone of each vertices in a 2-Dimensional Zonotope. (b)Normal Cones of all 2–dimensional vertices form the entire $\mathbb{R}^2$

It can be seen from Figure 2.5, that the normal cone of flat vertices have smaller share in forming space $\mathbb{R}^2$. The probability that a random vector $x \in \mathbb{R}^2$ lies inside the normal

cones of V_i is:

$$P_x\left(V_i\right) = \frac{\theta_{V_i}}{2\pi} \tag{2.35}$$

The Hausdorff distance, is a measure of the proximity of two sets. This measure can be used as a criterion of whether the convex hall of a subset of all vertices is a good approximation of the original Zonotope. Hausdorff distance between two non-empty sets X and Y, $h(X,Y)$, is defined as:

$$h(X,Y) = \max\left\{\sup_{x\in X}\inf_{y\in Y} d(x,y), \sup_{y\in Y}\inf_{x\in X} d(x,y)\right\} \tag{2.36}$$

where $d(x,y)$ is the metric distance between x and y. In our case when two polygonal subsets are identical in all vertices except one, the Hausdorff distance has a simpler form called as simplicity constant, $\alpha_Z(V)$, and has been used in [40]. As a criterion to identify how far the vertex is from the convex hull, we define:

$$\alpha_Z(V_i) = \inf_x\{\|V_i - x\|_2 \quad | \quad x \in conv(vert(z)\backslash\{V_i\})\} \tag{2.37}$$

Figure 2.6 shows a graphical interpretation of $\alpha_Z(V_1)$. It can be seen from Figure 2.6 that there is a direct relation between the angle of a vertex and its simplicity constant. In other words, the sharper the angle, the larger the simplicity constant.

These two definitions (Normal Cone and Hausdorff distance) have been well studied in [40]. Next, a simple form of the randomised algorithm [40] is outlined. For given $\epsilon \geq 0$ and $\delta \geq 0$ the algorithm approximates the Zonotope as:

$$V = Co\{\nu_1, \nu_2, ..., \nu_r\} \tag{2.38}$$

where $v_1, v_2, ..., v_r$ are the vertices selected by the randomised algorithm. A bound on the Hausdorff distance as an approximation criterion, between the original Zonotope and its approximate polytope is stated in Theorem 2.6.2 below [40]:

Theorem 2.6.2 ([40]). *Let $\mathcal{Z}(\omega_1, \ldots, \omega_n)$ to be a Zonotope with its generating vectors in general position. Given $\epsilon > 0$ and $\delta > 0$ choose $b > diam(\mathcal{Z})$ and p as*

$$p > \frac{\log\left(|vert(\mathcal{Z})|/\epsilon\right)}{\log\left(1/(1-k)\right)} \tag{2.39}$$

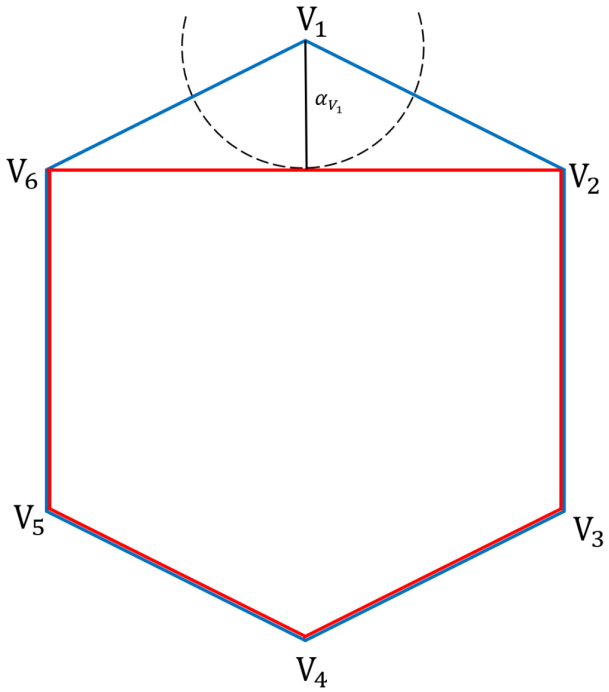

Figure 2.6: $\alpha_Z(V_i)$ is a Hausdorff distance between a vertex and the convex hall of rest of the vertices

where

$$k = \left(\frac{1}{2}\left(1 - \sin\left(\arctan(b/\delta)\right)\right)\right)^{\frac{m-1}{2}} \tag{2.40}$$

Let $\mathcal{V}$ be the subset of $\mathcal{Z}$'s vertices produced by the randomised algorithm after p iterations. Then, if $h(\mathcal{Z}, conv(\mathcal{V}))$ denotes the Hausdorff distance between sets $\mathcal{Z}$ and $conv(\mathcal{V})$ we have that:

$$h(\mathcal{Z}, conv(\mathcal{V})) \leq \frac{|vert(\mathcal{Z}) \setminus \mathcal{V}|}{2}\delta \tag{2.41}$$

with probability at least $1 - 2^a\epsilon$ where $a = |vert(\mathcal{Z}) \setminus \mathcal{U}_Z|/2$ and

$$\mathcal{U}_Z = \{v \in vert(\mathcal{Z}) : \alpha_Z(v) \geq \delta\}$$

Remark 2.6.1. *If $\Omega \in \mathbb{R}^{m\times n}$, $m \leq n$, is orthogonal then*

$$diam(\mathcal{Z}) = \max_{x\in[-1,1]^n} 2\|\Omega x\| \leq 2\sqrt{n}$$

and b in Theorem 2.6.2 may be selected as $b = 2\sqrt{n}$.

We can now prove the following Theorem which gives a probabilistic upper bound on γ. Here the generator matrix Ω of $\mathcal{Z}(\Omega)$ is identified with $\Omega = V'$ where V is defined in the spectral decomposition given in equation (2.13).

Theorem 2.6.3. *Let all variables be defined as in Theorem 2.6.2. Let also* $\mathcal{Z}(V') = \{V'x : x \in \{-1,\ 1\}^n\}$ *and assume that* $\mathcal{V}$ *is the set of vertices generated by the randomised algorithm in* p *iterations. (i) Suppose there is a vertex* $v \in \mathcal{V}$ *with* $\|v\| = \sqrt{n}$*. Then* $\gamma = \bar{\gamma}$*. (ii) Suppose that* $\tau = \max\{\|v\| : v \in \mathcal{V}\} < \sqrt{n}$*. Suppose that* δ *is selected such that*

$$\frac{|vert(\mathcal{Z}) \setminus \mathcal{V}|^2\delta^2}{4} + \tau^2 < n$$

and set

$$\hat{\gamma}_r = \frac{1}{n}\left(\frac{|vert(\mathcal{Z}) \setminus \mathcal{V}|^2\delta^2}{4} + \tau^2\right) < 1$$

Then

$$\gamma \leq \bar{\gamma} - n(1-\gamma_r)\underline{\lambda}(\Lambda_+) \leq \bar{\gamma} - n(1-\hat{\gamma}_r)\underline{\lambda}(\Lambda_+) \leq \bar{\gamma}$$

with probability at least $1 - 2^a\epsilon$*.*

Proof. (i) If the distance of an enumerated vertex from the origin is equal to $\sqrt{n}$ then this must be a vertex of $\mathcal{Z}$ furthest away from the origin and hence $\gamma = \bar{\gamma}$. (ii) Let v_0 be the vertex enumerated by the randomised algorithm which is furthest away from the origin so that $\tau = \|v_0\|$. Let also $\mathcal{P}$ be a compact superset of conv(V) such that

$$h(\mathcal{P}, \text{conv}(V)) = R(\delta) := \frac{|\text{vert}(\mathcal{Z}) \setminus V|\delta}{2} \tag{2.42}$$

$\mathcal{P}$ can be easily constructed as the intersection of hyperplanes drawn parallel to the faces of the polytope conv(V) displaced a distance $r(\delta)$ away from the origin with the hyperspheres $\{z : \|z - v\| \leq R(\delta)\}$ with centre the vertices of V and radius $R(\delta)$ (see Figure 2.7 for the construction of $\mathcal{P}$ in the case $r = 2$). It is also clear that the point on $\mathcal{P}$ furthest away from the origin is the intersection of the line connecting v_0 and the origin and the hypersphere $\{z : \|z - v_0\| = R(\delta)\}$ whose distance from the origin is $\sqrt{\tau^2 + R^2(\delta)}$. Since $\mathcal{P} \supseteq \mathcal{Z}$ with probability at least $1 - 2^a\epsilon$ we have that:

$$\begin{aligned}
\gamma_r :&= \frac{1}{n}\max_{x\in\mathcal{X}} \|V'x\|^2 = \frac{1}{n}\max_{z\in\mathcal{Z}} \|z\|^2 \\
&\leq \frac{1}{n}\max_{z\in\mathcal{P}} \|z\|^2 = \frac{1}{n}\left(\tau^2 + R^2(\delta)\right) \\
&= \hat{\gamma}_r < 1
\end{aligned}$$

where the inequality is valid with probability at least $1 - 2^a\epsilon$. The result then follows from Lemma 2.2.4. □

Remark 2.6.2. *Note that the inequality $\gamma_r \leq \hat{\gamma}_r$ is always in force and hence the probabilistic bound on γ cannot be tighter than the deterministic bound. The probabilistic bound, however, does not rely on the full enumeration of the vertices of the Zonotope and hence is potentially less computationally demanding. The accuracy of the approximation cannot be determined a-priori and depends on the particular set of vertices (and their number) that was selected by the randomised algorithm.*

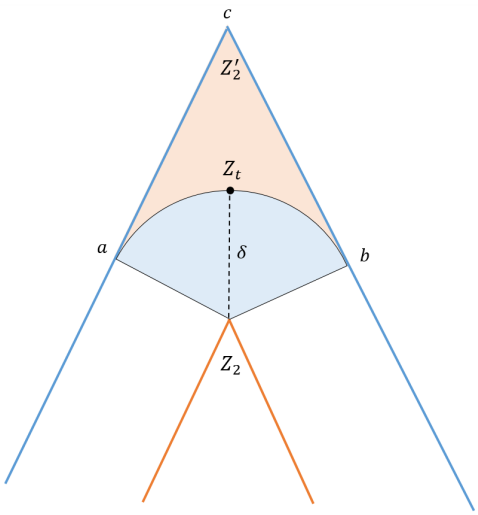

Figure 2.7: Estimation of $\max_{z \in \mathcal{Z}} \|z\|$

Remark 2.6.3. *The Hausdorff distance between $\mathcal{P}$ and $\mathcal{Z}$ can be made arbitrarily small by selecting δ sufficiently small. In this case, however, the required number of iterations p of the randomised algorithm may be too high, so that the algorithm may not be implementable in practice.*

A probabilistic version of Lemma 2.2.6 follows:

Theorem 2.6.4. *Let all variables be defined as in Lemma 2.2.6 and Theorem 2.6.2. Let also $\mathcal{Z}(V') = \{V'x : x \in \{-1,\ 1\}^n\}$ where V is as defined in Lemma 2.2.6 and assume that $\mathcal{V}$ is the set of vertices generated by the randomised algorithm in p iterations where p is defined in Theorem 4.4.1. Suppose that $\tau = \max\{\|v\| : v \in \mathcal{V}\} < 1$. Suppose also that δ is selected such that:*

$$\frac{|vert(\mathcal{Z}) \setminus \mathcal{V}|^2 \delta^2}{4} + \tau^2 < 1$$

and set

$$\hat{\phi}_i = \frac{|vert(\mathcal{Z}) \setminus \mathcal{V}|^2 \delta^2}{4} + \tau^2 < 1$$

Then

$$\gamma := \max_{x \in \mathcal{X}_1} x'Qx \leq \lambda_{i+1} + \hat{\phi}_i$$

with probability at least $1 - 2^a\epsilon$.

Proof. Similar to the proof of Theorem 2.6.3. □

2.6.1 Numerical Example 1

In this section the results of the previous section are illustrated with a numerical example. Matrix $Q \in \mathbb{R}^{10\times 10}$, $Q = Q' \geq 0$ has been chosen randomly as:

$$\begin{bmatrix}
4.07 & -0.46 & 0.25 & -2.79 & -1.50 & -0.41 & -1.54 & 0.04 & -1.604 & -1.03 \\
-0.46 & 4.77 & -4.56 & -1.36 & -0.70 & -0.45 & -0.49 & -2.15 & 0.33 & -1.19 \\
0.25 & -4.56 & 6.71 & 1.38 & 2.03 & 1.31 & -0.39 & 3.38 & 0.54 & -1.05 \\
-2.79 & -1.36 & 1.38 & 6.53 & 1.60 & -0.56 & 1.72 & -0.46 & -1.23 & 1.54 \\
-1.50 & -0.70 & 2.03 & 1.60 & 3.43 & -0.80 & 1.78 & 0.95 & -0.44 & -0.23 \\
-0.41 & -0.45 & 1.31 & -0.56 & -0.80 & 4.54 & -1.87 & 2.55 & 3.46 & 0.07 \\
-1.54 & -0.49 & -0.39 & 1.72 & 1.78 & -1.87 & 3.91 & -0.82 & -2.22 & 1.17 \\
0.04 & -2.15 & 3.38 & -0.46 & 0.95 & 2.55 & -0.82 & 3.73 & 1.61 & -0.22 \\
-1.60 & 0.33 & 0.54 & -1.23 & -0.44 & 3.46 & -2.22 & 1.61 & 5.27 & -0.11 \\
-1.03 & -1.19 & -1.05 & 1.54 & -0.23 & 0.07 & 1.17 & -0.22 & -0.11 & \bar{3.46}
\end{bmatrix}$$

(here truncated to two decimal places). The optimal value of the QIP can be calculated via all $2^{10} = 1024$ binary evaluations as $\gamma = 1.164243$. The SDR upper bound was obtained via MATLAB's LMI toolbox after 24 iterations as $\bar{\gamma} = 1.23753502$ with guaranteed relative accuracy $8.01 \cdot 10^{-11}$. The corresponding eigenvalues of $D - Q$ were obtained as:

$$\{10^{-8},\ 8 \times 10^{-8},\ 3.09,\ 4.31,\ 8.54,\ 11.00,\ 11.25,\ 11.62,\ 12.66,\ 14.77\}$$

indicating a nullity $r = 2$. From the spectral decomposition of $D - Q$ the matrix $V \in \mathbb{R}^{10\times 2}$ with orthonormal columns was constructed spanning the null-space of $D - Q$. The auxiliary QIP problem (2.9) was again solved via all possible (2^{10}) binary evaluations (for problems of higher complexity the reverse enumeration algorithm could have been used). The maximum was obtained as $\gamma_r = 0.882856$ which corresponds to the two vertices of the Zonotope $(Z) = \{\frac{1}{\sqrt{10}}\|V'x\| : x \in \text{co}(\mathcal{X})\}$ furthest away from the origin (see Figure 2.8). Note that the Zonotope has 20 vertices in agreement to equation (2.26), the remaining 1004 extreme points located in the interior of the Zonotope. With the achieved value of γ_r, equation (2.10) gives the improved bound on γ as 1.201261 which corresponds to a gap

reduction of approximately 50.51%.

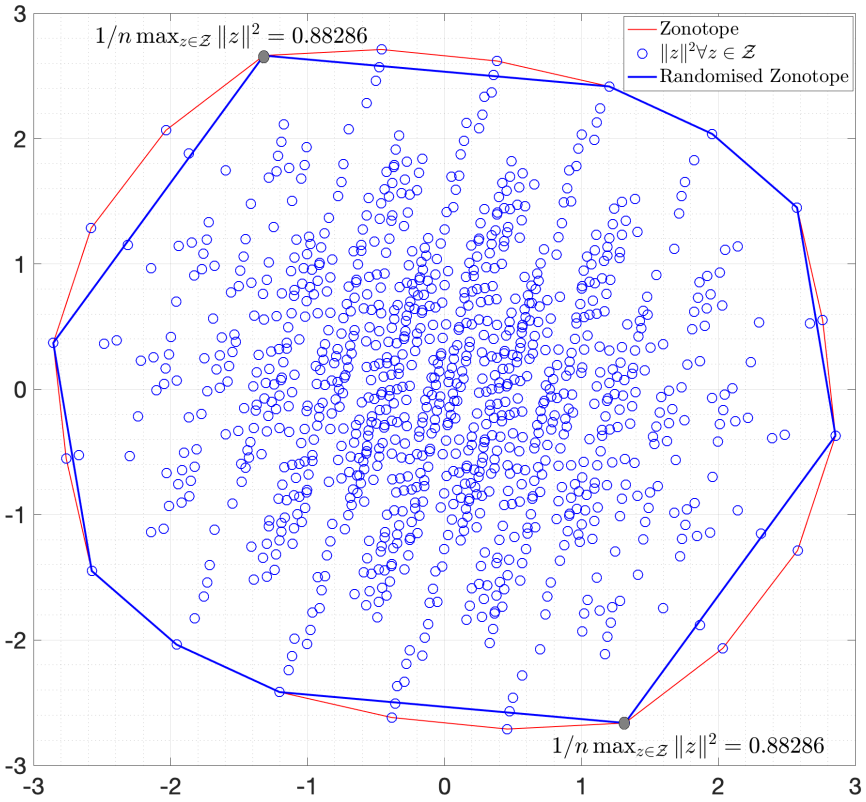

Figure 2.8: Zonotope and reduced polytope from randomised algorithm

Next, the randomised algorithm was tested by assuming that full enumeration of the vertices is not possible. The parameters of the algorithm were initially set as $\epsilon = 0.05$ and $\delta = 0.1$. This resulted in a randomised algorithm with $p = 10$ iterations. Random sampling from the standard normal bi-variate distribution was implemented with MATLAB's function *mvnrnd.m*. Note that in this case, a different subset of the Zonotope's vertices is obtained each time a set of $p = 10$ independent random samples is taken. Figure 2.8 shows the convex hull of the selected vertices on a particular occasion in which 14 vertices were obtained (i.e. 3 out of the 10 samples resulted in a vertex which was already listed). Three sets of results obtained in this way are summarised in the table below. This displays the number of polytope vertices selected (n_P), the Haussdorf distance estimate d_{H}, the bound $\bar{\gamma}_r$ on γ_r, the bound γ_u on γ and the corresponding confidence level $P[\gamma \leq \gamma_u]$.

n_P	d_{H}	$\bar{\gamma}_r$	γ_u	$P[\gamma \leq \gamma_u]$
14	0.30	0.891	120.404	$\geq 95\%$
12	0.40	0.898	120.621	$\geq 95\%$
10	0.50	0.907	120.900	$\geq 95\%$

To study the variability of the results produced by the randomised algorithm, the process was repeated 1000 times with the same parameters ($\epsilon = 0.05$ and $\delta = 0.1$). The empirical probability mass function of variables n_P, $\bar{\gamma}_r$ and γ_u are shown in Figures 2.9, 2.10 and 2.11, respectively. The mean value of γ_u is $\mathbb{E}[\gamma_u] = 120.6541$ and its standard deviation $\sigma = 0.2538$. Thus, on average, the randomised algorithm indicates that $\gamma \leq 120.6541$ with a probability of at least 95%.

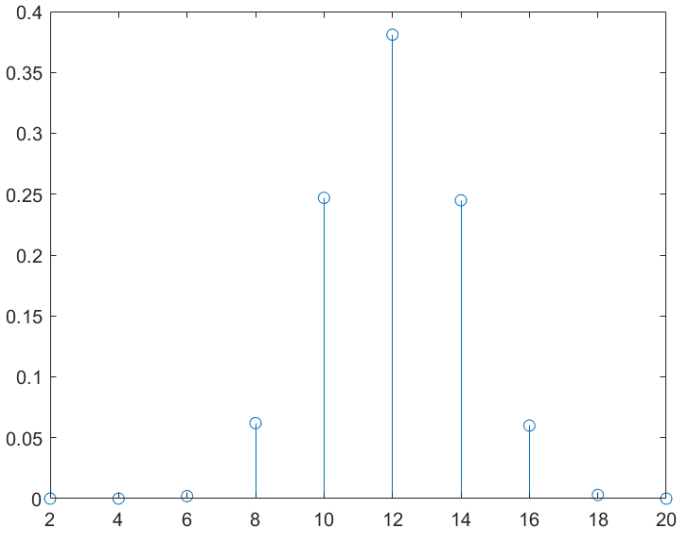

Figure 2.9: Empirical probability distribution: n_P

So far we have argued that the randomised algorithm generate a convex hull of the randomly enumerated vertices of the original vertices of the Zonotope. Those vertices which do not contribute to the approximation form the new reduced polytope, have a Hausdorff distance less than δ from the convex hall of the rest of the vertices with a certain probability. Eventually, the Hausdorff distance between the Zonotopes and its reduced polytope is as small as $\tau\delta$. We also discussed that the maximiser might not be included in set V when the algorithm is terminated, which could result in the Zonotope maximiser breach the optimal value of QIP problem. To overcome this issue, we extended the solution by Hausdorff distance, so we can form a convex set which, firstly, has the same number of extreme points as those obtained by the partial enumeration algorithm; and secondly, the maximiser

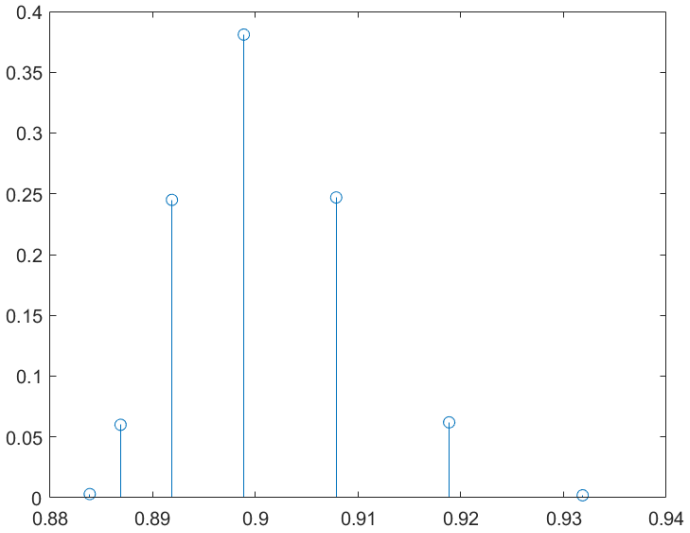

Figure 2.10: Empirical probability distribution: $\bar{\gamma}_r$

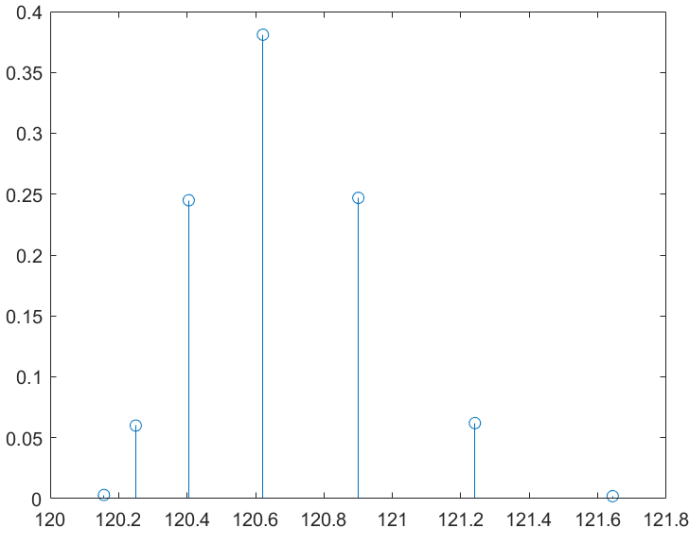

Figure 2.11: Empirical probability distribution: γ_u

over this set overbounds the maximum over the original Zonotope and hence the risk of underestimating the maximum is eliminated.

Another way of presenting this method is to extend the bases of the reduced polytope by $h(\mathcal{Z},\mathcal{U})$ to contain all vertices of original Zonotope. Figure 2.12 illustrates how the extended boundary includes all non-enumerated vertices. In Figure 2.12, a random rank 2 $Q_{7\times 7}$ has been generated in MATLAB and then the randomised algorithm has been programmed to enumerate the extreme points. As discussed previously, due to the fact that Q is positive semi-definite, the maximiser will occur on the boundary or more precisely on a

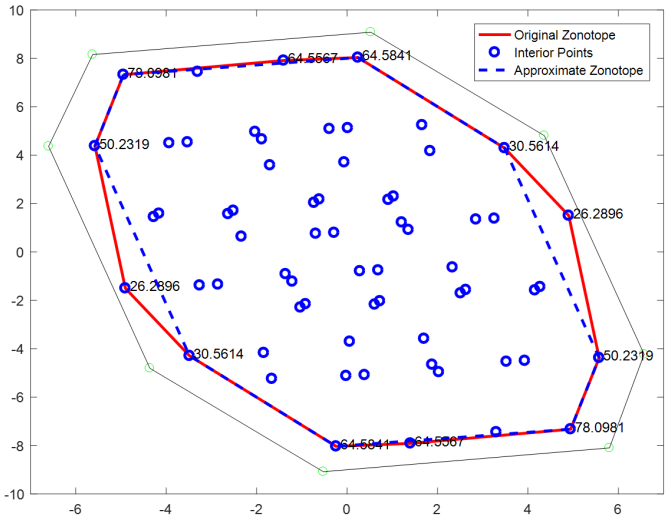

Figure 2.12: Expanding the boundary by Hausdorff distance will form a Zonotope which cover the original Vertices

vertex. In the example of Figure 2.12, the Zonotope has 12 vertices of which 4 have been omitted by the randomised algorithm. It can also be observed from Figure 2.12 that the omitted vertices are mainly flat ones. This is due to the fact that the probability of omitting a sharp vertex is small. The program then develops a method to extend the boundary of the reduced polytope by $h(\mathcal{Z},\mathcal{U})$ and form the new approximation (green circuits are new vertices and black lines are the boundary of approximation). It can be seen that the original Zonotope (red line) is within the extended approximation It follows immediately that:

If Zonotope A is inside convex polytope B, then the Maximise of the set associated with Zonotope A$\leq$ the Maximise of the set associated with polytope B

Note that the maximiser over the extended polytope gives an upper-bound for the RRQIP.

One point that worth mentioning is that the symmetry property of the Zonotope would normally push the maximiser toward a sharp vertex which has less probability to be omitted by the randomised algorithm.

Figure 2.13 shows examples of 2-dimension Zonotopes overlayed on a set of QIP level contours. As discussed, the maximiser tends to occur at the sharpest vertices of the Zonotope.

To demonstrate that the extreme points of a Zonotope tend to occur at its sharp vertices, a Matlab code was produced to randomly generate rank-2, 20 by 20 QIP and sort the vertices of the equivalent Zonotope from the sharpest to the widest angle. Then the extreme points were ordered according to their distance from the origin. Note that in a Zonotope of a rank-2, defined by a 10 by 10 matrix there were 20 pairs of vertices each corresponding to the same distance. This is due to the symmetry property of the Zonotope with respect to

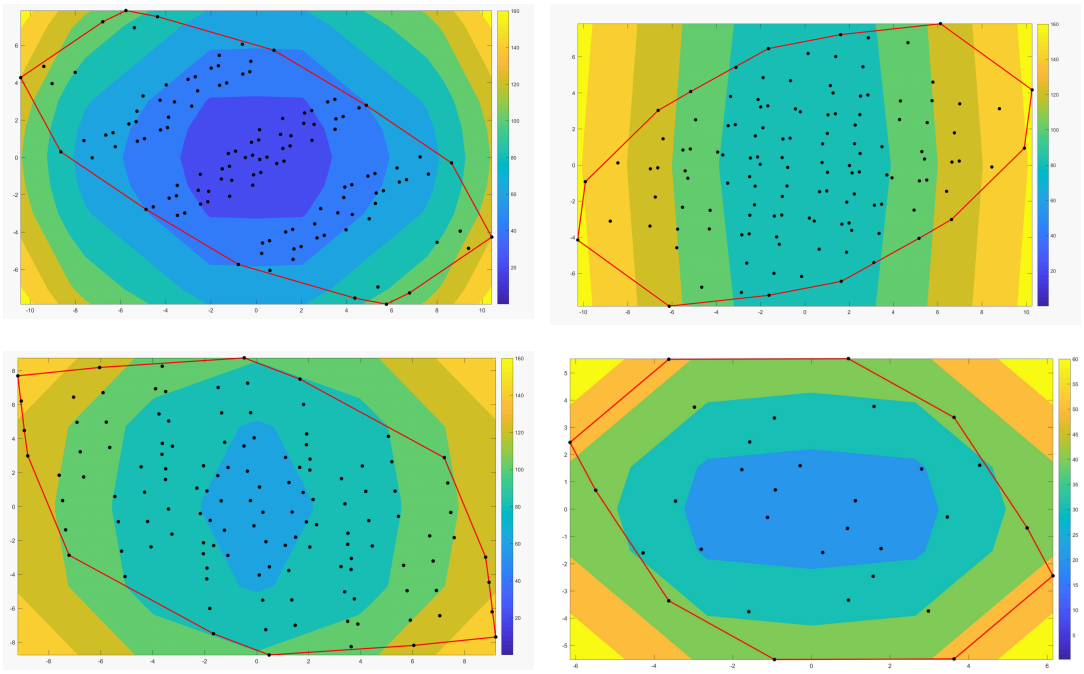

Figure 2.13: The maximiser of Zonotope is normally obtained in sharp vertices which the algorithm is unlikely to cancel them.

the origin. Hence the enumeration algorithm enumerates needs only half the vertices. In our example, these vertices are numbered 1 to 20 (the sharpest angle being number 20). The program has been run for 1000 random trials. Figure 2.14, shows the histogram of the results. In this case, more than 85% of the maximisers occur at the three sharpest vertices.

Next, a condition for which the reduced rank QIP guarantees that the duality gap of the original problem is breachable will be derived. Note that in [30], a sufficient condition for breaching the gap has been obtained if a bound on the corresponding RRQIP problem can be obtained. This condition is adapted here to derive a new probabilistic condition which guarantees that the duality gap is breachable. Theorem 2.6.5 derives this condition.

Theorem 2.6.5. *Let all variables be defined as in Lemma 2.2.6 and Theorem 2.6.2. Suppose $\mathcal{U}$ be the subset of the vertices of $\mathcal{Z}$ generated by randomised enumeration algorithm defined in Figure 2.3. For each $Z \in \mathcal{U}$, let ϕ_Z be the angle of vertex Z and define $\alpha = \dfrac{|vert(\mathcal{Z}) \setminus V|}{2}$. Then if*

$$\max_{Z \in \mathcal{U}} \|Z\| + \frac{\delta}{\sin(\phi/2)} \leq \sqrt{n} \tag{2.43}$$

an upper bound can be induced on γ that with probability of at least $1 - 2^{\alpha}\delta$, is tighter than

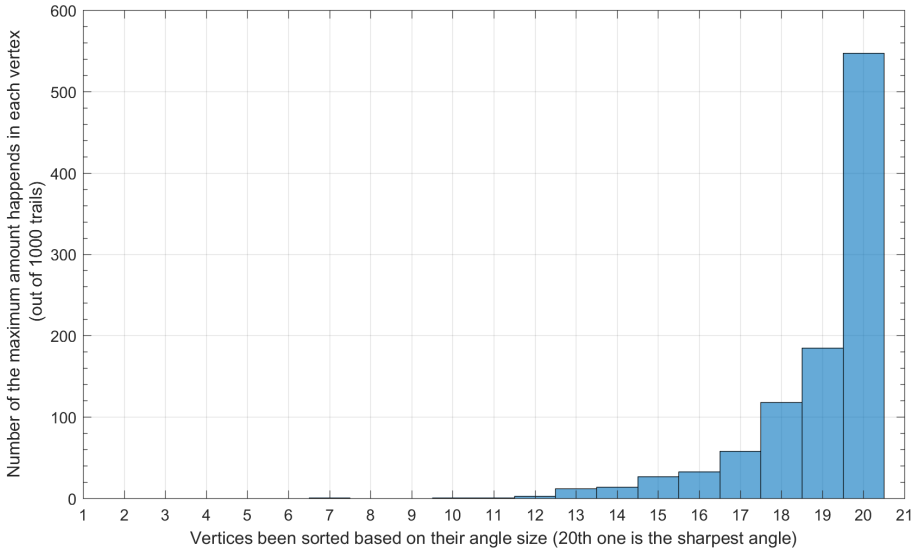

Figure 2.14: Histogram of the maximum amount in 1000 trails of randomly selected $Q_{20\times 20}$

semi definite relaxed bound $\overline{\gamma}$*:*

$$\gamma \leq \overline{\gamma} - n(1-\gamma_r)\underline{\lambda}(\Lambda_+) < \overline{\gamma} - n(1-\gamma_u)\underline{\lambda}(\Lambda_+) \leq \overline{\gamma} \tag{2.44}$$

where

$$\gamma_r := \frac{1}{n}\max_{Z\in\mathcal{Z}} Z'Z \tag{2.45}$$

and

$$\gamma_u := \frac{1}{n}\max_{Z\in\mathcal{U}} Z'Z \tag{2.46}$$

Proof. To distinguish between different sets, we define

$$\mathcal{U}^\delta = Co\{\nu_1^*, \nu_2^*, ..., \nu_r^*\} \tag{2.47}$$

as the convex hull of the extended vertices $\{\nu_1^*, \nu_2^*, ..., \nu_r^*\}$ where

$$\delta = h\left(\mathcal{Z},\mathcal{U}\right) = \frac{|Vert(Z)\backslash V|}{2}\delta_Z(V)$$

. Enumeration over the new vertices, therefore, gives an upper bound on $n\gamma_r$.

$$\mathcal{U} = Co\{\nu_1, \nu_2, ..., \nu_r\} \subseteq \mathcal{U}^\delta = Co\{\nu_1^*, \nu_2^*, ..., \nu_r^*\} \tag{2.48}$$

Assuming

$$Z_x \in \arg\max_{Z \in \mathcal{U}^\delta} \|Z\| \tag{2.49}$$

Therefore

$$\mathcal{U} \subseteq \mathcal{U}^\delta \Rightarrow \max_{Z \in \mathcal{Z}} \|Z\| = \|Z_x - \widehat{Z}_x + \widehat{Z}_x\| \leq \|Z_x - \widehat{Z}_x\| + \|\widehat{Z}_x\| \tag{2.50}$$

Let $\widehat{Z}_x$ is new vertex associate with Z_x in reduced polytope (outcome of enumeration algorithm). The geometric relation between two vertices can be illustrated as in Figure 2.15

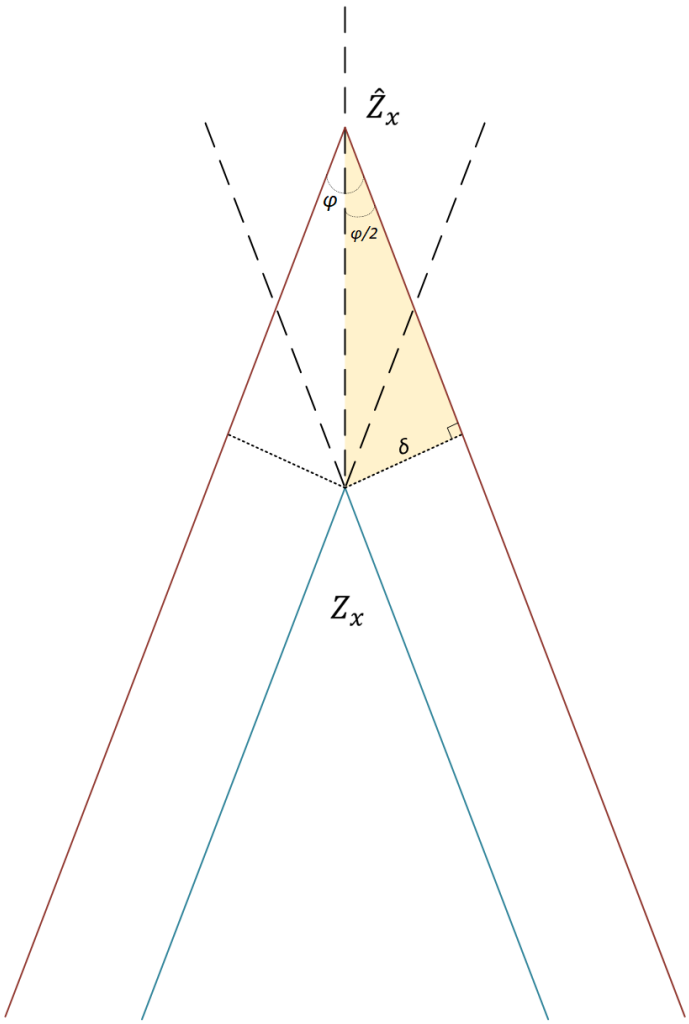

Figure 2.15: The geometric relation between associated vertices in the original Zonotope and reduced polytope

From Figure 2.15, it is easy to see that:

$$\|Z_x - \widehat{Z}_x\| = \frac{\delta}{\sin(\phi/2)} \tag{2.51}$$

Substituting (2.51) into (2.50), gives

$$\max_{Z\in\mathcal{Z}}\|Z\|\leq \frac{\delta}{\sin(\phi/2)}+\|\widehat{Z}_x\|\leq \frac{\delta}{\sin(\phi/2)}+\max_{Z\in\mathcal{U}}\|Z\| \tag{2.52}$$

$$\max_{Z\in\mathcal{Z}}\|Z\|\leq \frac{\delta}{\sin(\phi/2)}+\max_{Z\in\mathcal{U}}\|Z\| \tag{2.53}$$

$$\max_{Z\in\mathcal{Z}}\|Z\|-\max_{Z\in\mathcal{U}}\|Z\|\leq \frac{\delta}{\sin(\phi/2)} \tag{2.54}$$

Now recall the condition for which the gap between QIP and SDR program is zero:

$$\textit{If}\qquad \max_{Z\in\{-1,1\}^n} x'VV'x=\max_{x\in\mathcal{Z}} Z'Z\leq n \quad\Leftrightarrow\quad \gamma\leq\overline{\gamma} \tag{2.55}$$

This condition can be described as:

$$\max_{Z\in\mathcal{Z}}\|Z\|\leq\sqrt{n}\Rightarrow\gamma\leq\overline{\gamma} \tag{2.56}$$

Substituting (2.56) in (2.54), results in condition (2.43) in which the gap can be breached

$$\max_{Z\in\mathcal{U}}\|Z\|+\frac{\delta}{\sin(\phi/2)}\leq\sqrt{n}\Leftrightarrow\delta\leq\sin(\phi/2)\left(\sqrt{n}-\max_{Z\in\mathcal{U}}\|Z\|\right) \tag{2.57}$$

From Lemma 2.2.4 [30], we have

$$\gamma\leq\overline{\gamma}-n(1-\gamma_r)\underline{\lambda}(\Lambda_+)\leq\overline{\gamma}$$

Thus $\gamma_r<\gamma_u\Rightarrow n(1-\gamma_r)>n(1-\gamma_u)$ gives

$$\overline{\gamma}-n(1-\gamma_r)\underline{\lambda}(\Lambda_+)<\overline{\gamma}-n(1-\gamma_u)\underline{\lambda}(\Lambda_+)$$

which proves (2.44) □

The following example demonstrate the proposed method.

2.6.2 Numerical Example 2

A $Q\in\mathbb{R}^{7\times7}$ that the actual solution is known is selected:

$$Q = \begin{bmatrix} 0.7853 & -0.3957 & 0.2214 & -0.2242 & -0.0885 & -0.0952 & -0.4936 \\ -0.3957 & 1.2814 & -0.0511 & -0.1232 & 0.1041 & 0.6171 & 0.1491 \\ 0.2214 & -0.0511 & 0.8411 & 0.0477 & -0.3506 & 0.4144 & -0.5087 \\ -0.2242 & -0.1232 & 0.0477 & 0.7910 & -0.0847 & -0.0046 & 0.1315 \\ -0.0885 & 0.1041 & -0.3506 & -0.0847 & 0.6318 & -0.1260 & 0.3097 \\ -0.0952 & 0.6171 & 0.4144 & -0.0046 & -0.1260 & 1.2128 & -0.1793 \\ -0.4936 & 0.1491 & -0.5087 & 0.1315 & 0.3097 & -0.1793 & 0.8566 \end{bmatrix}$$

The solutiong of QIP for this example is $\gamma = 12.1306$. The Semidefinite Convex upper bound is calculated to be $\overline{\gamma} = 12.8382$ for

$$D = \begin{bmatrix} 1.9481 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1.9890 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2.1172 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.9683 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1.3880 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 2.0213 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 2.4062 \end{bmatrix}$$

which gives

$$D - Q = \begin{bmatrix} V_+ & V \end{bmatrix} \begin{bmatrix} \Lambda_+ & 0 \\ 0 & 0_r \end{bmatrix} \begin{bmatrix} V_+^T \\ V^T \end{bmatrix}$$

where

$$V_+ = \begin{bmatrix} -0.3567 & -0.3902 & -0.5564 & 0.4058 & -0.0334 \\ -0.0455 & 0.2126 & -0.6368 & -0.0884 & -0.0690 \\ -0.4327 & 0.6735 & 0.2152 & 0.2124 & 0.2669 \\ 0.0262 & -0.0356 & -0.1684 & 0.1887 & 0.8577 \\ 0.0637 & 0.1852 & 0.0971 & 0.7908 & -0.3673 \\ 0.0274 & -0.5226 & 0.4109 & 0.2813 & 0.1970 \\ -0.8234 & -0.2008 & 0.1789 & -0.2059 & -0.1165 \end{bmatrix}$$

$$V = \begin{bmatrix} -0.4122 & -0.2744 \\ 0.1381 & 0.7181 \\ -0.4322 & 0.0990 \\ 0.4306 & -0.1143 \\ 0.4379 & 0.0131 \\ -0.2353 & 0.6197 \\ 0.4378 & 0.0452 \end{bmatrix}$$

and

$$\lambda_+ = \begin{bmatrix} 2.0446 & 0 & 0 & 0 & 0 \\ 0 & 1.6893 & 0 & 0 & 0 \\ 0 & 0 & 1.5247 & 0 & 0 \\ 0 & 0 & 0 & 1.0531 & 0 \\ 0 & 0 & 0 & 0 & 0.1264 \end{bmatrix}$$

Since the nullity of $D - Q$ is 2, the equivalent RRQIP will be

$$\overline{\gamma} = \frac{1}{n} \max_{x \in \{-1,1\}^n} x' V V' x =$$

$$\frac{1}{7} \max_{x \in \{-1,1\}^7} x' \begin{bmatrix} 0.2452 & -0.2540 & 0.1510 & -0.1461 & -0.1841 & -0.0730 & -0.1928 \\ -0.2540 & 0.5347 & 0.0114 & -0.0226 & 0.0699 & 0.4125 & 0.0929 \\ 0.1510 & 0.0114 & 0.1966 & -0.1974 & -0.1880 & 0.1631 & -0.1847 \\ -0.1461 & -0.0226 & -0.1974 & 0.1985 & 0.1871 & -0.1722 & 0.1833 \\ -0.1841 & 0.0699 & -0.1880 & 0.1871 & 0.1919 & -0.0949 & 0.1923 \\ -0.0730 & 0.4125 & 0.1631 & -0.1722 & -0.0949 & 0.4394 & -0.0750 \\ -0.1928 & 0.0929 & -0.1847 & 0.1833 & 0.1923 & -0.0750 & 0.1937 \end{bmatrix} x$$

Enumeration over all the $x \in \{-1, 1\}^7$ gives that maximum value of $Z'Z = 6.5373$ at

$$x = \begin{bmatrix} -1 & -1 & -1 & 1 & 1 & -1 & 1 \end{bmatrix}$$

Hence, $\gamma_r := \frac{1}{n} \max_{Z \in (\mathcal{Z})} Z'Z = 0.9339$ which breach the convex upper bound by

$$\gamma < \overline{\gamma} - n(1 - \gamma_r)\underline{\lambda}(\Lambda_+) < \overline{\gamma} \Rightarrow$$

$$\gamma = 12.1306 < 12.8382 - 7 \times (1 - 0.9339) \times 0.1264 = 12.7797 < \overline{\gamma} = 12.8382$$

In next step, the proposed algorithm is applied with value of $\delta_Z = 0.07$ and $\epsilon = 0.04$. The algorithm successfully omitted two vertices, therefore

$$\delta = h\left(\mathcal{Z},\mathcal{U}\right) = \frac{|Vert(Z)\backslash V|}{2}\delta_Z(V) = \frac{2}{2} \times 0.07 = 0.07$$

The randomised algorithm did not omit the maximiser and it is remains as $\max_{Z\in(\mathcal{U})} Z'Z = 6.5373$. It is achieved at $\phi = 136°$, thus checking condition (2.43) gives

$$\sqrt{6.5373} + \frac{0.07}{\sin(136/2)} = 2.6323 < \sqrt{7}$$

The condition is satisfied and hence the gap is breatchable. Although, the randomised did not omit the maximiser of the original Zonotope, the extended polytope pushed the maximiser from $\max_{Z\in(\mathcal{Z})} Z'Z = 6.5373$ to $\max_{Z\in(\mathcal{U})} Z'Z = 6.7308$ which gives $\gamma_u = 0.9615$. Hence

$$\gamma < \overline{\gamma} - n(1-\gamma_u)\underline{\lambda}(\Lambda_+) < \overline{\gamma} \Rightarrow$$
$$\gamma = 12.1306 < 12.8382 - 7 \times (1 - 0.9615) \times 0.1264 = 12.8041 < \overline{\gamma} = 12.8382$$

The following facts can be observed from the example:

- The chosen Q matrix is full rank, however $D - Q$ matrix for the optimal D loses rank. This showing the validity of Lemma 2.2.2.
- Rank$(VV') = 2$, therefore the RRQIP form a 2-dimension Zonotope
- Since $\gamma_r < 1$, The convex upperbound is reduced by $n(1-\gamma_r)\underline{\lambda}(\Lambda_+)$ which shows the validity of Lemma 2.2.4
- Randomised algorithm omits two vertices
- Both omitted vertices have angle wider than $170°$ which contribute less in forming the original Zonotope
- The extended polytope pushes the maximiser from $Z'Z|_{Z\in\mathcal{Z}} = 6.5373$ to $Z'Z|_{Z\in\mathcal{U}} = 6.7308$. Therefore, $\gamma_u = \dfrac{6.7308}{7} < 1$ and the convex upper-bound is breached by $n(1-\gamma_u)\underline{\lambda}(\Lambda_+)$.
- If the Randomised algorithm omits more than two vertices, condition (2.43) will not be satisfied and hence the extended polytope fails to breach the convex bound. This is

due to the fact that the polytope extension, $\frac{|Vert(Z)\backslash V|}{2}$, is depending on the number of omitted vertices. Figure shows the Zonotope of this example where 2 and 4 vertices are omitted. Both polytopes shows good approximation of the general Zonotope. Nevertheless, the polytope with more omitted vertices gives $\gamma_u = 7.1197/7 = 1.017$.

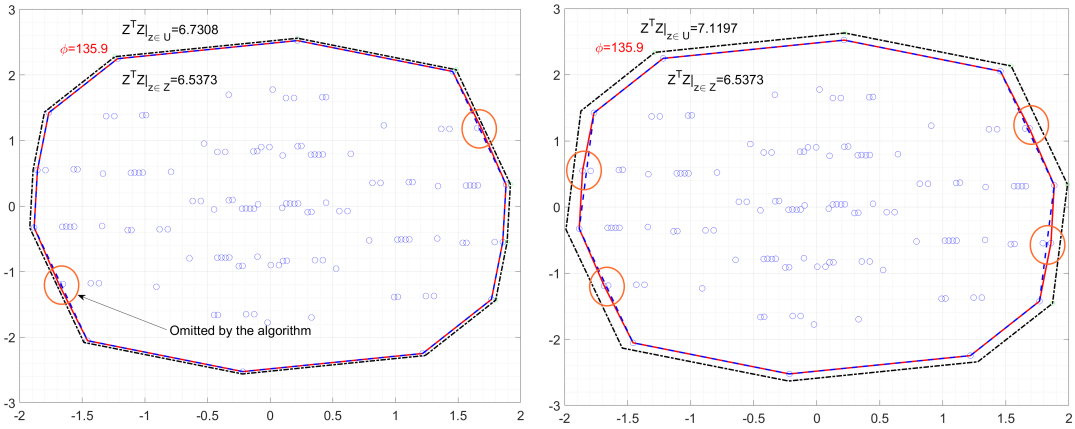

Figure 2.16: Zonotope of Section 2.6.2 example with two approximations

Table 2.1 shows a summary of the above example

Table 2.1: Summary of the numerical experiments 2

Measures	Maximiser	γ	Enumerated vertices	Upper bound
Original Zonotope	6.5373	0.9339	14	12.7797
Randomised polytope	6.5373	0.9339	12	12.7797
Extended polytope 1	6.7308	0.9615	12	12.8041
Extended polytope 2	7.1197	1.0171	10	N/A

2.7 Summary

In this chapter, a randomised algorithm is proposed for calculating an upper bound of the QIP problem. First, a convex relaxation of the problem was defined, the solution of which produces an upper bound to the original problem. The duality gap of the problem, i.e. the distance between the convex upper bound and the optimal solution, can be reduced provided an RRQIP problem can be solved which is equivalent to the full enumeration of the vertexes of a Zonotope. Although this can be achieved by a polynomial-time algorithm, the computation may still be intractable for problems of high-dimensionality. Hence, a randomisation algorithm is presented for breaching the duality gap when the full enumeration is computationally not feasible. It was shown that even with incomplete enumeration

improved probabilistic bounds may be obtained. The bound may be further improved by combining the proposed randomised approach with the solution of a sequence of deterministic QIP problems of increasing rank. This is useful in practice since it can exploit fully the available computational resources to obtain the tightest bound possible with a pre-specified probability.

Chapter 3

Preliminaries: The Complex Structured Singular Value and its Convex Upper Bound

In this chapter we provide an introduction to some of the material needed in the next Chapters. It is assumed that the reader is familiar with matrix theory, linear algebra, control theory and the fundamental concepts of linear systems theory. For more detailed studies, the reader is referred to [6] and [9]. The notation of this chapter is standard and is outlined here for convenience. $\mathbb{R}^n$ ($\mathbb{C}^n$) denote the spaces of n-dimensional real (complex) vectors, respectively. $\mathbb{R}^{n\times m}$ ($\mathbb{C}^{n\times m}$) denote the spaces of $n \times m$ (complex) matrices, respectively. If $z \in \mathbb{C}$ then $\Re(z)$ and $\Im(z)$ are the real and imaginary parts of z, respectively, and $|z|$ is the modulus of z. If $M \in \mathbb{C}^{m\times n}$ then M' denotes the transpose of M and M^* the complex conjugate transpose of M. The singular values of a matrix $M \in \mathbb{C}^{m\times n}$ are denoted as $\sigma_i(M)$, $i = 1, 2, \ldots, \min(m, n)$ and are indexed in non-increasing order of magnitude, i.e. $\sigma_1(M) \geq \sigma_2(M) \geq \ldots \geq \sigma_{\min(m,n)}(M) \geq 0$. $\|M\|$ denotes the spectral norm of M, i.e. the largest singular value of M. If $M \in \mathbb{C}^{n\times n}$ its spectrum, i.e. the set of its eigenvalues is denored as $\sigma(M)$. The eigenvalues of a Hermitian matrix $M \in \mathbb{C}^{n\times n}$, $M = M^*$ are denoted as $\lambda_i(M)$, $i = 1, 2, \ldots, n$ and are indexed in non-increasing order of magnitude, i.e. $\lambda_1(M) \geq \lambda_2(M) \geq \ldots \geq \lambda_n$. We also write $\lambda_1(M) = \lambda_{\max}(M)$. If $\Omega \in \mathbb{R}^{q\times p}$ with $p \geq q$, then the set $\mathcal{Z} = \{\Omega x : x \in [-1\ 1]^p\}$ defines a Zonotope in $\mathbb{R}^q$. If $\{\omega_i\}_{i=1}^p$ are the columns of Ω, this is also written as $\mathcal{Z}(\omega_1, \ldots, \omega_p)$ and the vectors $\{\omega_i\}_{i=1}^p$ are the generators of $\mathcal{Z}$. If $\mathcal{Z}$ is a polytope, then vert($\mathcal{Z}$) denotes the set of vertices of $\mathcal{Z}$ and conv($\mathcal{Z}$) the convex hull of $\mathcal{Z}$. If $\mathcal{Z}$ is a closed set, then int($\mathcal{Z}$) denotes the interior of $\mathcal{Z}$

and $\partial\mathcal{Z} = \mathcal{Z} \setminus \text{int}(\mathcal{Z})$ the boundary of $\mathcal{Z}$. If V is a finite set then $|V|$ denotes the number of elements of V. Further notation is introduces in the chapter when needed.

3.1 Introduction

Real-world systems are normally complex and nonlinear. However, in most industrial cases, simplified linear time-invariant models can sufficiently describe the behaviour of the system around an operating point and hence can be used to specify an appropriate control system. Such models usually contain errors and uncertainties which arise for the various factors [13]. Therefore, it is essential to ensure that the system is not destabilised due to allowable range of perturbations. In other words, the uncertain model of a plant needs to be analysed in order to find the maximum allowable perturbation for which the feedback system remains stable. Consider the feedback loop in Figure 3.1 in which $T \in \mathcal{H}_\infty$, i.e.

$$\|T\|_\infty := \max_{\omega \in \mathbb{R}} \overline{\sigma}\left[T(j\omega)\right]$$

and $\Delta \in \mathcal{H}_\infty$ is the plant structured uncertainty perturbations which are added according to the multiplicative perturbation in the system. The uncertainty matrix in a real system is in the form of diagonal block structure. An example at the end of this section illustrates the uncertainty matrix for a real system. Here T represents the nominal closed-loop system incorporating the nominal plant and the designed feedback controller.

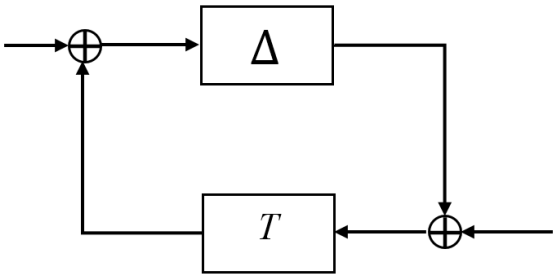

Figure 3.1: Problem of Close loop Stability

The problem is to find the minimum $\mathcal{H}_\infty$ norm of Δ for which the feedback loop is de-stablised. It can be shown from the Small Gain Theorem [69], that for any $\gamma > 0$, the interconnected system in Figure 3.1 is internally stable for all Δ with

- $\|\Delta\|_\infty \leq \frac{1}{\gamma}$ if and only if $\|T(s)\|_\infty < \gamma$
- $\|\Delta\|_\infty < \frac{1}{\gamma}$ if and only if $\|T(s)\|_\infty \leq \gamma$

The value of $r = \frac{1}{\gamma}$ is called the robust stability radius of T and is denoted by $r_\Delta(T)$. If $\Delta(s)$ represent an unstructured perturbation, then we can construct a specific $\Delta(s)$ of $r_\Delta(T) = \frac{1}{\gamma}$ such that the feedback system is unstable.

The problem of maximising the robust stability radius of systems subject to unstructured uncertainty has been widely investigated in the literature [15], [3]. Lehtomaki et. al, in [16] consider the problem of finding the robust stability radius when the perturbation is bounded away from the unstructured stability radius in the most vulnerable direction. The method in [16] is used to obtain a bound on the structured singular value.

Assuming Δ to be a full matrix will result in a relatively conservative bound. Considering a structured matrix instead has led to developing $\mu -$ analysis and synthesis techniques. In recent years, the structured singular value (μ) which is a generalisation of the singular value of a matrix has been the focus of significant research interest as an effective analytical and design tool in the area of robust control. The structured singular value was introduced by Doyle in [17] and by Safonov in [18]. Since the computation of μ is an NP-hard problem in the general case [18], the primary approach will become computing reliable bounds on $\mu_\Delta(M)$. The synthesis problem is also defined as the problem of finding a controller K that maximises the stability radius.

Equivalently, the objective of μ-synthesis problem is to find a stabilising controller in order to minimise the structured singular value of a system with respect to uncertainty in the form of a block diagonal structure Δ [14]. Unfortunately, the general solution to this problem is numerically intractable [5]. Nevertheless, solving special classes of μ problems can be relatively easy. Young in [70] proposed a complete solution for Rank-One problem. Smith in [71], proposed a low-rank algorithm to solve a special case of a rank-2 problem. Gungah et al. in [21], obtained bounds on μ by embedding the underlying block-structured uncertainty set within a larger set. Later on, they introduced a new upper bound on structured singular value [8]. Although the algorithm offered good performance results, only the complex structured singular value under mixed-type structured uncertainty was considered in this method.

Malik in [20] introduce a tighter bound on the structured singular value which has been developed by [9] and has resulted in a detailed study on the gap between the complex structured singular value and its convex upper bound [9]. In that paper, new necessary and sufficient conditions were derived for the absence of a duality gap. The study has also shown that if an upper bound on this reduced rank problem can be obtained, it can provide an upper bound on the original problem that is lower than the convex relaxation upper bound.

As a practical example, a Four-wheel steering (4WS) vehicle closed-loop control system can be described in Figure3.2, where $G_f(s)$ and $G_r(s)$ are the functions of velocity and mass which may vary under different running conditions[2].

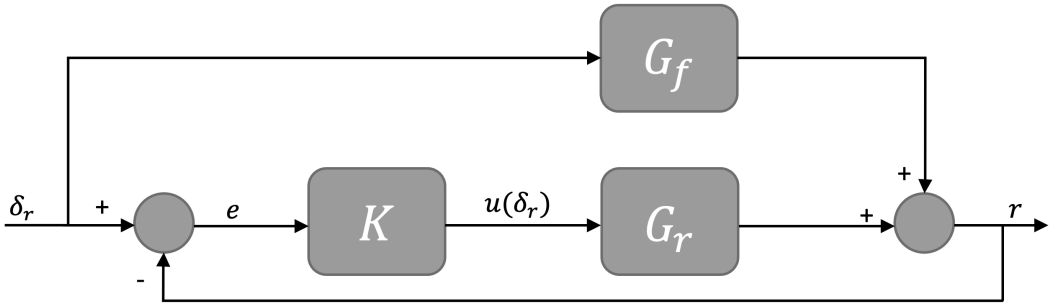

Figure 3.2: 4WS vehicle closed-loop control system[2]

The aim is to find a robust controller K, to stabilise the system in the presence of the uncertainties. Hence, to expand the model based on μ framework, two perturbation blocks Δ_f and Δ_r are added to the control loop. Figure 3.3 shows the 4WS vehicle model in the μ framework where $W_s = \text{diag}[W_{sf}, W_{sr}]$ is a frequency dependant weighting function matrix to represent all possible structured perturbations. Rearranging the feedback system

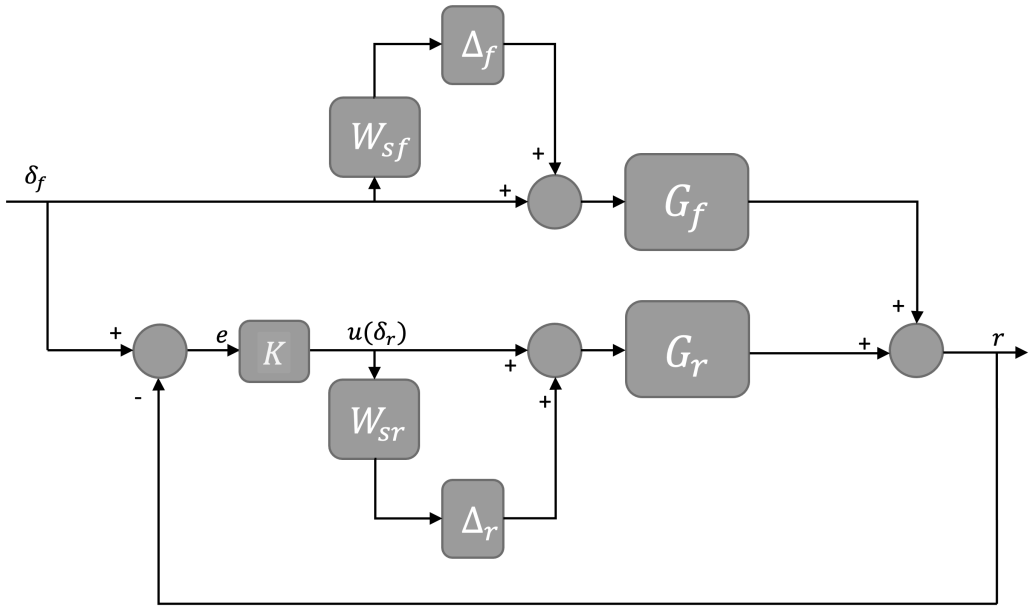

Figure 3.3: 4WS vehicle closed-loop control system in μ framework[2]

in Figure 3.3 leads to the general structure shown in Figure 3.4, where (u, r) are control and measure signals, (δ_f, e_r, e_β) constitute the performance variables and (d_1, d_2, z_f, z_r) connect the system $\boldsymbol{P}$ to perturbation $\boldsymbol{\Delta}$.

Compare to Figure 3.1

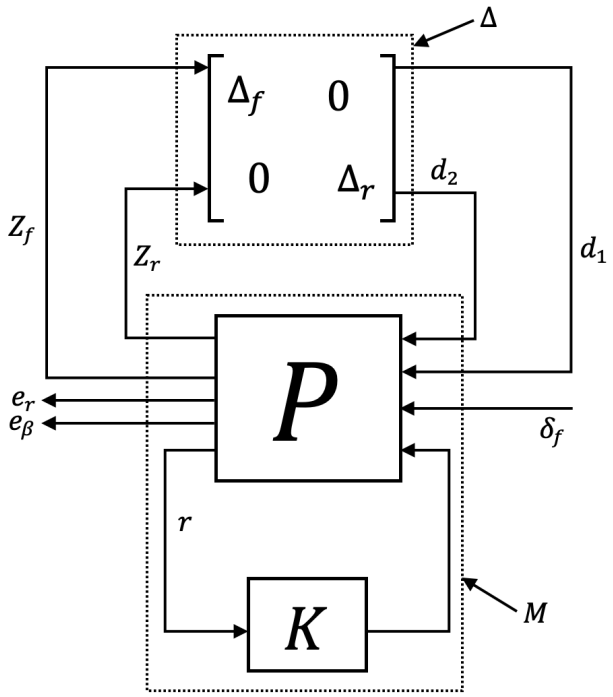

Figure 3.4: 4WS vehicle generalized system[2]

$$\mathbf{\Delta} = \text{diag}[\Delta_f, \Delta_r]$$

and the control system is stable when $\|\mathbf{\Delta}\|_\infty \leq 1$. This example shows how a real system can be formulated in μ framework.

In the following paragraph, we give a short description of the theory of the structured singular value.

Let $M \in \mathbb{C}^{n\times n}$ be a non-singular matrix. Define block-diagonal structured uncertainty set $\mathbf{\Delta}$ as:

$$\begin{aligned}\mathbf{\Delta} =&\{diag\left[\delta_1 I_{r1}, ..., \delta_s I_{rs}, \Delta_{s+1}, ..., \Delta_{s+f}\right] : \\ &\delta_i \in \mathbb{C}, \Delta_{s+j} \in \mathcal{C}^{m_j \times m_j}, 1 \leq i \leq s \leq j \leq f\}\end{aligned} \tag{3.1}$$

where $\sum_{i=1}^{s} r_i + \sum_{j=1}^{f} m_j = n$.
Let $\mathbf{B_\Delta}$ denote the unit ball of $\mathbf{\Delta}$ were

$$\mathbf{B_\Delta} = \{\Delta \in \mathbf{\Delta} : \overline{\sigma}(\Delta) \leq 1\}$$

Definition 3.1.1. *For $M \in \mathbb{C}^{n\times n}$, the Structured Singular Value of M corresponding to the*

structured set $\mathbf{\Delta}$ *is*

$$\mu_{\Delta}(M) := \frac{1}{\min\{\|\Delta\| : \Delta \in \mathbf{\Delta}, det(I - M\Delta) = 0\}} \tag{3.2}$$

If $I - M\Delta$ is nonsingular for every $\Delta \in \mathbf{\Delta}$ then $\mu_{\mathbf{\Delta}}(M) = 0$.

The general real structured singular value (μ) problem is known to be NP-hard. i.e. any algorithm to compute μ will fail to find the answer in polynomial time [9]. Hence, in practical applications, upper or lower bounds are introduced for special structured of $\mathbf{\Delta}$ which makes it possible to find a solution to the problem or derive countable upper bound. Such bounds are often sufficient for many applicable cases.

The following lemma shows the most constrained upper bound when Δ is allowed to be a full block, $\Delta = \mathbb{C}^{n\times n}$; and the most constrained lower bound when Δ is a diagonal matrix.

Lemma 3.1.1. *[6] For every* $M \in \mathbb{C}^{n\times n}$,

$$\rho(M) \leq \mu_{\Delta}(M) \leq \overline{\sigma}(M)$$

Proof. First take $\Delta = \{\delta I_n : \delta \in \mathbb{C}\}$, then

$$\mu_{\Delta}(M) = \max_{|\delta|\leq 1} \rho(\delta M) = \max_{|\delta|\leq 1} |\delta|\rho(M) = \rho(M)$$

Now take $\Delta = \Delta_1 = \mathbb{C}^{n\times n}$ and let $M = U\Sigma V^*$ be singular value decomposition where $UV^* = V^*U = I$ and $\Sigma = diag(\sigma_1, ..., \sigma_n)$ where $\sigma_1 \geq \cdots \geq \sigma_n$ are the n singular values. Thus

$$\det(I - \Delta M) = 0 \Rightarrow \det(I - \Delta U\Sigma V^*) = 0$$

This gives $\det(V^*U - V^*\Delta U\Sigma) = 0$. Since $V^*\Delta U \in \mathbf{\Delta}$ and $\|V^*\Delta U\| = \|\Delta\|$ then

$$\det\left(\Sigma^{-1} - \Delta\right) = 0$$

or equivalently

$$\det\left(\begin{bmatrix} \frac{1}{\sigma_1} & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \frac{1}{\sigma_n} \end{bmatrix} - \Delta\right) = 0$$

It is obvious that $\Delta = \Delta_2 = diag(1/\sigma_1, 0, \ldots, 0)$ has the smallest $\|\Delta\|$ for which $\det\left(\Sigma^{-1} - \Delta\right) = 0$. Hence for any given structured Δ in (3.1)

$$\Delta_2 \subseteq \Delta \subseteq \Delta_1 \Rightarrow \rho(M) \leq \mu_\Delta(M) \leq \overline{\sigma}(M)$$

□

Note 3.1.1. *The function $\mu : \mathbb{C}^{n\times n} \to \mathbb{R}$ is not a norm, as it does not satisfy the triangular inequality.*

Note 3.1.2. *The function $\mu : \mathbb{C}^{n\times n} \to \mathbb{R}$ is continuous.*

The bounds of

$$\rho(M) \leq \mu_\Delta(M) \leq \overline{\sigma}(M) \tag{3.3}$$

provide little information about the value of μ as the gap between them can be arbitrary large[6]. Hence, a transformation can be introduced on M which change the bound but not affect the $\mu_\Delta(M)$. Thus two subsets of $\mathbb{C}^{n\times n}$ can be defines

$$\mathbf{Q} = \{Q \in \mathbf{\Delta} : Q^*Q = I_n\} \tag{3.4}$$

$$\mathbf{D} = \{diag\left[D_1, ..., D_S, d_{S+1}I_{m1}, ..., d_{S+F}I_{mF}\right] : D_i \in \mathcal{C}^{r_i\times r_i}, D_i = D_i^* > 0, d_{s+j} > 0\} \tag{3.5}$$

For any $\Delta \in \mathbf{\Delta}, Q \in \mathbf{Q}$ and $D \in \mathbf{D}$,

$$Q^* \in \mathbf{Q}, Q\Delta \in \mathbf{\Delta}, \Delta Q \in \mathbf{\Delta}$$
$$\overline{\sigma}(Q\Delta) = \overline{\sigma}(\Delta Q) = \overline{\sigma}(\Delta)$$
$$D^{1/2}\Delta = \Delta D^{1/2}$$

Theorem 3.1.1. *[6] For all $Q \in \boldsymbol{Q}$ and $D \in \boldsymbol{D}$*

$$\mu_\Delta(MQ) = \mu_\Delta(QM) = \mu_\Delta(M) = \mu_\Delta\left(D^{1/2}MD^{-1/2}\right) \tag{3.6}$$

Proof. For all $D \in \mathbf{D}$ and $\Delta \in \mathbf{\Delta}$,

$$\det(I - M\Delta) = \det\left(I - MD^{-1/2}\Delta D^{1/2}\right) = \det\left(I - D^{1/2}MD^{-1/2}\Delta\right)$$

Since D commutes with Δ, $\mu_{\Delta}(M) = \mu_{\Delta}\left(D^{1/2}MD^{-1/2}\right)$. On the other hand, since $\Delta Q^* = \Delta$, then $\|\Delta Q^*\| = \|\Delta\|$. Therefore,

$$I - \Delta M = I - \Delta Q^* Q M \Rightarrow \mu_{\Delta}(M) = \mu_{\Delta Q^*}(QM) = \mu_{\Delta}(QM)$$

Same argument can be used for MQ.

□

Hence, the bound on (3.3) can be tightened to

$$\max_{Q \in \mathbf{Q}} \rho(QM) \leq \mu_{\Delta}(M) \leq \inf_{D \in \mathbf{D}} \overline{\sigma}\left(D^{1/2}MD^{-1/2}\right) \tag{3.7}$$

Without loss of generality, we assume that D is normalised ($d_{F+S} = 1$), Hermitian and positive definite. These conditions will not affect the infimum. The advantage of such upper bound in (3.7) is that with a certain convexity properties, this upper bound is computationally achievable [6]. In general

$$\begin{aligned}
&\overline{\sigma}\left(D^{1/2}MD^{-1/2}\right) < \beta \\
&\Leftrightarrow \overline{\lambda}\left(D^{-1/2}M^*D^{1/2}D^{1/2}MD^{-1/2}\right) < \beta^2 \\
&\Leftrightarrow D^{-1/2}M^*D^{1/2}D^{1/2}MD^{-1/2} - \beta^2 I < 0 \\
&\Leftrightarrow M^*DM - \beta^2 D < 0
\end{aligned}$$

The last equation is a Linear Matrix Inequality (LMI) which is clearly a convex condition for a given $M \in \mathbb{C}^{n \times n}$ and $\beta > 0$ in a scaling set $\boldsymbol{D}$. In the next section we describe Linear Fractional Transformations as a practical application of structured singular value in control theory.

3.1.1 Linear Fractional Transformations

Linear Fractional Transformations (LFTs) is a class of general linear feedback loops which are widely used in the application of Structured Singular value in control theory. To define LFTs, let $M \in \mathbb{C}^{n \times n}$ is partitioned as

$$M = \begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{bmatrix}$$

and let $\Delta_2 \in \boldsymbol{\Delta_2}$ be compatible is size with M_{22}. Consider the control block diagram

in Figure 3.5 with the following corresponding equations

$$\begin{bmatrix} e \\ z \end{bmatrix} = \begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{bmatrix} \begin{bmatrix} d \\ w \end{bmatrix} \tag{3.8}$$

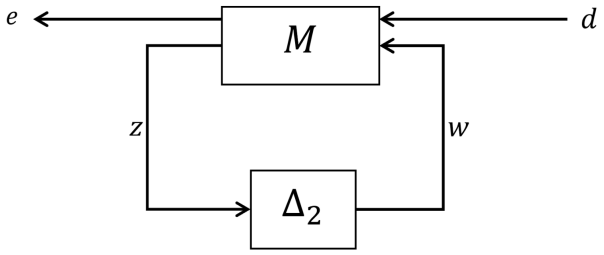

Figure 3.5: Linear Fraction Transformation

Solving equation (3.8) for z and e, we have

$$\begin{aligned} e &= M_{11}d + M_{12}w \\ z &= M_{21}d + M_{22}w \\ &\Rightarrow \begin{cases} z = (I - M_{22}\Delta_2)^{-1} M_{21} d \\ e = \left(M_{11} + M_{12}\Delta_2 (I - M_{22}\Delta_2)^{-1} M_{21}\right) d \end{cases} \end{aligned}$$

It is obvious that these equation for e and z are well-posed if and only if $(I - M_{22}\Delta_2)^{-1}$ exists.

3.2 Structured distance to singularity

The distance to singularity, $\gamma_\Delta(M)$, is defined as the inverse of structured singular value in definition 3.1.1. Note that both structured singular value and structured distance to singularity are identical notions which describe the behaviour of a control system in the face of uncertainty; therefore, we may use either of these definitions in this work. To formulise, $\gamma_\Delta(M)$, we first start with the most general form of notation structured uncertainty and establish a connection with the structured distance to singularity. The content of the following section summarises the results of references [6] and [9]. However, we expand the proofs and highlight the results that are extensively used in the following chapters. First, consider the following definition which is another formulation of the structured distance to singularity.

Definition 3.2.1. *Let A be*

$$A = diag(a_1, a_2, ..., a_n) \in \mathbb{C}^{n\times n} \tag{3.9}$$

where $a_1 > a_2 > ... > a_n$ and let $\zeta = \begin{bmatrix} I_m & 0_{m\times(n-m)} \end{bmatrix}^*$. For any set $\Delta_{11} \subseteq \mathbb{C}^{m\times m}$define the structured distance of A to singularity

$$\gamma_{\Delta_{11}} = \min\left\{\|\Delta\| : \det(A-\Delta) = 0, \zeta^*\Delta\zeta \in \mathbf{\Delta_{11}}\right\} \tag{3.10}$$

and the set of all optimal structured rank reducing perturbations

$$\mathcal{D}_{\Delta_{11}} := \left\{\Delta \in \mathbb{C}^{n\times n} : \|\Delta\| = 0, \det(A-\Delta) = 0, \zeta^*\Delta\zeta \in \mathbf{\Delta_{11}}\right\} \tag{3.11}$$

In this notation the (unstructured) distance to singularity will be denoted as

$$\begin{aligned}\gamma_{\mathbb{C}^{m\times m}} &= \min\left\{\|\Delta\| \in \mathbb{C}^{n\times n} : \det(A-\Delta) = 0, \zeta^*\Delta\zeta \in \mathbb{C}^{m\times m}\right\} \\ &= \min\left\{\|\Delta\| : \det(A-\Delta) = 0\right\}\end{aligned}$$

Note that the above definition is identical to the definition of matrix $\mu_\Delta(M)$. i.e.

$$\det(I-\Delta M) = 0 \Rightarrow \det(I-\Delta U\Sigma V^*) = 0 \Rightarrow \det(I - V^*\Delta U\Sigma) = 0 \Rightarrow \det\left(\Sigma^{-1} - V^*\Delta U\right) = 0 \tag{3.12}$$

where

$$M = U\Sigma V^*, \quad \Sigma = diag(\sigma_1, ..., \sigma_n), \quad \sigma_1 > .. > \sigma_n$$

is singular value decomposition. Since $UV^* = V^*U = I$, then

$$V^*\Delta U \in \mathbf{\Delta}, \quad \text{and} \quad \|V^*\Delta U\| = \|\Delta\|$$

Therefore by taking $A = \Sigma^{-1}$ and considering $\gamma_\Delta = \mu_\Delta^{-1}$, equation (3.2) can be reduced to (3.10); hence from now on we replace term $\det(I-\Delta M)$ by $\det(A-\Delta)$. Moreover, since for any $\alpha \in \mathbb{C}$, $\mu(\alpha M) = |\alpha|\mu(M)$ we can always normalise A to its largest element,a_1, thus simplifying it to

$$A = \Sigma^{-1} = diag(A_1, A_2), \quad \text{where} \quad A_1 = I_m, \text{ and }, A_2 = diag(a_{m+1}, ..., a_n).$$

where $1 < a_{m+1} \leq ... \leq a_n$.

Note that m in above equation is the multiplicity of the smallest singular value of A.

3.2.1 Lehtomaki's Result

Before we continue, we present a well-known theorem to calculate a bound on $\gamma_\Delta(M)$. This is in fact the main result of Lehtomaki et al at [16].

It is known that if $\|\Delta\| = 1/\sigma_{min} = a_1$, then $A + \Delta$ could be singular; and if $\|\Delta\| < 1/\sigma_{min}$ then $A + \Delta$ is non-singular [9]. In fact, the Singular Decomposition of Matrix A is

$$A = U\Sigma V^* = [u_1, u_2, ..., u_n] \begin{bmatrix} a_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & a_n \end{bmatrix} \begin{bmatrix} v_1^* \\ v_2^* \\ \vdots \\ v_n^* \end{bmatrix} = \sum_{i=1}^{n} a_i u_i v_i^*$$

where $a_1 < a_2 < \ldots < a_n$. For an arbitrary matrix Δ

$$\Delta = \sum_{i=1}^{n} \sum_{j=1}^{n} \langle u_i v_j^*, \Delta \rangle u_i v_j^*$$

where the inner product for matrix A, B is defined by $\langle A, B \rangle = tr(A^* B)$

The matrix $\langle u_i v_j^*, \Delta \rangle u_i v_j^*$ is the projection of Δ into the subspace spanned by $u_i v_i^*$ which has magnitude $|\langle u_i v_j^*, \Delta \rangle|$
If $A + \Delta$ is singular then

$$\det(A + \Delta) = 0 \Rightarrow \det(U\Sigma V^* + \Delta) = 0 \Rightarrow \det(\Sigma + V\Delta U^*) = 0$$

Now let $\tilde{\Delta} = V\Delta U^*$ then

$$\det(diag(a_1, a_2, ..., a_n) - \tilde{\Delta}) = 0$$

Selecting $\Delta = \begin{bmatrix} a_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & 0 \end{bmatrix}$ gives $\tilde{\Delta} = V\Delta U^* = a_1 v_1 u_1^* = \begin{bmatrix} a_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & 0 \end{bmatrix}$ and thus $\min \|\Delta\| = a_1$.

Now, if the projection of Δ in the most sensitive direction, i.e. $|u_1 \Delta v_1^*|$, is constrained to have a magnitude less than $\sigma_{min}(A) = a_1$, $A+\Delta$ cannot be singular along this direction and thus $\|\Delta\|$ needs to increase to make $A + \Delta$ singular [16]. This means that by constraining the uncertainty we can achieve a larger stability radius. Now the question is that how big can the perturbation be in this condition to make $A + \Delta$ singular. Therefore, the problem can be defined as

$$\gamma_\Delta(M) = \min \left\{ \|\Delta\| \colon \Delta \in \mathbb{C}^{n\times n}, |v_1^* \Delta u_1| \leq \phi < a_1, \phi \in \mathbb{R}, \det(A - \Delta) = 0 \right\} \quad (3.13)$$

Thus, we impose a constraint such that Δ cannot have a projection of magnitude a_1 in the most sensitive direction $u_1 v_1^*$. In other words $A + \Delta$ cannot become singular along the direction of $u_1 v_1^*$ and hence $\|\Delta\|$ have to increase in order to make $A + \Delta$ singular.

This problem has been evaluated in details by Lehtomaki et al at [16]. The main result of this section can be summarised via the following Theorem

Theorem 3.2.1. *[16] The optimiser of the problem* (3.13) *is given by* $\|\Delta\| = \sqrt{a_1 a_2 + \phi(a_1 - a_2)}$. *The set of* Δ *associated to the optimiser is also given by*

$$\Delta = U \left[\begin{array}{c:cc} P_s & 0 & \\ \hdashline 0 & \phi & \gamma \\ & \gamma^* & -\phi \end{array}\right] V^* \quad (3.14)$$

where P_s *is arbitrary and*

$$\|P_s\| \leq \sqrt{a_1 a_2 + \phi(a_1 - a_2)} = \|\Delta\| \quad (3.15)$$

and

$$\gamma = \sqrt{(\phi + a_2)(a_1 - \phi)} e^{j\theta}, \quad (3.16)$$

where θ *is arbitrary.*

Proof. The result is first established for the 2×2 matrix case and then generalised for general $n \times n$ matrices

Let $A = diag[a_1 \quad a_2]$ and assume $\Delta = \Delta^*$, thus

$$\Delta = \begin{bmatrix} a & b \\ b^* & d \end{bmatrix}$$

Then $\|\Delta\|$ is given by

$$\|\Delta\| = |\frac{a+d}{2}| + \sqrt{\left(\frac{a+d}{2}\right)^2 + |b|^2}$$

Using the fact that $A + \Delta$ is rank deficient $|b|^2 = (a_1 - a)(a_2 - d)$ and thus

$$\|\Delta\| = |\frac{a+d}{2}| + \sqrt{\left(\frac{a+d}{2}\right)^2 + (a_1 - a)(a_2 - d)}$$

Taking partials of $\|\Delta\|$ with respect to a and d, we have

$$\frac{\partial\|\Delta\|}{\partial a} = 1/2\left[sgn(a+d) + z_1\right], \quad \frac{\partial\|\Delta\|}{\partial d} = 1/2\left[sgn(a+d) + z_2\right]$$

Here, $sgn(\bullet)$ is sign function and

$$z_1 = \frac{\left(\frac{a+d}{2}\right) + a_2}{\sqrt{\left(\frac{a-d}{2}\right)^2 + (a_1 + a)(a_2 + d)}}$$

$$z_1 = \frac{\left(\frac{a+d}{2}\right) + a_1}{\sqrt{\left(\frac{a-d}{2}\right)^2 + (a_1 + a)(a_2 + d)}}$$

It follows via a straightforward calculation that for $|d| < a_2$ and $|a| < a_2$, $|z_1| > 1$ and $|z_2| < 1$. Thus $\frac{\partial\|\Delta\|}{\partial d}$ has the same sign as $(a+d)$, hence the global minimum is at $a = -d$. On the other hand, since$|z_2| > 1$, $\frac{\partial\|\Delta\|}{\partial a}$ is always positive, thus the minimum is at $d = -\phi$. Therefore, the optimal Δ has the form

$$\Delta = \begin{bmatrix} \phi & \gamma \\ \gamma^* & -\phi \end{bmatrix}$$

where $\gamma = \sqrt{(a_1 + \phi)(a_2 - \phi)}e^{j\theta}$, θ arbitrary.The proof for the structure of the optimum Δ in the general case can be found in [16]. Here we only calculate the optimal norm. This

is given by

$$\|\Delta\| = \sqrt{\lambda_{max}\left(\Delta^*\Delta\right)} = \left(\lambda_{max}\left[\begin{array}{c|cc} P_sP_s^* & \multicolumn{2}{c}{0} \\ \hline 0 & \phi(a_1-a_2)+a_1a_2 & 0 \\ & 0 & \phi(a_1-a_2)+a_1a_2 \end{array}\right]\right)^{1/2}$$

Thus

$$\|\Delta\|_2 = \|\begin{bmatrix} \phi(a_1-a_2)+a_1a_2 & 0 \\ 0 & \phi(a_1-a_2)+a_1a_2 \end{bmatrix}\|$$

This is equivalent to

$$\lambda_{max}(P_sP_s^*) \leq \lambda_{max}\begin{bmatrix} \phi(a_1-a_2)+a_1a_2 & 0 \\ 0 & \phi(a_1-a_2)+a_1a_2 \end{bmatrix}$$

Hence

$$\|\Delta\| = \sqrt{\lambda_{max}\left(\Delta^*\Delta\right)} = \left(\lambda_{max}\begin{bmatrix} \phi(a_1-a_2)+a_1a_2 & 0 \\ 0 & \phi(a_1-a_2)+a_1a_2 \end{bmatrix}\right)^{1/2} = \sqrt{a_1a_2+\phi(a_1-a_2)}$$

□

Note that the tightest result is achieved when $\phi = 0$ which gives $\|\Delta\| = \sqrt{a_1a_2}$. We can now generalise the Lehtomaki's idea of constraining the 1×1 entry of Δ to the case where the leading $m \times m$ block of a diagonal matrix Δ, i.e. Δ_{11} is constrained or is of different classes of uncertainty. Next we evaluate $\gamma_{\Delta_{11}}$ and $\mathcal{D}_{\Delta_{11}}$ for different sets of Δ_{11}.

3.2.2 A class of uncertainty where $\mathbf{\Delta_{11}} = \mathbb{C}$.

First we consider the simplest form of Δ_{11} where $m = 1$ and it is free to take any complex value.i.e.

$$\Delta = \left[\begin{array}{c|c} \delta_{11} & 0 \\ \hline 0 & \Delta_{22} \end{array}\right]$$

so that

$$\det(A-\Delta) = 0 \Leftrightarrow \det\left(\left[\begin{array}{c|c} a_{11} & 0 \\ \hline 0 & A_{22} \end{array}\right] - \left[\begin{array}{c|c} \delta_{11} & 0 \\ \hline 0 & \Delta_{22} \end{array}\right]\right) = 0$$

It is clear that $a_{11} = \delta_{11}$ gives $\det(A - \Delta) = 0$. Also $\|\Delta\| = \max\{\delta_{11}, \|\Delta_{22}\|\}$. Hence, taking $\|\Delta_{22}\|$ and $\delta_{11} = \underline{\sigma}(A) = a_{11}$ gives the minimiser

$$\gamma_{\mathbb{C}} = \min\{\|\Delta\| : \det(A - \Delta) = 0\} = \underline{\sigma}(A) = 1$$

and the set of all optimal rank reducing perturbations is given by

$$\mathcal{D}_{\mathbb{C}} = \left\{\Delta = diag(1, \Delta_{22}) \in \mathbb{C}^{n\times n}, \Delta_{22} \in \mathbb{C}^{(n-1)\times(n-1)}, \|\Delta_{22}\| \leq 1\right\}$$

Note that, if $\|\Delta\| = \sigma_{min}(A)$ then $A + \Delta$ is singular.

3.2.3 A class of uncertainty where $\mathbf{\Delta_{11}} = \{\delta \in \mathbb{C} : |\delta| \leq \phi\}, 0 \leq \phi < 1$

As mentioned earlier, if the perturbation of the $(1,1)$ entry is confined to be less than or equal to ϕ where $0 \leq \phi < 1$, $\|\Delta\|$ must increase to make $A + \Delta$ singular. This is the Lehtomaki result that has been reviewed in 3.2.1 and can be summarised to the following theorem.

Theorem 3.2.2. *[16] Let $A = diag(1, a_2, a_3, ..., a_n)$ with $1 < a_2 < ... < a_n$ be given and define $\zeta = \begin{bmatrix} 1 & 0_{1,n-1} \end{bmatrix}^*$ and $\mathbf{\Delta}_\phi = \{\delta \in \mathbb{C} : |\delta| \leq \phi\}$ for any $0 \leq \phi \leq 1$.Then*

$$\gamma_{\mathbf{\Delta}_\phi} := \min\{\|\Delta\| : det(A - \Delta) = 0, \zeta^*\Delta\zeta \in \mathbf{\Delta}_\phi\} = \sqrt{a_2 - \phi(a_2 - 1)}$$

and for $\phi < 1$, the set of all optimal rank reducing perturbation is given by

$$\mathcal{D}_{\mathbf{\Delta}_\phi} = \left\{ diag\left(\begin{bmatrix} \phi & \gamma_{\mathbf{\Delta}_\phi} e^{j\theta} \\ \gamma_{\mathbf{\Delta}_\phi} e^{-j\theta} & -\phi \end{bmatrix}\right), \Delta_{33} : \theta \in \mathbb{R}, \Delta_{33} \in \mathbb{C}^{(n-2)\times(n-2)}, \|\Delta_{33}\| \leq \gamma_{\mathbf{\Delta}_\phi} \right\}$$

Proof. Same as Theorem 3.2.1 by setting $a_1 = 1$. □

3.2.4 A class of uncertainty where $\mathbf{\Delta_{11}} = \mathbb{C}^{m\times m}$

To take one step further, we assume. $\Delta_{11} = \mathbb{C}^{m\times m}$. This gives generalisation of Theorem 3.2.2 for a case that $m > 1$. It has been shown in [9] that,

Lemma 3.2.1. *For A and ζ defined previously,*

$$\gamma_{\Delta\in\mathbb{C}^{m\times m}} = min\left\{\|\Delta\| : det\left(A-\Delta\right)=0\right\} = \sigma_{min}(A) = 1 \tag{3.17}$$

and the set of all optimal Δ is given by

$$\mathcal{D} = \left\{ \begin{bmatrix} W & 0 \\ 0 & I_{n-m} \end{bmatrix} \left[\begin{array}{cc|c} 1 & 0 & 0 \\ 0 & \Delta_{22} & \Delta_{23} \\ \hline 0 & \Delta_{32} & \Delta_{33} \end{array}\right] \begin{bmatrix} W^* & 0 \\ 0 & I_{n-m} \end{bmatrix} \in \mathbb{C}^{n\times n} \right\} \tag{3.18}$$

where

$$WW^* = W^*W = I_m, \text{ and }, \| \begin{bmatrix} \Delta_{22} & \Delta_{23} \\ \Delta_{32} & \Delta_{33} \end{bmatrix} \| \leq 1$$

Proof. See Appendix A.1

□

3.2.5 A class of uncertainty where $\mathbf{\Delta_{11}} = \{0_{m\times m}\}$

Similar to the condition when $\Delta_{11} \in \mathbb{C}$, we violate the constrain on Δ_{11} and assume that $\|\Delta_{11}\|$ cannot reach 1. We start with the simplest condition where $\Delta_{11} = \mathbf{0}$.

Lemma 3.2.2. *[9]*

Let $A = diag(A_1, A_2) \in \mathbb{R}^{n\times n}$ with

$$A_1 = \left[\begin{array}{c|c} I_{m1} & 0 \\ \hline 0 & A_{22} \end{array}\right] \in \mathbb{R}^{m\times m}$$

and

$$A_2 = \left[\begin{array}{c|c} a_3 I_{m3} & 0 \\ \hline 0 & A_{44} \end{array}\right] \in \mathbb{R}^{(n-m)\times(n-m)}$$

Assume that $1 < \underline{\sigma}(A_{22}), 0 < a_3 < \underline{\sigma}(A_{44})$, and $1 < a_3$. Then the structured distance to singularity is

$$\gamma_{0_{m\times m}} := \min\left\{\|\Delta\| : det\left(A-\Delta\right)=0, \zeta^*\Delta\zeta = 0\right\} = \sqrt{a_3} =: \sqrt{\underline{\sigma}(A_1)\underline{\sigma}(A_2)} \tag{3.19}$$

Furthermore all optimal rank reducing perturbations are generated by

$$W \left[\begin{array}{ccc||ccc} 0 & 0 & 0 & \sqrt{a_3} & 0 & 0 \\ 0 & 0 & 0 & 0 & \Delta_{13} & \Delta_{14} \\ \hline 0 & 0 & 0 & 0 & \Delta_{23} & \Delta_{24} \\ \hline\hline \sqrt{a_3} & 0 & 0 & 0 & 0 & 0 \\ 0 & \Delta_{31} & \Delta_{32} & 0 & \Delta_{33} & \Delta_{34} \\ \hline 0 & \Delta_{41} & \Delta_{42} & 0 & \Delta_{43} & \Delta_{44} \end{array}\right] W^* = W\overline{\Delta}W^* \tag{3.20}$$

where $W = diag(W_1, I_{m_2}, W_3, I_{m_4}) \in \mathbb{C}^{n\times n}$ *is unitary and*

$$\left\| \begin{bmatrix} 0 & 0 & \Delta_{13} & \Delta_{14} \\ 0 & 0 & \Delta_{23} & \Delta_{24} \\ \Delta_{31} & \Delta_{32} & \Delta_{33} & \Delta_{34} \\ \Delta_{41} & \Delta_{42} & \Delta_{43} & \Delta_{44} \end{bmatrix} \right\| \leq \sqrt{a_3} \tag{3.21}$$

Proof. See Appendix A.2 □

3.2.6 A class of uncertainty where $\mathbf{\Delta_{11}} \in \mathbb{C}^{m\times m}, \|\mathbf{\Delta_{11}}\| \leq 1$

Next we assume that $\Delta_{11} \in \mathbb{C}^{m\times m}$, i.e. Δ_{11} is a fixed complex-valued matrix with $\|\Delta_{11}\| \leq 1$.

The following lemma, from [9], summarises the optimal solution in this case.

Lemma 3.2.3. *[9] For matrices A and ζ, as defined previously, let $\Delta_{11} \in \mathbb{C}^{m\times m}$ has the same structure as section 3.2.6 where $\|\Delta_{11}\| \leq 1$ and det $(I - \Delta_{11}) \neq 0$. Then*

$$\min_{\substack{det(A-\Delta)=0 \\ \zeta^*\Delta\zeta=\Delta_{11}}} \|\Delta\| = \min_{\substack{\|\left(\gamma^2 I-\Delta_{11}\right)(I-\Delta_{11})^{-1}\|=a_{m+1} \\ \gamma>1}} \gamma \tag{3.22}$$

Proof. See Appendix A.3 □

Remark 3.2.1. *The optimisation on the right-hand side of (3.22) is an eigenvalue problem, as the following argument shows. Setting $\gamma^2 = 1 + \zeta, \zeta > 0$, gives*

$$
\begin{aligned}
(\gamma^2 I - \Delta_{11})(I - \Delta_{11})^{-1} &= ((1+\zeta)I - \Delta_{11})\,(I - \Delta_{11})^{-1} \\
&= (I + \zeta I - \Delta_{11})\,(I - \Delta_{11})^{-1} \\
&= (\zeta I + (I - \Delta_{11}))\,(I - \Delta_{11})^{-1} \\
&= \zeta (I - \Delta_{11})^{-1} + I \\
&\Rightarrow \|\gamma^2 I - \Delta_{11}(I - \Delta_{11})^{-1}\| = \|\zeta (I - \Delta_{11})^{-1} + I\|
\end{aligned}
$$

On the other hand

$$
\min \gamma = \min \sqrt{1+\zeta} = \sqrt{1 + \min \zeta}
$$

hence,

$$
\min_{\substack{\|(\gamma^2 I - \Delta_{11})(I-\Delta_{11})^{-1}\| = a_{m+1} \\ \gamma > 1}} \gamma = \sqrt{1 + \min_{\substack{\|I + \zeta (I-\Delta_{11})^{-1}\| = a_{m+1} \\ \zeta > 0}} \zeta} \tag{3.23}
$$

Note that

$$
a_{m+1} = \|I + \zeta\,(I - \Delta_{11})^{-1}\| \Rightarrow a_{m+1}^2 = \|I + \zeta\,(I - \Delta_{11})^{-1}\|^2
$$

Let $I + \zeta\,(I - \Delta_{11})^{-1} = \theta$, thus

$$
a_{m+1}^2 = \|\theta\|^2 \Rightarrow a_{m+1}^2 = \lambda_{max}\,(\theta\theta^*) \Rightarrow a_{m+1}^2 - \lambda_{max}\,(\theta\theta^*) = 0
$$

this can be express as

$$
\lambda_{min}\left(a_{m+1}^2 I - \theta\theta^*\right) = 0 \Rightarrow
$$

which gives

$$
\begin{aligned}
&\lambda_{min}\left(a_{m+1}^2 I - \left[I + \zeta\,(I - \Delta_{11})^{-1}\right]\left[I + \zeta\,(I - \Delta_{11})^{-1}\right]^*\right) = 0 \Rightarrow \\
&\lambda_{min}\left(a_{m+1}^2 I - \left[I + \zeta\,(I - \Delta_{11})^{-1}\right]\left[I + \zeta\,(I - \Delta_{11}^*)^{-1}\right]\right) = 0
\end{aligned}
$$

simple calculation gives

$$\lambda_{min}\left[a_{m+1}^2 I - I - \zeta\left(I-\Delta_{11}\right)^{-1} - \zeta\left(I-\Delta_{11}^*\right)^{-1} - \zeta^2\left(I-\Delta_{11}\right)^{-1}\left(I-\Delta_{11}^*\right)^{-1}\right] \Rightarrow$$
$$\lambda_{min}\left[I - \frac{\zeta(I-\Delta_{11})^{-1}}{a_{m+1}^2-1} - \frac{\zeta(I-\Delta_{11}^*)^{-1}}{a_{m+1}^2-1} - \frac{\zeta^2\left(I-\Delta_{11}\right)^{-1}\left(I-\Delta_{11}^*\right)^{-1}}{a_{m+1}^2-1}\right] = 0 \Rightarrow$$

Factorizing ζ gives

$$\lambda_{min}\left[\zeta^{-1}I - \frac{(I-\Delta_{11})^{-1}}{a_{m+1}^2-1} - \frac{(I-\Delta_{11}^*)^{-1}}{a_{m+1}^2-1} - \frac{\zeta\left(I-\Delta_{11}\right)^{-1}\left(I-\Delta_{11}^*\right)^{-1}}{a_{m+1}^2-1} -\right] = 0 \tag{3.24}$$

Using the matrix identity

$$\det\left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right] = \det(D)\det(A - BD^{-1}C)$$

It follows that

$$\det\left[\begin{array}{c|c} A & B \\ \hline B^* & \alpha I \end{array}\right] = \det(\alpha I)\det(A - B(\alpha I)^{-1}B^*) = \alpha.\det(A - \alpha^{-1}BB^* \text{ for any } \alpha \in \mathbb{R}$$

Comparing with equation(3.24) and taking

$$A = \zeta^{-1}I - \frac{(I-\Delta_{11})^{-1}}{a_{m+1}^2-1} \quad , B = -\frac{(I-\Delta_{11})^{-1}}{\sqrt{a_{m+1}^2-1}} \quad , D = \zeta^{-1}I$$

gives

$$\lambda_{min}\left[\zeta^{-1}I - \frac{(I-\Delta_{11})^{-1}}{a_{m+1}^2-1} - \frac{(I-\Delta_{11}^*)^{-1}}{a_{m+1}^2-1} - \frac{\zeta\left(I-\Delta_{11}\right)^{-1}\left(I-\Delta_{11}^*\right)^{-1}}{a_{m+1}^2-1}\right] = 0$$

which is equivalent to

$$\lambda_{min}\left[\begin{array}{cc} \zeta^{-1}I - \dfrac{(I-\Delta_{11})^{-1} - (I-\Delta_{11}^*)^{-1}}{a_{m+1}^2-1} & \dfrac{(I-\Delta_{11})^{-1}}{\sqrt{a_{m+1}^2-1}} \\ \dfrac{(I-\Delta_{11}^*)^{-1}}{\sqrt{a_{m+1}^2-1}} & \zeta^{-1}I \end{array}\right] = 0$$

or

$$\lambda_{min}\left[\begin{bmatrix}\zeta^{-1}I & 0\\ 0 & \zeta^{-1}I\end{bmatrix}-\begin{bmatrix}\dfrac{(I-\Delta_{11})^{-1}-(I-\Delta_{11}^{*})^{-1}}{a_{m+1}^{2}-1} & -\dfrac{(I-\Delta_{11})^{-1}}{\sqrt{a_{m+1}^{2}-1}}\\ -\dfrac{(I-\Delta_{11}^{*})^{-1}}{\sqrt{a_{m+1}^{2}-1}} & 0\end{bmatrix}\right]=0$$

Since $\zeta > 0$,

$$min\{\zeta^{-1}\}-\lambda_{max}\begin{bmatrix}\dfrac{(I-\Delta_{11})^{-1}-(I-\Delta_{11}^{*})^{-1}}{a_{m+1}^{2}-1} & -\dfrac{(I-\Delta_{11})^{-1}}{\sqrt{a_{m+1}^{2}-1}}\\ -\dfrac{(I-\Delta_{11}^{*})^{-1}}{\sqrt{a_{m+1}^{2}-1}} & 0\end{bmatrix}=0$$

Thus minimum in (3.23) is given by ζ_0, where

$$\zeta_0^{-1}=\lambda_{max}\begin{bmatrix}\dfrac{(I-\Delta_{11})^{-1}}{a_{m+1}^{2}-1} & -\dfrac{(I-\Delta_{11})^{-1}}{\sqrt{a_{m+1}^{2}-1}}\\ -\dfrac{(I-\Delta_{11}^{*})^{-1}}{\sqrt{a_{m+1}^{2}-1}} & 0\end{bmatrix}$$

To summarise, this section has focused on the problem with the following form

$$\gamma=\min_{\substack{\det(A-\Delta)=0\\ \Delta \text{ is a given diagonal positive definite matrix}\\ \text{Constraint on } \Delta}}\|\Delta\| \tag{3.25}$$

Different constraints were imposed on Δ to find optimal distance to singularity. Table 3.1 gives a summary of the main results:

Table 3.1: Summary the optimum γ for different class of uncertainties

Constraint on Δ	Optimum γ
No Constraint	a_1
$\delta_{11}=0$	$\sqrt{a_1a_2}$
$\lvert\delta_{11}\le\phi_1<a_1\rvert$	$\sqrt{a_1a_2-\phi_1(a_2-a_1)}$
$\Delta_{11}\in\mathbb{C}^{m\times m}$	a_1
$\Delta_{11}=\{0_{m\times m}\}$	$\sqrt{\underline{\sigma}(A_1)\underline{\sigma}(A_2)}$
$\Delta_{11}\in\mathbb{C}^{m\times m},\|\Delta_{11}\|\le 1$	answer of eigen value problem in (3.22)

One of the main result of this Chapter is Lemma 3.2.3 which reduces to eigenvalue problem in (3.22). This problem is used in next chapter to obtain a graphical method for achieving an upper bound on μ. Next, we propose several algorithms of varying computa-

tional complexity for calculating upper bounds of the structured singular value (equivalently lower bounds of the structured distance to singularity) of a matrix M subject to real parametric uncertainty. Our approach is based on the projection of the uncertainty set in the most critical direction along the lines summarised in the present Chapter.

3.3 Summary

This chapter is preliminary to chapter 4 where μ analysis is briefly described. Several definitions, equations, theorems and proofs are mentioned in this chapter. The results of this chapter was based on two main references, [6] and [9].

Chapter 4

Structured singular value of matrices with real parametric uncertainty: Deterministic and Probabilistic algorithms

The structured singular value (μ) is a powerful tool for measuring the stability margins of a system subject to parametric uncertainty. In robust control, μ can be used to establish conditions under which robust stability of feedback systems is guaranteed when the system is subject to structured or norm-bounded uncertainties. Combined with $\mathcal{H}_\infty$ optimal control, it provides a framework for robust control synthesis.

When real scalar perturbations describe the uncertainty, we have a real μ-problem. The solution of the general real (μ) problem is known to be NP-hard, i.e. one can always construct examples in which any algorithm used to compute μ fails to perform the calculation in polynomial time [9]. Hence, in practical applications, upper and lower bounds are often sought, or, alternatively, specialised algorithms are tuned to the specific class of problems under consideration. The most straightforward case is when uncertainty is unstructured, i.e. when the uncertain set $\mathbf{\Delta}$ consists of a single (norm-bounded) matrix block. In this case the structured singular value coincides with the standard singular value and can be easily calculated. Another category of problems is when μ coincides with its convex upper bound (problems with μ-simple structure). However, this class is too limited in practice to address problems of realistic complexity. For problems of medium or high complexity, obtaining a tight upper bound on μ is often needed to guarantee robust control design margins with

a minimal degree of conservativeness. In this chapter, structured singular value problems with real scalar (possibly repeated) uncertainty is considered. By adapting previous work results in this area, we obtain several deterministic and probabilistic bounds on μ. Thus, our approach addresses the trade-offs between accuracy and algorithmic complexity explicitly and is easily adaptable to the needs of practical control design problems.

The chapter extends the methodology of [9] to the case of real-parametric uncertainty. Reference [9] considered the complex $\mu(M)$ problem. It aimed to derive conditions for breaching the convex upper bound of $\mu(M)$, normally obtained via convex programming or the so-called D-iteration algorithm. It was shown that this bound could be breached if an auxiliary μ problem of a low-rank matrix could be solved (or a sufficiently tight bound could be obtained).

The approach followed in this chapter proceeds by characterising the projection of the uncertainty set in the most critical direction corresponding to unstructured perturbations. This direction is defined by the pair of singular vectors (u, v) corresponding to the maximum singular value of M. In the unstructured case, all minimum-norm singularising perturbations Δ have a projection of magnitude $\sigma_{\min}(M)$ in the worst-case direction. If the magnitude of the projection of the perturbation in this direction is not allowed to reach $\sigma_{\min}(M)$, i.e. $|(uv^*, \Delta)| \leq \phi\sigma_{\min}(M)$, $0 \leq \phi < 1$, the norm of nearest destabilising perturbations must increase for $I - \Delta M$ to lose rank. This constrained distance to singularity problem can be solved exactly in terms of the constraint parameter ϕ and the two smaller singular values of M. Note that imposing a constraint on the projection of Δ in the direction $(uv^*, \cdot)$ defines an implicit "structure". By embedding the actual (diagonal) uncertainty structure as tightly as possible within this implicit structure allows for the derivation of an upper bound on $\mu(M)$ with respect to the diagonal structure of real perturbations (equivalently a lower bound on the structured distance to singularity).

We consider two types of relaxation. The first bounds the magnitude of the projected uncertainty set. To obtain the tightest possible embedding within the artificial uncertainty structure we need to maximise a convex multivariable quadratic function with box constraints, for which the maximum occurs on the boundary of the constraint set. Several relaxation methods can be applied to this optimisation for problems of high dimensionality. The second relaxation method is computationally more demanding but produces a tighter bound. In this case we do not only want to calculate the maximum magnitude of the projected set in the most critical direction but to characterise fully the projected set and its boundary in the complex plane. In the standard real μ problem the uncertain parameters

enter linearly and the projected set takes the form of a Zonotope. It can be shown [8] that the calculation of the tightest lower bound to the structured distance to singularity reduces to a geometric problem, in particular the calculation of the minimum value of a parameter for which the intersection of an Apollonius circle and the Zonotope in the complex plane is a non-empty set. This can be obtained by enumerating the vertices of the Zonotope, a task that can be achieved efficiently using the "reverse enumeration" algorithm [42],[29].

Despite the efficient computational properties of this algorithm (in terms of speed and memory storage requirements) the full enumeration of the vertices becomes impractical for high dimensional problems. In this case a randomised enumeration algorithm may be used as an alternative. This generates sequential Zonotope vertices by random sampling the standard two-dimensional Gaussian distribution. The full Zonotope is approximated by the convex hull of the set of partially enumerated vertices that are generated by the algorithm. The method also produces a probabilistic upper bound of the Hausdorff distance between the original Zonotope and the convex polytope which approximates it. Embedding this approximating polytope within a broader convex set of minimum volume (using the estimated Hausdorff distance) allows us to conclude (with a specific high probability controlled by the designer) that this also contains the original Zonotope. Finally, calculating again the minimum value of the parameter for which the intersection of the Apollonius circle and the extended polytope is non-empty, allows us to obtain a lower bound of the structured distance to singularity (this time probabilistic).

An appealing aspect of our methodology is that it can be easily extended (at least conceptually) to structured perturbation problems with "correlated" parameters or, in general, to problems in which parameters enter the uncertainty set described in a nonlinear way. What is required in this case is the description of the boundary of the uncertainty set projected in a low dimensional space. We illustrate this method with a simple example. Of course, in the nonlinear case, the projected uncertainty set is not expected to be a Zonotope in general or even convex. We believe that this approach is promising for establishing robust stability conditions in nonlinear problems of this type and is worth investigating further in future work. For problems of high dimensionality in which the projected set boundary is challenging to determine deterministically, probabilistic estimation methods that involving random sampling techniques may be a viable alternative.

The structure of this chapter is as follows: Section 4.1 gives several definitions related to the structured singular value and the structured distance to singularity for real parametric diagonal uncertainty structures. In addition, the two basic relaxations techniques introduced

in [9] are reviewed and specialised to the real uncertainty structure considered here. The results of this section apply to the general case for which the multiplicity of the larger singular value m is an arbitrary integer. The generic case $m = 1$ is analysed in Section 4.2. This allows for a more concrete set of algorithms to be formulated and solved based on an interesting geometric interpretation of the problem. Extensions of the proposed methodology to correlated parameters and nonlinear uncertainty models are also presented here. Several algorithmic implementation aspects are discussed in Section 4.3. Section 4.4 describes a randomised algorithm for the partial enumeration of the Zonotope's vertices along with an estimation of the Hausdorff distance between the Zonotope and the convex hull of the polytope corresponding to the reduced vertices obtained from the randomisation algorithm. This can be used to obtain a probabilistic lowed bound on the structured distance to singularity. In section 4.5, a numerical example is presented to illustrate the performance of the algorithms described in previous sections.

The notation of this chapter is standard and is outlined here cor convenience. $\mathbb{R}^n$ ($\mathbb{C}^n$) denote the spaces of n-dimensional real (complex) vectors, respectively. $\mathbb{R}^{n\times m}$ ($\mathbb{C}^{n\times m}$) denote the spaces of $n \times m$ (complex) matrices, respectively. If $z \in \mathbb{C}$ then $\Re(z)$ and $\Im(z)$ are the real and imaginary parts of z, respectively, and $|z|$ is the modulus of z. If $M \in \mathbb{C}^{m\times n}$ then M' denotes the transpose of M and M^* the complex conjugate transpose of M. The singular values of a matrix $M \in \mathbb{C}^{m\times n}$ are denoted as $\sigma_i(M)$, $i = 1, 2, \ldots, \min(m, n)$ and are indexed in non-increasing order of magnitude, i.e. $\sigma_1(M) \geq \sigma_2(M) \geq \ldots \geq \sigma_{\min(m,n)}(M) \geq 0$. $\|M\|$ denotes the spectral norm of M, i.e. the largest singular value of M. If $M \in \mathbb{C}^{n\times n}$ its spectrum, i.e. the set of its eigenvalues is denored as $\sigma(M)$. The eigenvalues of a Hermitian matrix $M \in \mathbb{C}^{n\times n}$, $M = M^*$ are denoted as $\lambda_i(M)$, $i = 1, 2, \ldots, n$ and are indexed in non-increasing order of magnitude, i.e. $\lambda_1(M) \geq \lambda_2(M) \geq \ldots \geq \lambda_n$. We also write $\lambda_1(M) = \lambda_{\max}(M)$. If $\Omega \in \mathbb{R}^{q\times p}$ with $p \geq q$, then the set $\mathcal{Z} = \{\Omega x : x \in [-1\ 1]^p$ defines a Zonotope in $\mathbb{R}^q$. If $\{\omega_i\}_{i=1}^p$ are the columns of Ω, this is also written as $\mathcal{Z}(\omega_1, \ldots, \omega_p)$ and the vectors $\{\omega_i\}_{i=1}^p$ are the generators of $\mathcal{Z}$. If $\mathcal{Z}$ is a polytope, then vert($\mathcal{Z}$) denotes the set of vertices of $\mathcal{Z}$ and conv($\mathcal{Z}$) the convex hull of $\mathcal{Z}$. If $\mathcal{Z}$ is a closed set, then int($\mathcal{Z}$) denotes the interior of $\mathcal{Z}$ and $\partial\mathcal{Z} = \mathcal{Z} \setminus \text{int}(\mathcal{Z})$ the boundary of $\mathcal{Z}$. If V is a finite set then $|V|$ denotes the number of elements of V. Further notation is introduces in the chapter when needed.

4.1 Structured Singular Value for a real set of uncertainty

First, we describe the problem using some basic definitions from chapter 3. Let $M \in \mathbb{C}^{n\times n}$ be a square non-singular value with singular value decomposition

$$M = U\Sigma V^* = \left[\begin{array}{c|c} V_1 & V_2 \end{array}\right] \left[\begin{array}{c|c} I_m & \mathbf{0} \\ \hline \mathbf{0} & \Sigma_2 \end{array}\right] \left[\begin{array}{c} V_1^* \\ V_2^* \end{array}\right] \tag{4.1}$$

where

$$\Sigma = diag(\sigma_{m+1}, \sigma_{m+2}, \ldots, \sigma_n), \quad 1 > \sigma_{m+1} \geq \sigma_{m+2} \geq \cdots \geq \sigma_n > 0. \tag{4.2}$$

and $V^*V = U^*U = I_n$.

Note that the largest singular value is assumed equal to one. This involves no loss of generality and can be achieved for any non singular matrix M by the transformation $M/\sigma_1 \to M$. We also define

$$A = \Sigma^{-1} = \text{diag}(1, \ldots, 1, a_{m+1}, \ldots, a_n) \tag{4.3}$$

with $1 < a_{m+1} \leq \ldots \leq a_n$ and $a_i = \sigma_i^{-1}$ for $i = m+1, \ldots, n$.

We define $\mathbf{\Delta} \subseteq \mathbb{C}^{n\times n}$ to be a uncertainty set which defines the "structure" of the model. At present we consider $\mathbf{\Delta}$ be the set of (possibly repeated) real diagonal perturbations:

$$\mathbf{\Delta} = \left\{\text{diag}\left(\delta_1 I_{k_1} \ldots, \delta_p I_{k_p}\right), \delta_i \in \mathbb{R},\ i = 1, 2, \ldots, p\right\},\ \sum_{i=1}^{p} k_i = n \tag{4.4}$$

The structured distance to singularity of $M \in \mathbb{C}^{n\times n}$ with respect to $\mathbf{\Delta}$ is defined as

$$\gamma_\Delta(M) := \min\left\{\|\Delta\| : \Delta \in \mathbf{\Delta}, \det(I - M\Delta) = 0\right\} \tag{4.5}$$

Similarly, the structured singular value with respect to $\mathbf{\Delta}$, $\mu_{\mathbf{\Delta}}(M)$, is defined as the inverse of $\gamma_\Delta(M)$ (provided the inverse exists), i.e.

$$\mu_\Delta(M) = \gamma_\Delta(M)^{-1} := \frac{1}{\min\left\{\|\Delta\| : \Delta \in \mathbf{\Delta}, \det(I - M\Delta) = 0\right\}} \tag{4.6}$$

If $\det(I - M\Delta) \neq 0$ for every $\Delta \in \mathbf{\Delta}$, then $\mu_\Delta(M) = 0$. The following Lemma gives an alternative formulation of $\gamma_\Delta(M)$ which is used later.

Lemma 4.1.1. *[8] Let $M = U\Sigma V^*$ and A be as defined in equations (4.1) and (4.3),*

respectively, and let $\mathbf{\Delta}$ *be the structured uncertainty set defined in equation (4.4). Then:*

$$\gamma_{\mathbf{\Delta}}(M) := \min\{\gamma : E \in \boldsymbol{\mathcal{E}}, \textit{det}\,(A - \gamma E) = 0\} \tag{4.7}$$

where

$$\boldsymbol{\mathcal{E}} = \{V^*\Delta U : \Delta \in \mathbf{B}_{\mathbf{\Delta}}\} \quad \textit{and}\,; \mathbf{B}_{\mathbf{\Delta}} = \{\Delta \in \mathbf{\Delta} : \|\Delta\| \leq 1\}.$$

Proof. Equation (4.5) can be written as:

$$\begin{aligned}\gamma_{\Delta}(M) &:= \min\{\|\Delta\| : \Delta \in \mathbf{\Delta}, \det(I - M\Delta) = 0\} \\ &:= \min\{\|\Delta\| : \Delta \in \mathbf{\Delta}, \det(I - U\Sigma V^*\Delta) = 0\} \\ &:= \min\{\|\Delta\| : \Delta \in \mathbf{\Delta}, \det(A - V^*\Delta U) = 0\}\end{aligned} \tag{4.8}$$

where $A = \Sigma^{-1} = diag(1, \ldots, 1, a_{m+1}, \ldots, a_n)$ where $1 < a_{m+1} \leq \ldots, a_n$ and $a_i = 1/\sigma_i$ for $i = m+1, \ldots, n$

Hence

$$\gamma_{\Delta}(M) := \min\{\|\Delta\| : \Delta \in \mathbf{\Delta}, \det(A - V^*\Delta U) = 0\} \tag{4.9}$$

Now let

$$\mathbf{B}_{\mathbf{\Delta}} = \{\Delta \in \mathbf{\Delta} : \|\Delta\| \leq 1\}$$

be the unit ball and define

$$\boldsymbol{\mathcal{E}} = \{V^*\Delta U : \Delta \in \mathbf{B}_{\mathbf{\Delta}}\} \tag{4.10}$$

$\boldsymbol{\mathcal{E}}$ can also be written as

$$\boldsymbol{\mathcal{E}} = V^*\mathbf{B}_{\mathbf{\Delta}}U = \mathbf{B}_{V^*\mathbf{\Delta}U}$$

This follows from the fact that

$$\|V^*\Delta U\| = \|\Delta\| \text{ for all } \Delta \in \mathbf{\Delta}$$

Taking $\|V^*\Delta U\| = \|\Delta\| = \gamma$, equation (4.8) can be reformulated as

$$\begin{aligned}\gamma_{\Delta}(M) &:= \min\{\|\Delta\| : \Delta \in \mathbf{\Delta}, \det(I - M\Delta) = 0\} \\ &= \min\{\|\Delta\| : \Delta \in \mathbf{\Delta}, \det(A - V^*\Delta U) = 0\} \\ &= \min\left\{\|\Delta\| : \Delta \in \mathbf{\Delta}, \det\left(A - \|\Delta\|\frac{V^*\Delta U}{\|\Delta\|}\right) = 0\right\} \\ &= \min\{\gamma : E \in \boldsymbol{\mathcal{E}}, \det(A - \gamma E) = 0\}\end{aligned} \tag{4.11}$$

□

Note that if set $\boldsymbol{\mathcal{E}}$ has complex structure, calculation of the minimum in equation (4.11) may be difficult. However, a lower bound of γ_Δ could be easy to obtain if we embed $\boldsymbol{\mathcal{E}}$ within a conveniently-chosen super-set $\boldsymbol{\mathcal{E}_1}$. In this case:

$$\begin{aligned}\gamma_{\boldsymbol{\mathcal{E}_1}}(M) &:= \min\{\gamma : E \in \boldsymbol{\mathcal{E}_1}, \det(A-\gamma E)=0\} \\ &\leq \min\{\gamma : E \in \boldsymbol{\mathcal{E}}, \det(A-\gamma E)=0\} \\ &= \gamma_{\boldsymbol{\Delta}}(M)\end{aligned} \tag{4.12}$$

provided $\boldsymbol{\mathcal{E}_1} \supseteq \boldsymbol{\mathcal{E}}$. Thus, $\gamma_{\boldsymbol{\mathcal{E}_1}}(M)$ is a lower bound on $\gamma_{\boldsymbol{\Delta}}$. Equivalently $\gamma^{-1}_{\boldsymbol{\mathcal{E}_1}} = \mu_{\boldsymbol{\mathcal{E}_1}}$ is an upper bound on $\mu_\Delta(M)$.

Note that $\boldsymbol{\mathcal{E}_1}$ can be chosen as an arbitrary superset of $\boldsymbol{\mathcal{E}}$. At this point, $\boldsymbol{\mathcal{E}_1}$ is chosen to facilitate the process of finding an upper bound of $\mu_{\boldsymbol{\Delta}}(M)$:

First Relaxation: First let $\boldsymbol{\mathcal{E}_1}$ have the following structure

$$\boldsymbol{\mathcal{E}_1} = \left\{E \in \mathbb{C}^{n\times n} : Z^*EZ \in \boldsymbol{\mathcal{E}_{11}}\right\} \tag{4.13}$$

where $\boldsymbol{\mathcal{E}_{11}} = \{V_1^*\Delta U_1 : \Delta \in \mathbf{B}_{\boldsymbol{\Delta}}\}$ and $Z^* = \left[I_m | \mathbf{0}_{m\times(n-m)}\right]$.

Note further that if $E \in \boldsymbol{\mathcal{E}}$ then $E = V^*\Delta U$ for some $\Delta \in \mathbf{B}_{\boldsymbol{\Delta}}$ and

$$Z^*EZ = Z^*V^*\Delta UZ = \left[\begin{array}{c|c} I_m & 0 \end{array}\right] \left[\begin{array}{c} V_1^* \\ \hline V_2^* \end{array}\right] \Delta \left[\begin{array}{c|c} U_1 & U_2 \end{array}\right] \left[\begin{array}{c} I_m \\ \hline 0 \end{array}\right] = V_1^*\Delta U_1 \in \boldsymbol{\mathcal{E}_{11}}$$

Thus $E \in \boldsymbol{\mathcal{E}} \Rightarrow E \in \boldsymbol{\mathcal{E}}_1$ and hence $\boldsymbol{\mathcal{E}} \subseteq \boldsymbol{\mathcal{E}}_1$. For simplicity, we denote Z^*EZ by E_{11}.

We partition V_1 and U_1 compatibly with $\boldsymbol{\Delta}$, i.e.

$$V_1^* = \left[\begin{array}{cccc} V_{11}^* & V_{22}^* & \dots & V_{1p}^* \end{array}\right] \in \mathbb{C}^{m\times n}, \quad U_1 = \left[\begin{array}{c} U_{11} \\ U_{22} \\ \vdots \\ U_{p1} \end{array}\right] \in \mathbb{C}^{n\times m}$$

where $V_{1j}^* \in \mathbb{C}^{m\times k_j}$ and $U_{j1} \in \mathbb{C}^{k_j\times m}$ for $j = 1, 2, \dots, p$, then for $\Delta \in \boldsymbol{\Delta}$,

$$V_1^*\Delta U_1 = \begin{bmatrix} V_{11}^* & V_{22}^* & \dots & V_{1p}^* \end{bmatrix} \begin{bmatrix} \delta_1 I_{k1} & 0_{k_1,k_2} & \dots & 0_{k_1,k_p} \\ 0_{k_2,k_1} & \delta_2 I_{k_2} & \dots & 0_{k_2,k_p} \\ \vdots & \vdots & \ddots & \vdots \\ 0_{k_p,k_1} & 0_{k_2,k_2} & \dots & \delta_p I_{kp} \end{bmatrix} \begin{bmatrix} U_{11} \\ U_{22} \\ \vdots \\ U_{p1} \end{bmatrix}$$

Equivalently,

$$V_1^*\Delta U_1 = \sum_{i=1}^{p} \delta_i (V_{1j}^*.U_{j1}) = \sum_{i=1}^{p} \delta_i E_i, \quad E_i = V_{1j}^*.U_{i1} \in \mathbb{C}^{m\times m} \tag{4.14}$$

for $i = 1, 2, \dots, p$. Therefore, the lower bound of $\gamma_{\mathbf{\Delta}}(M)$ is given by:

$$\gamma_{\boldsymbol{\varepsilon_1}} = \min\left(\gamma : \det(A - \gamma E) = 0, E \in \boldsymbol{\mathcal{E}}_1\right) \tag{4.15}$$

where

$$\begin{aligned} \boldsymbol{\mathcal{E}_1} = \{ \left[\begin{array}{c|c} E_{11} & E_{12} \\ \hline E_{21} & E_{22} \end{array}\right] : E_{11} = \sum_{i=1}^{p} \delta_i E_i, \quad -1 \le \delta_i \le 1, i = 1, \dots, p \\ E_{12} \in \mathbb{C}^{m\times(n-m)}, E_{21} \in \mathbb{C}^{(n-m)\times m}, E_{22} \in \mathbb{C}^{(n-m)\times(n-m)} \ \} \end{aligned} \tag{4.16}$$

$E_{12} \in \mathbb{C}^{m\times(n-m)}$, $E_{21} \in \mathbb{C}^{(n-m)\times m}$ and $E_{22} \in \mathbb{C}^{(n-m)\times(n-m)}$ are free matrices of the indicated dimensions and $\mathcal{E}_1$ contains all the matrices of the form

$$\mathcal{E}_1 = \left\{ \left[\begin{array}{c|c} \sum_{i=1}^{p} \delta_i E_i & * \\ \hline * & * \end{array}\right] \in \mathbb{C}^{n\times m} : \quad |\delta_i| \le 1, i = 1, 2, \dots, p \right\} \tag{4.17}$$

Next let $\delta_i = \delta_i^0, \quad i = 1, 2, \dots, p$ be real fixed values such that $|\delta_i^0| \le 1$ and consider the minimisation:

$$\beta(\delta_1^0, \dots, \delta_p^0) = \min\left\{ \|\Delta\| : \det(A - \Delta) = 0, |\delta_i^0| \le 1, \quad Z^*\Delta Z = \sum_{i=1}^{p} \delta_i^0 E_i \right\} \tag{4.18}$$

It follows from (4.15) that:

$$\gamma_{\boldsymbol{\varepsilon}_1}(M) = \min\{\beta(\delta_1, \dots, \delta_p) : \delta_i \in \mathbb{R}, -1 \le \delta_i \le 1, i = 1, 2, \dots, p\} \tag{4.19}$$

The solution to problem (4.18) and subsequently problem (4.19) is provided later in this section. First we need the following three Lemmas:

Lemma 4.1.2. *[9] Let A be as defined in equation (4.3) and $\Psi \in \mathbb{C}^{m\times m}$ such that $\|\Psi\| \leq 1$ and $1 \notin \sigma(\Psi)$ Then:*

$$\beta := \min_{\substack{det(A-\Delta)=0 \\ \Delta_{11}=\Psi}} \|\Delta\| = \min_{\substack{\|\left(\gamma^2 I-\Psi\right)(I-\Psi)^{-1}\|=a_{m+1} \\ \gamma>1}} \gamma \tag{4.20}$$

where Δ_{11} is the leading $m \times m$ block of Δ. Further,

$$\beta^2 - 1 = \min\left\{\zeta : \|I_m + \zeta\left(I_m - \Psi\right)^{-1}\| = a_{m+1}, \zeta > 0\right\}$$

$$= \lambda_1^{-1}\left(\begin{bmatrix} \dfrac{(I_m-\Psi)^{-1}+(I_m-\Psi^*)^{-1}}{a_{m+1}^2-1} & -\dfrac{(I_m-\Psi)^{-1}}{\sqrt{a_{m+1}^2-1}} \\ -\dfrac{(I_m-\Psi^*)^{-1}}{\sqrt{a_{m+1}^2-1}} & 0 \end{bmatrix}\right) \tag{4.21}$$

Proof. See Appendix A.3. □

A matrix $E \in \mathbb{C}^{m\times m}$ is called radial if $\rho(E) = \|E\|$. Order the eigenvalues of E as:

$$\rho(E) = |\lambda_1| = |\lambda_2| = \ldots = |\lambda_s| > |\lambda_{s+1}| \geq \ldots \geq |\lambda_m|$$

where s is the number of eigenvalues of E on the circle: $|z| = \rho(E)$. The following Lemma characterises radial matrices:

Lemma 4.1.3. *[72] A matrix $E \in \mathbb{C}^{m\times m}$ is radial if and only if it is unitarily similar to a matrix of the form diag(Λ, B) in which:*

$$\Lambda = \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_s \end{bmatrix} \quad \textit{and} \quad B = \begin{bmatrix} \lambda_{s+1} & & 0 \\ & \ddots & \\ (B_{ij}) & & \lambda_m \end{bmatrix}$$

and $\rho^2(E)I_{m-s} - B^*B > 0$.

Proof. See [72]. □

Lemma 4.1.4. *Assume that $1 \in \sigma(V_1^*\Delta U_1)$ for some $\Delta \in \mathbf{B_\Delta}$. Then $\mu_\Delta(M) = 1$.*

Proof. Let $\Delta \in \mathbf{B}_{\mathbf{\Delta}}$ with $1 \in \sigma(V_1^*\Delta U_1)$. Then

$$V^*\Delta U = \begin{bmatrix} V_1^* \\ V_2^* \end{bmatrix} \Delta \begin{bmatrix} U_1 & U_2 \end{bmatrix} = \begin{bmatrix} V_1^*\Delta U_1 & V_1^*\Delta U_2 \\ V_2^*\Delta U_1 & V_2^*\Delta U_2 \end{bmatrix}$$

Since $1 \in \sigma(V_1^*\Delta U_1)$ we have that $\rho(V_1^*\Delta U_1) \geq 1$. However, since V and U are unitary and $\|\Delta\| \leq 1$ we have that $\|V^*\Delta U\| \leq 1 \Rightarrow \|V_1^*\Delta U_1\| \leq 1$ and hence, since the spectral radius of a matrix does not exceed its spectral norm, we conclude that $\rho(V_1^*\Delta U_1) = \|V_1^*\Delta U_1\| = 1$ and matrix $V_1^*\Delta U_1$ is radial. Let $V_1^*\Delta U_1$ have eigenvalues ordered in non-increasing order of magnitude as:

$$\rho(V_1^*\Delta U_1) = 1 = \lambda_1 = |\lambda_2| = \ldots = |\lambda_s| > |\lambda_{s+1}| \geq \ldots \geq |\lambda_m|$$

Then, from Lemma 4.1.3, there exist a unitary matrix $Q \in \mathbb{C}^{m\times m}$ such that $Q^*V_1^*\Delta U_1 Q = \text{diag}(1, T)$, where $T = \text{diag}(\Lambda_1, B)$ in which:

$$\Lambda_1 = \begin{bmatrix} \lambda_2 & & 0 \\ & \ddots & \\ 0 & & \lambda_s \end{bmatrix} \text{ and } B = \begin{bmatrix} \lambda_{s+1} & & 0 \\ & \ddots & \\ (B_{ij}) & & \lambda_m \end{bmatrix}$$

Let $Q_a = \text{diag}(Q, I_{n-m})$ and consider the product:

$$L := Q_a^*V^*\Delta U Q_a = \begin{bmatrix} Q^* & 0 \\ 0 & I_{n-m} \end{bmatrix} \begin{bmatrix} V_1^* \\ V_2^* \end{bmatrix} \Delta \begin{bmatrix} U_1 & U_2 \end{bmatrix} \begin{bmatrix} Q & 0 \\ 0 & I_{n-m} \end{bmatrix}$$

Define:

$$\tilde{V}_1^* = Q^*V_1^* = \begin{bmatrix} \tilde{v}_{11}^* \\ \tilde{V}_{12}^* \end{bmatrix} \text{ and } U_1Q = \begin{bmatrix} \tilde{u}_{11} & \tilde{U}_{12} \end{bmatrix}$$

where $\tilde{v}_{11}^* \in \mathbb{C}^{1\times n}$ and $\tilde{u}_{11} \in \mathbb{C}^{n\times 1}$. Then

$$L = Q_a^*V^*\Delta U Q_a = \begin{bmatrix} 1 & 0 & \tilde{v}_{11}^*\Delta U_2 \\ 0 & T & \tilde{V}_{12}^*\Delta U_2 \\ V_2^*\Delta\tilde{u}_{11} & V_2^*\Delta\tilde{U}_{12} & V_2^*\Delta U_2 \end{bmatrix}$$

Since $\|L\| = 1$ it follows that $\tilde{v}_{11}^*\Delta U_2 = 0$ and $V_2^*\Delta\tilde{u}_{11} = 0$. Note also that: $\Delta =$

$VQ_aLQ_a^*U^*$. Hence

$$\det(I_n - M\Delta) = \det(I_n - U\Sigma V^* VQ_aLQ_a^*U^*) = \det(I_n - Q_a^*\Sigma Q_a L)$$

Now

$$Q_a^*\Sigma Q_a = \begin{bmatrix} Q^* & 0 \\ 0 & I_{n-m} \end{bmatrix} \begin{bmatrix} I_m & 0 \\ 0 & \Sigma_2 \end{bmatrix} \begin{bmatrix} Q & 0 \\ 0 & I_{n-m} \end{bmatrix} = \begin{bmatrix} I_m & 0 \\ 0 & \Sigma_2 \end{bmatrix} = \Sigma$$

and hence

$$\det(I_n - M\Delta) = \det\left(\begin{bmatrix} 0 & 0 & 0 \\ 0 & I_{m-1} - T & -\tilde{V}_{12}^*\Delta U_2 \\ 0 & -\Sigma_2 V_2^*\Delta\tilde{U}_{12} & I_{n-m} - \Sigma_2 V_2^*\Delta U_2 \end{bmatrix} \right) = 0$$

Since $\Delta \in \mathbf{\Delta}$ this implies that $\gamma_{\mathbf{\Delta}}(M) \leq 1$ (equivalently $\mu_{\mathbf{\Delta}}(M) \geq 1$). Since however $\mu_{\mathbf{\Delta}}(M) \leq \|M\| = 1$ we conclude that $\mu_{\mathbf{\Delta}}(M) = 1$. □

Remark 4.1.1. *Note that since Q is unitary we can redefine the columns of U_1Q and V_1Q as the singular vector pairs corresponding to the largest singular value of M (normalised to 1), i.e, write $M = (U_1Q)(Q^*V_1^*) + U_2\Sigma_2V_2^*$.*

Remark 4.1.2. *In the cases when $\mu_{\mathbf{\Delta}}(M) = 1$ the upper bound $\|M\| = 1$ is immediate (and exact). Thus, to simplify the presentation, it is normally assumed in the sequel that $\mu_{\mathbf{\Delta}} < 1$ (for otherwise the problem of calculating an upper bound is trivial). Lemma 4.1.4 then implies that $1 \notin \sigma(V_1^*\Delta U_1)$.*

Using the two previous results the solution to the optimisation problem defined in equation (4.18) can now be obtained:

Theorem 4.1.1. *Let $\delta_i = \delta_i^0$, $i = 1, 2, \ldots, p$, be real fixed values such that $|\delta_i^0| \leq 1$ and consider the minimisation defined in equation (4.18). Assume also that $\mu_{\mathbf{\Delta}}(M) < 1$. Then:*

$$\beta(\delta_1^0, \ldots, \delta_p^0) = \min\left\{\gamma : \|\left(\gamma^2 I_m - \Psi\right)\left(I_m - \Psi\right)^{-1}\| = a_{m+1}, \gamma > 1\right\} \tag{4.22}$$

where $\Psi = \sum_{i=1}^p \delta_i^0 E_i$. Further,

$$\beta(\delta_1^0, \ldots, \delta_p^0) = \sqrt{1 + \zeta_0(\delta_1^0, \ldots, \delta_p^0)} \tag{4.23}$$

where $\zeta_0(\delta_1^0, \ldots, \delta_p^0)$ *is equal to:*

$$\lambda_1^{-1}\left(\begin{bmatrix} \dfrac{\Phi(\delta_1^0,\ldots,\delta_p^0)+\Phi^*(\delta_1^0,\ldots,\delta_p^0)}{a_{m+1}^2-1} & -\dfrac{\Phi(\delta_1^0,\ldots,\delta_p^0)}{\sqrt{a_{m+1}^2-1}} \\ -\dfrac{\Phi^*(\delta_1^0,\ldots,\delta_p^0)}{\sqrt{a_{m+1}^2-1}} & 0 \end{bmatrix}\right) \tag{4.24}$$

in which

$$\Phi(\delta_1^0, \ldots, \delta_p^0) = \left(I_m - \sum_{i=1}^{p} \delta_i^0 E_i\right)^{-1}$$

Proof. Let $\Psi = \sum_{i=1}^{p} \delta_i^0 E_i$. Then,

$$\|\Psi\| = \|\sum_{i=1}^{p} \delta_i^0 E_i\| = \|V_1^* \Delta U_1\| \leq \|V_1\| \, \|\Delta\| \, \|U_1\| \leq 1$$

Further the assumption $\mu_{\mathbf{\Delta}}(M) < 1$ implies from Lemma 4.1.4 that $1 \notin \sigma(\Psi)$. Thus Lemma 4.1.2 is applicable and the result follows. □

Next we consider the optimisation defined in equation (4.15) or equivalently (4.19). Note that in contrast to the optimisation defined in equation (4.18), here the leading $m \times m$ block is not fixed. The solution is straightforward and is stated in the following Lemma.

Lemma 4.1.5. $\gamma_{\boldsymbol{\mathcal{E}}_1} = \sqrt{1 + \hat{\zeta}_0}$ *in which* $\hat{\zeta}_0^{-1}$ *is obtained by maximising*

$$\hat{\zeta}_0^{-1} = \max \lambda_{max}\left(\begin{bmatrix} \dfrac{\Phi(\delta_1,\ldots,\delta_p)+\Phi^*(\delta_1,\ldots,\delta_p)}{a_{m+1}^2-1} & -\dfrac{\Phi(\delta_1,\ldots,\delta_p)}{\sqrt{a_{m+1}^2-1}} \\ -\dfrac{\Phi^*(\delta_1,\ldots,\delta_p)}{\sqrt{a_{m+1}^2-1}} & 0 \end{bmatrix}\right) \tag{4.25}$$

over the hypercube $|\delta_i| \leq 1$, $i = 1, 2, \ldots, p$.

Proof. Proof is immediate from the previous discussion. □

Remark 4.1.3. *Recall that* $\gamma_{\boldsymbol{\mathcal{E}}_1}(M) \leq \gamma_{\mathbf{\Delta}}(M)$ *and so* $\mu_{\mathbf{\Delta}}(M) \leq \frac{1}{\sqrt{1+\hat{\zeta}_0}}$.

Remark 4.1.4. *Problem (4.25) is a maximisation of a non-concave function over a convex set. Although several algorithms can be used in an attempt to solve it (e.g. penalty/barrier methods, feasible ascent direction algorithm using sub-gradients) there is no guarantee of convergence to a global optimum. In the following section we will develop a solution technique in the special case* $m = 1$ *which simplifies the problem considerably.*

Example 4.1.1 shows the maximisation problem (4.25) applied to a 4×4 matrix.

Example 4.1.1. *Consider the matrix $M = M_r + jM_i$ with*

$$M_r = \begin{bmatrix} -0.2497 & 0.0404 & 0.4750 & 0.4051 \\ 0.3904 & 0.1175 & 0.1743 & 0.1459 \\ 0.1511 & 0.0523 & 0.2814 & 0.0608 \\ 0.7479 & 0.0473 & -0.2059 & 0.1892 \end{bmatrix}$$

and

$$M_i = \begin{bmatrix} 0.3368 & 0.0342 & 0.4223 & 0.0138 \\ 0.1828 & 0.0251 & 0.3155 & 0.4993 \\ 0.1559 & 0.3591 & 0.1061 & 0.1162 \\ 0.0541 & -0.0390 & 0.1553 & -0.1005 \end{bmatrix}$$

The singular values of M are $\sigma_1(M) = \sigma_2(M) = 1$, $\sigma_3(M) = 0.5$ and $\sigma_4(M) = 0.25$. The diagonal structure of the problem is:

$$\mathbf{\Delta} = \{diag(\delta_1, \delta_1, \delta_2, \delta_2) : \delta_i \in \mathbb{R}, |\delta_i| \leq 1, i = 1, 2\}$$

The graph and the level contours of the largest eigenvalue of the matrix in equation (4.25) are shown in Figures 4.1 and 4.2 respectively. Note that the function has a saddle point and is maximised at the boundary of the constraint region (at point $(\delta_1, \delta_2) = (-1, -1)$).

Second Relaxation: Next we consider a second relaxation. Consider the set $\boldsymbol{\mathcal{E}}_1$ defined in equation (4.16). Suppose that

$$\left\| \sum_{i=1}^{p} \delta_i E_i \right\| \leq \phi$$

for all $\delta_i \in \mathbb{R}$ with $|\delta_i| \leq 1$, $i = 1, 2, \ldots, p$ and introduce the set $\mathcal{E}_2^{\phi}$, where

$$\boldsymbol{\mathcal{E}}_2^{\phi} = \left\{ \left[\begin{array}{c|c} E_{11} & E_{12} \\ \hline E_{21} & E_{22} \end{array} \right] : E_{11} \in \mathbb{C}^{m \times m},\ \|E_{11}\| \leq \phi \right\} \tag{4.26}$$

in which the remaining blocks $E_{12} \in \mathbb{C}^{m \times (n-m)}$, $E_{21} \in \mathbb{C}^{(n-m) \times m}$ and $E_{22} \in \mathbb{C}^{(n-m) \times (n-m)}$

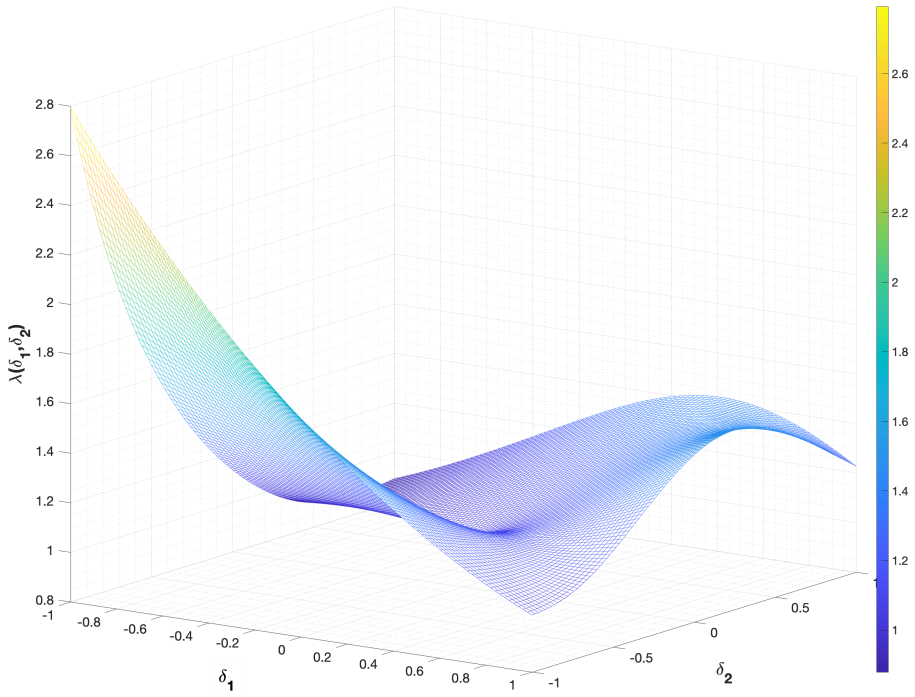

Figure 4.1: Graph of $\lambda_1(\delta_1, \delta_2)$, $(\delta_1, \delta_2) \in [-1\ 1]^2$

are free. Clearly $\boldsymbol{\mathcal{E}}_2^{\phi} \supseteq \boldsymbol{\mathcal{E}}_1 \supseteq \boldsymbol{\mathcal{E}}$ and hence

$$\begin{aligned} \gamma_{\mathbf{\Delta}}(M) \geq \gamma_{\boldsymbol{\mathcal{E}}_1}(M) &:= \min\{\gamma : \det(A - \gamma E) = 0, E \in \boldsymbol{\mathcal{E}}_1\} \\ &\geq \min\left\{\|\Delta\| \colon \det(A - \Delta) = 0, \|Z'\Delta Z\| \leq \phi\right\} \\ &:= \gamma_{\boldsymbol{\mathcal{E}}_2^{\phi}}(M) \end{aligned} \tag{4.27}$$

Intuitively, to minimise the gap $\gamma_{\mathbf{\Delta}}(M) - \gamma_{\boldsymbol{\mathcal{E}}_2^{\phi}}(M)$ variable ϕ should be selected to over-bound the set

$$\left\{\sum_{i=1}^{p} \delta_i E_i : \delta_i \in \mathbb{R}, |\delta_i| \leq 1, i = 1, 2, \ldots, p\right\}$$

as tightly as possible (in the spectral-norm sense). Thus ideally we would like to select ϕ which maximises:

$$\phi_0 = \max\left\{\left\|\sum_{i=1}^{p} \delta_i E_i\right\| : \delta_i \in \mathbb{R}, |\delta_i| \leq 1, i = 1, 2, \ldots, p\right\} \tag{4.28}$$

This is a maximisation of a convex function over a convex set and therefore the maximum is attained at some extreme point of the constraint region. It is also clear that $\phi_0 \leq 1$.

In the following section we specialise the problem to the (generic) case $m = 1$ and

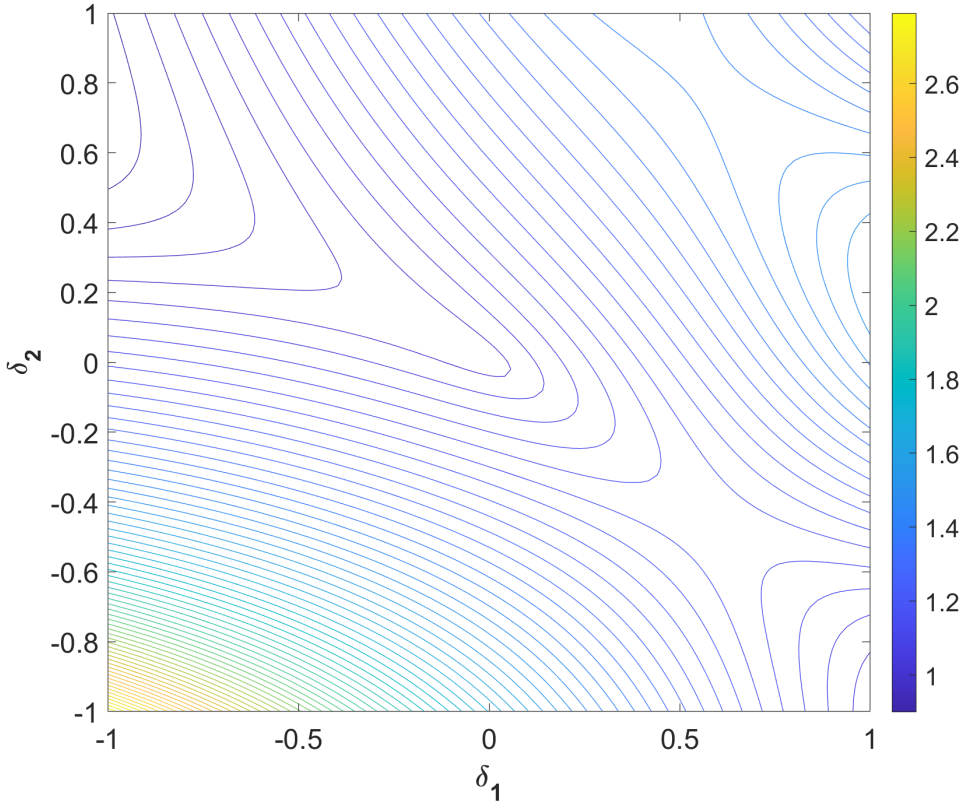

Figure 4.2: Level contours of $\lambda_1(\delta_1, \delta_2)$, $(\delta_1, \delta_2) \in [-1\ 1]^2$

develop solutions to the relaxed optimisation problems defined in this section, thereby establishing computable upper bounds to $\mu_{\mathbf{\Delta}}(M)$.

4.2 Solution of relaxed problems ($m = 1$)

In this section we describe several algorithms that can be used to provide solutions to the relaxed optimisation problems defined in the previous section in the case $m = 1$, i.e. when the largest singular value of M is non-repeated. To fix notation we denote the left and right singular vectors of M corresponding to the largest singular value (which is equal to 1) as u_1 and v_1, respectively. These are partitioned conformally with the diagonal structure $\mathbf{\Delta}$ defined in equation (4.4), i.e.

$$v_1^* = \begin{bmatrix} v_{11}^* & v_{12}^* & \dots & v_{1p}^* \end{bmatrix} \in \mathbb{C}^{1\times n}, \quad u_1 = \begin{bmatrix} u_{11} \\ u_{21} \\ \vdots \\ u_{p1} \end{bmatrix} \in \mathbb{C}^{n\times 1} \tag{4.29}$$

where $v_{1i}^* \in \mathbb{C}^{1\times k_i}$ and $u_{i1} \in \mathbb{C}^{k_i\times 1}$ for $i = 1, 2, \ldots, p$. The matrices E_i now reduce to complex numbers, i.e. $E_i := e_i = v_{1i}^* u_{i1} \in \mathbb{C}$, $i = 1, 2, \ldots, p$.

4.2.1 First relaxation: $m = 1$ case

We start the section by specialising Theorem 4.1.1 in the previous section to the case $m = 1$.

Lemma 4.2.1. *Let A be as defined above and let $\delta_{11} = \sum_{i=1}^{p} \delta_i^0 e_i$ where $\delta_i^0 \in \mathbb{R}$, $|\delta_i^0| \leq 1$, $i = 1, 2, \ldots p$. Then $|\delta_{11}| \leq 1$. Assume that $\mu_{\mathbf{\Delta}}(M) < 1$ so that $\delta_{11} \neq 1$ (see Remark 4.1.2). Then*

$$\min_{\substack{\Delta_{11}=\delta_{11}\\ \det(A-\Delta)=0}} \|\Delta\| = \min_{\substack{\gamma>1\\ |\gamma^2-\delta_{11}|=a_2|1-\delta_{11}|}} \gamma$$

Proof. The proof is immediate from Theorem 4.1.1. □

The following Theorem specialises Theorem 4.1.1 in the previous section to the case $m = 1$.

Theorem 4.2.1. *Let matrix A be as defined previously let*

$$\mathbf{\Delta}_{11} = \left\{ \delta_{11} = \sum_{i=1}^{p} \delta_i e_i, \delta_i \in \mathbb{R}, |\delta_i| \leq 1, i = 1, 2, \ldots, p \right\}$$

Then $|\delta_{11}| \leq 1$ for every $\delta_{11} \in \mathbf{\Delta}_{11}$. Assume also that $\mu_{\mathbf{\Delta}}(M) < 1$ so that $\delta_{11} \neq 1$ for every $\delta_{11} \in \mathbf{\Delta}_{11}$ (see Lemma 4.1.4). Then

$$\gamma_{\boldsymbol{\varepsilon}_1}(M) = \min_{\substack{\det(A-\Delta)=0\\ \Delta_{11}\in\mathbf{\Delta}_{11}}} \|\Delta\| = \sqrt{1+\zeta_0} \tag{4.30}$$

where ζ_0 is the solution of the following optimisation problem

$$\zeta_0 = (a_2^2 - 1)\ \min \left\{ \frac{1 - \delta' X + \delta' \Gamma \delta}{1 - \delta' X + a_2\sqrt{1 - \delta' X + \delta' \Gamma \delta}} \ \text{ such that } \ \|\delta\|_\infty \leq 1 \right\}$$

where

$$\delta = \begin{bmatrix} \delta_1 \\ \delta_2 \\ \vdots \\ \delta_p \end{bmatrix}, \quad X = \begin{bmatrix} \Re(e_1) \\ \Re(e_2) \\ \vdots \\ \Re(e_p) \end{bmatrix} \ \text{ and } \ Y = \begin{bmatrix} \Im(e_1) \\ \Im(e_2) \\ \vdots \\ \Im(e_p) \end{bmatrix}$$

and $\Gamma = XX' + YY'$.

Proof. Recall that from Theorem 4.1.1

$$\gamma_{\boldsymbol{\varepsilon}_1}(M) = \min_{\substack{\det(A-\Delta)=0 \\ \Delta_{11}\in\boldsymbol{\Delta}_{11}}} \|\Delta\| = \min_{\delta_{11}\in\boldsymbol{\Delta}_{11}} \min_{\substack{|(\gamma^2-\delta_{11})(1-\delta_{11})^{-1}|=a_2 \\ \gamma>1}} \gamma \tag{4.31}$$

$$= \sqrt{1 + \min_{\delta_{11}\in\boldsymbol{\Delta}_{11}} \min_{\substack{|1+\zeta(1-\delta_{11})^{-1}|=a_2 \\ \zeta>0}} \zeta} \tag{4.32}$$

where the minimum value of ζ (for fixed δ_{11}), ζ_0, satisfies $\zeta_0^{-1} = \lambda_1(\Psi)$ where

$$\Psi = \begin{bmatrix} \dfrac{(1-\delta_{11})^{-1} + (1-\overline{\delta}_{11})^{-1}}{a_2^2-1} & -\dfrac{(1-\delta_{11})^{-1}}{\sqrt{a_2^2-1}} \\ -\dfrac{\left(1-\overline{\delta}_{11}\right)^{-1}}{\sqrt{a_2^2-1}} & 0 \end{bmatrix} := \begin{bmatrix} \alpha & \beta \\ \bar{\beta} & 0 \end{bmatrix} \tag{4.33}$$

The maximum eigenvalue of Ψ is

$$\lambda_1(\Psi) = \zeta_0^{-1} = \frac{\alpha + \sqrt{\alpha^2 + 4|\beta|^2}}{2}$$

Write $\Re(\delta_{11}) = x$ and $\Im(\delta_{11}) = y$. Then

$$\alpha = \frac{2}{a_2^2-1}\,\frac{1-x}{1-2x+x^2+y^2} \quad \text{and} \quad |\beta|^2 = \frac{1}{a_2^2-1}\,\frac{1}{1-2x+x^2+y^2}$$

Also

$$\sqrt{\alpha^2 + 4|\beta|^2} = \frac{2a_2}{\left(a_2^2-1\right)\sqrt{(1-x)^2+y^2}}$$

after some algebra. Thus:

$$\zeta_0 = (a_2^2-1)\frac{(1-x)^2+y^2}{1-x+a_2\sqrt{(1-x)^2+y^2}} \tag{4.34}$$

Finally note that $x = \delta' X$, $y = \delta' Y$ and $x^2 + y^2 = \delta'(XX' + YY')\delta = \delta'\Gamma\delta$. □

We can now specialise the result of Lemma 4.1.2 as shown in the following Lemma:

Lemma 4.2.2. *Let A be as defined above and let $\delta_{11} = \sum_{i=1}^{p} \delta_i^0 e_i$ where $\delta_i^0 \in \mathbb{R}$, $|\delta_i^0| \leq 1$, $i = 1, 2, \ldots p$. Then $|\delta_{11}| \leq 1$ and $\delta_{11} \neq 1$. Further*

$$\min_{\substack{\Delta\in\boldsymbol{\Delta} \\ \det(A-\Delta)=0}} \|\Delta\| = \min_{\substack{\gamma>1 \\ |\gamma^2-\delta_{11}|=a_2|1-\delta_{11}|}} \gamma$$

where $\boldsymbol{\Delta} = \{\Delta \in \mathbb{C}^{n\times n},\ z'\Delta z = \delta_{11}\}$.

Proof. In the notation of Lemma 4.1.2 if $\Delta_{11} \in \mathbb{C}^{m\times m}$ with $\|\Delta_{11}\| \leq 1$ and

$$\det\left(I - \Delta_{11}\right) \neq 0$$

then

$$\min_{\substack{\det(A-\Delta)=0 \\ Z_1'\Delta Z_1=\Delta_{11}}} \|\Delta\| = \min_{\substack{\|\left(\gamma^2 I_m - \Delta_{11}\right)(I_m-\Delta_{11})^{-1}\|=a_{m+1} \\ \gamma>1}} \gamma \tag{4.35}$$

In the case $m = 1$, write $\Delta_{11} = \delta_{11} \in \mathbb{C}$. Then the constraint can be written as:

$$|\left(\gamma^2 - \delta_{11}\right)\left(1 - \delta_{11}\right)^{-1}| = a_2 \Rightarrow |\gamma^2 - \delta_{11}| = a_2|1 - \delta_{11}| \tag{4.36}$$

as required. □

The scaled version of the above result is the following:

Theorem 4.2.2. *[8] Let A be as defined previously, let $e_{11} \in \mathbb{C}$, $e_{11} \neq 1$, $|e_{11}| \leq 1$. Then:*

$$f(e_{11}) := \min_{\substack{E_{11}=e_{11} \\ \det(A-\gamma E)=0}} \gamma = \min_{\substack{\gamma>1 \\ |\gamma-e_{11}|=a_2|\gamma^{-1}-e_{11}|}} \gamma$$

Proof. Follows from Lemma 4.2.2 by carrying out the scaling $e_{11} = \gamma^{-1}\delta_{11}$. □

The above lemma can be generalised and used to compute $\gamma_{\boldsymbol{\mathcal{E}}_1}(M)$.

Lemma 4.2.3. *In the above notation:*

$$\gamma_{\boldsymbol{\mathcal{E}}_1}(M) = \min_{\substack{\gamma>1 \\ |\gamma-e_{11}|=a_2|\gamma^{-1}-e_{11}| \\ e_{11}\in\boldsymbol{\mathcal{E}}_{11}}} \gamma$$

where

$$\boldsymbol{\mathcal{E}}_{11} = \left\{ \sum_{i=1}^{p} \delta_i e_i : \delta_i \in \mathbb{R}, |\delta_i| \leq 1, i = 1, 2, \ldots, p \right\} \tag{4.37}$$

Proof. Note that in the above Lemma we can set $e_{11} = \sum_{i=1}^{p} \delta_i^0 e_i$ where $\delta_i^0 \in \mathbb{R}$, $|\delta_i| \leq 1$, $i = 1, 2, \ldots, p$. Allowing e_{11} to vary over the set $\boldsymbol{\mathcal{E}}_{11}$ defined in 4.37 above, we can write:

$$\gamma_{\boldsymbol{\mathcal{E}}_1}(M) = \min\{f(e_{11}) : e_{11} \in \boldsymbol{\mathcal{E}}_{11}\} = \min_{\substack{\gamma>1 \\ |\gamma-e_{11}|=a_2|\gamma^{-1}-e_{11}| \\ e_{11}\in\boldsymbol{\mathcal{E}}_{11}}} \gamma$$

as required. □

In this form, the problem has a nice geometric interpretation which allows us to calculate $\gamma_{\boldsymbol{\mathcal{E}}_1}(M)$, i.e. the optimal γ in the above equation corresponds to the intersection of a Zonotope and a parametric Apolonius circle (of variable centre and radius).

As defined previously in chapter 2, Zonotopes are convex, centrally symmetric polytopes. They can be thought of as linear projections of a high-dimensional hypercube. Specifically if $A \in \mathbb{R}^{p\times q}$ with $p < q$ then the set $\mathcal{Z} = \{Ax : x \in [-1, 1]^q\}$ defines a Zonotope in $\mathbb{R}^p$. An equivalent way to define Zonotope is via Minkowski sums of q line segments in $\mathbb{R}^p$. Let $\{a_i\}_{i=1}^q$, $a_i \in \mathbb{R}^p$, $i = 1, 2, \ldots, q$ be the columns of Ω. Then

$$\mathcal{Z} = \mathcal{A}_1 + \mathcal{A}_2 + \ldots \mathcal{A}_q = \left\{ \sum_{i=1}^{q} \omega_i : \omega_i \in \mathcal{A}_i, i = 1, 2, \ldots, q \right\}$$

where

$$\mathcal{A}_i = \{\lambda\omega_i : -1 \leq \lambda \leq 1\}, \quad i = 1, 2, \ldots, q$$

The vectors ω_i are called the generators of $\mathcal{Z}$. To emphasise the dependence of $\mathcal{Z}$ on its generators we will often write $\mathcal{Z} = \mathcal{Z}(\omega_1, \omega_2, \ldots, \omega_q)$. Note that identifying the complex plane $\mathbb{C}$ with $\mathbb{R}^2$ allows to write $\boldsymbol{\mathcal{E}}_{11} = \mathcal{Z}(\epsilon_1, \ldots, \epsilon_p)$ where $\epsilon_i = \begin{bmatrix} \Re(e_i) & \Im(e_i) \end{bmatrix}'$, $i = 1, 2, \ldots, p$, i.e. the set $\boldsymbol{\mathcal{E}}_{11} \subseteq \mathbb{R}^2$ is the Zonotope with generators the vectors $\omega_1, \omega_2, \ldots, \omega_p$.

It is clear that the whole set of δ_{11} on complex plane, can lies within a convex hull. i.e. since $\delta_{11} = f(\delta_1, \ldots, \delta_n) \in \mathbb{C}$ and $-1 \leq \delta_i \leq 1 : \forall i$, all the combinations of $f(\delta_i = \pm 1)$ make a set of points in complex plane. Thus a convexhull over these points forms the boundary that contains $f(-1 \leq \delta_i \leq 1 : \delta_i \in \mathbb{R}, \forall i)$. Figure 4.3 shows an example of convexhull for a randomly selected 8×8 matrix. As this figure shows, the boundary around the extreme point generate a convexhull.

Remark 4.2.1. *The analysis that follows is based on [8]. Consider (for fixed $\gamma \geq 1$) the locus of the points in the complex plane defined by equation $|\gamma - e_{11}| = a_2|\gamma^{-1} - e_{11}|$. Taking $e_{11} = x + jy$ gives:*

$$|\gamma - e_{11}|^2 = a_2^2|\gamma^{-1} - e_{11}|^2 \Leftrightarrow (\gamma - x)^2 + y^2 = a_2^2 \left\{ \left(\gamma^{-1} - x\right)^2 + y^2 \right\}$$

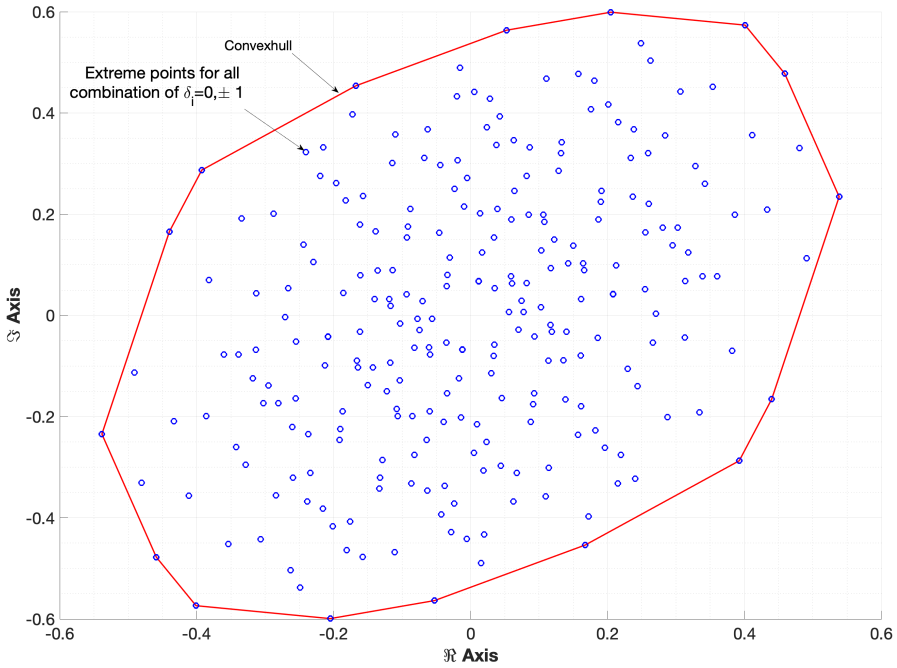

Figure 4.3: Example of convexhull over a set of point for all combinations of $f\,(\delta_i = 0, \pm 1)$ in an objective 8×8 matrix.

Equivalently:

$$x^2 - 2\gamma x + \gamma^2 + y^2 = a_2^2\left\{x^2 - 2\gamma^{-1}x + \gamma^{-2} + y^2\right\} \;\Leftrightarrow$$
$$\left(a_2^2 - 1\right)x^2 - 2x\left(\frac{a_2^2 - \gamma^2}{\gamma}\right) + \left(a_2^2 - 1\right)y^2 = \frac{\gamma^4 - a_2^2}{\gamma^2} \;\Leftrightarrow$$
$$x^2 - 2x\frac{a_2^2 - \gamma^2}{\gamma(a_2^2 - 1)} + y^2 = \frac{\gamma^4 - a_2^2}{(a_2^2 - 1)\gamma^2}$$

This eventually can be written as:

$$\left[x - \frac{a_2^2 - \gamma^2}{\gamma(a_2^2 - 1)}\right]^2 + y^2 = \left[a_2\gamma^{-1}\frac{\gamma^2 - 1}{a_2^2 - 1}\right]^2$$

and corresponds to a circle in the complex plane, $C(\gamma)$, with centre (x_c, y_c) in which

$$x_c(\gamma) = \frac{a_2^2 - \gamma}{\gamma\left(a_2^2 - 1\right)}, \quad y_c = 0$$

and radius

$$r(\gamma) = \frac{a_2}{\gamma}\frac{\gamma^2 - 1}{a_2^2 - 1}$$

respectively.

Note that $r(1) = 0$ and thus the circle collapses to the point $(1, 0)$ when $\gamma = 1$. Without loss of generality we assume that $1 \notin \mathcal{Z}$ for otherwise $\mu(M) = 1$. In addition

$$\frac{a_2^2 - 1}{a_2} \frac{dr(\gamma)}{d\gamma} = 1 + \gamma^{-2} > 0$$

and the radius $r(\gamma)$ becomes unbounded as $\gamma \to \infty$ (with linear asymptotic growth). We call this set of circle, Apollonius set. Further note that since:

$$\left(a_2^2 - 1\right) x_c = \frac{a_2^2}{\gamma} - \gamma \Rightarrow \frac{d}{d\gamma}\left(\left(a_2^2 - 1\right) x_c\right) = -a_2^2\gamma^{-2} - 1 < 0$$

the circle's centre moves to the left as γ increases. Further since $x_c \to -1$ in the limit as $\gamma \to \infty$ we can conclude by continuity (and compactness and convexity of $\mathcal{Z}$) that there is a unique minimal value of $\gamma = \gamma_0$ for which $C(\gamma_0) \cap \mathcal{Z} \neq \emptyset$ at which $C(\gamma_0)$ either includes a vertex of $\mathcal{Z}$ or is tangent to one of its sides. Figure 4.4 shows a set of Apollonius circles drawn for $1.1 \leq \gamma \leq 1.45$ with steps of 0.1. The figure shows how the Apollonius circle moves to the left with increasing diameter as γ growth.

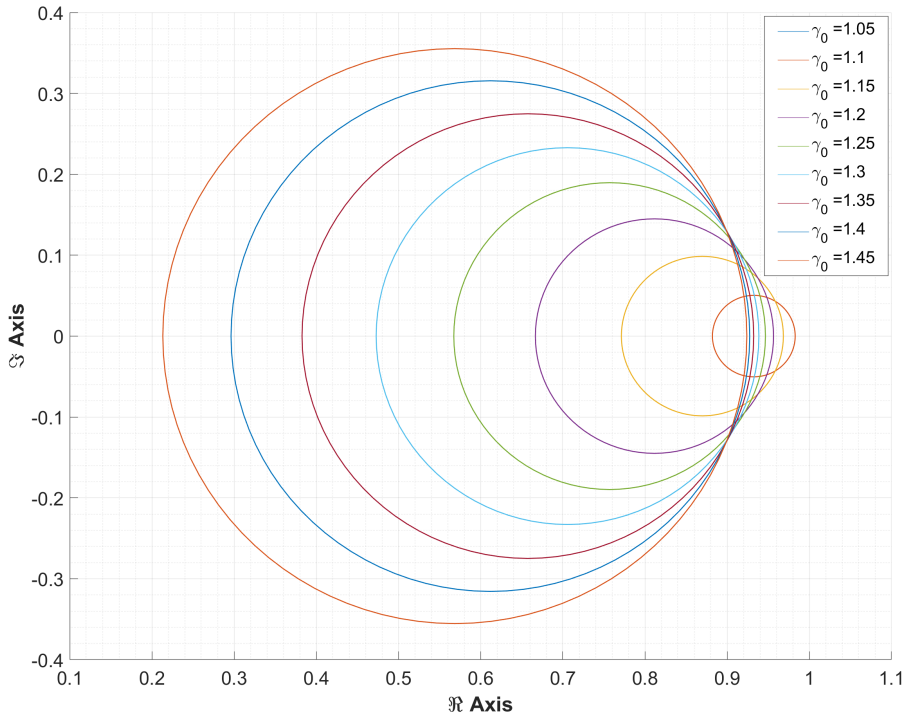

Figure 4.4: A set of Apollonius circles starting for $1.1 \leq \gamma \leq 1.45$ with steps of 0.1.

This is summarised by reformulating Lemma 4.2.2 as follows:

Lemma 4.2.4. $\gamma_{\boldsymbol{\epsilon}_1} = \min\{\gamma : C(\gamma) \cap \mathcal{Z}(\epsilon_1, \epsilon_2, \ldots, \epsilon_p) \neq \emptyset\}$.

Proof. Follows immediately from above discussion. □

In the following section we will present an algorithm for calculating $\boldsymbol{\mathcal{E}}_1$ using the Lemma 4.2.4 which relies on a (computationally efficient) enumeration of the vertices of $\mathcal{Z}(\omega_1, \omega_2, \ldots, \omega_p)$, along with a probabilistic bound when the enumeration is computationally prohibitive.

The approach can in principle be extended to the case of non-linear or "correlated" uncertainty structures. Let $\mathbf{B}\mathbb{R}^s_\infty$ be the unit ball in $\mathbb{R}^s$ with respect to the infinity norm, i.e. $\mathbf{B}\mathbb{R}^s_\infty = [-1 \quad 1]^s$ and assume that δ_1 be a vector of real uncertain parameters. Let $\delta := [\delta_1 \ldots \delta_s]' \in [-1 \quad 1]^s$ is a vector of real uncertain parameters. Let $\psi : \mathbf{B}\mathbb{R}^s_\infty \to \mathbb{R}^p$ be a contractive function, i.e. $\delta \in \mathbf{B}\mathbb{R}^s_\infty \Rightarrow \|\psi(\delta)\|_\infty \leq 1$, such that $\psi(0) = 0$. Let $\psi(\delta) := [\psi_1(\delta) \ldots \psi_p(\delta)]'$ define the diagonal uncertainty structure:

$$\boldsymbol{\Psi}_\delta = \left\{ \text{diag}\left(\psi_1(\delta) I_{k_1}, \ldots, \psi_p(\delta) I_{k_p}\right) : \delta \in \mathbf{B}\mathbb{R}^s_\infty \right\} \tag{4.38}$$

where $n = \sum_{i=1}^p k_i$. Note that the assumption $|\psi_i(\delta)| \leq 1$ for all $\delta \in \mathbf{B}\mathbb{R}^s_\infty = [-1 \quad 1]^s$ is not really restrictive and can be enforced by carrying out the maximisation $\omega_i := \max_{\|\delta\| \leq 1} |\psi_i(\delta)|$, scaling the $\psi' s$ as $\omega_i^{-1}\psi(\delta) \to \psi_i(\delta)$ and absorbing the scaling ω_i in the corresponding elements of M.

Let u_1 and v_i be the singular vector of M corresponding to the largest singular value of M as partitioned in equation (4.29) and define $e_i = v_{1i}^* u_{i1} \in \mathbb{C}, \quad i = 1, 2, \ldots, p$ and set $e_i = \alpha_i + j\beta_i$, $\alpha_i, \beta_i \in \mathbb{R}$, for $i = 1, 2, \ldots, p$ as before. Consider the set:

$$\mathcal{D} = \left\{ \sum_{i=1}^p \psi_i(\delta) e_i : \delta \in \mathbf{B}\mathbb{R}^s_\infty \right\} \subseteq \mathbb{C}$$

Assume that ψ is a continuous function, so that $\boldsymbol{\Psi}_\delta$ and $\mathcal{D}$ are compact subsets of $\mathbb{R}^s$ and $\mathbb{C}$ respectively. Let $\gamma_0(M)$ be the structured stability radius, i.e.

$$\gamma_0(M) = \min \{\gamma : \det(A - \gamma E) = 0, E \in \boldsymbol{\Psi}_\delta\}$$

Then we have the following theorem:

Theorem 4.2.3. *Let* $\gamma_l = \min \{\gamma : \mathcal{D} \cap C(\gamma) \neq 0\}$, *Then* $\gamma_l(M) \leq \gamma_0(M)$.

Proof. Similar to the proof of Lemma 4.2.4 □

Example 4.2.1. *Let*

$$M = \begin{bmatrix} 0.0991 + 0.1747i & 0.1889 + 0.1076i & 0.3233 + 0.2270i \\ 0.0501 + 0.3534i & 0.1925 + 0.2575i & 0.1470 + 0.3651i \\ 0.2549 + 0.3496i & 0.2778 + 0.2444i & 0.2779 + 0.1111i \end{bmatrix}$$

and consider the diagonal structure:

$$\mathbf{\Psi}(\delta_1, \delta_2) = \{diag\,(\delta_1, \delta_2, \delta_1\delta_2) : \delta_1, \delta_2 \in \mathbb{R}, -1 \leq \delta_2 \leq 1, -1 \leq \delta_2 \leq 1\}$$

The left and right singular vector associated to the largest singular value of M are

$$u_1 = \begin{bmatrix} -0.2598 - 0.3969i & -0.0709 - 0.6131i & -0.3691 - 0.5078i \end{bmatrix}^T$$
$$v_1^* = \begin{bmatrix} -0.5869 + 0.0000i & -0.4899 + 0.1976i & -0.5673 + 0.2337i \end{bmatrix}$$

so that $e_1 = 0.1525 + 0.2329i, e_2 = 0.1559 + 0.2864i, e_3 = 0.3281 + 0.2018i$. Thus

$$\mathcal{D} = \{\delta_1 e_1 + \delta_2 e2 + \delta_1\delta_2 e_3 : -1 \leq \delta_1 \leq 1, -1 \leq \delta_2 \leq 1\}$$

Set $\mathcal{D}$ is shown in Figure 4.5 along with a the minimum-area ellipse containing $\mathcal{D}$ and the touching Apollonius circle of minimum value $\gamma_l = 2.07$ which is therefore a lower bound to the structured distance to singularity of M. To verify that this is indeed a lower bound, the square $(\delta_1, \delta_2) \in [-1 \quad 1]^2$ was uniformly discretised (with a step equal to 0.01) and the smallest singular value of the matrix $I_3 - \gamma_{min}(\delta_1, \delta_2)EA^{-1}$ was calculated, where $E = V^*\Delta U$, $\Delta = diag(\delta_1, \delta_2, \delta_1\delta_2)$ and $\gamma_{min}(\delta_1, \delta_2)EA^{-1}$ is the minimising γ of $\sigma_{min}(I_3 - \gamma RA^{-1})$ in the interval $[0 \quad \gamma_l]$ (obtained approximately by discretion of this interval with a step equal to 0.01).

It can be observed from Figure 4.5 that unlike the linear structures, the shape of boundary enclosing all the pattern is not convex and highly depends on the values of M. It can also be noticed that the result of this method gives a considerably tighter bound for this example. This is specifically due to the position and the shape of the boundary that gives extra room for the Apollonius circle to grow bigger. This is of course due to nonlinear of Δ.

The graph of the function $(\delta_1, \delta_2) \to \sigma_{min}(I_3 - \gamma EA^{-1})$ is shown in Figure 4.6 below, along with the corresponding level contours of the function in Figure 4.7. The minimum value of the function is approximately 0.465 so the matrix singularity is never attained in this interval.

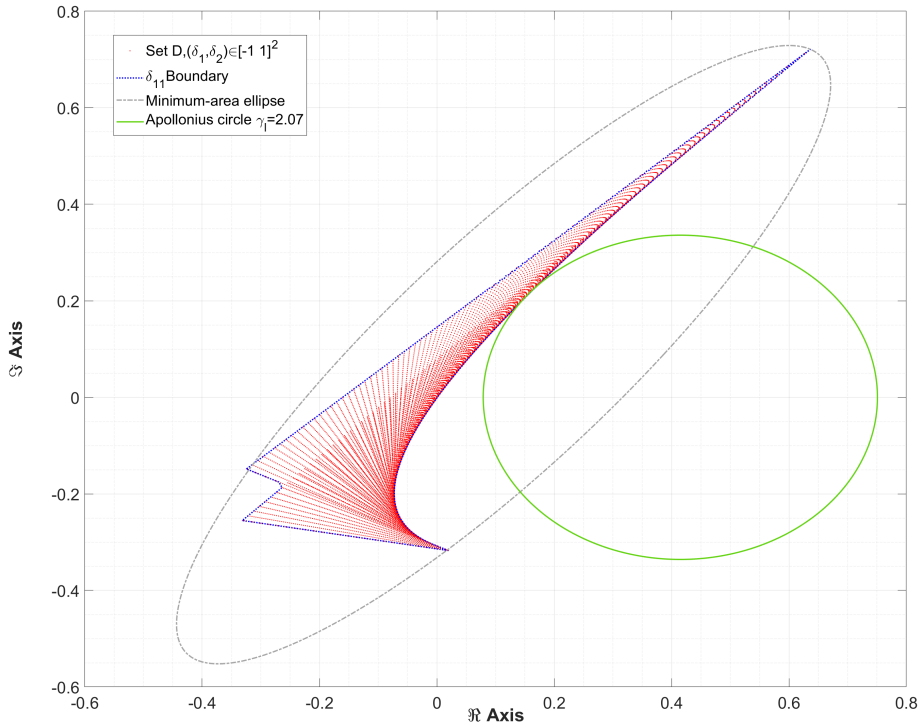

Figure 4.5: Set $\mathcal{E}_{11}$, minimum area ellipse and touching (minimum γ) Apollonius circle

4.2.2 Second relaxation: $m = 1$ case

We start by stating the expression for bound $\gamma_{\boldsymbol{\varepsilon}_2^\phi}(M)$ in the $m = 1$ case.

Lemma 4.2.5. *[9] Let $A = diag(1, a_2, \ldots, a_n)$ with $1 < a_2 \leq a_3 \leq \ldots \leq a_n$ and $\phi \in \mathbb{R}$, $0 \leq \phi < 1$, be given. Then*

$$\gamma_{\boldsymbol{\varepsilon}_2^\phi}(M) = \min\{\|\Delta\| : \det(A - \Delta) = 0, |\Delta_{11}| < \phi\} = \sqrt{a_2 - \phi(a_2 - 1)}$$

where Δ_{11} denotes the $(1,1)$ element of Δ. Further all optimal rank-reducing perturbations are given as

$$diag\left(\begin{bmatrix} \phi & \gamma_{\boldsymbol{\varepsilon}_2^\phi}(M)e^{j\theta} \\ \gamma_{\boldsymbol{\varepsilon}_2^\phi}(M)e^{j\theta} & -\phi \end{bmatrix}, \Delta_{33}\right) \tag{4.39}$$

where $0 \leq \theta < 2\pi$, $\Delta_{33} \in \mathbb{C}^{(n-2)\times(n-2)}$ and $\|\Delta_{33}\| < \gamma_{\boldsymbol{\varepsilon}_2^\phi}$.

Proof. See [9] □

As argued in the previous section the best choice for ϕ is obtained by carrying out the

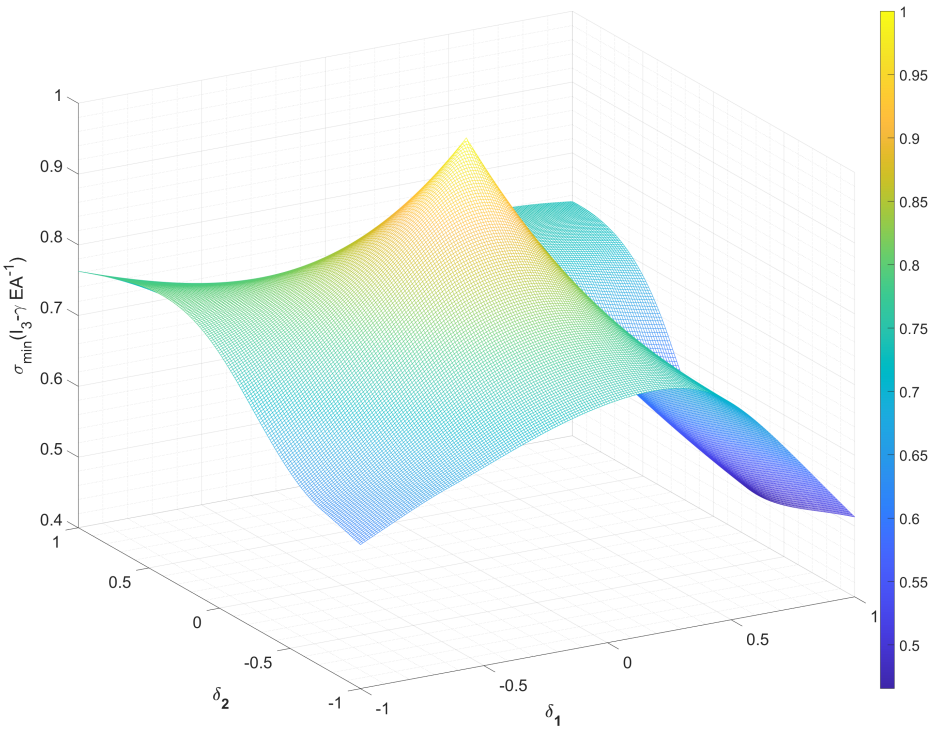

Figure 4.6: The graph of the function $(\delta_1, \delta_2) \Rightarrow \sigma_{min}(I_3 - \gamma E A^{-1})$

maximisation in equation (4.28). Let $e_i = \alpha_i + j\beta_i$, $\alpha_i, \beta_i \in \mathbb{R}$, for $i = 1, 2, \ldots, p$. Then

$$\begin{aligned}
\phi_0^2 &= \max\left\{\left|\sum_{i=1}^{m} \delta_i(\alpha_i + j\beta_i)\right|^2 : |\delta_i| \leq 1\right\} \\
&= \max\left\{\left|\left(\sum_{i=1}^{m} \delta_i\alpha_i\right) + j\left(\sum_{i=1}^{m} \delta_i\beta_i\right)\right|^2 : |\delta_i| \leq 1\right\} \\
&= \max\left\{\left(\sum_{i=1}^{m} \delta_i\alpha_i\right)^2 + \left(\sum_{i=1}^{m} \delta_i\beta_i\right)^2 : |\delta_i| \leq 1\right\}
\end{aligned} \tag{4.40}$$

Thus

$$\phi_0 = \max\left\{\sqrt{\delta'\Gamma\delta} : \|\delta\|_\infty \leq 1\right\} \tag{4.41}$$

where $\delta' = \begin{bmatrix} \delta_1 & \ldots & \delta_p \end{bmatrix} \in \mathbb{R}^p$ and

$$\Gamma = \Gamma' \in \mathbb{R}^{p\times p},\ (\Gamma)_{ij} = \alpha_i\alpha_j + \beta_i\beta_j,\ i = 1, 2, \ldots, p,\ j = 1, 2, \ldots, p$$

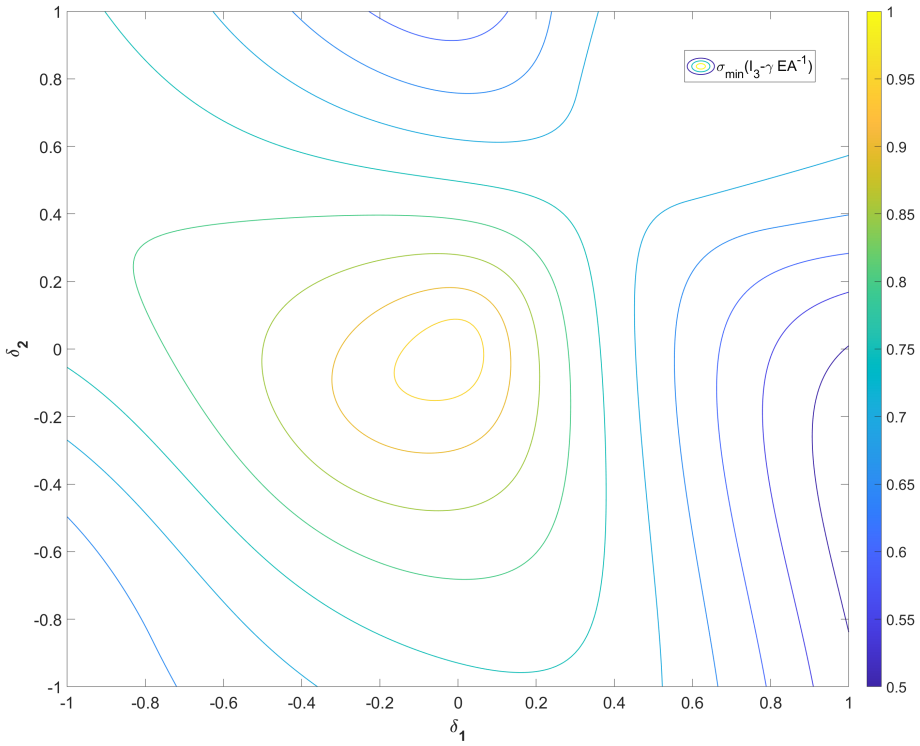

Figure 4.7: The contours of the function $(\delta_1, \delta_2) \Rightarrow \sigma_{min}(I_3 - \gamma EA^{-1})$

for example the $m = 2$ case gives

$$\Gamma = \begin{bmatrix} \alpha_1^2 + \beta_1^2 & \alpha_1\alpha_2 + \beta_1\beta_2 \\ \alpha_1\alpha_2 + \beta_1\beta_2 & \alpha_1^2 + \beta_1^2 \end{bmatrix}$$

Remark 4.2.2. *Maximisation of a convex function over convex set gives optimum at the boundary of the constrained region.*

Problem (4.41) is a maximisation of a rank-2 convex quadratic form with box-constraints. This gives us the tightest possible bound in this class:

$$\gamma_{\boldsymbol{\mathcal{E}}_2^\phi}(M) \leq \gamma_{\boldsymbol{\mathcal{E}}_1}(M) \leq \gamma_{\boldsymbol{\Delta}}(M)$$

or equivalently

$$\mu_{\boldsymbol{\mathcal{E}}_2^\phi}(M) \geq \mu_{\boldsymbol{\mathcal{E}}_1}(M) \geq \mu_{\boldsymbol{\Delta}}(M)$$

Remark 4.2.3. *A link between the two relaxations can be established by noting the scaling which allows us to transform the results in Lemma 4.2.2 to those of Theorem 4.2.2. Let*

$0 \leq \phi < 1$ *and consider the problem:*

$$\gamma_{\boldsymbol{\mathcal{E}_1}}(M) = \min_{\substack{|z'\Delta z| \leq \phi \\ det(A-\delta)=0}} \|\Delta\| = \min_{\delta_{11} \leq \phi} \min_{\substack{\gamma>1 \\ |\gamma^2 - \delta_{11}| = a_2|1-\gamma|}} \gamma \tag{4.42}$$

By using the change of variables $e_{11} = \gamma^{-1}\delta_{11}$ *and applying Lemma 4.2.2*

$$\gamma_{\boldsymbol{\mathcal{E}_1}}(M) = \min_{e_{11} \leq \gamma^{-1}\phi} \min_{\substack{\gamma>1 \\ |\gamma - e_{11}| = a_2|1-\gamma|}} \gamma \tag{4.43}$$

Thus, according to our geometric interpretation we seek to find the minimum γ *for which the Apollonius circle* $C(\gamma)$ *touches a circle with centre the origin and radius* $\phi\gamma^{-1}$*. Thus the condition for optimally is:*

$$\frac{a_2^2 - \gamma^2}{\gamma(a_2 - 1)} - \frac{a_2}{a_2^2 - 1}\frac{\gamma^2 - 1}{\gamma} = \frac{\phi}{\gamma} \tag{4.44}$$

which simplifies after some algebra to the condition $\dfrac{a_2 - \gamma^2}{a_2 - 1} = \phi$ *or equivalently* $\gamma = \sqrt{\phi + (1-\phi)a_2}$ *which is precisely the bound given in Lemma 4.2.5. Note that, as expected, the lower the value of* ϕ *the tighter the bound. Note also that for any set* $e_{11} \in \boldsymbol{\mathcal{E}_{11}}$ *the bound on the structured distance to singularity obtained via relaxation 2 cannot be tighter that the bound obtained via relaxation 1 provided the corresponding optimisation problem can be solved exactly.*

Remark 4.2.4. *Let*

$$\alpha = \begin{bmatrix} \alpha_1 & \alpha_2 & \dots & \alpha_p \end{bmatrix}', \ \beta = \begin{bmatrix} \beta_1 & \beta_2 & \dots & \beta_p \end{bmatrix}' \ \textit{and} \ \Omega = \begin{bmatrix} \alpha' \\ \beta' \end{bmatrix}$$

Then:

$$\Gamma = \alpha\alpha' + \beta\beta' \ \textit{ and } \ \phi_0 = \max_{\|\delta\|_\infty \leq 1} \sqrt{\delta'\Gamma\delta} = \max_{\|\delta\|_\infty \leq 1} \|\Omega\delta\|$$

Thus $\phi_0 = \max\{\|z\| : z \in \mathcal{Z}(\Omega)\}$*, i.e. the maximum distance of the vertices of the Zonotope with generator* Ω *from the origin. Thus, one possible way to calculate* ϕ_0 *is to enumerate all vertices of* $\mathcal{Z}(\Omega)\}$ *(e.g. via the "reverse enumeration algorithm" [42], [1], discussed in the next section), evaluate the distances of the vertices from the origin and select the largest one. An alternative approach, when the full enumeration of the vertices is not possible, is to solve the problem approximately via SDP relaxations. Let* $Q = -\Gamma$*,* $Q = Q' \leq 0$ *and consider the (primal) quadratic optimisation problem with box constraints*

[73]:

$$\min\{\delta' Q\delta : \|\delta\|_\infty \leq 1\} \tag{4.45}$$

The dual problem can be formulated as:

$$\max_{\lambda \geq 0} d(\lambda), \quad d(\lambda) = \inf_{\delta \in \mathbb{R}^p} \delta'(Q + diag(\lambda))\delta - 1'\lambda$$

(where $1'$ denotes the row vector of ones) or equivalently in SDP form:

$$\max -1'\lambda \;\; s.t. \;\; Q + diag(\lambda) \geq 0, \;\; \lambda \geq 0$$

Let $v(P)$ and $v(D)$ be the value of the primal and dual problem, respectively, so that $g = v(P) - v(D)$ is the duality gap. Reference [73] shows that a lower bound of g is given by $\xi_{r+1}\eta^2(\theta^*)$, where r is the nullity of $Q + diag(\lambda^0))$, ξ_{r+1} the smallest positive eigenvalue of $Q + diag(\lambda^0)$, λ^0 the optimal dual solution and θ^* a parameter in the interval $0 \leq \theta \leq 1$ which maximises the function $\eta(\theta)$. This leads to a tighter lower bound of the primal problem, $\nu = v(D) + \xi_{r+1}\eta^2(\theta^*) \leq v(P)$. The computation of $\eta(\theta^*)$ can be reduced to cell enumeration of hyperplane arrangement in discrete geometry which has complexity $O(p^{r+1})$ and hence can be computed in polynomial time for fixed r [73]. Thus, if $\nu \geq -1$ we have that $\phi_0 \leq \phi := \sqrt{-\nu} \leq 1$ and $\sqrt{a_2 - \phi(a_2 - 1)}$ is a lower bound of the structured distance to singularity.

Relaxation 2 can also be extended to non-linear or "correlated" structures. Define the nonlinear diagonal structure $\mathbf{\Psi}_\delta$ as in the previous section. Let u_1 and v_1 be the singular vector of M corresponding to the largest singular value of M as partitioned in equation (4.29) and define $e_i = v_{1i}^* u_{i1} \in \mathbb{C}, \quad i = 1, 2, \ldots, p$ and set $e_i = \alpha_i + j\beta_i$, $\alpha_i, \beta_i \in \mathbb{R}$, for $i = 1, 2, \ldots, p$ as before. Carry out the maximisation

$$\phi^2 = \max\left\{ \left(\sum_{i=1}^m \psi_i \alpha_i \right)^2 + \left(\sum_{i=1}^m \psi_i \beta_i \right)^2 : \mathbf{B}\mathbb{R}_\infty^s \right\} \tag{4.46}$$

or equivalently:

$$\phi_0 = \max\left\{ \sqrt{\psi(\delta)'\Gamma\psi(\delta)} : \|\delta\|_\infty \leq 1 \right\} \tag{4.47}$$

where

$$\Gamma = \Gamma^T \in \mathbb{R}^{m \times n}, \Gamma_{ij} = \alpha_i\alpha_j + \beta_i\beta_j$$

Then $\gamma_{\phi_0}(M) = \sqrt{a_2 + \phi_0(1 - a_2)}$ which give a bound to the structured distance to singularity that improves on the unstructured bound $a_1 = 1$ (In fact this still folds if ϕ_0 is substituted by any upper bound of ϕ_0 which is less that 1).The approach will be developed fully in future work. Here it is illustrated with a simple example:

Example 4.2.2. *Consider the same matrix M and diagonal structure as in previous Example . Then ϕ_0^2 can be found by maximising the quadratic function:*

$$\boldsymbol{B}_{\Psi(\delta_1,\delta_2)} = \{diag\,(\delta_1, \delta_2, \delta_1\delta_2) : \delta_1, \delta_2 \in \mathbb{R}, -1 \leq \delta_2 \leq 1, -1 \leq \delta_2 \leq 1\}$$

Let $e_i = \alpha_i + j\beta_i$, $i = 1, 2, 3$. Then ϕ_0^2 can be found by maximising the quadratic function:

$$F(\delta_1, \delta_2) = \begin{bmatrix} \delta_1 & \delta_2 & \delta_1\delta_2 \end{bmatrix} \begin{bmatrix} \alpha_1^2 + \beta_1^2 & \alpha_1\alpha_2 + \beta_1\beta_2 & \alpha_1\alpha_3 + \beta_1\beta_3 \\ * & \alpha_2^2 + \beta_2^2 & \alpha_2\alpha_3 + \beta_2\beta_3 \\ * & * & \alpha_3^2 + \beta_3^2 \end{bmatrix} \begin{bmatrix} \delta_1 \\ \delta_2 \\ \delta_1\delta_2 \end{bmatrix}$$

subjected to the constraints $-1 \leq \delta_1 \leq 1$ and $-1\delta_2 \leq 1$. Hence $e_1 = 0.1525 + 0.2329i, e_2 = 0.1559 + 0.2864i, e_3 = 0.3281 + 0.2018i$. Hence

$$\Gamma = \begin{bmatrix} 0.0775 & 0.0905 & 0.0970 \\ 0.0905 & 0.1063 & 0.1089 \\ 0.0970 & 0.1089 & 0.1484 \end{bmatrix}$$

Note that $F^2(\delta_1, \delta_2)$ may be expressed in full as:

$$\begin{aligned} F^2(\delta_1, \delta_2) = {} & (\alpha_1^2 + \beta_1^2)^2\delta_1^2 + (\alpha_2^2 + \beta_2^2)^2\delta_2^2 + (\alpha_1\alpha_2 + \beta_1\beta_2)\delta\delta_2 \\ & + (\alpha_1\alpha_3 + \beta_1\beta_3)\delta_1^2\delta_2 + (\alpha_2\alpha_3 + \beta_2\beta_3)\delta_1\delta_2^2 + (\alpha_3^2 + \beta_3^2)^2\delta_1^2\delta_2^2 \end{aligned}$$

In this case

$$\begin{aligned} F^2(\delta_1, \delta_2) = {} & 0.006\delta_1^2 + 0.0113\delta_2^2 + 0.0905\alpha_1\alpha_2 \\ & + 0.0970\delta_1^2\delta_2 + 0.1089\delta_1\delta_2^2 + 0.0220\delta_1^2\delta_2^2 \end{aligned}$$

The level contours of $F(\delta_1, \delta_2)$ are shown in Figure 4.8 below. $F(\delta_1, \delta_2)$ is maximised at $(\delta_1^0, \delta_2^0) = (1, 1)$, the maximum value being $F(\delta_1^0, \delta_2^0) = 0.9618$. Thus the corresponding value of $\gamma_{l2} = 1.0724$. As expected this is a less tight bound than γ_{l1}.

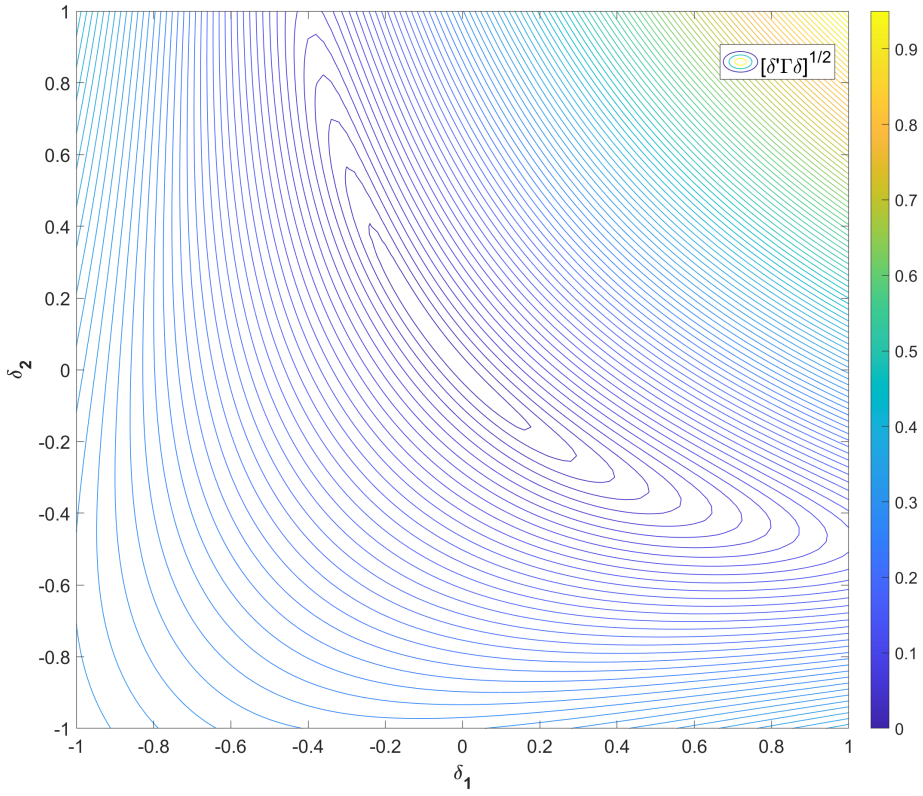

Figure 4.8: Level contours of f

4.3 Aspects of algorithmic implementation

In this section we discuss issues related to the efficient implementation of some methods outlined earlier. We start with the algorithmic implementation of the geometric method suggested in Lemma 4.2.4 which identifies the optimal bound as the intersection of a Zonotope and a parametric family of (Apollonius) circles. First we establish an easily computable upper bound on $\mu_{\boldsymbol{\mathcal{E}_1}}(M)$ which is obviously a lower bound of $\gamma_{\boldsymbol{\Delta}}(M)$:

Theorem 4.3.1. *Assume that* $\xi := \sum_{i=1}^{n} |\Re(e_i)| \leq 1$. *Then*

$$\gamma_l(M) := \frac{1}{2}\sqrt{\xi^2(1-a_2)^2 + 4a_2} - \frac{\xi}{2}(a_2 - 1) \leq \gamma_{\boldsymbol{\mathcal{E}_1}}(M) \leq \gamma_{\boldsymbol{\Delta}}(M) \tag{4.48}$$

Further $\gamma_l(M) = \gamma_{\boldsymbol{\mathcal{E}_1}}(M)$ *if and only if* $\xi \in \mathcal{Z}$.

Proof. Consider Figure 4.9 that gives a graphical interpretation of Theorem 4.3.1

Define

$$\xi = \max_{z \in \mathcal{Z}} \Re(z) = \max_{|\delta_i| \leq 1} \sum_{i=1}^{n} \delta_i \Re(e_i) = \sum_{i=1}^{n} |\Re(e_i)| \tag{4.49}$$

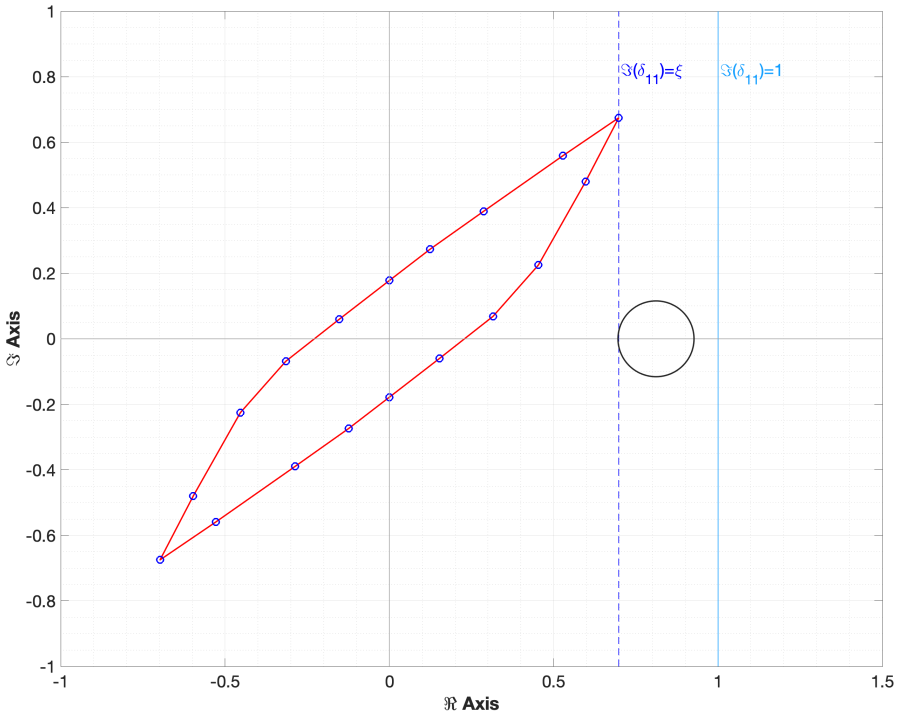

Figure 4.9: Graphical interpretation of Theorem 4.3.1

and assume that $\xi \leq 1$. Due to convexity there is at least one vertex of $\mathcal{Z}$ with real part equal to ξ. Consider the vertical line $\Re(z) = \xi$ which divides $\mathbb{C}$ into two half-planes. ξ is contained in the half-plane $\Re(z) \leq \xi$. Since

$$\gamma = \gamma_l := \frac{1}{2}\sqrt{\xi^2(1-a_2)^2 + 4a_2} - \frac{\xi}{2}(a_2 - 1) \geq 1 \tag{4.50}$$

is the positive root of the equation

$$\xi = x_c(\gamma) - r(\gamma) = \frac{a_2^2 - \gamma^2}{\gamma\left(a_2^2 - 1\right)} - \frac{a_2}{\gamma}\frac{\gamma^2 - 1}{a_2^2 - 1}$$

the circle $C(\gamma_l)$ with centre the point $(x_c(\gamma_l), 0)$ and radius $r(\gamma_l)$ lies in the half-plane $\Re(z) \geq \xi$. Thus $\mathcal{Z} \cap C(\gamma_l) = \emptyset$ except in the case $\xi \in \mathcal{Z}$ and hence $\gamma_l \leq \gamma_{\boldsymbol{\varepsilon_1}}(M)$ with $\gamma_l = \gamma_{\boldsymbol{\varepsilon_1}}(M)$ if and only if $\xi \in \mathcal{Z}$

□

One possible way of making use of Lemma 4.2.4 is to follow the following procedure:

(i) Enumerate all vertices of $\mathcal{Z}$.

(ii) Sort them in terms of their argument (between 0 and 2π, say).

(iii) List the Zonotope's edges (line-segments linking consecutive vertices).

(iv) Calculate the values of γ for which $C(\gamma)$ goes through a vertex or is tangent to an edge. The minimum of all these values is the required bound $\gamma_{\boldsymbol{\varepsilon_1}}$.

Note that the Zonotope vertices form a subset of all 2^p points of the form $\sum_{i=1}^{p} \delta_i e_i$ where the δ_i's take the extreme values ± 1 (in fact due to central symmetry 2^{p-1} evaluations are enough). Thus, in principle, the vertices of the Zonotope could be obtained by constructing the convex hull of these 2^p points. Although this approach works for small p is becomes computationally prohibitive as p increases. Thus a more efficient enumeration method is needed in this case. The main objective in vertex enumeration, is still to find an algorithm of low computational complexity (e.g. polynomial or linear-time)[59]. One of the first enumeration algorithms in graph-based problems was proposed by Read et al. [60] and is known as Backtracking. A linear-time algorithm was first proposed by Dyer [61]. This uses a depth-first search and a balanced tree data structure [59]. Other efficient enumeration algorithms include [59], [60], [61], [28], [29], [63] and [39]. Avis and Fukuda introduced an algorithm known as reverse search [42] for enumerating the vertices of a zonotope. It is known that for a zonotope $Z(\omega_1, \ldots, \omega_p)$ with p generating vector, $\omega_i \in \mathbb{R}^q, i = 1, 2, \ldots, q$ the number of vertices $|\text{vert}(\mathcal{Z})|$ is bounded as

$$|\text{vert}(\mathcal{Z})| \leq 2 \sum_{i=0}^{q-1} \binom{p-1}{i} \tag{4.51}$$

and the bound is tight if the generating vectors are in general position. The reverse enumeration algorithm is efficient both in terms of computational time and memory requirements and is highly parallelisable. It is typically implemented in the dual setting of central arrangement of p oriented hyper planes in $\mathbb{R}^q$ which are normal to the generating vectors ω_i. See [1] for details.

It is possible to avoid the solution of $2|\text{vert}(\mathcal{Z})| - 1$ scalar problems corresponding to step (iv) above after carrying out the Zonotope's vertex enumeration. This can be achieved by obtaining the ellipse of minimum area which contains all vertices of the Zonotope (and thus by convexity the entire Zonotope). In this way we can obtain a lower bound on $\gamma_{\boldsymbol{\varepsilon_1}}(M)$. Since Zonotope $\mathcal{Z}$ is centrally symmetric, the optimal ellipse will be centred at the origin and hence is described by an equation of the form:

$$\mathcal{H} = \left\{ z \in \mathbb{R}^2 | \quad z^T E z = 1 \text{ and } z^T E z \leq 1 \forall z \in \mathcal{Z} \right\} \tag{4.52}$$

where for $E \in \mathbb{R}^{2\times 2}$, $E = E^T > 0$. The area of the ellipse is $\frac{\pi}{\sqrt{\det(E)}}$ and hence the optimisation problem takes the form:

$$\min_{E=E^*>0} \frac{\pi}{\sqrt{\det(E)}} \quad \text{s.t.} \quad \begin{bmatrix} \Im\xi_i \\ \Re\xi_i \end{bmatrix} \in \mathcal{H}(E), \quad , i = 1, 2, \ldots, |\text{vert}(\mathcal{Z})| \tag{4.53}$$

where $\xi_i \in \mathbb{C}$ are the vertices of $\mathcal{Z}$. To turn this to a convex problem, set $Q = E^{1/2}$; then since $z^*Ez \leq 1 \Leftrightarrow \|Qz\| \leq 1$ an equivalent form of the optimisation problem is:

$$\min_{Q} log|Q^{-1}| \quad s.t. Q = Q^* > 0, \|Q\| \begin{bmatrix} \Im\xi_i \\ \Re\xi_i \end{bmatrix} \leq 1, i = 1, 2, \ldots, |\text{vert}(\mathcal{Z})| \tag{4.54}$$

The last set of constraint be written in the form of Linear Matrix Inequalities(LMI's:

$$\begin{bmatrix} I_2 & Q \begin{bmatrix} \Re e_i \\ \Im e_i \end{bmatrix} \\ \begin{bmatrix} \Re e_i & \Im e_i \end{bmatrix} Q & 1 \end{bmatrix} \geq 0, \quad i = 1, 2, \ldots, |\text{vert}(\mathcal{Z})|$$

This is now a convex optimisation problem and can be solved in a variety of ways. e.g. interior-point method or DRN algorithm [74]. Here we use a the Conditional Gradient Ascent algorithm and used the MATLAB code provided by [75]. To find the intersection point we make both ellipse and circle equation equal and will calculate the associated γ_l where gives a single intersection point. The solution is a polynomial equation of degree four which depending on the value of γ can have zero, one, two, three or four independent solutions, but only the γ associated with one real repeated solution (the smallest circle) will be accepted.

The above method relies on the enumeration of the vertices of $\mathcal{Z}$ which may be impractical for high dimensional problems. It is possible to find an ellipse containing $\mathcal{Z}$ directly from the generating matrix Ω without enumerating the vertices:

Lemma 4.3.1. *Let $\Omega \in \mathbb{R}^{2\times p}$ with rank$(\Omega) = 2$ and singular value decomposition $\Omega = U\Sigma V_1'$ with $\Sigma = diag(\Sigma)$, $UU' = U'U = I_2$ and $V_1'V_1 = I_2$. Then $\mathcal{Z} \subseteq \mathcal{E} := \{x \in \mathbb{R}^2 : x'(p^{-1}U\Sigma^{-2}U')x \leq 1\}$.*

Proof. First note that if $\delta \in \mathbb{R}^p$, $\|\delta\|_\infty \leq 1 \Rightarrow \|\delta\| \leq \sqrt{p}$ and hence:

$$\mathcal{Z}(\Omega) = \{\Omega\delta : \|\delta\|_\infty \leq 1\} \subseteq \{\sqrt{p}\Omega\delta : \|\delta\| \leq 1\} := \mathcal{P}_0$$

Hence, if V_2' is the orthogonal complement of V_1' so that $[V_1\ V_2]$ is square orthogonal,

$$\mathcal{P}_0 = \left\{\sqrt{p}U\Sigma V_1'\delta : \|\delta\| \leq 1\right\}$$

Changing variables $x = V'\delta$,

$$\mathcal{P}_0 = \left\{\sqrt{p}U\Sigma V_1'Vx : \|x\| \leq 1\right\} = \{\sqrt{p}U\Sigma x_1 : \|x_1\| \leq 1\}$$

where $x_1 = [I_2\ 0_{2,p-2}]x$. It now follows that

$$\mathcal{P}_0 = \mathcal{E} := \left\{y \in \mathbb{R}^2 : p^{-1}y'U\Sigma^{-2}U'y \leq 1\right\}$$

Take $y \in \mathcal{P}_0$ so that $y = \sqrt{p}U\Sigma x_1$, $\|x_1\| \leq 1$. Then $x_1 = \frac{1}{\sqrt{p}}\Sigma^{-1}U'y$ and

$$x_1'x_1 \leq 1 \ \Rightarrow\ p^{-1}y'U\Sigma^{-2}U'y \leq 1$$

so that $y \in \mathcal{E}$. Conversely if $y \in \mathcal{E}$ then $\|\frac{1}{\sqrt{p}}\Sigma^{-1}U'y\| \leq 1$, so y can be written as $y = \sqrt{p}U\Sigma x_1$ for some x_1 with $\|x_1\| \leq 1$, so that $y \in \mathcal{P}_0$. □

Remark 4.3.1. *If rank$(\Omega) = 1$ we get a degenerate ellipse consisting of a line segment with centre the origin.*

Remark 4.3.2. *Calculating the minimum value of γ such that the Apollonius circle $C(\gamma)$ intersects the ellipse $\mathcal{E}$ gives a lower bound of the structured distance to singularity. The bound may be conservative, although enumeration of the vertices of $\mathcal{Z}$ is avoided.*

Example 4.3.1. *Figure 4.10 shows a zonotope corresponding to a randomly generated matrix $M \in \mathbb{C}^{9\times 9}$ and a diagonal uncertainty structure with 9 real elements. This is approximated by the minimum-area ellipse containing the 18 vertices of the zonotope and the ellipse constructed via Lemma 4.3.1. The figure also shows the three Apollonius circles touching the Zonotope and the both ellipses. The corresponding lower bounds to the structured distance to singularity is $\gamma_l = 1.585$, $\gamma_l = 1.565$ and $\gamma_l = 1.530$, respectively.*

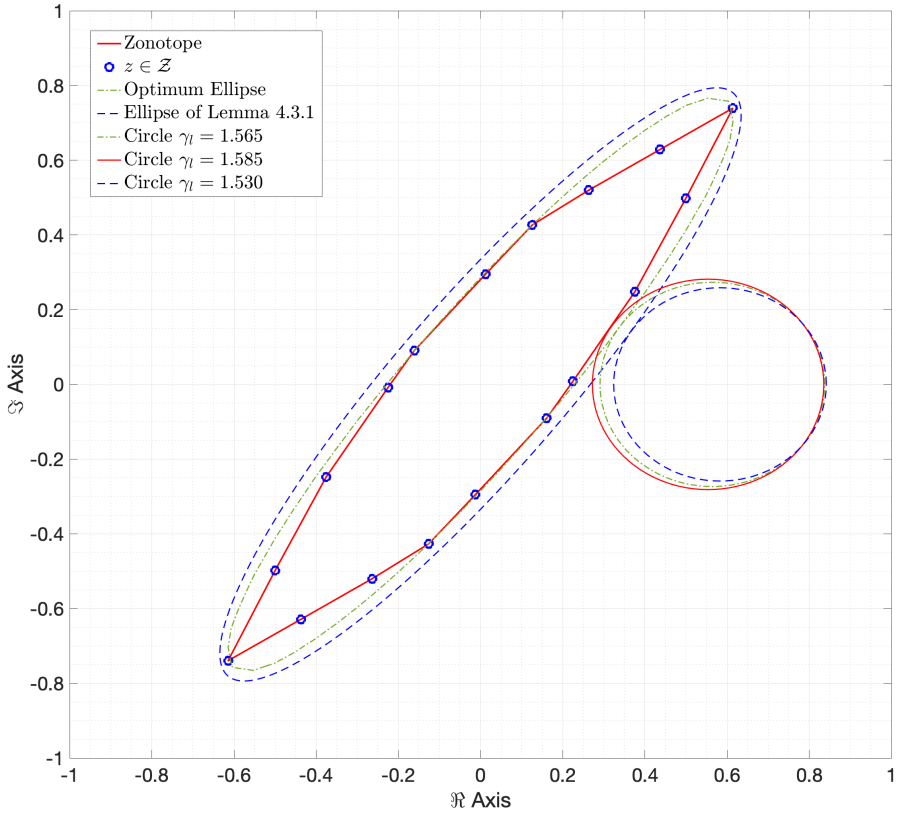

Figure 4.10: Using an enclosed ellipse to find a bound

4.4 A randomised algorithm to calculate an upper bound on the structured singular value

It is been argued previously that in the linear uncertainty case the calculation of a lower bound of the structured distance to singularity requires the enumeration of the vertices of a Zonotope. This can be effectively achieved using several deterministic algorithms, e.g. the reverse-enumeration algorithm described previously. In high dimensional problems, however, even efficient deterministic algorithms are impractical due to high computational complexity. In this case one can rely on randomised algorithms which approximate the Zonotope by generating only a subset of its vertices. Here we extend the method described in chapter 2 to obtain a probabilistic bound on γ_{l1}.

Consider a Zonotope with generator a matrix $\Omega \in \mathbb{R}^{q \times m}$ where $q \leq p$. In the special case considered in this section ($m = 1$) $q = 2$ and the Zonotope is two-dimensional. Recall also that p is the number of uncertainty parameters δ_i. Suppose that the columns of Ω ($\begin{bmatrix} \Re(e_i) & \Im(e_i) \end{bmatrix}'$) satisfy the following two (generic) conditions: (i) No column of Ω is

the zero vector, and (ii) no two columns of Ω are scalar multiples of each other. Then the Zonotope generated by Ω, i.e. $\mathcal{Z} = \{\Omega\delta : \delta \in [-1\ 1]^p\}$ is in general position and its number of its vertices is

$$|\text{vert}(\mathcal{Z})| = 2\sum_{i=0}^{q-1}\binom{p-1}{i}$$

($|\text{vert}(\mathcal{Z})| = 2p$ if $q = 2$). Then, under the above assumption, for $\delta \in \mathbb{R}^p$ such that $\Omega'\delta$ has all nonzero elements, the point v defined the mapping $v = m(\delta) := A\text{sign}(A'\delta)$ is a vertex of $\mathcal{Z}(\omega_1, \ldots, \omega_p)$. From central symmetry $v \in \text{vert}(\mathcal{Z}) \Rightarrow -v \in \text{vert}(\mathcal{Z})$. Moreover, if

$$H = \bigcup_{i=1}^{q}\left\{\delta \in \mathbb{R}^p : \hat{\omega}'\delta = 0\right\}$$

where $\hat{\omega}_i'$ are the rows of Ω, $i = 1, 2, \ldots, q$, then the mapping $m : \mathbb{R}^q \setminus H \rightarrow \text{vert}(\mathcal{Z})$ is well defined and onto [40]. A randomisation algorithm can now be used to enumerate the vertices of $\mathcal{Z}$. The algorithm updates a list of vertices (initialised as the empty list) by drawing independent samples $\delta \in \mathbb{R}^p$ from a p-dimensional standard Gaussian distribution, computing $v_+ = \Omega\text{sign}(\Omega'\delta)$ and $v_- = -v_+$ and adding them to the list (unless they are already listed). The algorithm can proceed until all vertices have been enumerated, or terminate after a fixed number of iterations. In the later case only a subset of vertices will be (in general) enumerated. The convex hull of these vertices will be a subset of $\mathcal{Z}$.

The Zonotope approximation of reference [40] proceeds by characterising the probability that any particular vector $\delta \in \mathbb{R}^p$ maps to a particular vertex via map m. This probability is characterised by a geometric feature, in particular, the vertex's normal cone. Specifically, it can be shown that under the previous assumption, the inverse function of m, $m^{-1} : \text{vert}(\mathcal{Z}) \rightarrow \mathbb{R}^q$ is $m^{-1}(v) = \text{inv}(N_{\mathcal{Z}}(v))$ where:

$$N_{\mathcal{Z}}(v) = \left\{\delta \in \mathbb{R}^p : \delta'(z - v) \leq 0 \ \text{ for all } \ z \in \mathcal{Z}\right\}$$

is the normal cone of v. This allows to assign a probability measure on the vertices of $v_i \in \mathcal{Z}$, i.e.

$$\mathbb{P}_v[v_i] := \mathbb{P}_\delta[N_{\mathcal{Z}}(v_i)] = \mathbb{P}_\delta[\text{int}(N_{\mathcal{Z}}(v_i))] = \mathbb{P}_\delta[\{\delta \in \mathbb{R}^p : \Omega\ \text{sign}(\Omega'\delta) = v_i\}]$$

Note that since the volume of the normal cone of a "sharp" vertex (weighted by the standard Gaussian density function) is larger than the weighted volume of a "flat" vertex, a "sharp" vertex gets mapped to more frequently relative to a "flat" vertex. This is important,

since "sharp" vertices contribute more significantly to the shape of the Zonotope than "flat" vertices and hence, if retained, result to a better approximation. This allows to estimate (probabilistically) the number of iterations of the randomised algorithm needed to generated any vertex of the Zonotope. It also allows for the derivation of a (probabilistic) bound on the Hausdorff distance between $\mathcal{Z}$ and its approximation, defined as the convex hull of the subset of vertices generated by the randomised algorithm executed with a fixed number of iterations. This is stated after the following definition:

Let $v \in \text{vert}(\mathcal{Z})$. We define the simplicial constant of v to be

$$\alpha_{\mathcal{Z}}(v) := \inf_{x} \{\|v - x\| : x \in \text{conv}(\text{vert}(\mathcal{Z}) \setminus \{v\})\}$$

Note that this is the Hausdorff distance between $\mathcal{Z}$ and its subset $\text{conv}(\text{vert}(\mathcal{Z}) \setminus \{v\})$. The concept Hausdorff distance and Normal cone has been discussed in details at chapter 2. The following Theorem now follows:

Theorem 4.4.1. *[40] Let $\mathcal{Z}(\omega_1, \ldots, \omega_p)$ to be a Zonotope with its generating vectors in general position. Given $\epsilon > 0$ and $\delta_1 > 0$ choose $b > diam(\mathcal{Z})$ and p_1 as*

$$p_1 > \frac{\log\left(|vert(\mathcal{Z})|/\epsilon\right)}{\log\left(1/(1-k)\right)} \tag{4.55}$$

where

$$k = \left(\frac{1}{2}\left(1 - \sin\left(\arctan(b/\delta)\right)\right)\right)^{\frac{1}{2}} \tag{4.56}$$

Let V be the subset of $\mathcal{Z}$'s vertices produced by the randomised algorithm after p_1 iterations. Let $h(\mathcal{Z}, conv(V))$ be the Hausdorff distance between sets $\mathcal{Z}$ and $conv(V)$. Then:

$$h(\mathcal{Z}, conv(V)) \leq \frac{|vert(\mathcal{Z}) \setminus V|}{2}\delta_1 \tag{4.57}$$

with probability at least $1 - 2^a\epsilon$ where $a = |vert(\mathcal{Z}) \setminus U_Z|/2$ and

$$U_Z = \{v \in vert(\mathcal{Z}) : \alpha_Z(v) \geq \delta_1\}$$

Proof. See [40]. □

We can now prove the following Corollary which gives a probabilistic lower bound to the structured distance to singularity.

Corollary 4.4.1. *Let all variables be defined as in Theorem 4.4.1 Let $\mathcal{P}$ be a compact superset of conv(V) such that*

$$h(\mathcal{P}, conv(V)) \geq \frac{|vert(\mathcal{Z}) \setminus V|}{2}\delta_1 \tag{4.58}$$

and $1 \notin \mathcal{P}$. Let also

$$\gamma_l(M) = \min\{\gamma : \gamma \geq 1, \mathcal{P} \cap C(\gamma) \neq \emptyset\} \tag{4.59}$$

Then $\gamma_l(M) \leq \gamma_{\mathbf{\Delta}}(M)$ with probability at least $1 - 2^a\epsilon$.

Proof. Conditions given by equations (4.57) and (4.58) imply that $\mathcal{P} \supseteq \mathcal{Z}$ with probability at least $1 - 2^a\epsilon$. The existence of the indicated minimum in equation (4.59) follows from the compactness of $\mathcal{P}$. The inequality $\gamma_l(M) \leq \gamma_{\mathbf{\Delta}}(M)$ follows from Theorem 4.2.4. □

Note that if we use the randomised enumeration algorithm defined in theorem 4.4.1, some of the vertices could be omitted. This makes the approximated polytope to have a smaller boundary. Therefore, the same circuit that touched the original Zonotope may leave some gap to the new one. In other words, γ will increase as the circle (with a larger radius) is pushed to the left before the tangency condition occurs. Note that this gives tighter upper bound on μ but there is a risk that the optimiser fall within the omitted area where the upper bound loses its validity. Figure 4.11 illustrates how the randomised algorithm could potentially eliminate the optimiser.

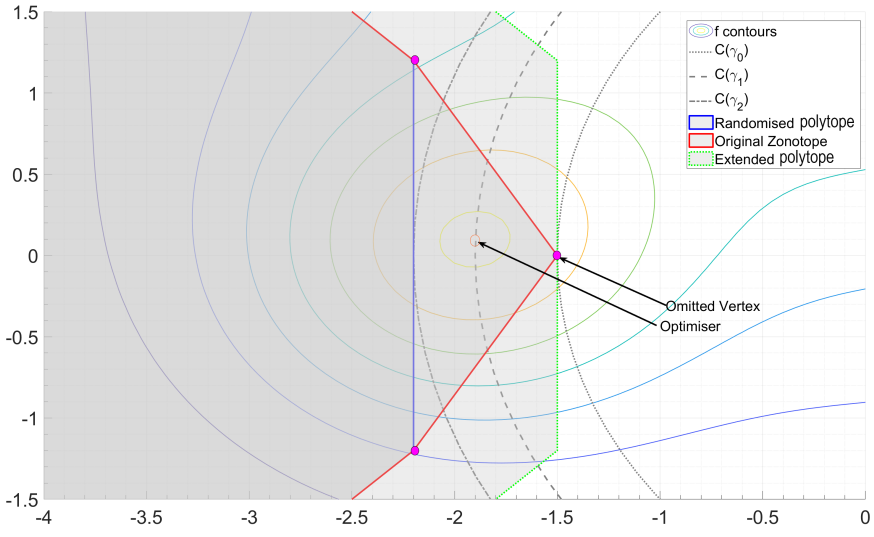

Figure 4.11: An example that the optimiser is omitted by the the randomised algorithm

In Figure 4.11, the red polygon shows the original Zonotope related to δ_{11} and blue boundary is the approximate polytope generated by the randomised algorithm. Thus the area between the Zonotope and the reduced polytope, i.e. the triangles, is omitted by the randomised algorithm. Assume that the contours show the actual optimisation of the main problem where the optimiser falls exactly in the omitted area. Now the $\gamma = \gamma_1$ related to middle circle is the exact solution of the μ-problem. Morever, the $\gamma = \gamma_0 = \mu_0^{-1}$ corresponding to the smaller circle which is tangential to the original Zonotope is smaller than the solution, $\gamma = \gamma_1 = \mu_1^{-1}$, i.e.

$$\mu_0 \leq \mu_1$$

Now imagine that the randomised algorithm omits the vertex shown in Figure 4.11. The result would be the blue Zonotope. The $C(\gamma)$ then will increase in area and shifted to the left in order to contact with the randomised polytope for a larger γ say γ_2. Using the same justification, $\gamma_2 \geq \gamma_0$ and thus

$$\mu_2 \leq \mu_0$$

This means that the upper bound could potentially be smaller than the actual solution and hence is not be a valid result. A solution proposed in this work to overcome this issue is to extend the randomised polytope so that the extended polytope constructed from the partially enumerated vertices contains the exact (but unknown) one with a high probability. To achieve this, however, we need a measure that identifies by how much the Zonotope should be extended to ensure all the original vertices are inside the extended polytope. Corollary 4.4.1 suggests that if we expand the randomised polytope by $\alpha\delta$, then with a probability of at least $1 - 2^{\alpha}\delta$ the new (extended) polytope includes the original Zonotope. The advantage of extended polytope is that it has the same number of vertices as the one obtained from the randomised algorithm. Since the extended polytope has a larger area, though, it pushes the tangential circle to the left and hence the corresponding γ is smaller than the one associated with the original Zonotope (green boundary in Figure 4.11). Thus

$$\mu_{\Delta}(M) \leq \{\gamma_0^{-1}, polytope(\mathcal{Z})\} \leq \{\gamma_0^{-1}, polytope(\mathcal{U}')\}$$

Remark 4.4.1. *The best bound is obtained when the Housdorff distance between conv(V) and the extended polytope $\mathcal{P}$ is equal to the bound given in the right-hand-side of equation* (4.57)*. The construction of $\mathcal{P}$ which achieves this is straightforward.*

4.5 Numerical experiment

To demonstrate this method a 9×9 matrix M is randomly selected and is normalised so that its largest singular value is equal to one. Here $M = M_r + jM_i$, where $M_r \in \mathbb{R}^{9\times 9}$ and $M_i \in \mathbb{R}^{9\times 9}$, where:

$$M_r = \begin{bmatrix} 0.02 & 0.08 & 0.13 & 0.06 & 0.05 & 0.01 & 0.07 & 0.12 & 0.07 \\ 0.06 & 0.14 & 0.13 & 0.11 & 0.12 & 0.10 & 0.01 & 0.03 & 0.09 \\ 0.05 & 0.11 & 0.12 & 0.12 & 0.037 & 0.09 & 0.08 & 0.10 & 0.05 \\ 0.07 & 0.10 & 0.14 & 0.09 & 0.13 & 0.11 & 0.03 & 0.01 & 0.01 \\ 0.09 & 0.07 & 0.03 & 0.01 & 0.11 & 0.15 & 0.08 & 0.12 & 0.13 \\ 0.02 & 0.04 & 0.08 & 0.14 & 0.08 & 0.06 & 0.05 & 0.10 & 0.02 \\ 0.13 & 0.08 & 0.09 & 0.01 & 0.04 & 0.04 & 0.05 & 0.14 & 0.05 \\ 0.09 & 0.04 & 0.14 & 0.08 & 0.04 & 0.13 & 0.11 & 0.01 & 0.09 \\ 0.02 & 0.05 & 0.10 & 0.02 & 0.11 & 0.12 & 0.08 & 0.13 & 0.15 \end{bmatrix}$$

and

$$M_i = \begin{bmatrix} 0.08 & 0.06 & 0.02 & 0.07 & 0.14 & 0.08 & 0.07 & 0.04 & 0.04 \\ 0.15 & 0.05 & 0.15 & 0.01 & 0.09 & 0.01 & 0.01 & 0.11 & 0.01 \\ 0.03 & 0.15 & 0.07 & 0.13 & 0.04 & 0.13 & 0.11 & 0.15 & 0.07 \\ 0.06 & 0.10 & 0.00 & 0.04 & 0.03 & 0.06 & 0.00 & 0.05 & 0.09 \\ 0.10 & 0.06 & 0.01 & 0.04 & 0.06 & 0.08 & 0.14 & 0.11 & 0.02 \\ 0.01 & 0.01 & 0.02 & 0.09 & 0.00 & 0.08 & 0.11 & 0.01 & 0.07 \\ 0.11 & 0.11 & 0.13 & 0.07 & 0.13 & 0.10 & 0.14 & 0.09 & 0.10 \\ 0.06 & 0.10 & 0.07 & 0.05 & 0.09 & 0.05 & 0.08 & 0.06 & 0.12 \\ 0.12 & 0.01 & 0.08 & 0.10 & 0.03 & 0.01 & 0.03 & 0.11 & 0.05 \end{bmatrix}$$

The uncertainty structure is assumed diagonal with 9 uncertain parameters. This gives rise to a Zonotope $\mathcal{Z}$ with $p = 18$ vertices shown in Figure 4.12. The figure also shows the minimum-area ellipse containing the Zonotope, the ellipse constructed via Lemma 4.3.1 and the vertical straight line through the vertex of the Zonotope with maximum real part.

In addition the Figure shows the Apollonius circles of minimal parameter values $\gamma = \gamma_l$ which touch the Zonotope $\mathcal{Z}$, each of the the two ellipses and the vertical straight line, respectively. The corresponding values of γ_l which corresponds to lower bounds on the structured distance to singularity of M are summarised in Table 4.1. Fig. 4.12 shows the results of several algorithms presented in the paper when applied to M.

Table 4.1: Summary of lower bounds

Method	Lower bound (γ_l)
Exact Zonotope $\mathcal{Z}$	1.520
Minimum-area Ellipse containing $\mathcal{Z}$	1.512
A-priori Ellipse (Lemma 7)	1.490
Extended 4-vertex Zonotope ($P = 88\%$)	1.420
Vertical line through vertex with max real part	1.117

Next the randomised algorithm was applied to the problem with parameters $\delta_1 = 0.04$ and $\epsilon = 0.02$. It was assumed that on termination the randomised algorithm enumerates only 4 out of 16 vertices whose convex hull $\mathcal{P}$ is shown in Figure 4.12. Note that all the omitted vertices have flat angles so $\mathcal{P}$ is the most probable outcome of the randomised algorithm. To obtain a probabilistic bound $\mathcal{P}$ was extended by the estimate of the Haussdorf distance $\upsilon\delta_1/2 = 0.14$, where $\upsilon = 14$ is the number of omitted vertices. Note that although in this case the extended polytope well covers the original one, this can be guaranteed only probabilistically when the vertices of $\mathcal{Z}$ cannot be fully enumerated. The corresponding probability in this case is $P = 1 - \epsilon \times 2^{\upsilon/2}$. In this example this is $P = 1 - 0.01 \times 2^{7/2} = 88\%$. Therefore, the result of this method, $\gamma_l = 1.420$ ($\mu_u = 0.704$) is guaranteed with a probability of 88%. Note that in general the choice of the most appropriate method is a compromise between accuracy (derivation of a tight bound) and computational complexity.

We conclude the chapter by illustrating the numerical performance of the algorithm. This time $M = M_r + jM_i$ (truncated to 3 decimal places) was defined as:

$$M_r = \begin{bmatrix} -0.166 & -0.059 & -0.116 & -0.000 & 0.109 & -0.097 & -0.066 & 0.039 & -0.096 \\ -0.009 & -0.029 & -0.005 & -0.018 & 0.005 & -0.030 & -0.018 & 0.005 & 0.001 \\ 0.040 & 0.098 & 0.023 & 0.060 & -0.018 & 0.110 & 0.059 & -0.028 & 0.004 \\ 0.026 & -0.013 & 0.022 & -0.014 & -0.020 & -0.008 & 0.001 & -0.000 & 0.021 \\ -0.087 & 0.033 & -0.065 & 0.044 & 0.061 & 0.015 & -0.002 & 0.008 & -0.073 \\ 0.126 & -0.107 & 0.091 & -0.102 & -0.091 & -0.082 & -0.023 & 0.005 & 0.117 \\ -0.171 & 0.015 & -0.121 & 0.049 & 0.118 & -0.020 & -0.029 & 0.024 & -0.122 \\ 0.006 & 0.027 & 0.002 & 0.017 & 0.000 & 0.028 & 0.016 & -0.006 & -0.004 \\ 0.000 & 0.163 & -0.004 & 0.115 & 0.013 & 0.171 & 0.080 & -0.040 & -0.047 \end{bmatrix}$$

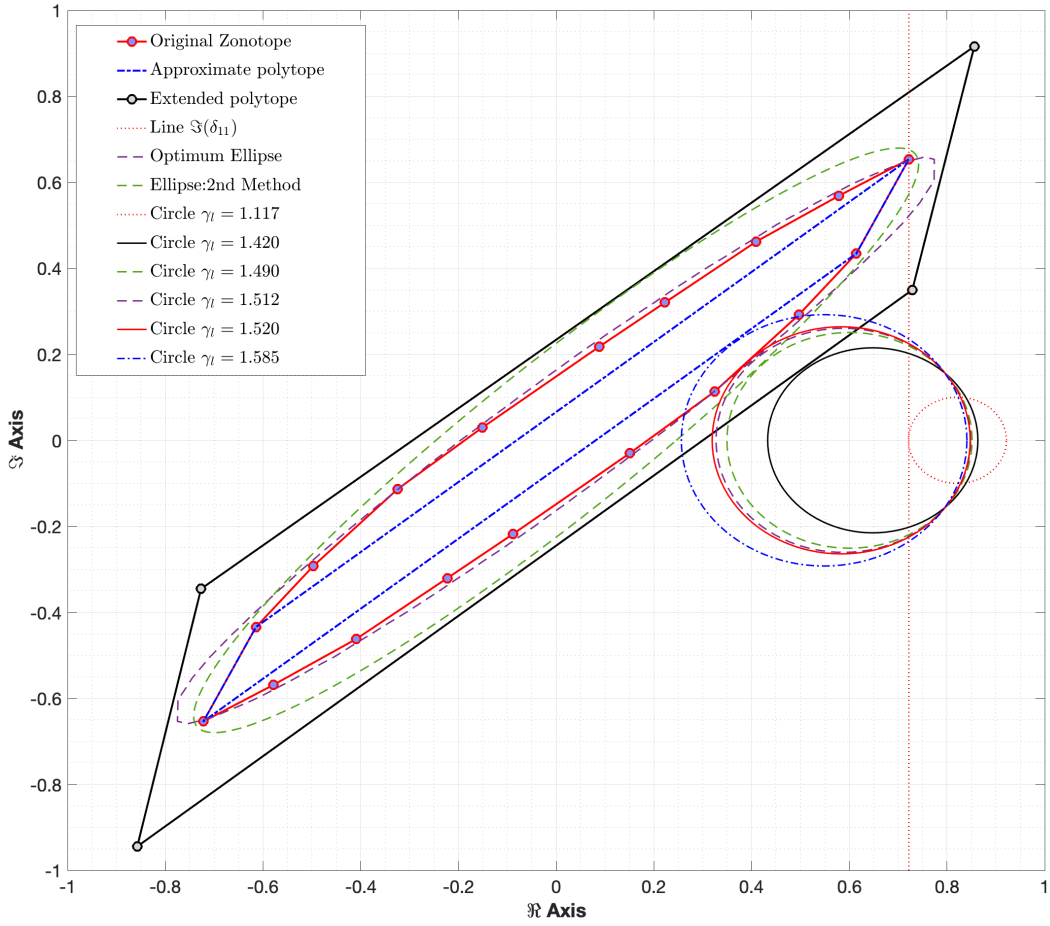

Figure 4.12: Numerical experiment with different methods

and

$$M_i = \begin{bmatrix} 0.125 & -0.256 & 0.097 & -0.207 & -0.103 & -0.238 & -0.093 & 0.038 & 0.156 \\ 0.031 & -0.023 & 0.020 & -0.022 & -0.021 & -0.020 & -0.005 & 0.002 & 0.026 \\ -0.105 & 0.102 & -0.078 & 0.093 & 0.074 & 0.081 & 0.024 & -0.007 & -0.099 \\ 0.001 & 0.025 & -0.000 & 0.020 & 0.001 & 0.028 & 0.015 & -0.006 & -0.007 \\ 0.011 & -0.105 & 0.008 & -0.071 & -0.014 & -0.105 & -0.048 & 0.021 & 0.032 \\ 0.043 & 0.123 & 0.023 & 0.076 & -0.013 & 0.138 & 0.067 & -0.034 & -0.008 \\ 0.061 & -0.229 & 0.053 & -0.174 & -0.060 & -0.218 & -0.095 & 0.043 & 0.107 \\ -0.027 & 0.020 & -0.016 & 0.022 & 0.021 & 0.012 & -0.000 & 0.001 & -0.023 \\ -0.144 & 0.073 & -0.104 & 0.083 & 0.100 & 0.044 & 0.000 & 0.009 & -0.121 \end{bmatrix}$$

The uncertainty structure in this case was defined as:

$$\mathbf{\Delta} = \{\text{diag}(\delta_1 I_3, \delta_2 I_3, \delta_3 I_3) : \delta_i \in \mathbb{R}, i = 1, 2, 3\} \tag{4.60}$$

The singular value decomposition $M = U\Sigma V^*$ was carried out via Matlab's *svd.m* function.

The corresponding Zonotope

$$\mathcal{Z} = \{v_1^* \Delta u_1 : \Delta = \text{diag}(\delta_1 I_3, \delta_2 I_3, \delta_3 I_3), -1 \leq \delta_i \leq 1, i = 1, 2, 3\} \tag{4.61}$$

(in which u_1 and v_1 denote the first columns of matrices U and V, respectively), has 6 vertices (and 2 internal points out of the total of $2^3 = 8$ extreme combinations of $v_1^* \Delta u_1$ when $\delta_i = \pm 1$). The convex (D-iteration) bound of $\mu_{\mathbf{\Delta}}(M)$ was calculated via Matlab's function *mussv.m* as $\bar{\mu}_1 = 0.282$. The corresponding bound obtained by the intersection of $\mathcal{Z}$ with the family of Apollonius circles is $\bar{\mu}_2 = 0.300$. The computational times needed to perform the calculation were $t_1 = 0.1976$ s and $t_2 = 0.0080$ s, respectively, corresponding to a speed-up factor of about 25 times when our method is applied. This is likely to further improve as the complexity of the problem increases. For the present problem, the largest computational load is the computation of the singular value decomposition. This could be avoided as what is really required is the pair of singular vectors corresponding to the largest singular value of M and this can be obtained efficiently by an alternative algorithm, e.g. the power method applied to MM^* and M^*M.

To compare the numerical efficiency of the two algorithms 1000 complex matrices of dimension 9×9 were generated randomly in Matlab and the ratio of the corresponding computational times was calculated in each case. The uncertainty structure remains the same as in equation (4.60). The histogram of the speed-up factor obtained with the proposed method is shown in Figure 4.13.

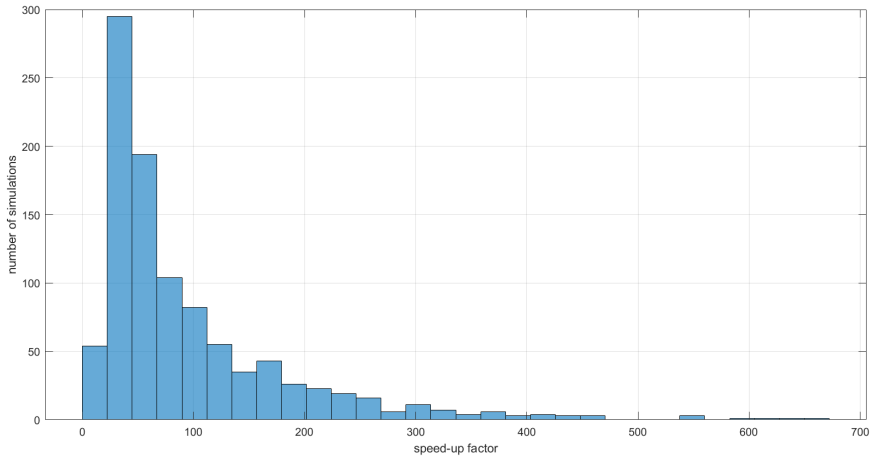

Figure 4.13: Speed-up factor

The average speed-up factor is 95.29 and the standard deviation of the distribution is 91.91. The minimum and maximum factors observed in the simulation were 5.24 and

669.69, respectively.

4.6 Summary

The problem of obtaining bounds on the structured singular value of a matrix M subject to real parametric uncertainty has been considered in this chapter. The method relies on two relaxation techniques involving the projection of the uncertainty in the "most critical direction" as defined by the set of optimal (maximum norm) unstructured perturbations. The first relaxation leads to the maximisation of a convex quadratic function with box constraints (for which convex bounds are possible). The second relaxation reduces to a geometric problem involving the intersection of a Zonotope and a family of Apollonius circles. This can be solved, provided the vertices of the Zonotope can be fully or partially enumerated. If full enumeration is possible (e.g. via an efficient algorithm based on "reverse-enumeration") a deterministic bound is obtained. For high dimensional problems, however, this may not be feasible computationally and we may need to rely on a randomisation algorithm producing only partial enumeration of the vertices. In this case, it may still be possible to obtain a probabilistic bound using the estimate of the Hausdorff distance between the true (but unknown) Zonotope and the convex hull of the polytope constructed from the partially enumerated vertices. Our approach is extended to the correlated-uncertainty case or when the uncertain perturbations enter the model in a nonlinear way. Several computational examples are included in this chapter to illustrate the main results of our work.

Chapter 5

Greatest Common Divisor(GCD): A Structured Singular Value Approach

5.1 Introduction

It is known that the computation of the Greatest Common Divisor (GCD) of two polynomials is a non-generic problem [76] with many application in engineering fields e.g. algebraic control methods, distance to controllability or observability, determinantal assignment problems, Robust Control, stability of dynamic systems subject to structured perturbations, Linear Systems, Numerical Analysis and other Engineering fields. This area is the subject of several researches in recent years who have attempted to introduce effective numerical calculations for the solution of the problem [77], [38] and [78]. Karcanias et al. in 1984 developed a non-generic techniques for calculating the nearest common root of polynomials with applications in the field of systems and control [77].

In [79] the notation of the "approximate coprimeness" of two polynomials has been introduced. This is defined as a minimum magnitude perturbation in polynomial's coefficient vectors such that the perturbed polynomials have a common root. Finding perturbations of this type is equivalent to the solution of the structured distance to singularity , or equivalently a structures singular value (μ) problem. The main advantage of this transformation is that it leads to various numerical techniques from the area of robust control which may be used to find a solution or approximate bounds to GCD problem.

Hence, the concept of "approximately coprimeness" can be defined as a distance from the nearest common divisor in an appropriate sense. Similar definition of "almost zeros" was first introduced in [31]. This definition has been reformulated to the notion of "approximate

GCD" which is, in fact, based on the relaxation of the conditions defining the exact GCD, see [32], [33], [34], [35], [36], [37] and references therein. Reference [80] proposed one of the most recent methods for calculating the distance of a set of co-prime polynomials to the set of polynomials sharing a common root. The technique is based on singular values to define and solve approximate GCD problems by converting the corresponding Sylvester matrix in the GCD problem to a diagonal matrix compatible to μ problem. A disadvantage of this method is that the size of equivalent objective matrix in the μ problem will increase considerably. To have a sense of how significant this could be, imagine that we want to solve a GCD problem of degree 10 by converting it to the μ problem. The equivalent M matrix would be 200×200. This will significantly increase the computational expense of the μ problem. We have mentioned in Chapter 4 that the structured singular value method is often computationally demanding. One way of overcoming this issue, which has been noted in most of references mentioned earlier, is to ignore the structure of the problem and use the singular value as an approximate measure of singularity (and thus loss of coprimeness). This, however, gives a loose bound which is far from the optimal solution in many cases. In this Chapter, we propose using the Lehtomaki approach [16] to perturbations matrix with a Sylvester structure. This gives an upper bound tighter than the largest singular value while avoiding exact calculations of high complexity. The advantages of this approach are illustrated via a numerical example. We use similar notation as in previous Chapters.

5.2 Minimum distance to common root of polynomials

A recent approach introduced in [79] proposes an algorithm for calculating the distance of two polynomials from non-coprimeness, i.e. the minimum size perturbations in the polynomial coefficients so that the two polynomials share a common root. It is shown that the problem is equivalent to the structured distance to singularity of the Sylvester matrix constructed from the nominal coefficients of the two polynomials.

The associated perturbation matrix inherits the structure of the Sylvester resultant matrix. Ignoring this structure, the smallest singular value of the Sylvester resultant matrix could be taken as an indicator for the distance to singularity. This, however, may be conservative. In the following subsection the problem of deriving a μ problem associated with the minimum distance to singularity is described.

5.2.1 Problem definition

Considering the definition of μ problem

$$\gamma_{\Delta} = \frac{1}{\mu_{\Delta}(M)} = \min_{\substack{\det(I-M\Delta)=0 \\ \Delta\in\mathbf{\Delta}}} \|\Delta\| \tag{5.1}$$

We first start with the following two Theorems from [79] that show how to define a Sylvester matrix related to two polynomials and how to formulate the problem of estimating the nearest common root in a μ framework.

Theorem 5.2.1. *[79] Let $S_{m,n}(a,b)$ be the Sylvester matrix related to two polynomial*

$$\begin{aligned} a(s) &= s^m + \alpha_{m-1}s^{m-1} + \alpha_{m-2}s^{m-2} + \cdots + \alpha_0 \\ b(s) &= s^n + \beta_{n-1}s^{n-1} + \beta_{n-2}s^{m-2} + \cdots + \beta_0 \end{aligned} \tag{5.2}$$

with $\partial a(s) = m$ and $\partial b(s) = n$, i.e.

$$S_{m,n}(a,b) = \begin{bmatrix} 1 & \alpha_{m-1} & \dots & \alpha_0 & 0 & \dots & 0 \\ 0 & 1 & \alpha_{m-1} & \dots & \alpha_0 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots & \\ 0 & 0 & \dots & 1 & \alpha_{m-1} & \dots & \alpha_0 \\ 1 & \beta_{m-1} & \dots & \beta_0 & 0 & \dots & 0 \\ 0 & 1 & \beta_{m-1} & \dots & \beta_0 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots & \\ 0 & 0 & \dots & 1 & \beta_{m-1} & \dots & \beta_0 \end{bmatrix} \in \mathbb{C}^{(m+n)\times(m+n)} \tag{5.3}$$

Then, $Rank(S_{m,n}(a,b)) = n + m - \partial\phi(s)$. Here, ∂ denotes the degree of the polynomial and ϕ is the GCD of $a(s), b(s)$.

Proof. see [79] and references therein. □

Theorem 5.2.1 shows that nullity of Sylvester matrix $S_{m,n}(a,b)$, identifies the degree of the GCD of $a(s), b(s)$. The following theorem establishes a connection between GCD and μ problems.

Theorem 5.2.2. *[79]*

Let

1. *$a_0(s)$ and $b_0(s)$ be two co-prime polynomial with $\partial a(s) = m$ and $\partial b(s) = n$ respectively.*

2. $S_{m,n}(a_0, b_0)$ *be the Sylvester matrix corresponding to* $a_0(s)$ *and* $b_0(s)$.

3. $\mathbf{\Delta}$ *be the following set of block diagonal perturbations: following structured*

$$\mathbf{\Delta} = \{diag\left(\delta_{m-1}I_n, \ldots, \delta_0 I_n, \epsilon_{n-1}I_m, \ldots, \epsilon_0 I_m\right) \ : \delta_i, \epsilon_i \in \mathbb{C}\} \tag{5.4}$$

4. $M = -ZS_{m,n}^{-1}\left(a_0, b_0\right)\Theta$, *where*

$$\begin{aligned} \Theta &= \left[\begin{array}{ccc|ccc} I_n & \ldots & I_n & 0_{n,m} & \ldots & 0_{n,m} \\ \hline 0_{m,n} & \ldots & 0_{m,n} & I_m & \ldots & I_m \end{array}\right] \in \mathbb{R}^{n+m,2nm} \\ Z' &= \left[\begin{array}{ccc|ccc} \left(Z_{nm}^0\right)' & \ldots & \left(Z_{nm}^{m-1}\right)' & \left(Z_{nm}^0\right)' & \ldots & \left(Z_{nm}^{m-1}\right)' \end{array}\right] \in \mathbb{R}^{n+m,2nm} \end{aligned} \tag{5.5}$$

in which

$$Z_{nm}^k = \left[\begin{array}{ccc} 0_{n,k+1} & I_n & 0_{n,m-k-1} \end{array}\right]$$

for $k = 0, 1, \ldots, m-1$.

And also lets $a(s) =$ and $b(s)$ be the perturbed polynomials,

$$\begin{aligned} a(s) =& a_0(s) + \delta(s) = s^m + (\alpha_{m-1} + \delta_{m-1})s^{m-1} + \cdots + \alpha_0 + \delta_0 \\ b(s) =& b_0(s) + \epsilon(s) = s^m + (\beta_{m-1} + \epsilon_{m-1})s^{m-1} + \cdots + \beta_0 + \epsilon_0 \end{aligned}$$

Then

$$\|\mu\| = \max\left\{|\delta_0|, \ldots, |\delta_{m-1}|, |\epsilon_0|, \ldots, |\epsilon_{m-1}|\right\}$$

and the minimum norm of $\|\Delta\|$ such that $a(s)$ and $b(s)$ have a common route is $\mu_{\Delta)^{-1}(M)}$.

Proof. The Sylvester equivalent matrix $S_{m,n}(a_0, b_0)$ is nonsingular due to the assumption that $a_0(s)$ and $b_0(s)$ are coprime. Perturbing both polynomials as. $a(s) = a_0(s) + \delta(s)$ and $b(s) = b_0(s) + \epsilon(s)$ result in a perturbation on the Sylvester matrix $S_{m,n}(a, b) = S_{m,n}(a_0, b_0) + E$ where E denotes the 'perturbation matrix':

$$E = \begin{bmatrix} 0 & \delta_{m-1} & \dots & \delta_0 & 0 & \dots & 0 \\ 0 & 0 & \delta_{m-1} & \dots & \delta_0 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots & \\ 0 & 0 & \dots & 0 & \delta_{m-1} & \dots & \delta_0 \\ 0 & \epsilon_{m-1} & \dots & \epsilon_0 & 0 & \dots & 0 \\ 0 & 0 & \epsilon_{m-1} & \dots & \epsilon_0 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots & \\ 0 & 0 & \dots & 0 & \epsilon_{m-1} & \dots & \epsilon_0 \end{bmatrix}$$

Matrix E can now be factored as $E = \Theta \Delta Z$, where Θ and Z are defined in (5.5) and $\mathbf{\Delta}$ is defined in (5.4).

Clearly $\mathbf{\Delta}$ is diagonal and

$$\max\{|\delta_0|, \dots, |\delta_{m-1}|, |\epsilon_0|, \dots, |\epsilon_{m-1}|\} = \|\Delta\|$$

Since $S_{m,n}(a_0, b_0)$ loses rank if and only if the polynomials $a(s)$ and $b(s)$ have a common root, the problem is equivalent to

$$\min \|\Delta\| \text{ such that } \det(S_{m,n}(a_0, b_0) + \Theta \Delta Z) = 0 \text{ and } \Delta \in \mathbf{\Delta} \tag{5.6}$$

Using the matrix identity,

$$\det(I + BC) = \det(I + CB) \tag{5.7}$$

we conclude that

$$\det\left(S_{m,n}(a_0, b_0) + \Theta \Delta Z\right) = 0 \Leftrightarrow \det\left(I + Z S_{m,n}^{-1}(a_0, b_0) \Theta \Delta\right) = 0 \Leftrightarrow \det\left(I - M\Delta\right) = 0 \tag{5.8}$$

Hence the problem becomes

$$\min\{\|\Delta\| : \det(I - M\Delta) = 0, \Delta \in \mathbf{\Delta}\} = \mu_{\mathbf{\Delta}}^{-1}(M)$$

□

Remark 5.2.1. *Note that GCD problem with a Sylvester matrix of dimension* $(n + m) \times (n + m)$ *is equivalent to a* μ *problem of a matrix of dimensions* $2nm \times 2nm$

Converting the GCD to a μ problem allows us to use existing algorithms for the solution of the approximate GCD problem at the disadvantage that the dimensionality of the problem increases significantly. Therefore, conversion to a μ problem may not be a cost efficient method for large scale systems. In the following section, we adapt the Lehtomaki approach to find a tight bound to the μ problem which can be used to estimate the value of the structured singular value using straightforward calculations.

As a practical example of GCD in control systems we can name the "Distance to Uncontrollability", which refers to controllablity of a given linear time-invariant system [81]. A system represented by state space with parameters A and B is state controllable if and only if the matrix

$$\mathcal{S}(A,B) := \begin{bmatrix} A & AB & \dots & A^{n-1}B \end{bmatrix}$$

is full rank. This involves rank computation. In fact, checking the controllability property requires a numerical test on the parameters of a specific representation of the system. Now consider two polynomials $a(s)$ and $b(s)$ in frequency domain. The single-input single-output system $\mathcal{S}(a(s), b(s))$ is controllable if an only if $a(s)$ and $b(s)$ are co-prime[82]. This is exactly the GCD problem of two polynomials.

5.2.2 Lehtomaki Approach

Considering $S_{m,n}(a,b)$ in (5.3), let the corresponding perturbation matrix, Δ belongs to the following structured set:

$$\mathbf{\Delta} = \{ \begin{bmatrix} 0 & \delta_{m-1} & \dots & \delta_0 & 0 & \dots & 0 \\ 0 & 0 & \delta_{m-1} & \dots & \delta_0 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots & \\ 0 & 0 & \dots & 0 & \delta_{m-1} & \dots & \delta_0 \\ 0 & \epsilon_{n-1} & \dots & \epsilon_0 & 0 & \dots & 0 \\ 0 & 0 & \epsilon_{n-1} & \dots & \epsilon_0 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots & \\ 0 & 0 & \dots & 0 & \epsilon_{n-1} & \dots & \epsilon_0 \end{bmatrix} : \delta_i, \epsilon_i \in \mathbb{C} \} \tag{5.9}$$

Thus, we need to solve the following structured distance to singularity problem:

$$\min_{\substack{\det(S_{m,n}(a,b)+\Delta)=0 \\ \Delta\in\mathbf{\Delta}}} \|\Delta\| \tag{5.10}$$

Let $S_{m,n}(a,b)$ have a singular value decomposition

$$S_{m,n}(a,b) = U\Sigma V^* = \left[\begin{array}{c|c|c|c} u_1 & u_2 & \dots & u_n \end{array}\right] \begin{bmatrix} \sigma_1 & 0 & \dots & 0 \\ 0 & \sigma_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sigma_n \end{bmatrix} \left[\begin{array}{c} v_1^* \\ \hline v_2^* \\ \hline \vdots \\ \hline u_n \end{array}\right]$$

where $u_i \in \mathbb{C}^{n\times i}$, $v_i \in \mathbb{C}^{n\times i}$ for $i = 1 \dots n$.

Then

$$\begin{aligned} &\det\left(S_{m,n}(a,b) + \Delta\right) = 0 \Leftrightarrow \\ &\det\left(U\Sigma V^* + \Delta\right) = 0 \Leftrightarrow \\ &\det(\Sigma + U^*\Delta V) = 0 \end{aligned} \tag{5.11}$$

The last statement is due to the orthogonality of U and V. Now we impose a constrain on Δ which does not allow a projection equal to σ_n in the most sensitive direction, $u_n v_n^*$. Problem (5.10) is thus relaxed to

$$\min_{\substack{(\Sigma+U^*\Delta V)=0 \\ \Delta\in\mathbf{\Delta} \\ |u_n\Delta v_n^*|\leq\phi<\sigma_n}} \|\Delta\| \tag{5.12}$$

The optimal value of Problem (5.12) is [16]:

$$\|U^*\Delta V\| = \sqrt{\sigma_n\sigma_{n-1} + \phi(\sigma_n - \sigma_{n-1})} \tag{5.13}$$

which implies that if the norm of Δ is less than the value in the right hand side of (5.12) then $\Sigma + U^*\Delta V$ is nonsingular,

Note that ϕ is an arbitrary non-negative parameter between 0 and σ_n. Setting $\phi = 0$ gives the maximum distance to singularity $\sqrt{\sigma_n\sigma_{n-1}}$, while setting $\phi = \sigma_n$ recovers the unstructured distance to singularity equal to σ_n. Hence, the more $|u_n\Delta v_n^*|$ is restricted, the better the bound achieved.

Now, assume that Δ varies in the structured set of unit ball $\mathbf{B}_{\mathbf{\Delta}}$. What is the minimum value of ϕ compatible with this structure, i.e. the smallest possible ϕ for which $|v_n^* \Delta u_n| \leq \phi$ for all Δ in $\mathbf{B}_{\mathbf{\Delta}}$? Clearly, the optimal ϕ is obtained by solving:

$$\max_{\Delta \in \mathbf{B}_{\mathbf{\Delta}}} |v_n^* \delta u_n| = \phi \tag{5.14}$$

Note that values of ϕ is the largest of the optimal value could be used and would result in a lower bound on the structured distance to singularity (equivalent upper bound to the structured singular value) by applying equation (5.13). The tightest possible bound, however, is obtained by using the optimal value of ϕ given by (5.14).

The above discussion is summarised in the following Lemma:

Lemma 5.2.1. *Let M, σ_n, σ_{n-1}, Σ and U be as defined as above and define*

$$\mu_{\Delta}(M) := \frac{1}{\min\{\|\Delta\| : \Delta \in \mathbf{\Delta}, \det((A - U^* \Delta V)) = 0\}} \tag{5.15}$$

where $\mathbf{\Delta}$ has the structure defined in (5.9)*. Then*

$$\mu_{\mathbf{\Delta}}(M) \leq \left(\sigma_n \sqrt{\frac{\sigma_{n-1}}{\sigma_n} + \tilde{\phi}_0 \left(1 - \frac{\sigma_{n-1}}{\sigma_n}\right)} \right)^{-1} \tag{5.16}$$

where

$$\tilde{\phi}_0 = \max_{\Delta \in \boldsymbol{B}_{\mathbf{\Delta}}} |u_n \Delta v_n^*| \tag{5.17}$$

Proof. Suppose $\mathbf{B}_{\mathbf{\Delta}}$ denote the unit ball as defined in chapter 4,

$$\mathbf{B}_{\mathbf{\Delta}} = \{\Delta \in \mathbf{\Delta} : \|\Delta\| \leq 1\}$$

Thus (5.14) implies that

$$\begin{aligned}
&\max_{\Delta \in \mathbf{\Delta}} |v_n^* \Delta u_n| = \phi \\
&\Rightarrow \max_{\Delta \in \mathbf{\Delta}} |u_n \frac{\Delta}{\|\Delta\|} v_n^*| = \phi \|\Delta\|^{-1} \\
&\Rightarrow \max_{\tilde{\Delta} \in \mathbf{B}_{\mathbf{\Delta}}} |u_n \tilde{\Delta} v_n^*| = \tilde{\phi} \sigma_n \|\Delta\|^{-1}
\end{aligned}$$

where $\tilde{\Delta} = \dfrac{\Delta}{\|\Delta\|}$ and $0 \leq \tilde{\phi} = \dfrac{\phi}{\sigma_n} \leq 1$.

Thus, if

$$|u_n \tilde{\Delta} v_n^*| \leq \tilde{\phi} \sigma_n \|\Delta\|^{-1} \quad \forall \Delta \in \mathbf{B}_{\mathbf{\Delta}}, 0 \leq \tilde{\phi} \leq 1$$

then from (5.13)

$$\min_{\substack{\Delta\in\mathbf{\Delta}\\ \det(A+U^*\Delta V)=0}} \|\Delta\| \geq \sigma_n \sqrt{\frac{\sigma_{n-1}}{\sigma_n} + \tilde{\phi}\left(1 - \frac{\sigma_{n-1}}{\sigma_n}\right)} \tag{5.18}$$

Equivalently, if

$$\max_{\tilde{\Delta}\in\mathbf{B_\Delta}} \det(u_n \tilde{\Delta} v_n^*) = \tilde{\phi}_0 \leq 1 \tag{5.19}$$

then μ_Δ has an upper bound of

$$\left(\sigma_n \sqrt{\frac{\sigma_{n-1}}{\sigma_n} + \tilde{\phi}_0\left(1 - \frac{\sigma_{n-1}}{\sigma_n}\right)}\right)^{-1} \tag{5.20}$$

□

Although Lemma 5.2.1 provide a better upper bound to the structured singular value compare to the norm of M, it still requires solving an optimisation problem in (5.19). This optimisation is not an easy problem to solve due to the structure of $\mathbf{B_\Delta}$. Thus we can define a set with a more convenient structure which contains $\mathbf{B_\Delta}$. Define:

$$\bar{\mathbf{\Delta}} = \{\Delta \in \mathbf{\Delta} : \quad |\delta_i| \leq 1, \quad |\epsilon_j| \leq 1, \forall \quad 0 \leq i \leq m-1, \quad 0 \leq j \leq n-1\} \tag{5.21}$$

Solving (5.19) over $\Delta \in \bar{\mathbf{\Delta}}$ is now straightforward, however, the question that is raised is whether a tighter set could be found that includes $\mathbf{B_\Delta}$? In other words, can we find a tighter subset, $\boldsymbol{\theta}\bar{\mathbf{\Delta}}$ where $\mathbf{B_\Delta} \subseteq \theta\bar{\mathbf{\Delta}}$ for any $0 < \theta \leq 1$? Figure 5.1 gives a visual interpretation of the set inclusion where $\theta_0 = \{\min \theta : \text{s.t.} \boldsymbol{\beta}\mathbf{\Delta} \subseteq \theta\overline{\Delta}\}$.

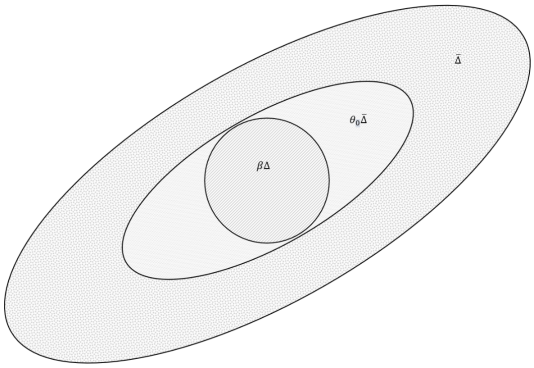

Figure 5.1: Different structured perturbation subsets

The following lemma gives a negative answer to the question posed above.

Lemma 5.2.2. *Let $\overline{\Delta}$ and $\boldsymbol{B}_{\boldsymbol{\Delta}}$ are as defined previously; Then*

$$\theta_0 = \left\{\min \theta : s.t. \quad \boldsymbol{B}_{\boldsymbol{\Delta}} \subseteq \theta\overline{\boldsymbol{\Delta}}\right\} = 1$$

Proof. Assume there exist a $0 \leq \theta < 1$ in which $\mathbf{B}_{\boldsymbol{\Delta}} \subseteq \theta\overline{\boldsymbol{\Delta}}$ this means that for each $\Delta_1 \in \mathbf{B}_{\boldsymbol{\Delta}}$ there exists $\Delta_2 \in \overline{\boldsymbol{\Delta}}$ such that $\Delta_1 \subseteq \Delta_2$. This implies $\|\Delta_1\| < \|\Delta_2\|$. Now consider the perturbation matrix, $\Delta \in \boldsymbol{\Delta}$, in 5.9 and let $\Delta_1, \Delta_2 \in \boldsymbol{\Delta}$ where $\Delta_1 = \{\Delta \in \mathbf{B}_{\boldsymbol{\Delta}}, \|\Delta\| = 1\}$ and $\Delta_2 = \left\{\Delta \in \overline{\boldsymbol{\Delta}}, \Delta_2 = 1, \delta_i = 0, \epsilon_j = 0 \text{ for other } i, j\right\}$. It is obvious that $\|\Delta_2\| = 1$ which is in contradiction with $\|\Delta_1\| < \|\Delta_2\|$. Hence, $\theta = 1$. □

Lemma 5.2.2 shows that the set defined in (5.21) is the tightest set with the given structure that includes $\mathbf{B}_{\boldsymbol{\Delta}}$. Note that the problem over the set in 5.2.1 is more convenient and can be readily be solved.

Lemma 5.2.3. *Assume all the variables of Lemma 5.2.1 and let*

$$\tilde{\phi}_0 = \max_{\Delta \in \bar{\boldsymbol{\Delta}}} |u_n \tilde{\Delta} v_n^*| = \max_{\substack{|\delta_i| \leq 1 \\ |\epsilon_i| \leq 1}} |\sum_{i=1}^{m} \delta_i x_i + \sum_{j=1}^{n} \epsilon_j y_j| \tag{5.22}$$

where $x_i, y_i \in \mathbb{C}$ are defined explicitly from the term $u_n \tilde{\Delta} v_n^*$ and $\bar{\Delta}$ is defined in (5.21)*. Then if*

$$\sum_{i=1}^{m} |x_i| + \sum_{j=1}^{n} |y_j| < 1 \tag{5.23}$$

μ_Δ has an upper bound of

$$\mu_\Delta \leq \left(\sqrt{\sigma_n \sigma_{n-1} + \left(\sum_{i=1}^{m} |x_i| + \sum_{j=1}^{n} |y_j| \right) (\sigma_n^2 - \sigma_n \sigma_{n-1})} \right)^{-1} \tag{5.24}$$

Proof. The maximiser in (5.22) occurs when all δ_i and ϵ_i are at their maximum absolute values with a sign that makes x_i and y_i positive, hence

$$\begin{aligned} \tilde{\phi}_0 &= \max_{\Delta \in \bar{\boldsymbol{\Delta}}} |u_n \tilde{\Delta} v_n^*| \\ &= \max_{\substack{|\delta_i| \leq 1 \\ |\epsilon_i| \leq 1}} |\sum_{i=1}^{m} \delta_i x_i + \sum_{j=1}^{n} \epsilon_j y_j| \\ &= \sum_{i=1}^{m} |x_i| + \sum_{j=1}^{n} |y_j| \end{aligned} \tag{5.25}$$

Thus, the distance to the singularity is bounded by

$$\sqrt{\sigma_n\sigma_{n-1} + \left(\sum_{i=1}^{m}|x_i| + \sum_{j=1}^{n}|y_j|\right)(\sigma_n^2 - \sigma_n\sigma_{n-1})} \tag{5.26}$$

□

As required, Lemma 5.2.3 gives an estimate to the distance of the two polynomials from having a common root without the need to undergo the expensive computation of the structured singular value of a high dimensional matrix. This is demonstrated in the next section with a numerical example.

5.2.3 Numerical Example

We select two simple polynomial of degree 2 and 3 which have a common root.

$$a(s) = s^3 - 8.29s^2 + 21.12s - 15.68$$
$$b(s) = s^2 - 8.70s + 18.77$$

The roots of each polynomial are

$$r_a = [3.95, 3.02, 1.31]$$
$$r_b = [4.74, 3.95]$$

We can see that 3.95 is a common root of the two polynomials. We now perturbed the coefficients by small amounts

$$\delta = \begin{bmatrix} 0 & -0.10 & 0.10 & -0.08 \end{bmatrix}$$
$$\epsilon = \begin{bmatrix} 0 & -0.08 & -0.05 \end{bmatrix}$$

corresponding to a Δ matrix in the form of (5.9)

$$\mathbf{\Delta} = \begin{bmatrix} 0 & -0.10 & 0.10 & -0.08 & 0 \\ 0 & 0 & -0.10 & 0.10 & -0.08 \\ 0 & -0.08 & -0.05 & 0 & 0 \\ 0 & 0 & -0.08 & -0.05 & 0 \\ 0 & 0 & 0 & -0.08 & -0.05 \end{bmatrix} \tag{5.27}$$

Note that the norm of the above perturbation matrix is $\|\Delta\| = 0.1903$. Nonetheless, this is not the smallest norm of a perturbation matrix that result in singularity. In other words a Δ with smaller norm may cause singularity. However, it can be taken as an index to test the final result in this example. Perturbing these two polynomials result in two co-prime polynomials,

$$a_0(s) = s^3 - 8.39s^2 + 21.05s - 15.60$$
$$b_0(s) = s^2 - 8.78s + 18.74$$

whose roots are

$$r_{a_0} = [4.44, 2.59, 1.35]$$
$$r_{b_0} = [5.13, 3.65]$$

The corresponding Sylvestor matrix is

$$S_{3,2}(a_0, b_0) = \begin{bmatrix} 1 & -8.39 & 21.05 & -15.60 & 0 \\ 0 & 1 & -8.39 & 21.05 & -15.60 \\ 1 & -8.78 & 18.74 & 0 & 0 \\ 0 & 1 & -8.78 & 18.74 & 0 \\ 0 & 0 & 1 & -8.78 & 18.74 \end{bmatrix}$$

which has eigenvalues

$$\lambda(S_{3,2}(a_0, b_0)) = \{30.89, 13.66 \mp j4.46, 0.05, -0.03\}$$

As we can see, the Sylvester matrix has some eigenvalues close to 0. This is due to the fact that the Sylvester matrix is close to singularity and a small perturbation ($\|\Delta\| \leq 0.1903$) may result in singularity. Now referring to theorem 5.2.1, the transformation $M = -ZS_{m,n}^{-1}(a_0, b_0)\,\Theta$ where

$$Z^T = \left[\begin{array}{cc|cc|cc|ccc|ccc} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 \end{array}\right]$$

and

$$\Theta = \left[\begin{array}{cc|cc|cc|ccc|ccc} 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 \end{array}\right]$$

gives

$$M = \left[\begin{array}{cccccccccccc} -31.7 & 141.9 & -31.7 & 141.9 & -31.7 & 141.9 & 31.7 & -130.5 & 118.1 & 31.8 & -130.5 & 118.1 \\ -7.6 & 34.7 & -7.6 & 34.7 & -7.6 & 34.7 & 7.6 & -31.7 & 28.9 & 7.5 & -31.7 & 28.9 \\ -7.6 & 34.7 & -7.6 & 34.7 & -7.6 & 34.7 & 7.6 & -31.7 & 28.9 & 7.5 & -31.7 & 28.9 \\ -1.8 & 8.7 & -1.8 & 8.7 & -1.8 & 8.7 & 1.8 & -7.9 & 7.2 & 1.8 & -7.9 & 7.2 \\ -1.8 & 8.7 & -1.8 & 8.7 & -1.8 & 8.7 & 1.8 & -7.9 & 7.2 & 1.8 & -7.9 & 7.2 \\ -0.5 & 2.2 & -0.5 & 2.2 & -0.5 & 2.2 & 0.5 & -2.0 & 1.8 & 0.5 & -2.0 & 1.8 \\ -31.7 & 141.9 & -31.7 & 141.9 & -31.7 & 141.9 & 31.7 & -130.5 & 118.1 & 31.8 & -130.5 & 118.1 \\ -7.6 & 34.7 & -7.6 & 34.7 & -7.6 & 34.7 & 7.6 & -31.7 & 28.9 & 7.5 & -31.7 & 28.9 \\ -1.8 & 8.7 & -1.8 & 8.7 & -1.8 & 8.7 & 1.8 & -7.9 & 7.2 & 1.8 & -7.9 & 7.2 \\ -7.6 & 34.7 & -7.6 & 34.7 & -7.6 & 34.7 & 7.6 & -31.7 & 28.9 & 7.5 & -31.7 & 28.9 \\ -1.8 & 8.7 & -1.8 & 8.7 & -1.8 & 8.7 & 1.8 & -7.9 & 7.2 & 1.8 & -7.9 & 7.2 \\ -0.5 & 2.2 & -0.5 & 2.2 & -0.5 & 2.2 & 0.5 & -2.0 & 1.8 & 0.5 & -2.0 & 1.8 \end{array}\right]$$

which is the objective matrix for the equivalent μ problem. First, it can be observed that the size of M matrix grows compared to the Sylvester matrix of the original GCD problem.

We first use D-iteration techniques [6] to calculate the convex upper bound on μ using Matlab software. The result is

$$0.0719 \leq \gamma_{\Delta}(M) = \mu_{\Delta}^{-1}(M) \leq 0.0748$$

Note that this value is considerable better than the smallest singular value of $S_{3,2}(a_0, b_0) = 0.0010$. However, as mentioned earlier, it is been achieved at the cost of solving a μ problem over 12×12 matrix. The improved estimation offers the unstructured distance to singularity, while avoiding the expensive calculation required by the computation of μ. To apply the

proposed method, first we check condition (5.23):

$$
\begin{aligned}
&|v_n^* \Delta u_n| = \\
&| \begin{bmatrix} -0.1411 & 0.6141 & 0.1402 & -0.5676 & 0.5110 \end{bmatrix} \\
&\begin{bmatrix} 0 & \delta_2 & \delta_1 & \delta_0 & 0 \\ 0 & 0 & \delta_2 & \delta_1 & \delta_0 \\ 0 & \epsilon_1 & \epsilon_0 & 0 & 0 \\ 0 & 0 & \epsilon_1 & \epsilon_0 & 0 \\ 0 & 0 & 0 & \epsilon_1 & \epsilon_0 \end{bmatrix} \begin{bmatrix} -0.3257 \\ 0.6806 \\ 0.6554 \\ -0.0332 \\ -0.0036 \end{bmatrix} | = \\
&|0.306\delta_2 - 0.113\delta_1 + 0.002\delta_0 - 0.294\epsilon_1 + 0.109\epsilon_0|
\end{aligned}
$$

Thus

$$
\begin{aligned}
\tilde{\phi}_0 &= \max_{\Delta \in \tilde{\mathbf{\Delta}}} |u_n \tilde{\Delta} v_n^*| \\
&= \sum_{i=1}^{m} |x_i| + \sum_{j=1}^{n} |y_j| \\
&= 0.306 + 0.113 + 0.002 + 0.294 + 0.109 \\
&= 0.824 < 1
\end{aligned}
$$

Applying (5.26) then gives

$$
\gamma_\Delta = \sqrt{0.001 \times 0.78 + 0.824\,(0.001^2 - 0.001 \times 0.78)} = 0.012
$$

It is seen that with a trivial calculation, an improved bound is achieved. Although the result is not as tight as the convex bound, it is considerably tighter compared to the singular value without having to apply the costly D-iteration method. To compare the distance to singularity of the proposed method with the greatest singular value and D-iteration estimate of μ, a comparison over 70 random examples with the the same value range fulfilling condition (5.19) has been carried out.

The hardware and software used to run the algorithm have the following specification:

- CPU: 1 (4 Cores) Processor Intel Core i5 2400s
- RAM: 4GB DDR3
- Graphic: Radeon HD 6770M

- Software: Matlab 9.5 R2018a

Figure 5.2 shows the result.

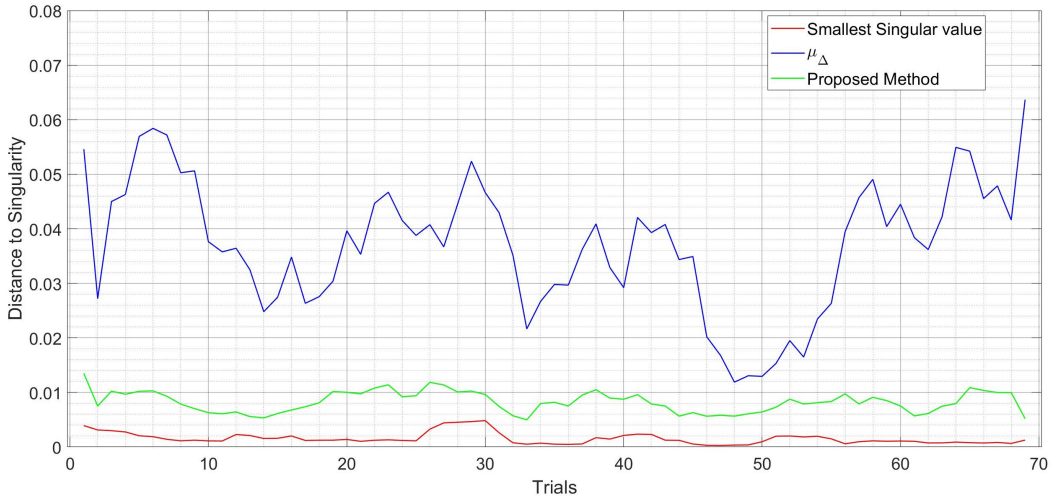

Figure 5.2: 70 other examples

Table 5.1: The proposed method in compare to the other conventional methods

Measures	Distance to singularity	CPU time(ms)
D-iteration	0.0719	302.03
Greatest singular value	0.0010	0.20
Proposed Method	0.0120	95.34

Table 5.1 shows the D-iteration method gives the wider distance to singularity (tighter upper bound on μ_Δ). However, the corresponding CPU time is significantly longer than those of the other methods. On the other hand, calculating the largest singular value is the fastest method but it offers the loosest bound. The bound obtained by the proposed method offers a reasonable compromise between these two extremes.

Zonotope can also be applied when polynomial coefficients (and M) are complex and perturbations as assumed real. Thus all the proposed method in chapter 4 are applicable. This remains to be studied in the future works.

5.3 Summary

The Greatest Common Divisor (GCD) of two polynomials is considered in this chapter where the GCD problem is converted to an equivalent structured singular value problem. This involves a significant increase in the size of the problem and thus may not be the most cost efficient solution. A method is proposed in this chapter that is directly applicable to

Sylvester uncertainty structures before converting to a μ problem. This method will avoid solving a costly high dimentional μ problem whilst it offers a bound tighter than the largest singular value. A numerical example compares the proposed method with the μ bounds obtained by the largest singular value and the convex relaxation ("D-iteration") method.

Chapter 6

Conclusions

In this chapter we summarise the results of this work . We also propose new areas of research that follow naturally from the results presented in this thesis. Two classical NP-hard problems, i.e. the general real μ problem and Quadratic Integer Programming problem are the main subjects of this thesis. The conventional methods to approach these problems are mainly deterministic. Time and space complexity is always a fundamental issue in these methods which could make them impractical if the control system has a large scale with many entries. This thesis deal with this part of the problem and aimed to generalise a convex approach to find an a probabilistic cost efficient bounds on both problems. The main motivation to study the QIP and μ problems is their unique applications in control systems. In addition, the various striking similarities existing in both solutions motivates their simultaneous study. Hence, the solutions of both approaches are similar and that one set of results can be applied to the other problem. Some of these similarities between the two problems that have been studying in this thesis is summarised in Table 6.1.

6.1 Randomised Algorithm to obtain an upper bound on QIP

A Randomised algorithm is proposed in chapter 2 for calculating an upper bound of the QIP problem. It is known that the convex relaxation of the problem reduces to a Reduced-Rank QIP (RRQIP) problem which is equivalent to the enumeration of the vertices of a Zonotope. The proposed approach derives an estimate of the Hausdorff distance between the Zonotope and its approximation, defined as the convex hull of the partially enumerated vertices of the Zonotope that are obtained by random sampling.

In current methods, the duality gap of the problem, i.e. the distance between the convex upper bound and the optimal solution, could be reduced provided the RRQIP problem can

μ-problem	**QIP problem**
$\mu^{-1} = \min_{\substack{\Delta \in \mathbf{\Delta} \\ det(I-\Delta M)=0}} \|\Delta\|$	$\gamma := \max_{x \in \{-1,1\}^n} x'Qx$
Formulate and solve dual: Get (μ_0, D_0) s.t. $\min_{D \in \mathbf{\Delta}} \|D^{1/2} M D^{-1/2}\|$ $M \leftarrow \mu_0^{-1} D_0^{1/2} M D_0^{-1/2}$	Formulate and solve dual: Get$(\overline{\gamma}, D_0)$ $s.t. \quad \overline{\gamma} = min \quad trace(D)$ $s.t. \quad D = diag(D), D - Q \geq 0$
Sufficient condition for zero duality gap $(\mu = \overline{\mu} \quad \text{if} \quad null(I - MM^*) = 1)$	Sufficient condition for zero duality gap $(\gamma = \overline{\gamma} \quad \text{if} \quad null(D_0 - Q) = 1)$
Duality gap can be breached if we can solve a rank-m, μ problem or get a bound less than 1: $\max_{\Delta \in \mathbf{\Delta}} \rho(\Delta U_1 v_1^*)$	Duality gap can be breached if we can solve a rank-r QIP problem of the form $\gamma_r = \frac{1}{n} \max_{x \in \{-1,1\}^n} x'VV'x$
Improved bound obtained via the solution of an eigenvalue problem	Improved bound can be obtained by calculating the eigenvalues of a symmetric matrix

Table 6.1: Common characteristic of μ-problem and QIP

be solved. This is equivalent to the full enumeration of the vertices of the corresponding Zonotope. Although this can be achieved by a polynomial-time algorithm, the computation may still be intractable for problems of high-dimensionality. It was shown that even with incomplete enumeration when the full enumeration is computationally infeasible, an improved probabilistic bounds could be obtained. The bound may be further improved by combining the proposed randomised approach with the solution of a sequence of deterministic QIP problems of increasing rank. In other words, the proposed method derives a decreasing sequence of upper bounds to the solution of the QIP problem by solving a sequence of auxiliary QIP problems of increasing rank. This is useful in practice since it can exploit fully the available computational resources to obtain the tightest bound possible with a pre-specified probability. By identifying the optimal transition between deterministic (full) and randomised (partial) vertex enumeration, it is also possible to obtain the tightest bound compatible with the available computational resources. This was the first major contribution of this work which is titled as "Propose a cost-efficient method for breaching the gap between the QIP problem andits convex relaxation" in introduction. The proposed method is demonstrated by the graphical interpretation of solution in a the form of 2 dimensional Zonotope.

First the algorithm features was evaluated and it was shown that the algorithm tend to enumerate those key vertices which contribute more in forming the Zonotope with higher probability. The proposed method is further improved to overcome the feature that the

approximate Zonotope may omit the optimiser and hence the upper bound breaches the solution of QIP. This was based on extending the gap between the QIP problem and its upper bound just sufficiently to cover all vertices which could potentially be selected by the algorithm, even with low probability. The overall result thus is valid with a certain probability. It was also shown that the trade-off between accuracy, speed and probability is always adjustable by selecting the algorithm's predefined parameters appropriately. Furthermore, a new probabilistic condition is derived for which the reduced rank QIP guarantees that the duality gap of the original problem is breachable. The validity of the proposed method has been illustrated by several numerical examples.The bound may also be further improved by combining the proposed randomised approach with the solution of a sequence of deterministic QIP problems of increasing rank. This is a minor contribution titled as "Improved the convex bound by combining the proposed randomised approach with the solution of a sequence of deterministic QIP problems of increasing rank" in the introduction.

6.2 Semidefinite upper bounds on μ-problem

In chapter 3 and chapter 4, μ-problem was the focus of the research.

In chapter 3, the problem of finding the smaller distance to the singularity,i.e.

$$\gamma = \min \{\|\Delta\| : det\,(A - \Delta) = 0\}$$

where A is a positive definite diagonal matrix was considered. This problem has a well known solution. The solution is novel if you constraint it further, e.g. require that the (1,1) block is a fixed matrix. The problem has been studied under different conditions imposed on the $k \times k$ leading block of Δ, Δ_{11} representing different classes of uncertainty. When we impose the constraint that $\|\Delta\| \leq 1$ the optimum value of γ increases and the optimisation reduces to an eigenvalue problem. This result has been used in chapter 4 where the problem of obtaining bounds on the structured singular value of a matrix M subject to real parametric uncertainty has been considered. The method relies on two relaxation techniques involving the projection of the uncertainty in the "most critical direction" as defined by the set of optimal (maximum norm) unstructured perturbations. The first relaxation leads to the maximisation of a convex quadratic function with box constraints (for which convex bounds are possible).

The second relaxation reduces to a geometric problem involving the intersection of a

Zonotope and a family of Apollonius circles. This can be solved, provided the vertices of the Zonotope can be enumerated. If full enumeration is possible (e.g. via an efficient algorithm based on "reverse-enumeration") a deterministic bound is obtained. For high dimensional problems, however, this may not be feasible computationally and we may need to rely on a randomisation algorithm producing only partial enumeration of the vertices. In this case, it may still be possible to obtain a probabilistic bound using the estimate of the Hausdorff distance between the true (but unknown) Zonotope and the convex hull of the polytope constructed from the partially enumerated vertices.

Both relaxations have been further studied in the special case where the multiplicity of the largest singular value, m, is equal to 1. This results in proposing a probabilistic method of calculating a convex upper bound on μ which is one of the main contributions as mentioned in introduction. The optimisation problem associated with the first relaxation has been specifically derived. Our approach is extended to the correlated-uncertainty case or when the uncertain perturbations enter the model in a nonlinear way. This was another main contribution as titled "Extend the main results of the proposed method to the distance to singularity problems with "correlated" or nonlinear descriptions of uncertainty" in introduction.

Approximate methods for estimating γ have also been considered. First an immediate low-cost upper bound on γ has been obtained by considering the intersection of an Apollonius circle with the line $\xi = \max_{z\in\mathcal{Z}} \Re(z)$. This bound could be obtained without enumerating the vertices of the corresponding Zonotope, $\mathcal{Z}$, which is equivalent to solving $2|\text{vert}(\mathcal{Z})| - 1$ scalar problems. It was also shown that avoiding the solution of these scalar problems is possible by constructing the ellipse of minimum area which contains all vertices of the Zonotope (and thus by convexity the entire Zonotope). In this way obtaining a lower bound on the structured distance to singularity is possible. The method of finding this ellipse however requires the enumeration of the vertices of $\mathcal{Z}$ which may be impractical for high dimensional problems. Hence, a method is proposed which constructs an ellipse containing $\mathcal{Z}$ directly from the generating matrix without enumerating the vertices.

A link between the two relaxations has been established by noting the scaling which allows us to transform the results between the two relaxation methods. Several computational examples are included to illustrate the main results of our work.

6.3 A method applicable directly to the GCD problem without converting it to an equivalent μ-problem

In chapter 5, the Greatest Common Divisor (GCD) of two polynomials is considered. The aim was to identify the minimum-magnitude perturbations in the coefficients of the polynomials so that the perturbed polynomials have a common root. It was shown that in this approximation, GCD problem is equivalent to the calculation of a structured singular value of a matrix. This however involves a significant increase in the size of the problem and thus may not be the most cost efficient solution. A method is proposed that is directly applicable to Sylvester uncertainty structures before converting to a μ problem. A numerical example compares the proposed method with the μ bounds obtained by the largest singular value and the convex relaxation ("D-iteration") method.

6.4 Suggestions for Future Research

In this section some suggestions for future research work related to the the results of this thesis are suggested.

In the proposed randomised algorithm, the number of iterations is calculated based on some adjustable parameters that identify the accuracy and the probability of the method. Note that the number of iterations does not necessarily coincide with the number of the selected vertices. To clarify this point, imagine a case in which all n iteration return the same vertex. The probability of randomised algorithm can be altered depending on how tight we need the upperbound to be (i.e. the required accuracy). An idea is to calculate the probability based both on accuracy and the number of iterations i.e. define a number of vertices (for example 75% of the total) and find the probability that the algorithm can identify the maximiser by enumerating this number of vertices only. This approach can give better flexibility to define a trade-off between accuracy and computational cost. For example, for an application in which 50% validity is acceptable, the corresponding number of required iterations is calculated. The maximiser is then obtained by the algorithm after this number of iterations and so the result is acceptable with 50% probability. Figure 6.1 illustrate the idea.

Alternatively, an algorithm could be designed based on the number of vertices, instead of the number of iterations. This could also lead to additional flexibility by developing clearer trade-offs between accuracy, computational cost and probability.

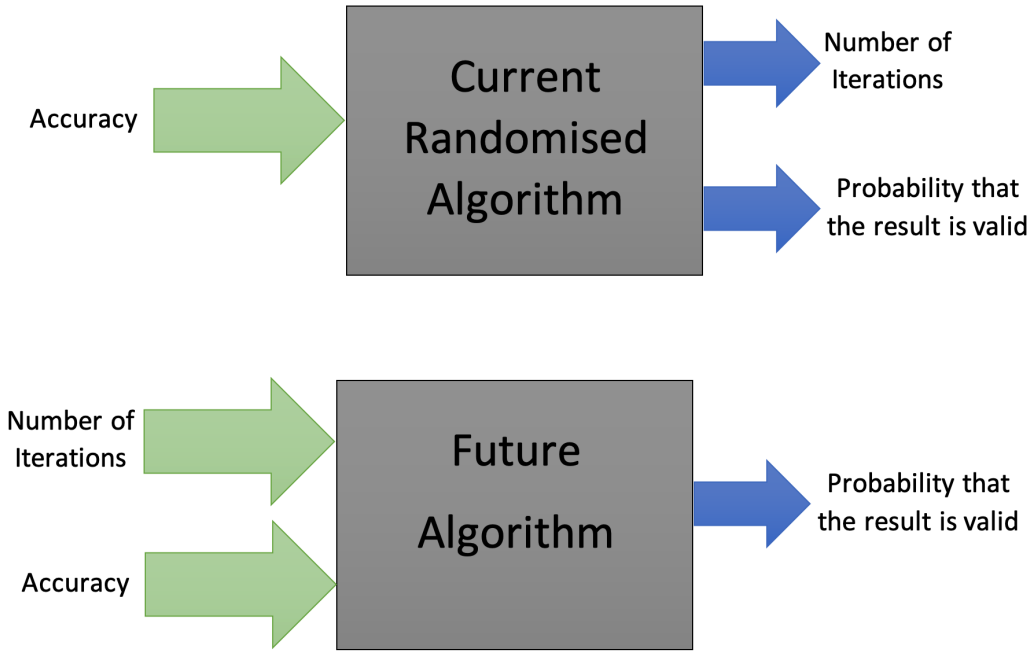

Figure 6.1: Calculating the probability based on the number of iteration and accuracy

It is also mentioned in chapter 2 that the probability that the maximiser occurs at a flat vertex is low in general. Depending on the direction of the eigenvectors this probability will vary. The feasibility and the cost of such calculation in each case remain to be studied in future work.

A limitation of the proposed method for μ-problem, was the assumption that the largest singular value of M is non repeated, i.e. $m = 1$. The case $m > 1$ involves the solution of a non-convex eigenvalue problem which is likely to prove computationally demanding. Extension of our method to address the general case will be the focus of future work.

The method of finding a convex upper bound on μ-problem was based on the idea of imposing artificial structure by bounding the magnitude of the projection of the perturbation in the most critical direction to be less than $\sigma_{\min}(M)$. This will then increase the norm of nearest destabilising perturbations for $I - \Delta M$ to lose rank. With the same justification, imposing an additional constraint on the projection of Δ in the second most critical direction could also increase the stability radius. This is an interesting part which is worth more consideration in future work.

The approximate GCD problem in Chapter 5 involves two polynomials. However, our method can be further extended to the cases where more than two polynomials are involved.

This will make the Sylvester matrix non-square and hence the standard Lehtomaki approach is not applicable. We have also considered real polynomial coefficients in studying the problem. This can further extended to a case where the coefficients are complex while the perturbations are real, i.e. the complex coefficients-real perturbation case. Is this case the proposed method of Chapter 4 is applicable. These parts remained to be studied in future work.

Appendix A

Appendix

A.1 Proof of Lemma 3.2.1

Lemma A.1.1. *For a class of uncertainty where* $\Delta_{11} = \mathbb{C}^{m\times m}$

$$\gamma_{\Delta\in\mathbb{C}^{m\times m}} = min\left\{\|\Delta\| : det\left(A-\Delta\right)=0\right\} = \sigma_{min}(A) = 1 \tag{A.1}$$

and the set of all optimal Δ *is given by*

$$\mathcal{D} = \left\{ \begin{bmatrix} W & 0 \\ 0 & I_{n-m} \end{bmatrix} \left[\begin{array}{cc|c} 1 & 0 & 0 \\ 0 & \Delta_{22} & \Delta_{23} \\ \hline 0 & \Delta_{32} & \Delta_{33} \end{array}\right] \begin{bmatrix} W^* & 0 \\ 0 & I_{n-m} \end{bmatrix} \in \mathbb{C}^{n\times n} \right\} \tag{A.2}$$

where

$$WW^* = W^*W = I_m,\ \textit{and}\ , \| \begin{bmatrix} \Delta_{22} & \Delta_{23} \\ \Delta_{32} & \Delta_{33} \end{bmatrix} \| \leq 1$$

Proof. It is obvious that

$$\Delta = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \in \mathbb{C}^{n\times n}$$

is the simplest form that can make $A-\Delta$ singular which prove that $\gamma_{\mathbb{C}^{m\times m}} = 1$. To prove (A.2), let $\|\Delta\| = 1$ and $A-\Delta$ is singular. Thus $(A-\Delta)\,x = 0$ where $\|x\| = 1$. Assume $x = \begin{bmatrix} x_1^* & x_2^* \end{bmatrix}$ where $x_1 \in \mathbb{C}^m$, hence,

$$(A-\Delta)x = 0 \Rightarrow \begin{cases} Ax = \Delta x \\ x^*A^* = x^*\Delta^* \end{cases} \Rightarrow x^*A^*Ax = x^*\Delta^*\Delta x \Rightarrow x^*A^2x = x^*\Delta^*\Delta x \Rightarrow \|Ax\|^2 = \|\Delta x\|^2$$

from $\begin{cases} 1 = a_1 < a_2 < ... < a_n \\ \|x\|^2 = \sum x_i^2 = 1 \end{cases} \Rightarrow \|Ax\|^2 = \sum_{i=1}^{n} a_i^2 x_i^2 \geq 1 \quad ①$

on the other hand:

$$\|Ax\| \leq \|A\| \cancel{\|x\|}^{1} \Rightarrow \|Ax\| \leq \|A\| = 1 \Rightarrow \|Ax\| \leq 1 \quad ②$$

$$① \text{ and } ② \Rightarrow \begin{cases} \|Ax\| \leq 1 \\ \|Ax\| \geq 1 \end{cases} \Rightarrow \|Ax\| = 1 \Rightarrow x^* A^2 x = 1 \Rightarrow x^* \Delta^* \Delta x = 1$$

$$x_2 = 0 \Rightarrow (A - \Delta) \begin{bmatrix} x_1 \\ 0 \end{bmatrix} = \begin{bmatrix} A_1 - \Delta_{11} & -\Delta_{12} \\ \Delta_{21} & A_2 - \Delta_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \Rightarrow$$

$$\begin{bmatrix} (A_1 - \Delta_{11}) x_1 \\ \Delta_{21} x_1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$(A_1 - \Delta_{11}) x_1 = 0 \Rightarrow A_1 x_1 = \Delta_{11} x_1$$

The equality in ① hold only if $x_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$ any other combination makes $\|Ax\|^2 > 1$.

Therefore

$$A_1 x_1 = \Delta_{11} x_1 = \zeta' \Delta \zeta x_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = x_1$$

$$\left.\begin{array}{l} \Delta_{11} x_1 = x_1 \Rightarrow \|\Delta_{11} x_1\| = \|x_1\| \\ \qquad\qquad\qquad\quad \|x_1\| = 1 \end{array}\right\} \Rightarrow \|\Delta_{11}\| \geq 1$$

$$\left.\begin{array}{l} \|\Delta_{11}\| \geq 1 \\ \|\Delta_{11}\| \leq \|\Delta\| = 1 \end{array}\right\} \Rightarrow \|\Delta_{11}\| = 1$$

All such Δ can be represented by (A.2).

□

A.2 Proof of Lemma 3.2.2

Lemma 3.2.2

Let $A = diag(A_1, A_2) \in \mathbb{R}^{n\times n}$ *with*

$$A_1 = \left[\begin{array}{c|c} I_{m1} & 0 \\ \hline 0 & A_{22} \end{array}\right] \in \mathbb{R}^{m\times m}$$

and

$$A_2 = \left[\begin{array}{c|c} a_3 I_{m3} & 0 \\ \hline 0 & A_{44} \end{array}\right] \in \mathbb{R}^{(n-m)\times(n-m)}$$

Assume that $1 < \underline{\sigma}(A_{22}), 0 < a_3 < \underline{\sigma}(A_{44})$, *and* $1 < a_3$. *Then the structured distance to singularity is*

$$\gamma_{0_{m\times m}} := \min\{\|\Delta\| : det\,(A-\Delta) = 0, \zeta^*\Delta\zeta = 0\} = \sqrt{a_3} =: \sqrt{\underline{\sigma}(A_1)\underline{\sigma}(A_2)} \quad \text{(A.3)}$$

Furthermore all optimal rank reducing perturbations are generated by

$$W \left[\begin{array}{ccc||ccc} 0 & 0 & 0 & \sqrt{a_3} & 0 & 0 \\ 0 & 0 & 0 & 0 & \Delta_{13} & \Delta_{14} \\ \hline 0 & 0 & 0 & 0 & \Delta_{23} & \Delta_{24} \\ \hline\hline \sqrt{a_3} & 0 & 0 & 0 & 0 & 0 \\ 0 & \Delta_{31} & \Delta_{32} & 0 & \Delta_{33} & \Delta_{34} \\ \hline 0 & \Delta_{41} & \Delta_{42} & 0 & \Delta_{43} & \Delta_{44} \end{array}\right] W^* = W\overline{\Delta}W^* \quad \text{(A.4)}$$

where $W = diag(W_1, I_{m_2}, W_3, I_{m_4}) \in \mathbb{C}^{n\times n}$ *is unitary and*

$$\left\| \begin{bmatrix} 0 & 0 & \Delta_{13} & \Delta_{14} \\ 0 & 0 & \Delta_{23} & \Delta_{24} \\ \Delta_{31} & \Delta_{32} & \Delta_{33} & \Delta_{34} \\ \Delta_{41} & \Delta_{42} & \Delta_{43} & \Delta_{44} \end{bmatrix} \right\| \leq \sqrt{a_3} \quad \text{(A.5)}$$

Proof. First we verify that all Δ of the form of (A.4) have norm $\sqrt{a_3}$:

$$\Delta\Delta^* = \begin{bmatrix} 0 & 0 & 0 & \sqrt{a_3} & 0 & 0 \\ 0 & 0 & 0 & 0 & \Delta_{13} & \Delta_{14} \\ 0 & 0 & 0 & 0 & \Delta_{23} & \Delta_{24} \\ \sqrt{a_3} & 0 & 0 & 0 & 0 & 0 \\ 0 & \Delta_{31} & \Delta_{32} & 0 & \Delta_{33} & \Delta_{34} \\ 0 & \Delta_{41} & \Delta_{42} & 0 & \Delta_{43} & \Delta_{44} \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 & \sqrt{a_3} & 0 & 0 \\ 0 & 0 & 0 & 0 & \Delta^*_{31} & \Delta^*_{41} \\ 0 & 0 & 0 & 0 & \Delta^*_{32} & \Delta^*_{42} \\ \sqrt{a_3} & 0 & 0 & 0 & 0 & 0 \\ 0 & \Delta^*_{13} & \Delta^*_{23} & 0 & \Delta^*_{33} & \Delta^*_{43} \\ 0 & \Delta^*_{14} & \Delta^*_{24} & 0 & \Delta^*_{34} & \Delta^*_{44} \end{bmatrix}$$

$$= \begin{bmatrix} a_3 & 0 & 0 & 0 & 0 & 0 \\ 0 & \Phi_1 & \Phi_2 & 0 & \Phi_3 & \Phi_4 \\ 0 & \Phi_5 & \Phi_6 & 0 & \Phi_7 & \Phi_8 \\ 0 & 0 & a_3 & 0 & 0 & 0 \\ 0 & \Phi_9 & \Phi_{10} & 0 & \Phi_{11} & \Phi_{12} \\ 0 & \Phi_{13} & \Phi_{14} & 0 & \Phi_{15} & \Phi_{16} \end{bmatrix}$$

where

$\Phi_1 = \Delta_{13}\Delta^*_{13} + \Delta_{14}\Delta^*_{14}$ $\quad \Phi_4 = \Delta_{13}\Delta^*_{43} + \Delta_{14}\Delta^*_{44}$ $\quad \Phi_7 = \Delta_{23}\Delta^*_{33} + \Delta_{24}\Delta^*_{34}$

$\Phi_2 = \Delta_{13}\Delta^*_{23} + \Delta_{14}\Delta^*_{24}$ $\quad \Phi_5 = \Delta_{23}\Delta^*_{13} + \Delta_{24}\Delta^*_{14}$ $\quad \Phi_8 = \Delta_{23}\Delta^*_{43} + \Delta_{24}\Delta^*_{44}$

$\Phi_3 = \Delta_{13}\Delta^*_{33} + \Delta_{14}\Delta^*_{34}$ $\quad \Phi_6 = \Delta_{23}\Delta^*_{23} + \Delta_{24}\Delta^*_{24}$ $\quad \Phi_9 = \Delta_{33}\Delta^*_{13} + \Delta_{34}\Delta^*_{14}$

$\Phi_{10} = \Delta_{33}\Delta^*_{23} + \Delta_{34}\Delta^*_{24}$ $\quad \Phi_{13} = \Delta_{43}\Delta^*_{13} + \Delta_{44}\Delta^*_{14}$ $\quad \Phi_{14} = \Delta_{43}\Delta^*_{23} + \Delta_{44}\Delta^*_{24}$

$\Phi_{11} = \Delta_{31}\Delta^*_{31} + \Delta_{32}\Delta^*_{32} + \Delta_{33}\Delta^*_{33} + \Delta_{34}\Delta^*_{34}$

$\Phi_{12} = \Delta_{31}\Delta^*_{41} + \Delta_{32}\Delta^*_{42} + \Delta_{33}\Delta^*_{43} + \Delta_{34}\Delta^*_{44}$

$\Phi_{15} = \Delta_{31}\Delta^*_{31} + \Delta_{42}\Delta^*_{32} + \Delta_{43}\Delta^*_{33} + \Delta_{44}\Delta^*_{34}$

$\Phi_{16} = \Delta_{41}\Delta^*_{41} + \Delta_{42}\Delta^*_{42} + \Delta_{43}\Delta^*_{43} + \Delta_{44}\Delta^*_{44}$

If we do the following transpose in rows/columns, the norm will not change

$$\text{Transpose} \begin{cases} Row2 \Leftrightarrow Row4 \\ Coloum2 \Leftrightarrow Coloum4 \\ Coloum3 \Leftrightarrow Coloum4 \\ Row3 \Leftrightarrow Row4 \end{cases}$$

$\Delta\Delta^*$ then will be

$$\Delta\Delta^* = \begin{bmatrix} a_3 & 0 & 0 & 0 & 0 & 0 \\ 0 & a_3 & 0 & 0 & 0 & 0 \\ 0 & 0 & \Phi_1 & \Phi_2 & \Phi_3 & \Phi_4 \\ 0 & 0 & \Phi_5 & \Phi_6 & \Phi_7 & \Phi_8 \\ 0 & 0 & \Phi_9 & \Phi_{10} & \Phi_{11} & \Phi_{12} \\ 0 & 0 & \Phi_{13} & \Phi_{14} & \Phi_{15} & \Phi_{16} \end{bmatrix}$$

Thus

$$\|\Delta\|_2^2 = \|\Delta\Delta^*\| = \max\left\{ \| \begin{bmatrix} a_3 & 0 \\ 0 & a_3 \end{bmatrix} \|, \| \begin{bmatrix} \Phi_1 & \Phi_2 & \Phi_3 & \Phi_4 \\ \Phi_5 & \Phi_6 & \Phi_7 & \Phi_8 \\ \Phi_9 & \Phi_{10} & \Phi_{11} & \Phi_{12} \\ \Phi_{13} & \Phi_{14} & \Phi_{15} & \Phi_{16} \end{bmatrix} \| \right\}$$

Note that

$$\begin{bmatrix} \Phi_1 & \Phi_2 & \Phi_3 & \Phi_4 \\ \Phi_5 & \Phi_6 & \Phi_7 & \Phi_8 \\ \Phi_9 & \Phi_{10} & \Phi_{11} & \Phi_{12} \\ \Phi_{13} & \Phi_{14} & \Phi_{15} & \Phi_{16} \end{bmatrix} = \begin{bmatrix} 0 & 0 & \Delta_{13} & \Delta_{14} \\ 0 & 0 & \Delta_{23} & \Delta_{24} \\ \Delta_{31} & \Delta_{32} & \Delta_{33} & \Delta_{34} \\ \Delta_{41} & \Delta_{42} & \Delta_{43} & \Delta_{44} \end{bmatrix} \begin{bmatrix} 0 & 0 & \Delta_{31}^* & \Delta_{41}^* \\ 0 & 0 & \Delta_{32}^* & \Delta_{42}^* \\ \Delta_{13}^* & \Delta_{23}^* & \Delta_{33}^* & \Delta_{43}^* \\ \Delta_{14}^* & \Delta_{24}^* & \Delta_{34}^* & \Delta_{44}^* \end{bmatrix} =$$
$$\begin{bmatrix} 0 & 0 & \Delta_{13} & \Delta_{14} \\ 0 & 0 & \Delta_{23} & \Delta_{24} \\ \Delta_{31} & \Delta_{32} & \Delta_{33} & \Delta_{34} \\ \Delta_{41} & \Delta_{42} & \Delta_{43} & \Delta_{44} \end{bmatrix} \begin{bmatrix} 0 & 0 & \Delta_{13} & \Delta_{14} \\ 0 & 0 & \Delta_{23} & \Delta_{24} \\ \Delta_{31} & \Delta_{32} & \Delta_{33} & \Delta_{34} \\ \Delta_{41} & \Delta_{42} & \Delta_{43} & \Delta_{44} \end{bmatrix}^* \text{ therefore:}$$

$$\|\Delta\|_2^2 = \|\Delta\Delta^*\| = \max\left\{ \| \begin{bmatrix} a_3 & 0 \\ 0 & a_3 \end{bmatrix} \|, \| \begin{bmatrix} 0 & 0 & \Delta_{13} & \Delta_{14} \\ 0 & 0 & \Delta_{23} & \Delta_{24} \\ \Delta_{31} & \Delta_{32} & \Delta_{33} & \Delta_{34} \\ \Delta_{41} & \Delta_{42} & \Delta_{43} & \Delta_{44} \end{bmatrix} \|^2 \right\}$$

We know from (A.5) that

$$\left\| \begin{bmatrix} 0 & 0 & \Delta_{13} & \Delta_{14} \\ 0 & 0 & \Delta_{23} & \Delta_{24} \\ \Delta_{31} & \Delta_{32} & \Delta_{33} & \Delta_{34} \\ \Delta_{41} & \Delta_{42} & \Delta_{43} & \Delta_{44} \end{bmatrix} \right\| \leq \sqrt{a_3}$$

therefore $\|\Delta\|_2^2 = \|\Delta\Delta^*\| = \| \begin{bmatrix} a_3 & 0 \\ 0 & a_3 \end{bmatrix} \| = a_3 \Rightarrow \|\Delta\| = \sqrt{a_3}$

Now we show all Δ of the form (A.4) satisfy $det(A - \Delta) = 0$.Note that $det(A - W\overline{\Delta}W^*) = 0 \Leftrightarrow det(W^*AW - \overline{\Delta}) = 0$ and

$$W^*AW = \left[\begin{array}{c|c||c|c} W_1^* I_{m_1} W_1 & 0 & 0 & 0 \\ \hline 0 & I_{m_2} A_{22} I_{m_2} & 0 & 0 \\ \hline\hline 0 & 0 & a_3 W_3^* W_3 & 0 \\ \hline 0 & 0 & 0 & I_{m_4} A_{44} I_{m_4} \end{array}\right] = \left[\begin{array}{c|c||c|c} I_{m_1} & 0 & 0 & 0 \\ \hline 0 & A_{22} & 0 & 0 \\ \hline\hline 0 & 0 & a_3 & 0 \\ \hline 0 & 0 & 0 & A_{44} \end{array}\right] = A$$

Therefore $det(A - \Delta) = det(W^*AW - \overline{\Delta}) = det(A - \overline{\Delta})$. On the other hand $det(A - \Delta) = 0$ means that there exists a non-zero vector x for which $(A - \Delta)x = \underline{0}$. The block diagonal structure of A and $\overline{\Delta}$ gives

$$A - \Delta = \left[\begin{array}{ccc||ccc} 1 & 0 & 0 & -\sqrt{a_3} & 0 & 0 \\ 0 & I_{m_1-1} & 0 & 0 & -\Delta_{13} & -\Delta_{14} \\ \hline 0 & 0 & A_{22} & 0 & -\Delta_{23} & -\Delta_{24} \\ \hline\hline -\sqrt{a_3} & 0 & 0 & a_3 & 0 & 0 \\ 0 & -\Delta_{31} & -\Delta_{32} & 0 & a_3 I_{m_3-1} - \Delta_{33} & -\Delta 34 \\ \hline 0 & -\Delta 41 & -\Delta 42 & 0 & -\Delta 43 & A_{44} - D_{44} \end{array}\right] \quad \text{(A.6)}$$

Define $x = \left[\begin{array}{cc|c||c|c} a_3 & 0 & 0 & -\sqrt{a_3} & 0 \end{array}\right.$ $\left.\begin{array}{c} 0 \end{array}\right]^T$ gives

$$(A-\Delta)x = \left[\begin{array}{ccc||ccc} 1 & 0 & 0 & -\sqrt{a_3} & 0 & 0 \\ 0 & I_{m_1-1} & 0 & 0 & -\Delta_{13} & -\Delta_{14} \\ \hline 0 & 0 & A_{22} & 0 & -\Delta_{23} & -\Delta_{24} \\ \hline\hline -\sqrt{a_3} & 0 & 0 & a_3 & 0 & 0 \\ 0 & -\Delta_{31} & -\Delta_{32} & 0 & a_3 I_{m_3-1} - \Delta_{33} & -\Delta 34 \\ \hline 0 & -\Delta 41 & -\Delta 42 & 0 & -\Delta 43 & A_{44} - D_{44} \end{array}\right] \left[\begin{array}{c} a_3 \\ \hline 0 \\ \hline 0 \\ \hline\hline -\sqrt{a_3} \\ \hline 0 \\ \hline 0 \end{array}\right] = \left[\begin{array}{c} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{array}\right]$$

Hence $det(A-\Delta)=0$. Since $\|\Delta\|=\sqrt{a_3}$ this is an upper bound.

Now let $x=\begin{bmatrix} x_1' \\ x_2' \end{bmatrix} \in \mathbb{C}^n$ which $x_1 \in \mathbb{C}^m$. It is clear that $x_2 \neq 0$ because

If $x_2=0 \Rightarrow \begin{bmatrix} A_1 & * \\ * & * \end{bmatrix}\begin{bmatrix} x_1' \\ 0 \end{bmatrix}=\underline{0} \Rightarrow A_1x_1'=\underline{0}$

However A_1 is full rank and thus $x_1'=0$ is in contradiction of $x\neq 0$. On the other hand x_1' is also non-zero vector because if $x_1'=0$ then

$(A-\Delta)x=\underline{0}$, let $\Delta=\begin{bmatrix} 0 & D_{12} \\ D_{21} & D_{22} \end{bmatrix}$, thus

$(A-\Delta)x=\begin{bmatrix} A_1 & -D_{12} \\ -D_{21} & A_2-D_{22} \end{bmatrix}\begin{bmatrix} 0 \\ x_2' \end{bmatrix} \Rightarrow (A_2-D_{22})x_2'=\underline{0} \Rightarrow det(A_2-D_{22})=0$

To fulfil this condition $min\{\|D_{22}\|\} \leq \underline{\sigma}(A_{22}) \leq a_3$ that gives $\|\Delta\| \geq \|D_{22}\| \geq a_3 > \sqrt{a_3}$ which is contradiction of $\|\Delta\| \leq \sqrt{a_3}$. Hence, $x_1 \neq 0$ and $x_2 \neq 0$.

Now let $Z=diag(z_1,z_2)$ where $z_1=x_1/\|x_1\|$ and $z_2=x_2/\|x_2\|$. Then

$$(A-\Delta)x=\underline{0} \Rightarrow (A-\Delta)Z\begin{bmatrix} \|x_1\| \\ \|x_2\| \end{bmatrix}=\underline{0} \Rightarrow det\left(\begin{bmatrix} z_1^*A_1z_1 & 0 \\ 0 & z_2^*A_2z_2 \end{bmatrix}-Z^*\Delta Z\right)$$

Since $z_1 \in \mathbb{C}^{m\times m}$ and $z_2 \in \mathbb{C}^{(n-m)\times(n-m)}$, then $\begin{bmatrix} z_1^*A_1z_1 & 0 \\ 0 & z_2^*A_2z_2 \end{bmatrix} \in \mathbb{C}^{2\times 2}$.This can be consider as a case which $det(A-\Delta)=0$, s.t. $\Delta_{11}=\delta \in \mathbb{C} : |\delta| \leq \phi$ where $\phi=0$, thus base on sec 3.2.

$$\{\|Z^*\Delta Z\|: det\left(\begin{bmatrix} z_1^*A_1z_1 & 0 \\ 0 & z_2^*A_2z_2 \end{bmatrix}-Z^*\Delta Z\right)=0\} \geq \sqrt{|z_1^*A_1z_1||z_2^*A_2z_2|} \geq \sqrt{a_3} \tag{A.7}$$

This is due to the fact that z_1 and z_2 are orthogonal and $\|z_1^*A_1z_1\|=\|A_1\|=1$. Similarly $\|z_2^*A_2z_2\|=\|A_2\|=a_3$. On the other hand $\|Z^*\|\|\Delta\|\|Z\| \geq \|Z^*\Delta Z\|$ and $\|Z^*\|=\|Z\|=\|z_1\|=\|z_2\|=0$ which gives $\|\Delta\| \geq \|Z^*\Delta Z\| \geq \sqrt{a_3}$. Since the $(1,1)$ entry of $Z^*\Delta Z$ is zero(from the definition), $Z^*\Delta Z$ will be in the form of

$$\mathcal{D}_{\Delta_\phi} = \left\{ \begin{bmatrix} 0 & \sqrt{a_3}e^{j\theta} \\ \sqrt{a_3}e^{-j\theta} & 0 \end{bmatrix} : \theta \in \mathbb{R} \right\} \tag{A.8}$$

Note that since $z_1^* A_1 z_1$ and $z_2^* A_2 z_2$ are scaler, $\Delta \in \mathbb{C}^{2\times 2}$ and thus Δ_{33} is not available. In addition, if $z_1 = \begin{bmatrix} z_{11}^* & z_{22}^* \end{bmatrix}^*$ and $z_2 = \begin{bmatrix} z_{33}^* & z_{44}^* \end{bmatrix}^*$, we have

$$z_1^* A_1 z_1 = \begin{bmatrix} z_{11}^* & z_{22}^* \end{bmatrix} \begin{bmatrix} I_{m_1} & 0 \\ 0 & A_{22} \end{bmatrix} \begin{bmatrix} z_{11} \\ z_{22} \end{bmatrix} = z_{11}^* z_{11} + z_{22}^* A_{22} z_{22}$$

$$z_2^* A_2 z_2 = \begin{bmatrix} z_{33}^* & z_{44}^* \end{bmatrix} \begin{bmatrix} a_3 I_{m_3} & 0 \\ 0 & A_{44} \end{bmatrix} \begin{bmatrix} z_{33} \\ z_{44} \end{bmatrix} = a_3 z_{33}^* z_{44} + z_{44}^* A_{44} z_{44}$$

Because $z_{22}^* A_{22} z_{22}$ and $z_{44}^* A_{44} z_{44}$ are scaler we replace them by positive real number

$$z_1 A_1 z_1^* = z_{11}^* z_{11} + \alpha$$

$$z_2^* A_2 z_2 = a_3 z_{33}^* z_{44} + \beta$$

where $\alpha, \beta \in \mathbb{R} \geq 0$. It is obvious that the minimum happens when both α and β are equal to zero which requires $z_{22} = \underline{0}$ and $z_{44} = \underline{0}$. In other word the equality in (A.7) happens if and only if $z_1 = \begin{bmatrix} z_{11}^* & 0 \end{bmatrix}^*$ and $z_2 = \begin{bmatrix} z_{33}^* & 0 \end{bmatrix}^*$.
Therefore $\sqrt{a_3}$ is the lower bound and so far (A.3) has been proved. Now we show that all the structured Δ which fulfil (A.3) are in form of (A.4).Let any minimiser in (A.3) be as

$$\Delta = \begin{bmatrix} 0_{m\times m} & \Delta_{12} \\ \Delta_{21} & \Delta_{22} \end{bmatrix} \in \mathbb{C}^{n\times n}$$

And let each partition be defined as follow

$$\Delta_{12} = \begin{bmatrix} \hat{\Delta}_{11} & \hat{\Delta}_{12} \\ \hat{\Delta}_{21} & \hat{\Delta}_{22} \end{bmatrix}, \quad \Delta_{21} = \begin{bmatrix} \bar{\Delta}_{11} & \bar{\Delta}_{12} \\ \bar{\Delta}_{21} & \bar{\Delta}_{22} \end{bmatrix}, \quad \Delta_{22} = \begin{bmatrix} \tilde{\Delta}_{11} & \tilde{\Delta}_{12} \\ \tilde{\Delta}_{21} & \tilde{\Delta}_{22} \end{bmatrix}$$

Where $\hat{\Delta}_{11} \in \mathbb{C}^{m_1\times m_3}$, $\bar{\Delta}_{11} \in \mathbb{C}^{m_3\times m_1}$ and $\tilde{\Delta}_{11} \in \mathbb{C}^{m_3\times m_3}$

From previous argument

$$det\left(\left[\begin{array}{cc} z_1^* A_1 z_1 & 0 \\ 0 & z_2^* A_2 z_2 \end{array}\right] - \left[\begin{array}{cc} 0 & z_1^* \Delta_{12} z_2 \\ z_2^* \Delta_{21} z_1 & z_2^* \Delta_{22} z_2 \end{array}\right]\right) = 0 \tag{A.9}$$

Note that

$$\left[\begin{array}{cc} z_{11}^* & 0 \end{array}\right] \left[\begin{array}{cc} I_{m_1} & 0 \\ 0 & A_{22} \end{array}\right] \left[\begin{array}{c} z_{11} \\ 0 \end{array}\right] = \|z_{11}\|^2 = 1$$

$$\left[\begin{array}{cc} z_{33}^* & 0 \end{array}\right] \left[\begin{array}{cc} a_3 I_{m_3} & 0 \\ 0 & A_{44} \end{array}\right] \left[\begin{array}{c} z_{33} \\ 0 \end{array}\right] = a_3 \|z_{11}\|^2 = a_3$$

Thus

$$det\left(\left[\begin{array}{cc} 1 & 0 \\ 0 & a_3 \end{array}\right] - \left[\begin{array}{cc} 0 & z_1^* \Delta_{12} z_2 \\ z_2^* \Delta_{21} z_1 & z_2^* \Delta_{22} z_2 \end{array}\right]\right) = 0 \tag{A.10}$$

Parametrize Δ based on (A.8)

$$\left[\begin{array}{cc} 0 & z_1^* \Delta_{12} z_2 \\ z_2^* \Delta_{21} z_1 & z_2^* \Delta_{22} z_2 \end{array}\right] = \left[\begin{array}{cc} 0 & \sqrt{a_3} e^{j\theta} \\ \sqrt{a_3} e^{-j\theta} & 0 \end{array}\right]$$

Therefore

$$z_1^* \Delta_{12} z_2 = \left[\begin{array}{cc} z_{11}^* & 0 \end{array}\right] \left[\begin{array}{cc} \hat{\Delta}_{11} & \hat{\Delta}_{12} \\ \hat{\Delta}_{21} & \hat{\Delta}_{22} \end{array}\right] \left[\begin{array}{c} z_{33} \\ 0 \end{array}\right] = z_{11}^* \hat{\Delta}_{11} z_{33} = e^{j\theta} \sqrt{a_3}$$

$$z_2^* \Delta_{21} z_1 = \left[\begin{array}{cc} z_{33}^* & 0 \end{array}\right] \left[\begin{array}{cc} \bar{\Delta}_{11} & \bar{\Delta}_{12} \\ \bar{\Delta}_{21} & \bar{\Delta}_{22} \end{array}\right] \left[\begin{array}{c} z_{11} \\ 0 \end{array}\right] = z_{33}^* \bar{\Delta}_{11} z_{11} = e^{-j\theta} \sqrt{a_3}$$

$$z_2^* \Delta_{22} z_2 = \left[\begin{array}{cc} z_{33}^* & 0 \end{array}\right] \left[\begin{array}{cc} \tilde{\Delta}_{11} & \tilde{\Delta}_{12} \\ \tilde{\Delta}_{21} & \tilde{\Delta}_{22} \end{array}\right] \left[\begin{array}{c} z_{33} \\ 0 \end{array}\right] = z_{33}^* \tilde{\Delta}_{11} z_{33} = 0$$

for some $\theta \in (-\pi, \phi]$. Next define any two unitary completions $U_\perp$ and $V_\perp$ of $e^{j\theta} z_{11}$ and $z_3 3$, respectively, to construct unitary matrices $W_1 = \left[\begin{array}{cc} e^{j\theta} z_{11} & U_\perp \end{array}\right]$ and $W_3 = \left[\begin{array}{cc} z_{33} & V_\perp \end{array}\right]$ and consider the product

$$\begin{bmatrix} W_1^* & 0 \\ 0 & I_{m_2} \end{bmatrix} \Delta_{12} \begin{bmatrix} W_3 & 0 \\ 0 & I_{m_4} \end{bmatrix} = \begin{bmatrix} e^{-j\theta} z_{11}^* & 0 \\ U_\perp^* & 0 \\ 0 & I_{m_2} \end{bmatrix} \begin{bmatrix} \hat{\Delta}_{11} & \hat{\Delta}_{12} \\ \hat{\Delta}_{21} & \hat{\Delta}_{22} \end{bmatrix} \begin{bmatrix} z_{33} & V_\perp & 0 \\ 0 & 0 & I_{m_4} \end{bmatrix}$$

$$= \begin{bmatrix} e^{-j\theta} z_{11}^* \hat{\Delta}_{11} z_{33} & e^{-j\theta} z_{11}^* \hat{\Delta}_{11} V_\perp & e^{-j\theta} z_{11}^* \hat{\Delta}_{12} \\ U_\perp^* \hat{\Delta}_{11} z_{33} & U_\perp^* \hat{\Delta}_{11} V_\perp & U_\perp^* \hat{\Delta}_{12} \\ \hat{\Delta}_{21} z_{33} & \hat{\Delta}_{12} V_\perp & \hat{\Delta}_{22} \end{bmatrix}$$

$$= \begin{bmatrix} \cancelto{\sqrt{a_3}}{e^{-j\theta} z_{11}^* \hat{\Delta}_{11} z_{33}} & \cancelto{0}{e^{-j\theta} z_{11}^* \hat{\Delta}_{11} V_\perp} & \cancelto{0}{e^{-j\theta} z_{11}^* \hat{\Delta}_{12}} \\ \cancelto{0}{U_\perp^* \hat{\Delta}_{11} z_{33}} & U_\perp^* \hat{\Delta}_{11} V_\perp & U_\perp^* \hat{\Delta}_{12} \\ \cancelto{0}{\hat{\Delta}_{21} z_{33}} & \hat{\Delta}_{12} V_\perp & \hat{\Delta}_{22} \end{bmatrix}$$

Defining

$$= \begin{bmatrix} U_\perp^* \hat{\Delta}_{11} V_\perp & U_\perp^* \hat{\Delta}_{12} \\ \hat{\Delta}_{12} V_\perp & \hat{\Delta}_{22} \end{bmatrix} = \begin{bmatrix} E_{13} & E_{14} \\ E_{23} & E_{24} \end{bmatrix}$$

gives

$$\begin{bmatrix} \hat{\Delta}_{11} & \hat{\Delta}_{12} \\ \hat{\Delta}_{21} & \hat{\Delta}_{22} \end{bmatrix} = \begin{bmatrix} W_1 & 0 \\ 0 & I_{m_2} \end{bmatrix} \begin{bmatrix} \sqrt{a_3} & 0 & 0 \\ 0 & E_{13} & E_{14} \\ 0 & E_{23} & E_{24} \end{bmatrix} \begin{bmatrix} W_3^* & 0 \\ 0 & I_{m_4} \end{bmatrix}$$

similar argument for Δ_{21} and Δ_{22} gives

$$\begin{bmatrix} \bar{\Delta}_{11} & \bar{\Delta}_{12} \\ \bar{\Delta}_{21} & \bar{\Delta}_{22} \end{bmatrix} = \begin{bmatrix} W_3 & 0 \\ 0 & I_{m_4} \end{bmatrix} \begin{bmatrix} \sqrt{a_3} & 0 & 0 \\ 0 & E_{31} & E_{41} \\ 0 & E_{41} & E_{42} \end{bmatrix} \begin{bmatrix} W_1^* & 0 \\ 0 & I_{m_2} \end{bmatrix}$$

$$\begin{bmatrix} \tilde{\Delta}_{11} & \tilde{\Delta}_{12} \\ \tilde{\Delta}_{21} & \tilde{\Delta}_{22} \end{bmatrix} = \begin{bmatrix} W_3 & 0 \\ 0 & I_{m_4} \end{bmatrix} \begin{bmatrix} 0 & \tilde{E}_{23} & \tilde{E}_{24} \\ \tilde{E}_{32} & E_{33} & E_{34} \\ \tilde{E}_{42} & E_{43} & E_{44} \end{bmatrix} \begin{bmatrix} W_3^* & 0 \\ 0 & I_{m_4} \end{bmatrix}$$

Combine all together gives

$$\Delta = \left[\begin{array}{c|c} 0_{m\times m} & \Delta_{12} \\ \hline \Delta_{21} & \Delta_{22} \end{array}\right] = W^* \left[\begin{array}{cc|cc} 0 & 0 & \hat{\Delta}_{11} & \hat{\Delta}_{12} \\ 0 & 0 & \hat{\Delta}_{21} & \hat{\Delta}_{22} \\ \hline \bar{\Delta}_{11} & \bar{\Delta}_{12} & \tilde{\Delta}_{11} & \tilde{\Delta}_{12} \\ \bar{\Delta}_{21} & \bar{\Delta}_{22} & \tilde{\Delta}_{21} & \tilde{\Delta}_{22} \end{array}\right] W$$

$$= \left[\begin{array}{cc||cc} W_1^* & 0 & 0 & \\ 0 & I_{m_2} & 0 & 0 \\ \hline\hline 0 & 0 & W_3^* & 0 \\ 0 & 0 & 0 & I_{m_4} \end{array}\right] \left[\begin{array}{cc|c||ccc} 0 & 0 & 0 & \sqrt{a_3} & 0 & 0 \\ 0 & 0 & 0 & 0 & E_{13} & E_{14} \\ \hline 0 & 0 & 0 & 0 & E_{23} & E_{24} \\ \hline\hline \sqrt{a_3} & 0 & 0 & 0 & \tilde{E}_{23} & \tilde{E}_{24} \\ 0 & E_{31} & E_{32} & \tilde{E}_{32} & E_{33} & E_{34} \\ \hline 0 & E_{41} & E_{42} & \tilde{E}_{42} & E_{43} & E_{44} \end{array}\right] \left[\begin{array}{cc||cc} W_1 & 0 & 0 & \\ 0 & I_{m_2} & 0 & 0 \\ \hline\hline 0 & 0 & W_3 & 0 \\ 0 & 0 & 0 & I_{m_4} \end{array}\right]$$

since$\|\Delta\| = \sqrt{a_3}$ any non-zero entry in the row/column which contain $\sqrt{a_3}$ will result in a norm greater than $\sqrt{a_3}$. Therefore all $\tilde{E}_{ij} = 0$ and thus all optimal rank reducing perturbation are generated by

$$\left[\begin{array}{cc||cc} W_1^* & 0 & 0 & \\ 0 & I_{m_2} & 0 & 0 \\ \hline\hline 0 & 0 & W_3^* & 0 \\ 0 & 0 & 0 & I_{m_4} \end{array}\right] \left[\begin{array}{cc|c||ccc} 0 & 0 & 0 & \sqrt{a_3} & 0 & 0 \\ 0 & 0 & 0 & 0 & E_{13} & E_{14} \\ \hline 0 & 0 & 0 & 0 & E_{23} & E_{24} \\ \hline\hline \sqrt{a_3} & 0 & 0 & 0 & 0 & 0 \\ 0 & E_{31} & E_{32} & 0 & E_{33} & E_{34} \\ \hline 0 & E_{41} & E_{42} & 0 & E_{43} & E_{44} \end{array}\right] \left[\begin{array}{cc||cc} W_1 & 0 & 0 & \\ 0 & I_{m_2} & 0 & 0 \\ \hline\hline 0 & 0 & W_3 & 0 \\ 0 & 0 & 0 & I_{m_4} \end{array}\right]$$

which is identical to (A.4). □

A.3 Proof of Lemma 3.2.3

Lemma 3.2.3 *For matrices A and ζ, as defined previously, let $\Delta_{11} \in \mathbb{C}^{m\times m}$ has the same structure as section 3.2.6 where $\|\Delta_{11}\| \leq 1$ and $det\,(I - \Delta_{11}) \neq 0$. Then*

$$\min_{\substack{det(A-\Delta)=0 \\ \zeta^*\Delta\zeta=\Delta_{11}}} \|\Delta\| = \min_{\substack{\|(\gamma^2 I-\Delta_{11})(I-\Delta_{11})^{-1}\|=a_{m+1} \\ \gamma>1}} \gamma \tag{A.11}$$

Proof. First consider the following lemma which can be deduced from [83]:

Lemma A.3.1. *[9] Let the following are true:*

- $H = \begin{bmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{bmatrix}$
- U *is complex matrix*
- $(I - H_{11}U)^{-1}$, H^-1_{12} *and* H^-1_{12} *exist*

Thus

1. *If* $det(H) \neq 0$, *then* $\Phi = \mathcal{F}_u(H, U) \Leftrightarrow U = \mathcal{F}_u\left(H^{-1}, \Phi\right)$
2. *If* $det(H) \neq 0$ *and* $det(U) \neq 0$, *then* $[\mathcal{F}_u(H, U)]^{-1} = \mathcal{F}_u\left(H^{-1}, U^{-1}\right)$
3. *If* H *is* $\gamma-$*unitary, i.e* $HH^* = H^*H = \gamma^2 I$ *for some* $\gamma > 0$, *then* $\|\mathcal{F}_u(H, U)\| = \gamma \Leftrightarrow \|U\| = \gamma^{-1}$
4. *If* H *is* $\gamma-$*unitary,then* $\|\mathcal{F}_u(H, U)\| < \gamma \Leftrightarrow \|U\| < \gamma^{-1}$
5. *If* H *is* $\gamma-$*unitary,* $det(I - H_{11}) \neq 0$, H_{11} *is an square matrix and* $\|H_{11}\| \leq 1$, *then* $\|\mathcal{F}_u(H, I)\| = \|\left(\gamma^2 I - H_{11}\right)(I - H_{11})^{-1}\|$

Proof. For the complete proof refer to [83]. □

Since $\|\Delta_{11}\| \leq 1$, for any $\gamma > 1$ there exist $\gamma-$unitary completion of Δ_{11} of the form

$$\Delta_1^\gamma = \begin{bmatrix} \Delta_{11} & \Delta_{13}^\gamma \\ \Delta_{31}^\gamma & \Delta_{33} \end{bmatrix} \in \mathbb{C}^{2m \times 2m} \quad \Delta_0^\gamma = \begin{bmatrix} \Delta_{11} & 0 & \Delta_{13}^\gamma & 0 \\ 0 & 0 & 0 & \gamma I_{n-m} \\ \Delta_{31}^\gamma & 0 & \Delta_{33} & 0 \\ 0 & \gamma I_{n-m} & 0 & 0 \end{bmatrix} \in \mathbb{C}^{2n \times 2n}$$

Taking $\Delta = \begin{bmatrix} \Delta_{11} & 0 \\ 0 & a_{m+1} I_{n-m} \end{bmatrix}$, gives

$$det(A - \Delta) = det\left(\left[\begin{array}{c|c} A_1 - \Delta_{11} & 0 \\ \hline 0 & A_2 - a_{m+1} I_{n-m} \end{array}\right]\right) = 0$$

Since $\|\Delta_{11}\| \leq 1 \Rightarrow \|\Delta\| = max\left\{\|\Delta_{11}\|, \|a_{m+1}I_{n-m}\|\right\} = a_{m+1}$ and hence a_{m+1} is an upper bound on equation (3.22) we can assume $\gamma \leq a_{m+1}$.
Define $X_A^\gamma = \mathcal{F}_u\left[(\Delta_0^\gamma)^{-1}, A\right]$, thus

$$X_A^\gamma = \mathcal{F}_u\left[(\Delta_0^\gamma)^{-1}, A\right] = \mathcal{F}_u\left[(\Delta_0^\gamma)^*, A\right] = \mathcal{F}_u\left(\begin{bmatrix} \Delta_{11}^* & 0 & (\Delta_{31}^\gamma)^* & 0 \\ 0 & 0 & 0 & \gamma I_{n-m} \\ (\Delta_{13}^\gamma)^* & 0 & (\Delta_{33})^* & 0 \\ 0 & \gamma I_{n-m} & 0 & 0 \end{bmatrix}, A\right) \Rightarrow$$

$$X_A^\gamma = \begin{bmatrix} \Delta_{33}^* & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} (\Delta_{13}^\gamma)^* & 0 \\ 0 & \gamma I_{n-m} \end{bmatrix} \begin{bmatrix} A_1 & 0 \\ 0 & A_2 \end{bmatrix} \times$$

$$\left\{\begin{bmatrix} I_m & 0 \\ 0 & I_{n-m} \end{bmatrix} - \begin{bmatrix} \Delta_{11}^* & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} A_1 & 0 \\ 0 & A_2 \end{bmatrix}\right\}^{-1} \times \begin{bmatrix} (\Delta_{31}^\gamma)^* & 0 \\ 0 & \gamma I_{n-m} \end{bmatrix}$$

$$= \begin{bmatrix} \Delta_{33}^* & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} (\Delta_{13}^\gamma)^* A_1 & 0 \\ 0 & \gamma A_2 \end{bmatrix} \begin{bmatrix} I_m - \Delta_{11}^* A_1 & 0 \\ 0 & I_{n-m} \end{bmatrix}^{-1} \begin{bmatrix} (\Delta_{31}^\gamma)^* & 0 \\ 0 & \gamma I_{n-m} \end{bmatrix}$$

$$= \begin{bmatrix} \Delta_{33}^* & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} (\Delta_{13}^\gamma)^* A_1 & 0 \\ 0 & \gamma A_2 \end{bmatrix} \begin{bmatrix} (I - \Delta_{11}^* A_1)^{-1}\Delta_{31}^* & 0 \\ 0 & \gamma I_{n-m} \end{bmatrix}$$

$$= \begin{bmatrix} \Delta_{33}^* + (\Delta_{13}^\gamma)^* A_1 (I - \Delta_{11}^* A_1)^{-1} (\Delta_{31}^\gamma)^* & 0 \\ 0 & \gamma^2 A_2 \end{bmatrix} = \begin{bmatrix} \mathcal{F}_u\left[(\Delta_1^\gamma)^{-1}, A_1\right] & 0 \\ 0 & \gamma^2 A_2 \end{bmatrix} =$$

$$diag\left(\mathcal{F}_u\left[(\Delta_1^\gamma)^{-1}, A_1\right], \gamma^2 A_2\right) \Rightarrow X_A^\gamma =: diag\left(X_1^\gamma, X_2^\gamma\right)$$

Note that since $A_1 = 1/\sigma_1 . I_m$ and $\sigma_1 = 1$ (normalized), we replace A_1 by I, hence, from the above definition

$$X_1^\gamma = \mathcal{F}_u\left[(\Delta_1^\gamma)^{-1}, I\right]$$
$$X_2^\gamma = \gamma^2 A_2$$

referring to part 2 of the lemma A.3.1

$$X_1^\gamma = \mathcal{F}_u\left[(\Delta_1^\gamma)^{-1}, I\right] = \left[\mathcal{F}_u\left(\Delta_1^\gamma, I\right)\right]^{-1} \Rightarrow (X_1^\gamma)^{-1} = \mathcal{F}_u\left[(\Delta_1^\gamma), I\right]$$

From part 3 of lemma A.3.1, also, since Δ_1^γ is $\gamma-$unitary and $\|I\| = 1 < \gamma$, thus

$$\|\,(X_1^\gamma)^{-1}\,\| = \|\mathcal{F}_u\left[(\Delta_1^\gamma), I\right]\| > \gamma \Rightarrow \frac{1}{\|\,(X_1^\gamma)\,\|} > \gamma \Rightarrow$$
$$\frac{1}{\underline{\sigma}(X_1^\gamma)} > \frac{1}{\overline{\sigma}(X_1^\gamma)} > \gamma \Rightarrow \underline{\sigma}\,(X_1^\gamma) < \gamma^{-1}$$

On the other hand

$$\underline{\sigma}\left(X_2^\gamma\right) = \underline{\sigma}\left(\gamma^{-2}A_2\right) = \gamma^{-2}\underline{\sigma}\left(A_2\right) = \gamma^{-2}a_{m+1} \text{ , since } \gamma \leq a_{m+1} \Rightarrow \begin{cases} \underline{\sigma}\left(X_2^\gamma\right) \geq \dfrac{1}{a_{m+1}} \\ \underline{\sigma}\left(X_1^\gamma\right) < \dfrac{1}{a_{m+1}} \end{cases}$$

which gives

$$\underline{\sigma}\left(X_2^\gamma\right) > \underline{\sigma}\left(X_1^\gamma\right)$$

Now define

$$\Phi_\Delta^\gamma := \mathcal{F}_u\left[\left(\Delta_0^\gamma\right)^{-1}, \Delta\right]$$

for any $\Delta \in \mathbb{C}^{n\times n}$.

From part 4 of the lemma A.3.1

$$\|\mathcal{F}_u\left[\left(\Delta_0^\gamma\right)^{-1}, \Delta\right]\| = \|\Phi_\Delta^\gamma\| = \gamma^{-1} \text{ if and only if } \|\Delta\| = \gamma \tag{A.12}$$

On the other hand, from part 1 of the lemma A.3.1

$$\Phi_\Delta^\gamma := \mathcal{F}_u\left[\left(\Delta_0^\gamma\right)^{-1}, \Delta\right] \Leftrightarrow \Delta := \mathcal{F}_l\left[\Delta_0^\gamma, \Phi_\Delta^\gamma\right] \tag{A.13}$$

Hence, Φ_Δ^γ and Δ_{11} need to be of the same size which gives $\Phi_\Delta^\gamma \in \mathbb{C}^{m\times m}$.
Assume $\Phi_\Delta^\gamma = \begin{bmatrix} \mathcal{A} & \mathcal{B} \\ \mathcal{C} & \mathcal{D} \end{bmatrix}$, we have

$$\Delta = \mathcal{F}_l\left[\Delta_0^\gamma, \Phi_\Delta^\gamma\right] \Rightarrow$$

$$\begin{bmatrix} \Delta_{11} & \Delta_{12} \\ \Delta_{21} & \Delta_{22} \end{bmatrix} = \mathcal{F}_l\left(\left[\begin{array}{cc|cc} \Delta_{11}^* & 0 & (\Delta_{31}^\gamma)^* & 0 \\ 0 & 0 & 0 & \gamma I_{n-m} \\ \hline (\Delta_{13}^\gamma)^* & 0 & (\Delta_{33})^* & 0 \\ 0 & \gamma I_{n-m} & 0 & 0 \end{array}\right], \begin{bmatrix} \mathcal{A} & \mathcal{B} \\ \mathcal{C} & \mathcal{D} \end{bmatrix}\right) \Rightarrow$$

$$
\begin{bmatrix} \Delta_{11} & \Delta_{12} \\ \Delta_{21} & \Delta_{22} \end{bmatrix} =
$$

$$
\begin{bmatrix} \Delta_{11} & 0 \\ 0 & 0 \end{bmatrix} +
$$

$$
\begin{bmatrix} \Delta_{13}^{\gamma} & 0 \\ 0 & \gamma I_{n-m} \end{bmatrix} \begin{bmatrix} \mathcal{A} & \mathcal{B} \\ \mathcal{C} & \mathcal{D} \end{bmatrix} \left(I - \begin{bmatrix} \Delta_{33} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \mathcal{A} & \mathcal{B} \\ \mathcal{C} & \mathcal{D} \end{bmatrix} \right)^{-1} \begin{bmatrix} \Delta_{31}^{\gamma} & 0 \\ 0 & \gamma I_{n-m} \end{bmatrix} =
$$

$$
\begin{bmatrix} \Delta_{11} & 0 \\ 0 & 0 \end{bmatrix} +
$$

$$
\begin{bmatrix} \Delta_{13}^{\gamma} & 0 \\ 0 & \gamma I_{n-m} \end{bmatrix} \begin{bmatrix} \mathcal{A} & \mathcal{B} \\ \mathcal{C} & \mathcal{D} \end{bmatrix} \left(\begin{bmatrix} I_m - \Delta_{33}\mathcal{A} & -\Delta_{33}\mathcal{B} \\ 0 & I_{n-m} \end{bmatrix} \right)^{-1} \begin{bmatrix} \Delta_{31}^{\gamma} & 0 \\ 0 & \gamma I_{n-m} \end{bmatrix} =
$$

$$
\begin{bmatrix} \Delta_{11} & 0 \\ 0 & 0 \end{bmatrix} +
$$

$$
\begin{bmatrix} \Delta_{13}^{\gamma}\mathcal{A} & \Delta_{13}^{\gamma}\mathcal{B} \\ \gamma\mathcal{C} & \gamma\mathcal{D} \end{bmatrix} \begin{bmatrix} (I_m - \Delta_{33}\mathcal{A})^{-1} & (I_m - \Delta_{33}\mathcal{A})^{-1}\Delta_{33}\mathcal{B} \\ 0 & I_{n-m} \end{bmatrix} \begin{bmatrix} \Delta_{31}^{\gamma} & 0 \\ 0 & \gamma I_{n-m} \end{bmatrix} =
$$

$$
\begin{bmatrix} \Delta_{11} & 0 \\ 0 & 0 \end{bmatrix} +
$$

$$
\begin{bmatrix} \Delta_{13}^{\gamma}\mathcal{A}(I_m - \Delta_{33}\mathcal{A})^{-1}\Delta_{31}^{\gamma} & \Delta_{13}^{\gamma}\mathcal{A}(I_m - \Delta_{33}\mathcal{A})^{-1}\Delta_{33}\mathcal{B} + \Delta_{13}^{\gamma}\mathcal{B} \\ \gamma\mathcal{C}(I_m - \Delta_{33}\mathcal{A})^{-1}\Delta_{31}^{\gamma} & \gamma\mathcal{C}\Delta_{13}^{\gamma}\mathcal{A}(I_m - \Delta_{33}\mathcal{A})^{-1}\Delta_{33}\mathcal{B} + \gamma\mathcal{D} \end{bmatrix} \Rightarrow
$$

$$
\Delta_{11} = \Delta_{11} + \Delta_{13}^{\gamma}\mathcal{A}(I_m - \Delta_{33}\mathcal{A})^{-1}\Delta_{31}^{\gamma} \tag{A.14}
$$

For that, $\Delta_{13}^{\gamma}A(I_m - \Delta_{33}A)^{-1}\Delta_{31}^{\gamma}$ has to be identical to zero matrix. Since Δ_{13}^{γ} and Δ_{31}^{γ} are $\gamma-$unitary, they cannot be zero matrices. Similarly, $(I_m - \Delta_{33}A)$ shall be invertable and hence cannot be identical to zero. Therefor $A = 0$ and thus

$$
\zeta^{*}\Delta\zeta = \Delta_{11} \Leftrightarrow \zeta^{*}\Phi_{\Delta}^{\gamma}\zeta = 0 \tag{A.15}
$$

From part 2 of the lemma A.3.1

$$
X_A^{\gamma} := \mathcal{F}_u\left[(\Delta_0^{\gamma})^{-1}, A\right] \Leftrightarrow A = \mathcal{F}_l\left[(\Delta_0^{\gamma}), X_A^{\gamma}\right] \tag{A.16}
$$

Thus, from (A.13) and (A.16),

$$\begin{aligned} det\,(A-\Delta) &= det\left(\mathcal{F}_l\left[\Delta_0^\gamma, X_A^\gamma\right] - \mathcal{F}_l\left[\Delta_0^\gamma, \Phi_\Delta^\gamma\right]\right) \\ &= det\left(\mathcal{F}_u\left[\left(\Delta_0^\gamma\right)^{-1}, A\right] - \mathcal{F}_u\left[\left(\Delta_0^\gamma\right)^{-1}, \Delta\right]\right) \\ &= det\left(X_A^\gamma - \Phi_\Delta^\gamma\right) \end{aligned} \tag{A.17}$$

Hence, from (A.12), (A.15) and (A.17), we have

$$\min_{\substack{det(A-\Delta)=0 \\ \gamma=\|\Delta\| \\ \zeta^*\Delta\zeta=\Delta_{11}}} \gamma = \min_{\substack{det\left(X_A^\gamma-\Phi_\Delta^\gamma\right)=0 \\ \|\Phi_\Delta^\gamma\|=\gamma^{-1} \\ \zeta^*\Phi_\Delta^\gamma\zeta=0}} \gamma \tag{A.18}$$

Now all the conditions in right hand side, $\begin{cases} det\left(X_A^\gamma - \Phi_\Delta^\gamma\right) = 0 \\ \|\Phi_\Delta^\gamma\| = \gamma^{-1} \\ \zeta^*\Phi_\Delta^\gamma\zeta \end{cases}$, can be replaced by $\sqrt{\underline{\sigma}\left(X_1^\gamma\right)\underline{\sigma}\left(X_2^\gamma\right)} \leq \gamma^{-1}$, i.e. based on Lemma 3.2.2, for $\Delta := \Phi_\Delta^\gamma$ and $A := X_A^\gamma$, and since $X_A^\gamma = diag\left(X_1^\gamma, X_2^\gamma\right)$, we have

$$min\left\{\|\Phi_\Delta^\gamma\|\text{: } det\left(X_A^\gamma - \Delta_\Phi^\gamma\right) = 0, \zeta^*\Phi_\Delta^\gamma\zeta = 0\right\} = \sqrt{\underline{\sigma}\left(X_1^\gamma\right)\underline{\sigma}\left(X_2^\gamma\right)}$$

From the second condition, $\|\Phi_\Delta^\gamma\| = \gamma^{-1}$, thus

$$min\left\{\gamma^{-1} : det\left(X_A^\gamma - \Delta_\Phi^\gamma\right) = 0, \zeta^*\Phi_\Delta^\gamma\zeta = 0\right\} = \sqrt{\underline{\sigma}\left(X_1^\gamma\right)\underline{\sigma}\left(X_2^\gamma\right)} \Rightarrow \sqrt{\underline{\sigma}\left(X_1^\gamma\right)\underline{\sigma}\left(X_2^\gamma\right)} \leq \gamma^{-1}$$

Hence, (A.18), can be reduced to

$$\min_{\substack{det(A-\Delta)=0 \\ \gamma=\|\Delta\| \\ \zeta^*\Delta\zeta=\Delta_{11}}} \gamma = \min_{\substack{det\left(X_A^\gamma-\Phi_\Delta^\gamma\right)=0 \\ \|\Phi_\Delta^\gamma\|=\gamma^{-1} \\ \zeta^*\Phi_\Delta^\gamma\zeta=0}} \gamma = \min_{\sqrt{\underline{\sigma}\left(X_1^\gamma\right)\underline{\sigma}\left(X_2^\gamma\right)} \leq \gamma^{-1}} \gamma = \min_{\|(X_1^\gamma)^{-1}\| \leq a_{m+1}} \gamma \tag{A.19}$$

□

References

[1] J. A. Ferrez, K. Fukuda, and T. Liebling. *Cuts, zonotopes and arrangements*. The Sharpest Cut: The Impact of Manfred Padberg and His Work (MPS-SIAM Series on Optimization). 2004.

[2] Guodong Yin, Nan Chen, and Pu Li. Improving handling stability performance of four-wheel steering vehicle via μ-synthesis robust control. *IEEE transactions on vehicular technology*, 56(5):2432–2439, 2007.

[3] K. Glover. Robust stabilization of linear multivariable systems: relations to approximation. *International Journal of Control*, 43(3):741–766, 1986. https://doi.org/10.1080/00207178608933499.

[4] M. Vidyasagar. *Control system synthesis: a factorization approach*, volume 2. Morgan and Claypool Publishers, San Rafael, California, 2011.

[5] O. Toker and H. Ozbay. On the np-hardness of solving bilinear matrix inequalities and simultaneous stabilization with static output feedback. In *Proceedings of 1995 American Control Conference - ACC'95*, volume 4, pages 2525–2526 vol.4, 1995.

[6] A. Packard and J. Doyle. The complex structured singular value. *Automatica*, 29(1):71–109, 1993.

[7] A. Megretski. On the gap between structured singular values and their upper bounds. In *Proceedings of 32nd IEEE Conference on Decision and Control*, pages 3461–3462 vol.4, 1993.

[8] S. K. Gungah, U. Malik, I.M Jaimoukha, and G.D Halikias. A new upper bound for the real structured singular value. volume 1, pages 247–248 vol.1. IEEE, 2001.

[9] I.M Jaimoukha, G.D Halikias, U. Malik, and S.K. Gungah. On the gap between the

complex structured singular value and its convex upper bound. *SIAM Journal on Control and Optimization*, 45(4):1251–1278, 2006.

[10] D. Axehill. *Applications of Integer Quadratic Programming in Control and Communication*. PhD thesis, 2005.

[11] Dingzhu Du and Ker-I Ko. *Theory of computational complexity*. John Wiley & Sons, Hoboken, New Jersey, second edition, 2014.

[12] Walter Dean. Computational complexity theory and the philosophy of mathematics. *Philosophia mathematica*, 27(3):381–439, 2019.

[13] S. Skogestad and I. Postlethwaite. *Multivariable feedback control: analysis and design*. Wiley, Chichester, 2nd edition, 2005.

[14] S. K. Gungah. *Maximally Robust Controllers for a class of Structured Uncertainty*. PhD thesis, 1998.

[15] M. Vidyasagar. Control system synthesis: A factorization approach, part ii. *Automatica*, 22(6):767–767, 2011.

[16] N. Lehtomaki, D. Castanon, B. Levy, G. Stein, N. Sandell, and M. Athans. Robustness and modeling error characterization. *IEEE Transactions on Automatic Control*, 29(3):212–220, 1984.

[17] J. Doyle. Analysis of feedback systems with structured uncertainties. *IEE Proceedings D - Control Theory and Applications*, 129(6):242–250, November 1982.

[18] M. G. Safonov. Stability margins of diagonally perturbed multivariable feedback systems. *IEE Proceedings D Control Theory and Applications*, 129(6):251, 1982.

[19] O. Toker and H. Ozbay. On the complexity of purely complex mu computation and related problems in multidimensional systems. *IEEE Transactions on Automatic Control*, 43(3):409–414, 1998.

[20] U. Malik. *New Bounds for the Structured Singular Value*. PhD thesis, 2003.

[21] S. Gungah, G.D Halikias, and I.M Jaimoukha. Maximally robust controllers for multivariable systems. *SIAM Journal on Control and Optimization*, 38(6):1805–1829, 2000.

[22] K. Allemand, K. Fukuda, T. M. Liebling, and E. Steiner. A polynomial case of unconstrained zero-one quadratic optimization. *Mathematical Programming*, 91(1):49–52, 2001.

[23] J. C. Picard and H. D. Ratliff. Minimum cuts and related problems. *Networks*, 5(4):357–370, 10/01; 2020/06 1975. doi: 10.1002/net.3230050405; 26.

[24] M. X. Goemans and D. P. Williamson. Approximation algorithms for max-3-cut and other problems via complex semidefinite programming. *Journal of Computer and System Sciences*, 68(2):442–470, March 2004 2004. ID: 272574.

[25] Y. Crama, P. Hansen, and B. Jaumard. The basic algorithm for pseudo-boolean programming revisited. *Discrete Applied Mathematics*, 29(2):171–185, 1990.

[26] F. Barahona. A solvable case of quadratic 0–1 programming. *Discrete Applied Mathematics*, 13(1):23–26, January 1986. ID: 271602.

[27] D. Li and X. Sun. *Nonlinear Integer Programming*. Springer US, 2006. 2006922525.

[28] J. A Ferrez, K. Fukuda, and Th M. Liebling. Solving the fixed rank convex quadratic maximization in binary variables by a parallel zonotope construction algorithm. *European Journal of Operational Research*, 166(1):35–50, 10/1 2005.

[29] K. Fukuda. From the zonotope construction to the minkowski addition of convex polytopes. *Journal of Symbolic Computation*, 38(4):1261–1272, 2004.

[30] U. Malik, I. M. Jaimoukha, G. D. Halikias, and S. K. Gungah. On the gap between the quadratic integer programming problem and its semidefinite relaxation. *Mathematical Programming*, 107(3):505–515, 2006.

[31] N. Karcanias, C. Giannakopoulos, and M. Hubbard. Almost zeros of a set of polynomials of $R(s)$. *International Journal of Control*, 38(6):1213–1238, 1983.

[32] M. Mitrouli and N. Karcanias. Computation of the GCD of polynomials using gaussian transformations and shifting. *International Journal of Control*, 58(1):211–228, 1993.

[33] N. Karcanias and M. Mitrouli. A matrix pencil based numerical method for the computation of the GCD of polynomials. *IEEE Transactions on Automatic Control*, 39(5):977–981, 1994.

[34] S. Fatouros, N. Karcanias, D. Christou, and P. Papadopoulos. Approximate greatest common divisor of many polynomials and pseudo-spectrum. *IFAC Proceedings Volumes*, 46(2):623–628, 2013.

[35] G. Halikias, G. Galanis, N. Karcanias, and E. Milonidis. Nearest common root of polynomials, approximate greatest common divisor and the structured singular value. *IMA Journal of Mathematical Control and Information*, 30(4):423–442, 2012.

[36] N. Karcanias. The greatest common divisor of a set of polynomials, control theory and approximate algebraic computations. *Bulletin of the Greek Mathematical Society*, 54:29–46, 2007.

[37] M.T. Noda and T. Sasaki. Approximate GCD and its application to ill-conditioned equations. *Journal of Computational and Applied Mathematics*, 38(1-3):335–351, 1991.

[38] S. Barnett. *Polynomials and Linear Control Systems*. Marcel Dekker, Inc, USA, 1983.

[39] R. Seidel. Constructing higher-dimensional convex hulls at logarithmic cost per face. In *Proceedings of the Eighteenth Annual ACM Symposium on Theory of Computing*, STOC '86, pages 404–413, New York, NY, USA, 1986. ACM.

[40] K. Stinson, D. F. Gleich, and P. G. Constantine. A randomized algorithm for enumerating zonotope vertices. *ArXiv*, abs/1602.06620, 2016.

[41] G.D Halikias, I.M Jaimoukha, U. Malik, and S. K. Gungah. New bounds on the unconstrained quadratic integer programming problem. *Journal of Global Optimization*, 39(4):543–554, 2007.

[42] D. Avis and K. Fukuda. Reverse search for enumeration. *Discrete Applied Mathematics*, 65(1):21–46, 1996.

[43] M. Garey and D. S. Johnson. *Computers and intractability: a guide to the theory of NP–completeness*. Freeman, San Francisco, CA, 1979.

[44] S. Poljak, F. Rendl, and H. Wolkowicz. A recipe for semidefinite relaxation for (0,1)-quadratic programming. *Journal of Global Optimization*, 7(1):51–73, Jul 1995.

[45] G. Finke, R. E. Burkard, and F. Rendl. Quadratic assignment problems. *North-Holland Mathematics Studies*, 132:61–82, 1987. ID: 275453.

[46] C. Delorme and S. Poljak. Laplacian eigenvalues and the maximum cut problem. *Mathematical Programming*, 62(1):557–574, 02/01 1993. ID: Delorme1993.

[47] E. Balas, S. Ceria, and G. Cornuéjols. A lift–and–project cutting plane algorithm for mixed 0–1 programs. *Mathematical Programming*, 58:295–324, 1993.

[48] H. Wolkowicz, R. Saigal, and L. Vandenberghe. *Handbook of semidefinite programming: theory, algorithms and applications*. Springer Science & Business Media, Boston, 2012.

[49] L. Lovasz and A. Schrijver. Cones of matrices and set-functions and 0-1 optimization. *SIAM Journal on Optimization*, 1(2):166–190, 1991.

[50] S. Poljak and H. Wolkowicz. Convex relaxations of (0, 1)-quadratic programming. *Mathematics of Operations Research*, 20(3):550–561, 1995.

[51] Y. Yang. A polynomial arc-search interior-point algorithm for convex quadratic programming. *European Journal of Operational Research*, 215(1):25–38, 2011.

[52] Y. Nesterov. Semidefinite relaxation and nonconvex quadratic optimization. *Optimization Methods and Software*, 9(1-3):141–160, 1998.

[53] J. Zhang, Y. Ye, and Q. Han. Improved approximations for max set splitting and max nae sat. *Discrete Applied Mathematics: Boolean and Pseudo-Boolean Functions*, 142(1):133, 2004.

[54] H. Edelsbrunner. *Algorithms in Combinatorial Geometry*. Springer-Verlag Berlin Heidelberg, 1987.

[55] R. Creighton Buck. Partition of space. *The American Mathematical Monthly*, 50(9):541–544, 1943.

[56] H. Edelsbrunner, J. O'Rourke, and R. Seidel. Constructing arrangements of lines and hyperplanes with applications. *SIAM Journal on Computing*, 15(2):341–363, 05/01; 2020/06 1986. doi: 10.1137/0215024; 27.

[57] S. Onn and L. Schulman. The vector partition problem for convex objective functions. *Mathematics of Operations Research*, 26, 05/03 2001.

[58] T. Zaslavsky. *Facing up to Arrangements: Face-Count Formulas for Partitions of Space by Hyperplanes: Face-count Formulas for Partitions of Space by Hyperplanes*. American Mathematical Society, 1975.

[59] K. Fukuda, T. M. Liebling, and F. Margot. Analysis of backtrack algorithms for listing all vertices and all faces of a convex polyhedron. *Computational Geometry: Theory and Applications*, 8(1):1–12, 1997.

[60] R. C. Read and R. E. Tarjan. Bounds on backtrack algorithms for listing cycles, paths, and spanning trees. *Networks: An International Journal*, 5:237–252, 07 1975.

[61] M. E. Dyer. The complexity of vertex enumeration methods. *Mathematics of Operations Research*, 8(3):381–402, 1983.

[62] A. Gurung and R. Ray. An efficient algorithm for vertex enumeration of two-dimensional projection of polytopes. *CoRR*, abs/1611.10059, 2016.

[63] T. H. Mattheiss. An algorithm for determining irrelevant constraints and all vertices in systems of linear inequalities. *Operations research*, 21(1):247–260, 1973.

[64] D. Avis. *A Revised Implementation of the Reverse Search Vertex Enumeration Algorithm*, pages 177–198. Polytopes Combinatorics and Computation. Birkha user Basel, Basel, 2000.

[65] C. Bradford Barber, D. P. Dobkin, and H. Huhdanpaa. The quickhull algorithm for convex hulls. *ACM Transactions on Mathematical Software*, 22(4):469–483, dec 1996.

[66] M. Cerný. Goffin's algorithm for zonotopes. *Kybernetika*, 48:890–906, 2012.

[67] S. Sahni and T. Gonzalez. P-complete approximation problems. *J.ACM*, 23(3):555–565, July 1976.

[68] M. Laurent and F. Rendl. *Semidefinite Programming and Integer Programming*, volume 12 of *Discrete Optimization*, page 393. Elsevier, 2005.

[69] G. Zames. On the input-output stability of time-varying nonlinear feedback systems-part ii: Conditions involving circles in the frequency plane and sector nonlinearities. *IEEE Transactions on Automatic Control*, 11(3):465–476, 1966.

[70] P. M. Young. The rank one mixed μ problem and 'kharitonov-type' analysis. *Automatica*, 30(12):1899–1911, 1994.

[71] R. Smith. A study of the gap between the structured singular value and its convex upper bound for low-rank matrices. *IEEE Transactions on Automatic Control*, 42(8):1176–1179, 1997.

[72] M. Goldberg and G. Zwas. On matrices having equal spectral radius and spectral norm. *Linear Algebra and Its Applications*, 8(5):427–434, 1974.

[73] Y. Xia, X. Sun, D. Li, and X. Zheng. On the reduction of duality gap in box constrained nonconvex quadratic program. *SIAM Journal on Optimization*, 21(3):706–729, 2011.

[74] L. Vandenberghe, S. Boyd, and S. Wu. Determinant maximization with linear matrix inequality constraints. *SIAM Journal on Matrix Analysis and Applications*, 19(2):499–533, 1998.

[75] N. Moshtagh. Minimum volume enclosing ellipsoid, Retrieved April 30, 2020. MATLAB Central File Exchange.

[76] N. Karcanias and G.D Halikias. Approximate zero polynomials of polynomial matrices and linear systems. *Linear Algebra and Its Applications*, 439(4):1091–1103, 2013.

[77] S. Barnett. Greatest common divisor of several polynomials. *Mathematical proceedings of the Cambridge Philosophical Society*, 70(2):263–268, 1971.

[78] M.T. Noda and T. Sasaki. Approximate gcd and its application to ill-conditioned equations. *Journal of Computational and Applied Mathematics*, 38(1):335–351, 1991.

[79] G. Halikias, G. Galanis, N. Karcanias, and E. Milonidis. Nearest common root of polynomials, approximate greatest common divisor and the structured singular value. *IMA Journal of Mathematical Control and Information*, 30(4):423–442, 2013.

[80] O. Limantseva, G. Halikias, and N. Karcanias. Nearest common root of a set of polynomials: A structured singular value approach. *Linear Algebra and its Applications*, 584:233–256, January 2020.

[81] Frank M. Callier. Introduction to mathematical systems theory: A behavioural approach, by jan willem polderman and jan c. willems, springer, new york, ny, 1998, texts in applied mathematics vol. 26, 424pp, 2002.

[82] Ivan Markovsky, Antonio Fazzi, and Nicola Guglielmi. *Applications of Polynomial Common Factor Computation in Signal Processing*, pages 99–106. Springer International Publishing, Cham, 2018.

[83] R. M. Redheffer. On a certain linear fractional transformation. *Journal of Mathematics and Physics*, 39(1-4):269–286, 04/01; 2020/04 1960. doi: 10.1002/sapm1960391269; 04.

[84] N. N. Abdelmalek and T. Kasvand. Digital image restoration using quadratic programming. *Appl.Opt.*, 19(19):3407–3415, Oct 1980.

[85] J. W. Adams. Quadratic programming approaches to new problems in digital filter design. volume 1, page 307, 1990.

[86] F. Rendl. Semidefinite programming and combinatorial optimization. *Applied Numerical Mathematics*, 29(3):255–281, March 1999.

[87] D. Bienstock. Computational study of a family of mixed-integer quadratic programming problems. *Mathematical Programming*, 74(2):121–140, Aug 1996.

[88] A. Billionnet and S. Elloumi. Using a mixed integer quadratic programming solver for the unconstrained quadratic 0-1 problem. *Mathematical Programming*, 109(1):55–68, 2007.

[89] R.M. Karp. *Reducibility among Combinatorial Problems*, volume 40 of *Complexity of computer computations*, chapter 4, pages 618–619. R. e. miller, j. w. thatcher, jean d. bohlinger edition, 1975.

[90] E. Çela, B. Klinz, and C. Meyer. Polynomially solvable cases of the constant rank unconstrained quadratic 0-1 programming problem. *Journal of Combinatorial Optimization*, 12(3):187–215, 2006.

[91] O. Goldreich. *P, NP, and NP-Completeness: The Basics of Computational Complexity*. Cambridge University Press, Cambridge, 2010.

[92] D. J. Haglin and S. M. Venkatesan. Approximation and intractability results for the maximum cut problem and its variants. *IEEE Transactions on Computers*, 40(1):110–113, 1991.

[93] E. Hamadi and G. Halikias. Structured singular value of matrices with real parametric uncertainty: Deterministic and probabilistic algorithms. *International Journal of Control is an international, 2020. Under review.*

[94] C. Helmberg and F. Rendl. Solving quadratic (0,1)-problems by semidefinite programs and cutting planes. *Mathematical Programming*, 82(3):291–315, 1998.

[95] T. Hofmeister and H. Lefmann. A combinatorial design approach to maxcut. *Random Structures and Algorithms*, 9(1-2):163–175, 1996.

[96] P. L. Ivănescu. Some network flow problems solved with pseudo-boolean programming. *Operations research*, 13(3):388–399, 06/01; 2018/01 1965. doi: 10.1287/opre.13.3.388; 17.

[97] J. W. Adams. Quadratic programming approaches to new optimal windows and antenna arrays. In *24th Asilomar Conference on Signals, Systems and Computers*, volume 1, page 69, 1990. ID: 1.

[98] H. R. Lewis. Review: Michael r. garey, david s. johnson, computers and intractability. a guide to the theory of np-completeness. *J.Symbolic Logic*, 48(iss. 2):498–500, 1983.

[99] C. Lu, Z. Wang, and W. Xing. An improved lower bound and approximation algorithm for binary constrained quadratic programming problem. *Journal of Global Optimization*, 48(3):497–508, 2010.

[100] G. Meinsma, Y. Shrivastava, and M. Fu. Some properties of an upper bound for mu. *IEEE Transaction on Automatic Control*, 41(9):1326–1330, 1996.

[101] P. M. Young. The rank one mixed /spl mu/ problem and "kharitonov-type" methods. In *Proceedings of 32nd IEEE Conference on Decision and Control*, pages 523–528 vol.1, 1993. ID: 1.

[102] P. M. Pardalos and S. Jha. Graph separation techniques for quadratic zero-one programming. *Computers and Mathematics with Applications*, 21(6):107–113, 1991.

[103] A. Rantzer. On the kalman-yakubovich-popov lemma for positive systems. *IEEE Transactions on Automatic Control*, 61(5):1346–1349, 2016.

[104] A. T. Redpath, B. L. Vickery, and D. H. Wright. A new technique for radiotherapy planning using quadratic programming. *Physics in Medicine and Biology*, 21(5):781–791, 1976.

[105] S. Poljak and D. Turzik. A polynomial algorithm for constructing a large bipartite subgraph with an application to a satisability problem. *Canadian Journal of Mathematics*, 34:519-524, 1982.

[106] P. Sun and R. M. Freund. Computation of minimum-volume covering ellipsoids. *Operations research*, 52(5):690–706, 2004.

[107] P. M. B. Vitányi. How well can a graph be n-colored? *Discrete Mathematics*, 34(1):69–80, 1981.

[108] H. Wolkowicz, R. Saigal, and L. Vandenberghe. *Handbook of semidefinite programming: theory, algorithms, and applications*, volume 27. Springer Science & Business Media, 2012.

[109] Y. Xia and W. Xing. Parametric lagrangian dual for the binary quadratic programming problem. *Journal of Global Optimization*, 61(2):221–233, 2015.

[110] P. M. Young. Controller design with mixed uncertainties. volume 2, pages 2333–2337 vol.2, 1994.

[111] K. Zhou, J. C. Doyle, and K. Glover. *Robust and optimal control*. Prentice Hall, Upper Saddle River, N.J; London, 1996.

[112] N. Karcanias and C. Giannakopoulos. Grassmann invariants, almost zeros and the determinantal zero, pole assignment problems of linear multivariable systems. *International Journal of Control*, 40(4):673–698, 10/01 1984. doi: 10.1080/00207178408933300.