# Invariant conic optimization with basis-dependent cones: scaled diagonally dominant matrices and real *-algebra decomposition

by
Khashayar Neshat Taherzadeh
B.Sc., Azad University, Department of Electrical Engineering, 2016
M.Sc., Sharif University of Technology, Department of Electrical Engineering and Computer Science, 2019

A Thesis Submitted in Partial Fulfillment of the
Requirements for the Degree of

MASTER OF SCIENCE

in the Department of Mathematics and Statistics

# Invariant conic optimization with basis-dependent cones: scaled diagonally dominant matrices and real *-algebra decomposition

by
Khashayar Neshat Taherzadeh
B.Sc., Azad University, Department of Electrical Engineering, 2016
M.Sc., Sharif University of Technology, Department of Electrical Engineering and Computer Science, 2019

**Supervisory Committee**

Dr. David Goluskin
(Department of Mathematics and Statistics, University of Victoria)

---
**Supervisor**

Dr. Heath Emerson
(Department of Mathematics and Statistics, University of Victoria)

---
**Committee Member**

# Abstract

Symmetry reduction for a semidefinite program (SDP) with symmetries makes computational solution of the SDP easier by decomposing the semidefiniteness constraint into multiple smaller semidefiniteness constraints. This decomposition requires changing to a symmetry-adapted basis that block diagonalizes the matrix variable, but this does not change the optimum value of the SDP because the semidefinite cone is basis-independent. For other cones that are basis-dependent, if optimization problems over those cones have symmetries one can still change to a symmetry-adapted basis that block diagonalizes the matrix. However, this change of basis generally changes the constraint cone and can change the optimum. In this thesis, we develop a framework for determining when symmetry reduction for basis-dependent conic optimization makes the optimum increase, decrease, or stay the same. The aim is to determine this using general features such as the symmetry group of the optimization problem, without having to solve the problem computationally. We then use our framework to prove various results of this type for scaled diagonally dominant programs (SDDPs), which are convex optimization problems over the cone of scaled diagonally dominant matrices. These results depend on the orbital structure of the underlying representation of invariant SDDPs. Using the regular representation, we demonstrate that analysis of SDDPs of any size can be confined to a smaller SDDP that is invariant under a particular representation. Our approach uses real *-algebra decomposition of equivariant maps, which is not needed for existing symmetry reduction of SDPs. Because polynomial optimization problems with sum-of-squares and sum-of-binomial-squares can be represented as SDPs and SDDPs, respectively, our results on SDDPs have implications for polynomial optimization. Using several polynomial optimization problems as examples, we give computational results that illustrate our theorems. For polynomial optimization subject to sum-of-binomial-squares, our examples include cases in which symmetry reduction causes the optimum to increase, decrease, or stay the same.

# Contents

# List of Tables

# Acknowledgments

I am deeply grateful to my supervisor, David Goluskin, for trusting me to work with him despite my background in engineering. He has given me the freedom to pursue my passion for algebra. Now, after more than 30 years of living, I am confident that I can find joy until the last day of my life by satisfying my curiosity about algebra. I also want to thank Heath Emerson, whose expertise in algebra shed light on my algebraic path, and Cordian Reiner who generously dedicated his time to serve as the external examiner for this thesis.

I am immensely thankful to my wife, Neda, who respects my love for mathematics and continually motivates me, even while enduring my selfishness.

Last but not least, I am profoundly grateful to my father for his unwavering support throughout these years, helping me reach my goals, even though there is still a long journey ahead.

# Chapter 1

# Introduction

Cone programming, or conic optimization, is the task of minimizing a linear functional over the intersection of a convex cone $\mathcal{K}$ and an affine subspace $\mathcal{L}$ within a vector space $\mathbb{M}$. When $\mathbb{M}$ is a finite-dimensional real space, each linear functional endows $\mathbb{M}$ with an inner product $\langle \cdot, \cdot \rangle$ and acts as $\langle C, \cdot \rangle$ for some $C \in \mathbb{M}$. These programs have the following general form

$$f := \min_{X \in \mathbb{M}} \quad \langle C, X \rangle \qquad \text{subject to} \quad \begin{array}{l} X \in \mathcal{K} \\ X \in \mathcal{L}. \end{array} \tag{1.1}$$

The linear functional and the minimization region are known as the objective function and the feasible region. If both are invariant under some action of a group $G$, cone program (1.1) has a *symmetry* of the group $G$ and is termed *$G$-invariant*. In such programs, symmetry can sometimes be exploited to reduce the computational cost of solving the optimization problems numerically. The initial step to exploit symmetry is to confine the feasible region into the fixed-point subspace $\mathbb{M}^G$ of all elements that remain invariant under the action of $G$. As explained in the next chapter, the optimal value $f$ of an invariant (1.1) remains unchanged when the feasible region is confined to $\mathbb{M}^G$, resulting in

$$f = \min_{X \in \mathbb{M}} \quad \langle C, X \rangle \qquad \text{subject to} \quad \begin{aligned} & X \in \mathcal{K} \cap \mathbb{M}^G \\ & X \in \mathcal{L} \cap \mathbb{M}^G. \end{aligned} \tag{1.2}$$

Although this new program has a smaller feasible region, it does not give immediate computational advantages. Further steps to exploit symmetry rely on the underlying vector space $\mathbb{M}$ and cone $\mathcal{K}$.

When $\mathbb{M}$ is the space $\mathcal{S}(N)$ of $N \times N$ symmetric matrices and $\mathcal{K}$ is the cone $\mathcal{S}(N)_+$ of positive semidefinite (PSD) matrices, program (1.1) is termed semidefinite programming (SDP). The methodology for exploiting symmetry in invariant SDPs, known as symmetry reduction, dates back at least to the late 1970s when [Schrijver, 1979] reduced the size of PSD matrices based on their underlying structure. Gaterman and Parrilo in [Gatermann and Parrilo, 2004] initiated the systematic way of exploiting symmetry for SDPs by finding a basis on which the problem turns into a series of smaller-sized SDPs. Their work utilizes representation theory to find an orthonormal matrix $U$, known as a symmetry-adapted basis, that transforms each element of $\mathcal{S}(N)^G$ into a particular block diagonal form through the matrix similarity transformation. An alternative methodology for exploiting symmetry in SDPs, utilizing the Artin-Wedderburn decomposition of the semisimple algebra $\mathcal{S}(N)^G$, is explored in [de Klerk et al., 2011, Bachoc et al., 2012]. Advances in using symmetry for SDP relaxations of moment problems are highlighted in [Riener et al., 2013]. For a contemporary examination of these constructions, along with essential background on representation theory and algebra decomposition within the settings of SDP and polynomial optimization, refer to [Moustrou et al., 2023]. While there have been extensive advancements in exploiting symmetry in SDPs, the potential for doing so in other classes of cone programs remains largely unexplored, except in relative entropy programming [Moustrou et al., 2022].

In this work, we aim to explore symmetry reduction methods for the cases when $\mathbb{M} = \mathcal{S}(N)$ but $\mathcal{K}$ is a non-PSD cone—of most relevance here, cone of scaled diagonally dominant matrices. The second step, here, is to use the matrix similarity transformation by a symmetry-adapted basis $U$ as a change of variable in (1.2), resulting in

$$f = \min_{X \in \mathcal{S}(N)} \quad \langle U^T C U, X \rangle \qquad \text{subject to} \quad \begin{aligned} & X \in U^T \mathcal{K} U \cap U^T \mathcal{S}(N)^G U \\ & X \in U^T \mathcal{L} U \cap U^T \mathcal{S}(N)^G U. \end{aligned} \tag{1.3}$$

Although matrices in the feasible region of (1.3) have block diagonal form, cone programming solvers do not utilize this form. When a solver is designed to check the membership of elements within $\mathcal{K}$, it cannot directly check the membership of elements in the transformed cone $U^T \mathcal{K} U$. Therefore, conic solvers treat the constraint $X \in U^T \mathcal{K} U$ in (1.3) as $U X U^T \in \mathcal{K}$, which offers no computational advantage. This issue does not occur in the SDP case, where $\mathcal{K}$ is the PSD cone. In this case, $U^T \mathcal{K} U = \mathcal{K}$ for any orthogonal matrix $U$, and the cone program

$$f_U := \min_{X \in \mathcal{S}(N)} \quad \langle U^T C U, X \rangle \qquad \text{subject to} \quad \begin{aligned} & X \in \mathcal{K} \cap U^T \mathcal{S}(N)^G U \\ & X \in U^T \mathcal{L} U \cap U^T \mathcal{S}(N)^G U \end{aligned} \tag{1.4}$$

has the same optimal value as program (1.3) (i.e., $f_U = f$). This property of the PSD cone in particular underlies the symmetry exploitation method studied by [Gatermann and Parrilo, 2004] and subsequent authors, where one can change the basis to block diagonalize $X$ without changing the representation of the underlying cone. When $\mathcal{K} = \mathcal{S}(N)_+$, the first constraint in (1.4) imposes a block structure on $X$ while ensuring $X$ maintains its positive semidefinite nature. This arrangement effectively decomposes the PSD cone into a direct sum of smaller individual PSD cones for each block, reducing the problem to a series of smaller-sized SDPs. Consequently, program (1.4), termed the symmetry-reduced version of the SDP problem, offers computational efficiency while retaining the same optimal values as the original problem.

Unlike the PSD cone, other cones generally do not remain invariant under similarity transformations. Consequently, programs (1.3) and (1.4) do not necessarily have the same optimal value for non-PSD cones $\mathcal{K}$. In such a situation, one approach as the third step would be to develop computational methods that directly enforce membership in the transformed cone $U^T \mathcal{K} U$ when solving (1.3) numerically. This approach is a numerical task and is not pursued in this thesis. Another approach, which we pursue here, is to consider program (1.4) with optimal value $f_U$ as the symmetry-reduced version of program (1.1) (when $\mathbb{M} = \mathcal{S}(N)$) for any arbitrary cone $\mathcal{K}$ and exploring the relation between $f_U$ and $f$ for different symmetry-adapted bases $U$. This approach is consistent with the practical implementation of symmetry reduction. In fact, when existing symmetry reduction packages such as SymbolicWedderburn.jl [Kaluba et al., 2019] in Julia programming language [Bezanson et al., 2017] are applied to cones other than the PSD cone, the solution returned will be the optimal value $f_U$ of (1.4), rather than $f$; this may not be immediately evident to users.

Establishing the equality $f = f_U$ *a priori* guarantees that symmetry reduction reduces the computational complexity without changing the optimal value. However, an inequality relation between $f$ and $f_U$ is often more desirable than equality for the following reason; when program (1.1) is considered as an inner (resp. outer) approximation of an SDP, its symmetry-reduced version (1.4) is also an inner (resp. outer) approximation of the SDP. This is because whenever $\mathcal{K}$ is the subcone (resp. supercone) of the PSD cone, $U^T \mathcal{K} U$ for any orthogonal matrix $U$ is also a subcone (resp. supercone) of the PSD cone. In this situation, one hopes that symmetry reduction not only reduces computational complexity but also provides a closer upper or lower bound for the optimal value of the SDP. It is indeed possible in some cases to show that $f \leq f_U$ or $f \geq f_U$ without computing the values $f$ or $f_U$. Here, we prove some such results that depend only on the cone $\mathcal{K}$ and the symmetry group $G$, and sometimes also on the choice of the symmetry-adapted basis $U$. The present work focuses on the cases where $\mathcal{K}$ is the cone of scaled diagonally dominant (SDD) matrices. Program (1.1) over this cone is called scaled diagonally dominant programming (SDDP). [Ahmadi and Majumdar, 2019] showed that SDDP problems can be solved using second-order cone programming, a computationally faster and more tractable optimization method than SDP. Since the SDD cone is a subcone of the PSD cone, SDDP serves as an efficient inner approximation for SDP, offering a trade-off between computational complexity and solution quality. Approximating SDP with SDDP is particularly relevant in polynomial optimization problems. Membership in the cone of sum-of-squares (SOS) polynomials serves as a certification of nonnegativity. This leads to solving an SDP problem [Parrilo, 2003]. Alternatively but more conservatively, membership in the cone of sum-of-binomial-squares (SBS) provides such a certificate leading to solving an SDDP problem [Ahmadi and Majumdar, 2019].

We begin by establishing geometric conditions for a general cone $\mathcal{K}$ and a symmetry-adapted basis $U$ to guarantee equality or inequality relation between $f$ and $f_U$. Focusing on SDDPs, we identify group representations under which the SDD cone is invariant. Among such representations, we characterize representations that result in $f_U = f$ and $f_U \leq f$ for some or all symmetry-adapted bases. Our methodology relies on the geometric structures of the SDD cone and invokes the Wedderburn-Artin decomposition of the semisimple real algebras. We illustrate the effect of symmetry reduction on polynomial optimization problems which result in SDDP problems.

The rest of this thesis is structured as follows. Chapter 2 provides background on invariant cone programming and SDDPs, and further characterizes the SDD cone and invariant SDDPs. Chapter 3 develops the theory of *-algebra decomposition of equivariant maps and offers an overview of representation theory. Chapter 4 presents the characterization of finite groups and their representations for the relation between $f$ and $f_U$ in SDDPs.

Chapter 5 applies our results to polynomial optimization with numerical experiments. We conclude this work in chapter 6.

# Chapter 2

# Invariant cone programming

Invariant optimization problems in this work are defined in terms of group representations. For our purposes, the reader can refer to any standard textbook on representation theory, such as [Serre, 1977]. Let $\mathbb{F}$ be a field. A linear $\mathbb{F}$-representation of $\boldsymbol{G}$ is a group homomorphism $\sigma : \boldsymbol{G} \to \mathrm{GL}(\mathbb{M})$ from $\boldsymbol{G}$ to the general linear group of a $\mathbb{F}$-vector space $\mathbb{M}$. Each $\sigma(g)$, which is usually denoted by $\sigma_g$, is an element of $\mathrm{GL}(\mathbb{M})$ and so is a linear bijection from $\mathbb{M}$ to $\mathbb{M}$. The group $\boldsymbol{G}$ acts on $\mathbb{M}$ through $\sigma$ via $g \cdot X := \sigma_g(X)$ for $X \in \mathbb{M}$, endowing $\mathbb{M}$ with the structure of left $\boldsymbol{G}$-module, and so $\mathbb{M}$ is termed representation module. We shall also refer to $\mathbb{M}$ simply as representation. Defining $\sigma_g(\mathfrak{X}) = \{\sigma_g(X) \mid X \in \mathfrak{X}\}$ for any $\mathfrak{X} \subset \mathbb{M}$, the subset $\mathfrak{X}$ is $\boldsymbol{G}$-invariant with respect to $\sigma$ if $\sigma_g(\mathfrak{X}) \subset \mathfrak{X}$ for all $g \in \boldsymbol{G}$. A functional $F(\cdot) : \mathbb{M} \to \mathbb{F}$ is $\boldsymbol{G}$-invariant with respect to $\sigma$ if $F\left(\sigma_g(X)\right) = F(X)$ for any $X \in \mathbb{M}$ and all $g \in \boldsymbol{G}$. We shall omit "with respect to $\sigma$" when the underlying representation $\sigma$ has already been specified. The optimization problems discussed remain invariant under real representations, which requires representation theory over the reals $\mathbb{R}$, as opposed to the simpler theory over such algebraically closed fields as the complex numbers $\mathbb{C}$. For our purpose, the underlying field $\mathbb{F}$ designated as $\mathbb{R}$ or $\mathbb{C}$. When a real (resp. complex) representation module $\mathbb{M}$ is additionally equipped with an inner product preserved by each

$\sigma_g$, $(\mathbb{M}, \sigma)$ is termed an orthogonal (resp. unitary) representation. We frequently consider three sub-types of orthogonal representations: signed permutation or orthogonal monomial representations, permutation representations, and signed diagonal representations. In signed permutation representations, each $\sigma_g$ is a signed permutation matrix, similar to a permutation matrix but allowing the non-zero entries to be $-1$ as well as 1. When these operators are permutation matrices, $\sigma$ is a permutation representation. Such representations are distinct from "sign representations" in representation theory [Serre, 1977], but optimization problems invariant under signed diagonal representations are said to possess sign symmetry.

## 2.1 General formulation

Given a group $G$ and representation $\sigma : G \to \mathrm{GL}(\mathbb{M})$, program (1.1) is $G$-invariant (with respect to $\sigma$) if its objective function and feasible region are $G$-invariant. The fixed-point or $G$-stable subspace of $\mathbb{M}$ associated with $\sigma$ is defined

$$\mathbb{M}^G := \{X \in \mathbb{M} \mid \sigma_g(X) = X \quad \forall g \in G\}. \tag{2.1}$$

When $G$ is finite with cardinality $|G|$, the group average (also known as the Reynolds operator) of $X \in \mathbb{M}$ is well defined by

$$X_\sigma := \frac{1}{|G|} \sum_{g \in G} \sigma_g(X). \tag{2.2}$$

Every averaged $X_\sigma$ lies in $\mathbb{M}^G$, and if $X$ was already in $\mathbb{M}^G$ then its average leaves it fixed. Thus $\mathbb{M}^G$ is exactly the group average of $\mathbb{M}$,

$$\mathbb{M}^G = \left\{ X_\sigma \,\middle|\, X \in \mathbb{M} \right\}. \tag{2.3}$$

The key step for exploiting symmetry in invariant (1.1) is the following lemma.

**Lemma 2.1.1.** *Given a finite group $G$ and a $G$-invariant program* (1.1) *with optimal value $f$, its fixed-point restricted version* (1.2) *has the same optimal value $f$.*

*Proof.* The proof is a symmetrization argument that is standard in convex optimization. The SDP case is given by [Gatermann and Parrilo, 2004, Theorem 3.3], and the same argument works for cone programs more generally. Since the feasible region of (1.2) is a subset of that of (1.1), it suffices to show that, for every element in the feasible region of (1.1), there exists an element in the restricted feasible region such that the objective function attains the same value at both points. Since taking group average is a convex combination and the feasible region of (1.1) is convex, for every $X$ in the feasible region of (1.1), $X_\sigma$ is in the feasible region of (1.2). Additionally,

$$\langle C, X_\sigma \rangle = \left\langle C, \frac{1}{|G|} \sum_{g \in G} \sigma_g(X) \right\rangle = \frac{1}{|G|} \sum_{g \in G} \left\langle C, \sigma_g(X) \right\rangle = \frac{1}{|G|} \sum_{g \in G} \langle C, X \rangle = \langle C, X \rangle. \tag{2.4}$$

□

As mentioned in the introduction, program (1.2) does not offer immediate computational advantages. Further steps are required to decompose the conic constraint $\mathcal{K} \cap \mathbb{M}^G$ into simpler sub-constraints. This can be done when the underlying vector space $\mathbb{M}$ is the space $\mathrm{End}(\mathbb{V})$ of linear maps from a vector space $\mathbb{V}$ to itself, which possesses the structure of an (associative) algebra, and $\sigma : G \to \mathrm{Aut}(\mathrm{End}(\mathbb{V}))$ is a homomorphism where each $\sigma_g$ is not only a linear transformation but also an algebra homomorphism. Note that $\mathbb{M} = \mathcal{S}(N)$ in this context is a specific case, representing the algebra $\mathrm{End}(\mathbb{R}^N)^{sa}$ of self-adjoint operators of $\mathrm{End}(\mathbb{R}^N)$.

We aim to decompose the fixed-point subspace $\mathbb{M}^G = \mathrm{End}(\mathbb{V})^G$. By the Skolem-Noether theorem, Aut (End($\mathbb{V}$)) is isomorphic to the projective general linear group PGL($\mathbb{V}$) of $\mathbb{V}$ [Gille and Szamuely, 2017, Corrollary 2.4.2]. Thus, $\sigma$ is indeed a projective representation $\sigma : G \to \mathrm{PGL}(\mathbb{V})$. If $\sigma$ lifts to a linear representation $\rho : G \to \mathrm{GL}(\mathbb{V})$, i.e.,

$$\sigma_g(X) = \rho_g^{-1} X \rho_g \quad \forall X \in \mathrm{End}(\mathbb{V}) \text{ and } \forall g \in G, \tag{2.5}$$

then $G$ will act on $\mathbb{V}$, endowing $\mathbb{V}$ with the structure of $G$-module. Consequently, the fixed-point subspace will be the algebra of $G$-equivariant maps

$$\mathrm{End}_G(\mathbb{V}) := \{X \in \mathrm{End}(\mathbb{V}) \mid X\rho_g = \rho_g X \ \ \forall g \in G\} \tag{2.6}$$

If this lift is not possible, $\sigma$ can instead lift to a linear representation $\tilde{\rho} : \tilde{G} \to \mathrm{GL}(\mathbb{V})$, where $\tilde{G}$ is a central extension of $G$ [Gannon, 2007, Chapter 3]. Here, the fixed-point subspace equals to $\mathrm{End}_{\tilde{G}}(\mathbb{V})$ associated with $\tilde{\rho}$. Henceforth, without loss of generality, we assume the projective representation $\sigma$ is induced by $\rho$ via (2.5), and so the fixed-point subspace is $\mathrm{End}_G(\mathbb{V})$. Additionally, if the underlying space $\mathbb{M}$ is the algebra $\mathrm{End}(\mathbb{V})^{sa}$ (i.e. the linear representation $\rho$ is orthogonal or unitary), then the fixed-point subspace will be $\mathrm{End}_G(\mathbb{V})^{sa}$. A key property of $\mathrm{End}_G(\mathbb{V})$ (resp. $\mathrm{End}_G(\mathbb{V})^{sa}$) is that its elements have a particular block diagonal form in a suitable basis (resp. orthonormal basis), known as a symmetry-adapted basis. The fact that such a basis exists leads to the simplification of the conic constraint. A comprehensive discussion of this basis is provided in the subsequent chapter.

## 2.2 Effects of symmetry reduction

In some programs, the underlying cone $\mathcal{K}$ is replaced with a simpler cone in terms of computation. The new program is termed the relaxation of the initial one. Later, we aim to investigate $G$-invariant (1.1) whose relaxation remains $G$-invariant as well, which, in turn, involves the following definition.

**Definition 2.2.1.** A strongly $G$-invariant program (1.1) is a $G$-invariant program whose underlying cone $\mathcal{K}$ and affine subspace $\mathcal{L}$ are individually $G$-invariant.

The invariance of $\mathcal{K}$ and $\mathcal{L}$ is a stronger condition than their intersection being invariant. If the underlying cone of a strongly $G$-invariant program is replaced with another $G$-invariant cone, the new program is still strongly $G$-invariant. We focus on strongly $G$-invariant programs (1.1) with $\mathbb{M} = \mathcal{S}(N) = \text{End}(\mathbb{R}^N)^{sa}$ and finite group $G$. Here, $\sigma$ is induced by orthogonal representation $\rho : G \to O(N)$, where $O(N)$ is the group of $N \times N$ orthogonal matrices. Thus,

$$\sigma_g(X) = \rho_g^T X \rho_g \quad \forall X \in \mathcal{S}(N) \text{ and } \forall g \in G. \tag{2.7}$$

**Remark 2.2.2.** *Every $G$-invariant SDP is strongly $G$-invariant. This is because the PSD cone is basis independent and always $G$-invariant. The $G$-invariance of the feasible region, hence, necessitates the $G$-invariance of affine subspace $\mathcal{L}$.*

There exists an orthogonal matrix $U$ such that, for every $A \in \mathcal{S}(N)^G$, $U^T A U$ has a particular block diagonal form. The existence of matrix $U$ and detailed discussion about the structure of blocks will be explained in the next chapter. Matrix $U$ is also known as a symmetry-adapted basis. In our context, it will be clear when referring to a symmetry-adapted basis, we mean a genuine basis or matrix $U$. The change of variable $X$ in program (1.2) to $UXU^T$ results in program (1.3). Elements of the feasible region of this new optimization have the desired block-diagonal form. However, the underlying cone may no longer be $\mathcal{K}$. This does not offer computational advantages as explained in the introduction. Thus, $U^T \mathcal{K} U$ is replaced with $\mathcal{K}$, resulting in program (1.4). This optimization has the same underlying cone as the original problem with the desired block-diagonal form for the elements of its feasible region. The following example illustrates these steps.

**Example 2.2.3.** Let $\mathbb{M} = \mathcal{S}(3)$ with inner product $\langle C, X \rangle = \text{Trace}(CX)$. Consider

$$\mathcal{L} = \left\{ \begin{bmatrix} 1 & 1 & 1 \\ 1 & a_1 & b \\ 1 & b & a_2 \end{bmatrix} \middle| \, a_1, a_2, b \in \mathbb{R} \right\},$$

$$C = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

and a proper cone $\mathcal{K} \subset \mathcal{S}(N)$ which is invariant under simultaneous permutation of rows and columns. The program (1.1) gets the form

$$f = \min_{X \in \mathcal{S}(3)} a_1 + a_2 \quad \text{s.t.} \quad X = \begin{bmatrix} 1 & 1 & 1 \\ 1 & a_1 & b \\ 1 & b & a_2 \end{bmatrix} \in \mathcal{K}. \tag{2.8}$$

This problem is invariant under the action of the cyclic group of order two given by the simultaneous permutation of the last two rows and columns. Restricting the feasible region to the fixed-point subspace results in $a_1 = a_2 = a$, leading to the following form of program (1.2):

$$f = \min_{X \in \mathcal{S}(3)} 2a \quad \text{s.t.} \quad X = \begin{bmatrix} 1 & 1 & 1 \\ 1 & a & b \\ 1 & b & a \end{bmatrix} \in \mathcal{K}. \tag{2.9}$$

A symmetry-adapted basis associated with the fixed-point subspace $\mathcal{S}(3)^G$ is

$$U = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix}.$$

Here, $U^T C U = C$, thus program (1.3), obtained by similarity transform with $U$ as a change of variable, has the form

$$f = \min_{X \in \mathcal{S}(3)} 2a \quad \text{s.t.} \quad X = \begin{bmatrix} 1 & \sqrt{2} & 0 \\ \sqrt{2} & a+b & 0 \\ 0 & 0 & a-b \end{bmatrix} \in U^T \mathcal{K} U. \tag{2.10}$$

Since the membership in the transferred cone $U^T\mathcal{K}U$ is not defined for optimization solvers, we replace it with $\mathcal{K}$, resulting in program (1.4) with the form

$$f_U = \min_{X \in \mathcal{S}(3)} 2a \quad \text{s.t.} \quad X = \begin{bmatrix} 1 & \sqrt{2} & 0 \\ \sqrt{2} & a+b & 0 \\ 0 & 0 & a-b \end{bmatrix} \in \mathcal{K}. \tag{2.11}$$

This final optimization allows for checking $2 \times 2$ and $1 \times 1$ matrices within the desired cones, rather than $3 \times 3$ matrices. In the SDP case where $U^T\mathcal{K}U = \mathcal{K}$ for any matrix $U$, $f_U = f$. This, however, is not true in generality.

In the cases where $\mathcal{K} = \mathcal{S}(N)_+$, since the PSD cone is basis independent (i.e. invariant under any similarity transform), the optimal value $f_U$ of (1.4) equals the optimal value $f$ of the original SDP. However, this is not generally true, and we aim to characterize the relation between $f_U$ and $f$ based on the cone $\mathcal{K}$ and the symmetry-adapted basis $U$. This characterization is possible in certain instances.

**Lemma 2.2.4.** *The following are sufficient conditions for the optimum $f$ of* (1.3) *and the optimum $f_U$ of* (1.4) *to be related by equalities or inequalities.*

1. *The relation $f_U = f$ holds, if*

$$U^T \mathcal{K} U = \mathcal{K}. \tag{2.12}$$

2. *The relation $f_U \leq f$ holds, if*

$$U^T \left(\mathcal{K} \cap \mathcal{S}(N)^G\right) U \subset \mathcal{K}. \tag{2.13}$$

3. *The relation $f \leq f_U$ holds, if*

$$\mathcal{K} \cap U^T \mathcal{S}(N)^G U \subset U^T \mathcal{K} U. \tag{2.14}$$

*Proof.* Given that programs (1.3) and (1.4) share the same objective function, it is straightforward to show the first statement. If the minimizer of program (1.3) belongs to the feasible region of program (1.4), then $f_U \leq f$; conversely, if the minimizer of (1.4) belongs to the feasible region of (1.3), then $f \leq f_U$. We will demonstrate that the inclusion $U^T \left(\mathcal{K} \cap \mathcal{S}(N)^G\right) U \subset \mathcal{K}$ ensures the feasible region of (1.4) includes that of (1.3), thereby guaranteeing that the minimizer of (1.3) lies within the feasible region of (1.4).

$$U^T \left(\mathcal{K} \cap \mathcal{S}(N)^G\right) U \subset \mathcal{K} \Rightarrow U^T \left(\mathcal{K} \cap \mathcal{S}(N)^G\right) U \subset \mathcal{K} \cap U^T \mathcal{S}(N)^G U \tag{2.15}$$

$$\Rightarrow U^T \left(\mathcal{K} \cap \mathcal{S}(N)^G\right) U \cap U^T \mathcal{L} U \subset \mathcal{K} \cap U^T \mathcal{S}(N)^G U \cap U^T \mathcal{L} U \tag{2.16}$$

This completes the proof of the second statement. The proof of the third statement follows a similar argument. □

This lemma provides sufficient conditions for various cones $\mathcal{K}$ and orthogonal matrices $U$. Our interest lies in determining when these sufficient conditions hold, particularly when $U$ is a symmetry-adapted basis associated with $\mathcal{S}(N)^G$. As mentioned in the introduction, it is especially useful to verify the second part of lemma 2.2.4 for some symmetry-adapted basis $U$ when an SDP is approximated by program (1.3) with $\mathcal{K} \subset \mathcal{S}(N)_+$, and the third part of lemma 2.2.4 when an SDP is approximated by program (1.3) with $\mathcal{S}(N)_+ \subset \mathcal{K}$. In what follows, we focus on the cases where $\mathcal{K}$ is the cone of SDD matrices.

## 2.3 Gemoetry of the SDD cone

SDDPs are a class of cone programming whose underlying cone $\mathcal{K}$ compromises the SDD matrices. Compared to SDPs, the algorithm to solve this class is more tractable with less computational cost, making them an effective inner approximation of SDPs. For this reason, we further characterize SDD matrices in this section. Symmetric matrix $A = (a_{ij})$ is diagonally dominant if $a_{ii} \geq \sum_{j \neq i} |a_{ij}|$ for all $i$. A symmetric matrix $A$ is scaled diagonally dominant if there exists a diagonal matrix $D$, with positive diagonal entries, such that $DAD$ is diagonally dominant. We denote the set of $N \times N$ SDD matrices as $\mathcal{S}(N)_{2,+}$. The choice for this notation is explained after the following lemma which recalls two more characterizations of the SDD cone.

**Lemma 2.3.1** ([Ahmadi and Majumdar, 2019, Boman et al., 2005]). *The following are equivalent characterizations that a real symmetric matrix $A \in \mathcal{S}(N)$ belonging to $\mathcal{S}(N)_{2,+}$:*

1. *There exists a positive definite diagonal matrix $\mathbf{D}$ such that $DAD$ is diagonally dominant.*

2. *Considering $e_i$ to be the ith member of the standard basis of $\mathbb{R}^N$. For some $A_0^{ij} \in \mathcal{S}(2)_+$,*

$$A = \sum_{1 \le i < j \le N} [e_i \; e_j] A_0^{ij} [e_i \; e_j]^T \tag{2.17}$$

3. *$com(A) := A - |A - Diag(A)| \in \mathcal{S}(N)_+$, where $|\cdot|$ denotes elementwise absolute value.*

For the proof of the equivalence of this lemma's first and second parts, see [Ahmadi and Majumdar, 2019] and for that of the second and third parts, seec[Boman et al., 2005]. The matrix com($A$) is called the comparison matrix of $A$, which has the same diagonal as $A$ while giving all off-diagonal entries a nonpositive sign.

**Remark 2.3.2.** *Part 3 of lemma 2.3.1 implies $\mathcal{S}(2)_{2,+} = \mathcal{S}(2)_+$. It also implies that PSD matrices are SDD if they have a block diagonal structure (up to reordering of indices) with no blocks larger than $2 \times 2$. We will make use of this fact for the proof of theorem 2.3.3 and also in chapter 3 and chapter 4.*

The factor width of a real symmetric matrix $A$ is the smallest integer $k$ such that there exists a real (rectangular) matrix $B$ where $A = BB^T$ and each column of $B$ contains at most $k$ non-zeros. We denote the set of $N \times N$ factor-width-$k$ matrices with $\mathcal{S}(N)_{k,+}$. Since the set of SDD matrices coincides with the set of symmetric matrices with factor-width two [Boman et al., 2005], this set is denoted by $\mathcal{S}(N)_{2,+}$. We note that $\mathcal{S}(N)_{2,+} \subset \mathcal{S}(N)_{3,+} \subset \cdots \subset \mathcal{S}(N)_{N,+} = \mathcal{S}(N)_+$ (the more general case of remark 2.3.2). The set $\mathcal{S}(N)_{2,+}$ forms a full dimensional proper cone within the space $\mathcal{S}(N)$, meaning that it is convex, pointed, closed, and solid (with interior). Characterization 2 of lemma 2.3.1 establishes that $\mathcal{S}(N)_{2,+}$ is representable as a second-order cone. Consequently, SDDPs can effectively be solved by second-order cone programming. Existing algorithms for solving these problems are more cost-effective than SDPs at the same matrix size; for more details on second-order cone programming, see [Boyd and Vandenberghe, 2004].

Investigating the relation between programs(1.3) and (1.4) when $\mathcal{K} = \mathcal{S}(N)_{2,+}$ requires understanding how conjugation by $U$ transforms this cone. Since $\mathcal{S}(N)_{2,+} = \mathcal{S}(2)_+$ and SDP cones are basis-invariant, $\mathcal{S}(N)_{2,+} = U^T \mathcal{S}(N)_{2,+} U$ for any orthogonal $U$. When $N \geq 3$, such a relation holds only for certain $U$. These $U$ are characterized by the following theorem, which is used repeatedly in what follows.

**Theorem 2.3.3.** *Let $N \geq 3$ and $Q \in O(N)$. If and only if $Q$ is a signed permutation matrix,*

$$Q^T \mathcal{S}(N)_{2,+} Q = \mathcal{S}(N)_{2,+}. \tag{2.18}$$

*Proof.* For the 'only if' part, assume $Q$ is any signed permutation matrix. Note that the properties of being a diagonally dominant matrix and being a diagonal matrix with positive diagonal entries are preserved under conjugation by $Q$. For an arbitrary $A \in \mathcal{S}(N)_{2,+}$, there exists a diagonal matrix $D$ with positive diagonal entries such that $DAD$ is diagonally dominant. This ensures $Q^T DADQ$ remains diagonally dominant. Defining $\tilde{D} := Q^T DQ$, which is also a diagonal matrix with positive diagonal entries, $\tilde{D} Q^T AQ \tilde{D} = Q^T DAQ^T DQ$ is diagonally dominant. Consequently, $Q^T AQ \in \mathcal{S}(N)_{2,+}$, leading to the inclusion $Q^T \mathcal{S}(N)_{2,+} Q \subset \mathcal{S}(N)_{2,+}$. Since conjugation by $Q$ is bijective, it follows that $Q^T \mathcal{S}(N)_{2,+} Q = \mathcal{S}(N)_{2,+}$.

For the 'if' part, we aim to demonstrate that for every orthogonal matrix $Q$ which is not a signed permutation matrix, there exists $A \in \mathcal{S}(N)_{2,+}$ such that $Q^T AQ \notin \mathcal{S}(N)_{2,+}$. Let $m$ denote the maximum number of nonzero elements amongst the columns of $Q$. If $m = 1$, then $Q$ is a signed permutation matrix. Therefore, we need to consider only the cases where $m \geq 3$ and $m = 2$.

**Case 1** ($m \geq 3$)**:** Suppose $a_i$ is the i-th column of $Q^T$. Consider $\hat{a}_i$ to be the $m \times 1$ vector whose elements are nonzero elements of $a_i$. Choose $A = e_i e_i^T$. Clearly, $A \in \mathcal{S}(N)_{2,+}$, and

$$Q^T A Q = a_i a_i^T \notin \mathcal{S}(N)_{2,+} \iff \hat{a}_i \hat{a}_i^T \notin \mathcal{S}(m)_{2,+} \tag{2.19}$$

$$\iff \text{com}(\hat{a}_i \hat{a}_i^T) \notin \mathcal{S}(m)_+ \tag{2.20}$$

The last equivalence is a consequence of lemma 2.3.1 Part 4. To prove $\text{com}(\hat{a}_i \hat{a}_i^T)$ is not positive semidefinite, we show its determinant is negative. Let $\prod \hat{a}_i$ represent the multiplication of all elements of $\hat{a}_i$, $1_m$ represent $m \times m$ matrix of all ones, and $I_m$ represent $m \times m$ identity matrix.

$$|\text{com}(\hat{a}_i \hat{a}_i^T)| = (\prod \hat{a}_i)^2 |1_m - 2I_m| \tag{2.21}$$

$$= -(m-2)2^{m-1}(\prod \hat{a}_i)^2 \tag{2.22}$$

where the last equality can be computed by elementary row and column operations on $|1_m - 2I_m|$.

**Case 2** ($m = 2$)**:** We also divide this case into two subcases:

I: $Q$ has a column with one nonzero element.

II: All the columns of $Q$ have two nonzero elements.

I) Suppose $Q$ has a column with one nonzero element. Then, after appropriately signed permutation, there exists a $3 \times 3$ principal submatrix of the form

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos(\theta) & \sin(\theta) \\ 0 & \sin(\theta) & -\cos(\theta) \end{bmatrix},$$

for some $0 < \theta < \frac{\pi}{2}$. We, then, can consider $A$ to have only nonzero elements corresponding to the $3 \times 3$ principal submatrix. For this reason, without loss of generality, we consider

$N = 3$ and

$$Q = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos(\theta) & \sin(\theta) \\ 0 & \sin(\theta) & -\cos(\theta) \end{bmatrix}$$

for some $0 < \theta < \frac{\pi}{2}$. Choose

$$A = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

and define $B(\theta) := Q^T AQ$. It is straightforward to show that while $A \in \mathcal{S}(3)_{2,+}$, $B(\theta) \neq \mathcal{S}(3)_{2,+}$ for all $0 < \theta < \frac{\pi}{2}$ since $\left|\mathrm{com}\big(B(\theta)\big)\right| = -4\sin(\theta)^2\cos(\theta)^2 < 0$.

II) Suppose all the columns of $Q$ have two nonzero elements. Then, with similar reasoning, without loss of generality, we can consider $N = 4$ and

$$Q = \begin{bmatrix} \cos(\theta_1) & \sin(\theta_1) & 0 & 0 \\ \sin(\theta_1) & -\cos(\theta_1) & 0 & 0 \\ 0 & 0 & \cos(\theta_2) & \sin(\theta_2) \\ 0 & 0 & \sin(\theta_2) & -\cos(\theta_2) \end{bmatrix}$$

for some $0 < \theta_1 < \frac{\pi}{2}$ and $0 < \theta_2 < \frac{\pi}{2}$. Choose

$$A = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

and define $B(\theta_1, \theta_2) := Q^T AQ$ for all $0 < \theta_1 < \frac{\pi}{2}$ and $0 < \theta_2 < \frac{\pi}{2}$. It is straightforward to show that while $A \in \mathcal{S}(4)_{2,+}$, $B(\theta_1, \theta_2) \neq \mathcal{S}(4)_{2,+}$ for all $0 < \theta_1 < \frac{\pi}{2}$ and $0 < \theta_2 < \frac{\pi}{2}$

since $\left|\text{com}\big(B(\theta_1,\theta_2)\big)\right| = -16\sin(\theta_1)^2\cos(\theta_2)^2\sin(\theta_2)^2\cos(\theta_2)^2 < 0$. This completes the proof and shows $Q^T\mathcal{S}(N)_{2,+}Q \neq \mathcal{S}(N)_{2,+}$ for every orthogonal $Q$ which is not a signed permutation matrix. □

**Remark 2.3.4.** *Let $N \geq N'$ and $P$ be $N \times N'$ matrix containing only* 0, 1, *or* −1 *such that each column contains exactly one nonzero element and each row contains at most one nonzero element, then $P^T\mathcal{S}(N)_{2,+}P = \mathcal{S}(N')_{2,+}$ and $P\mathcal{S}(N')_{2,+}P^T \subset \mathcal{S}(N)_{2,+}$.*

## 2.4 Invariant SDDP

Part 2 of lemma 2.3.1 is equivalent to saying that

$$\Sigma_0^N := \left\{[e_i\ e_j]A_0[e_i\ e_j]^T \mid 1 \leq i < j \leq N,\ \ A_0 \in \mathcal{S}(2)_+\right\} \tag{2.23}$$

contains all extreme rays of $\mathcal{S}(N)_{2,+}$. The symmetrized part $\mathcal{S}(N)_{2,+} \cap \mathcal{S}(N)^G$ remains in $\mathcal{S}(N)_{2,+}$ after conjugation by $U$ if and only if its symmetrized extreme rays remain in $\mathcal{S}(N)_{2,+}$, as stated by the following lemma.

**Lemma 2.4.1.** *Let*

$$\Sigma^N := \left\{ A_\sigma \;\middle|\; A \in \Sigma_0^N \right\}. \tag{2.24}$$

*The inclusion* (2.13) *with $\mathcal{K} = \mathcal{S}(N)_{2,+}$ holds if and only if*

$$U^T\Sigma^N U \subset \mathcal{S}(N)_{2,+}. \tag{2.25}$$

*Proof.* The 'only if' part is straightforward since $\Sigma^N \subset \mathcal{S}(N)_{2,+} \cap \mathcal{S}(N)^G$. For the reverse, $\Sigma^N$ generates $\mathcal{S}(N)_{2,+} \cap \mathcal{S}(N)^G$ by the fact that elements of $\mathcal{S}(N)_{2,+} \cap \mathcal{S}(N)^G$ are conic combinations of elements of $\Sigma^N$. Given that the SDD cone is closed under conic combinations, the result follows. □

It is straightforward to show $\mathcal{S}(N)_{2,+} \cap \mathcal{S}(N)^G$ forms a proper cone. Given this property, $\Sigma^N$ indeed contains all extreme rays of $\mathcal{S}(N)_{2,+} \cap \mathcal{S}(N)^G$. To certify $f_U \leq f$ in invariant SDDPs, it is sufficient to check (2.25) rather than (2.13). This verification process can be further simplified when the problem is strongly $G$-invariant. As demonstrated in theorem 2.3.3, $\mathcal{S}(N)_{2,+}$ is preserved under conjugation by $\rho_g$ if and only if $\rho_g$ is a signed permutation matrix. For this reason, we focus on the cases where $\rho$ is a signed permutation representation.

Let $\mathfrak{B} := \{e_1, e_2 \cdots, e_N\}$ be the standard basis of $\mathbb{R}^N$. A signed permutation representation $\rho : G \to O(N)$ is induced by an action of $G$ on a set whose elements are either $e_i$ or $-e_i$ for each $1 \leq i \leq N$. This action partitions that set into a disjoint union of $L$ different orbits. Taking the absolute value of elements of each orbit, each orbit corresponds to a subset $\mathfrak{O}_l$ of $\mathfrak{B}$ such that

$$\mathfrak{B} = \bigsqcup_{l=1}^{L} \mathfrak{O}_l. \tag{2.26}$$

Let

$$\Sigma_0^N(p,q) := \left\{ \begin{bmatrix} e_i & e_j \end{bmatrix} A_0 \begin{bmatrix} e_i^T \\ e_j^T \end{bmatrix} \mid e_i \in \mathfrak{O}_p, e_j \in \mathfrak{O}_q, A_0 \in \mathcal{S}(2)_+ \right\} \tag{2.27}$$

denote the subset of $\Sigma_0^N$ whose elements are matrices with non-zero rows and columns corresponding only to $\mathfrak{O}_p \cup \mathfrak{O}_q$. Defining

$$\Sigma^N(p,q) := \left\{ A_\sigma \mid A \in \Sigma_0^N(p,q) \right\}, \tag{2.28}$$

whose elements are the symmetrized version of the elements of $\Sigma_0^N(p,q)$, the following theorem results.

**Theorem 2.4.2.** *Given a representation $\sigma$ induced by a signed permutation representation, the inclusion* (2.13) *with $\mathcal{K} = \mathcal{S}(N)_{2,+}$ holds if and only if*

$$U^T \Sigma^N(p,q) U \subset \mathcal{S}(N)_{2,+} \quad \forall p,q \in \{1,\cdots,L\} \,:\, p \le q. \tag{2.29}$$

*Proof.* Owing to lemma 2.4.1, it is sufficient to prove that (2.25) holds if and only (2.29) holds. We have

$$\Sigma_0^N = \bigsqcup_{1\le p\le q\le L} \Sigma_0^N(p,q). \tag{2.30}$$

This implies

$$\Sigma^N = \bigsqcup_{1\le p\le q\le L} \Sigma^N(p,q), \tag{2.31}$$

and the result follows. □

As a consequence of theorem 2.4.2, $f_U \le f$ if (2.29) holds. We use this result in chapter 4 to find, in the case of the SDD cone, various symmetry groups $G$ and their representations for which $f_U \le f$ must hold. We close this chapter with an example illustrating the possibility of $f_U \ne f$ in SDDP.

**Example 2.4.3.** Consider example 2.2.3 with $\mathcal{K} = \mathcal{S}(3)_+$. This program is an SDP, yielding $f = 2$ with $a_1 = a_2 = b = 1$. Its symmetry-reduced version (2.11) also yields $f_U = 2$ with $a = b = 1$. Consequently, $f_U = f = 2$.

Now, consider example 2.2.3 with $\mathcal{K} = \mathcal{S}(3)_{2,+}$. This program is an SDDP, resulting in $f = 4$ with $a_1 = a_2 = 2$ and $b = 0$. However, its symmetry-reduced version yields $f_U = 2$ with $a = b = 1$. Thus, in this case, $f_U \ne f$.

# Chapter 3

# Involutive algebra of equivariant maps

Chapter 2 provides sufficient conditions to explore the relationship $f_U \leq f$ between the optima of programs (1.3) and (1.4), particularly when $\mathcal{K} = \mathcal{S}(N)_{2,+}$, for any orthogonal $U$. In chapter 4, we introduce more refined conditions to be capable of being checked computationally specifically tailored to cases where $U$ is a symmetry-adapted basis associated with the fixed-point subspace $\mathcal{S}(N)^G = \mathrm{End}_G(\mathbb{R}^N)^{sa}$. These conditions apply for any orthogonal $U$. This section aims to provide prerequisites of Chapter 4, decomposing the algebra $\mathrm{End}_G(\mathbb{V})$ into simple algebras, where $\mathbb{V}$ is a module representation. Our methodology relies on the Wedderburn–Artin theorem on semisimple algebras. While much of this chapter involves reviewing existing concepts available in noncommutative algebra textbooks such as [Lam, 2001], we introduce minor modifications to adapt these concepts to our specific case. These details are typically not discussed in the context of SDPs, as the precise size of each block primarily affects computational complexity. However, for our purposes, determining block sizes is crucial for investigating the impact of symmetry reduction on invariant SDDPs. The construction of symmetry-adapted bases within the general framework of representation theory is outlined in [Serre, 1977], with a more detailed approach provided in [Heaton et al., 2021]. In the context of $C^*$-algebra of matrices, such construction is outlined in [Polak, 2022, De Klerk and Sotirov, 2010, Bachoc et al., 2012]. However, these works do not address the construction of symmetry-adapted bases in the presence of quaternion-type real representations. [Murota et al., 2010, Maehara and Murota, 2010, Kaluba et al., 2019] offer methods for constructing symmetry-adapted bases in the presence of any type of real representations. Our description of symmetry-adapted bases is fundamentally similar to [Kaluba et al., 2019] but takes a different perspective. We construct each block through block diagonalization, relying on the real *-algebra decomposition of equivariant maps, which lets us analyze the relationship between symmetry-adapted bases associated with $\mathrm{End}_G(\mathbb{V})$ and those associated with $\mathrm{End}_G(\mathbb{S})$ for submodule representation $\mathbb{S}$ of $\mathbb{V}$. This lets us derive, in chapter 4, situations where each block is an SDD matrix.

## 3.1 Representation theory over reals

We recall the basics of real representations, which can be found in standard references such as [Fulton and Harris, 2004, Lang, 2002]. For two $G$-modules $\mathbb{V}_1$ and $\mathbb{V}_2$, a $G$-homomorphism $\psi : \mathbb{V}_1 \to \mathbb{V}_2$ is a linear map such that $g \cdot \psi(v) = \psi(g \cdot v)$. The $G$-modules $\mathbb{V}_1$ and $\mathbb{V}_2$ are $G$-isomorphic if there exists a $G$-isomorphism (bijective $G$-homomorphism) from $\mathbb{V}_1$ onto $\mathbb{V}_2$. Any $G$-invariant subspace $\mathbb{S}$ of $\mathbb{V}$ is a $G$-submodule or a subrepresentation of $\mathbb{V}$. A nontrivial representation whose $G$-submodules are only $\{0\}$ and itself is called an irreducible representation or simple module. For every finite group $G$ and any field, there are a finite number of irreducible representations over the given field, the real or complex numbers in this article. Given a finite group $G$ and its representation $\mathbb{V}$, the non-isomorphic irreducible representations of $G$ over the same field as $\mathbb{V}$ are denoted by $\mathbb{W}_i$, where $1 \leq i \leq m$. Schur's lemma asserts that there does not exist a nontrivial $G$-homomorphism between two non-isomorphic irreducible representations, and therefore $\mathrm{Hom}_G(\mathbb{W}_i, \mathbb{W}_j) = \{0\}$ whenever $i \neq j$. It also reveals that $\mathrm{End}_G(\mathbb{W})$ is a division algebra for an irreducible representation $\mathbb{W}$ [Lang, 2002, Chapter XVII, Proposition 1.1].

If $\mathbb{W}$ is a complex irreducible representation, then $\mathrm{End}_G(\mathbb{W}) \cong \mathbb{C}$. If $\mathbb{W}$ is a real irreducible representation, Frobenius' theorem on real (associative) division algebras implies that $\mathrm{End}_G(\mathbb{W}) \cong \mathbb{D}$, where $\mathbb{D}$ is either $\mathbb{R}$, $\mathbb{C}$, or $\mathbb{H}$, the division ring of quaternions. In the three cases, $\mathbb{W}$ is referred to as a real-type, complex-type, or quaternion-type irreducible representation, respectively.

Let $d := \dim(\mathbb{W})$ and $k := \dim\left(\mathrm{End}_G(\mathbb{W})\right)$ denote the dimensions of $\mathbb{W}$ and $\mathrm{End}_G(\mathbb{W})$, respectively, as an algebra over the ground field of $\mathbb{W}$. In the case of real representations, $k$ is 1, 2, or 4 for real-type, complex-type, or quaternion-type irreducible representation, respectively, and $d$ is a multiple of $k$. We use the same notation when $\mathbb{W}$ is an irreducible representation over complex numbers, with $k$ being always one. We refer to representations that are a direct sum of only one irreducible representation (i.e., $\mathbb{W}^n$) as isotypic representations. In the next section, we decompose the algebra $\mathrm{End}_G(\mathbb{W}^n)$ and extend this to $\mathrm{End}_G(\mathbb{V})$, where $\mathbb{V}$ is any representation.

## 3.2 Orthogonal decomposition of G-equivariant maps: isotypical case

Endowing an irreducible representation $\mathbb{W}$ with a $G$-invariant inner product makes $\mathbb{W}$ orthogonal if it is a real representation or unitary if it is a complex representation. Consequently, the adjoint acts as an involution in $\mathrm{End}_G(\mathbb{W})$, and the isomorphism $\mathrm{End}_G(\mathbb{W}) \cong \mathbb{D}$ preserves involution. For discussion on the correspondence of adjoint and involution in algebras, see [Jacobson, 2009, Chapter V]. For real division algebras with involution, see [Badger, 2006]. The *-algebra isomorphism $\mathrm{End}_G(\mathbb{W}) \cong \mathbb{D}$ induces another *-algebra isomorphism,

$$\mathrm{Mat}_n\left(\mathrm{End}_G(\mathbb{W})\right) \cong \mathrm{Mat}_n(\mathbb{D}), \tag{3.1}$$

where $\mathrm{Mat}_n\big(\mathrm{End}_G(\mathbb{W})\big)$ is the algebra of $n \times n$ matrices whose elements are operators in $\mathrm{End}_G(\mathbb{W})$, and $\mathrm{Mat}_n(\mathbb{D})$ is the algebra of $n \times n$ matrices over the division ring $\mathbb{D}$. When $\mathbb{W}$ is a real-type irreducible representation, $\iota_{\mathbb{W}}$ is defined to be the isomorphism in (3.1). For complex-type irreducible representations, define $\iota_{\mathbb{W}}$ to be the composition of a natural *-algebra embedding $\mathrm{Mat}_n(\mathbb{C}) \hookrightarrow \mathrm{Mat}_{2n}(\mathbb{R})$ with the isomorphism (3.1), where the embedding is

$$A \mapsto \begin{bmatrix} A_R & A_I \\ -A_I & A_R \end{bmatrix} \text{ for } A = A_R + iA_I. \tag{3.2}$$

For quaternion-type irreducible representations, define $\iota_{\mathbb{W}}$ to be the composition of a natural *-algebra embedding $\mathrm{Mat}_n(\mathbb{H}) \hookrightarrow \mathrm{Mat}_{4n}(\mathbb{R})$ with the isomorphism (3.1), where the embedding is

$$A \mapsto \begin{bmatrix} A_R & -A_I & -A_J & -A_K \\ A_I & A_R & -A_K & A_J \\ A_J & A_K & A_R & -A_I \\ A_K & -A_J & A_I & A_R \end{bmatrix} \text{ for } A = A_R + iA_I + jA_J + kA_K. \tag{3.3}$$

In summary,

$$\iota_{\mathbb{W}} : \mathrm{Mat}_n\Big(\mathrm{End}_G(\mathbb{W})\Big) \to \mathrm{Mat}_{kn}(\mathbb{R}). \tag{3.4}$$

is a *-algebra isomorphism if $\mathbb{W}$ is a real-type irreducible representation, and a *-algebra embedding if $\mathbb{W}$ is a complex-type or quaternion-type irreducible representation. There are infinitely many ways to define these embeddings; we have chosen one natural way. Indeed, any conjugation of the image of these embeddings by an orthogonal matrix with appropriate size is a *-algebra embedding.

In the following, we show there exists an *-algebra embedding

$$\mathrm{End}_G(\mathbb{W}^n) \hookrightarrow \mathrm{Mat}_{dn}(\mathbb{R}) \tag{3.5}$$

whose image consists of block diagonal matrices with a particular form. To this end, a *-algebra isomorphism between $\mathrm{End}_G(\mathbb{W}^n)$ and $\mathrm{Mat}_n\big(\mathrm{End}_G(\mathbb{W})\big)$ is first established in the following lemma.

**Lemma 3.2.1.** *Let $\mathbb{W}$ be an orthogonal or unitary irreducible representation of $G$, let $\pi_i$ be the natural projection of $\mathbb{W}^n$ onto its $i^{th}$ copy of $\mathbb{W}$, and let $\eta_i$ be the natural injection from the $i^{th}$ copy of $\mathbb{W}$ into $\mathbb{W}^n$. Then*

$$\theta : End_G(\mathbb{W}^n) \cong Mat_n\big(End_G(\mathbb{W})\big), \tag{3.6}$$

*defined by $\theta : \tau \mapsto (\tau_{ij}) := (\pi_i \tau \eta_j)$ is an *-algebra isomorphism and so preserves self-adjoitness and positive semidefiniteness.*

*Proof.* [Passman, 2004, Lemma 4.2] shows $\theta$ is a ring isomorphism. From this, it is straightforward to show that $\theta$ is an algebra isomorphism. To show $\theta$ preserves involution, note that $\eta_i$ is the Hilbert space adjoint of $\pi_i$ for each $i$. Hence,

$$\theta(\tau)^* = (\tau_{ij})^* = (\tau_{ji}^*) = ((\pi_j \tau \pi_i^*)^*) = (\pi_i \tau^* \pi_j^*) = \theta(\tau^*). \tag{3.7}$$

This implies $\theta$ preserves self-adjoitness. In addition, if $\tau$ is positive semidefinite, then $\tau = \tau_0^* \tau_0$ for some $\tau_0$, and so $\theta(\tau) = \theta(\tau_0)^* \theta(\tau_0)$ is positive semidefinite. □

As a result of Lemma 3.2.1,

$$\iota_{\mathbb{W}} \circ \theta : \mathrm{End}_G(\mathbb{W}^n) \to \mathrm{Mat}_{kn}(\mathbb{R}) \tag{3.8}$$

is a *-algebra isomorphism when $\mathbb{W}$ is a real-type and a *-algebra embedding when $\mathbb{W}$ is a complex or quaternion-type. For $\tau \in \mathrm{End}_G(\mathbb{W}^n)$, $\iota_{\mathbb{W}} \circ \theta(\tau)$ is a $kn \times kn$ matrix. The Kronecker product of this matrix with $I_{d/k}$, the identity matrix of integer size $d/k$, results in a $dn \times dn$ matrix. This construction provides the desired *-algebra embedding in (3.5), as summarized in the following lemma.

**Lemma 3.2.2.** *The map*

$$\tilde{\theta}_{\mathbb{W}} : \tau \mapsto \iota_{\mathbb{W}} \circ \theta(\tau) \otimes I_{d/k} \tag{3.9}$$

*is a *-algebra embedding* (3.5) *whose image consists of block diagonal matrices with $d/k$ identical square blocks of size $kn$.*

A basis $\epsilon$ for a representation $\mathbb{W}^n$ is called symmetry-adapted if, for every $\tau \in \mathrm{End}_G(\mathbb{W}^n)$, the matrix $[\tau]_\epsilon$ has the form $A \otimes I_{d/k}$ for some $kd \times kd$ matrix $A$. When $\mathbb{W}^n$ is identified with $\mathbb{R}^{dn}$, a matrix $U$ is called a symmetry-adapted basis if, for every $A \in \mathrm{End}_G(\mathbb{R}^{dn})$, the conjugate $U^{-1}AU$ has the form $A_0 \otimes I_{d/k}$ for some $kd \times kd$ matrix $A_0$. Context will clarify whether a symmetry-adapted basis refers to this matrix or the basis itself. For an orthogonal or unitary representation $(\mathbb{W}^n, \rho)$, by $\mathfrak{E}(\mathbb{W}^n)$ or $\mathfrak{E}(\rho)$, we denote the set of orthonormal bases themselves, and by $\mathfrak{U}(\mathbb{W}^n)$ or $\mathfrak{U}(\rho)$, we denote the set of orthogonal symmetry-adapted bases in the sense of matrices. The next lemma illustrates the relation between symmetry-adapted bases and the embedding (3.5).

**Lemma 3.2.3.** *Let $\mathbb{W}$ be an orthogonal irreducible representation, and let $n$ be a nonzero integer. There is a one-to-one correspondence between each *-algebra embedding* (3.5) *and each symmetry-adapted basis (in $\mathfrak{E}(\mathbb{W}^n)$ or $\mathfrak{U}(\mathbb{W}^n)$).*

*Proof.* The algebra $\mathrm{End}_G(\mathbb{W}^n)$ is isomorphic to a simple subalgebra of the central simple $\mathbb{R}$-algebra $\mathrm{Mat}_{dn}(\mathbb{R})$. Due to the Skolem-Noether theorem, the $\mathbb{R}$-embedding (3.5) extends to an inner automorphism of $\mathrm{Mat}_{dn}(\mathbb{R})$. Since (3.5) is a *-algebra embedding, the inner automorphism should preserve involution, which corresponds to transpose in the real case and conjugate transpose in the complex case. Hence, the automorphism is of the form $U^T(\cdot)U$ for an orthogonal symmetry-adapted basis. Conversely, identifying $\mathbb{W}^n$ with $\mathbb{R}^{dn}$, for every $U \in \mathfrak{U}(\mathbb{W}^n)$, $U^T(\cdot)U$ corresponds to an *-algebra embedding

$$\mathrm{End}_G\left(\mathbb{R}^{dn}\right) \hookrightarrow \mathrm{Mat}_{dn}(\mathbb{R}). \tag{3.10}$$

□

The Skolem-Noether theorem is a standard result in central simple algebras–e.g., [Knus et al., 1998, Theorem 1.4]–but we are not aware of its previous application in the context of symmetry-adapted bases. For discussion of simple algebras, we refer to [Jacobson, 2009].

## 3.3 Orthogonal decomposition of G-equivariant maps: general case

Lemma 3.2.2 and lemma 3.2.3, whose proofs assume isotopic representations, can be extended to all representations. By Maschke's theorem, every real or complex representation $\mathbb{V}$ is the direct sum of isotypic representations, meaning

$$\mathbb{V} = \bigoplus_{i=1}^{m} \mathbb{W}_i^{n_i} \tag{3.11}$$

Moreover, if $\mathbb{V}$ is an orthogonal (resp. unitary) representation, $\mathbb{W}_i$ can be chosen to be orthogonal (resp. unitary) irreducible representations.

**Lemma 3.3.1.** *Let $\chi_i$ be the character of $\mathbb{W}_i$, and let*

$$\Pi_i := \frac{d_i}{k_i|G|} \sum_{g \in G} \chi_i(g^{-1})\rho_g. \tag{3.12}$$

1. *$\sum_i \Pi_i$ is a resolution of the identity associated with the decomposition* (3.11).

2. *If $\rho$ is an orthogonal representation, $\sum_i \Pi_i$ will be an orthogonal resolution of the identity.*

*Proof.* In the case of complex representations (where $k_i = 1$ for each $i$), this argument is a standard result of character theory and can be found in representation theory books; including [Serre, 1977, Fulton and Harris, 2004]. Assume then that $\mathbb{V}$ is a real representation. If $\mathbb{W}_i$ is real-type, then it is also realizable as an irreducible representation over complex numbers with the same character, so $\Pi_i$ is the projection from $\mathbb{V}$ onto the isotypic component $\mathbb{W}_i^{n_i}$. If $\mathbb{W}_i$ is complex-type, then there are nonisomoprhic irreducible representations $\mathbb{W}_i^0$ and $(\mathbb{W}_i^0)^*$ over $\mathbb{C}$ with dimension $d_i^0$ and character $\chi_i^0$ and $\bar{\chi}_i^0$ such that $\mathbb{W}_i \otimes \mathbb{C} = \mathbb{W}_i^0 \oplus (\mathbb{W}_i^0)^*$, $\chi_i = \chi_i^0 + \bar{\chi}_i^0$, and $d_i^0 = \frac{d_i}{2}$. Hence,

$$\Pi_i = \frac{d_i^0}{|G|} \sum_{g \in G} \bar{\chi}_i^0(g^{-1})\rho_g + \frac{d_i^0}{|G|} \sum_{g \in G} \chi_i^0(g^{-1})\rho_g = \frac{d_i^0}{|G|} \sum_{g \in G} (\bar{\chi}_i^0 + \chi_i^0)(g^{-1})\rho_g = \frac{d_i}{2|G|} \sum_{g \in G} \chi_i(g^{-1})\rho_g. \tag{3.13}$$

If $\mathbb{W}_i$ is quaternion-type, then there is irreducible representation $\mathbb{W}_i^0$ over $\mathbb{C}$ with dimension $d_i^0$ and character $\chi_i^0$ such that $\mathbb{W}_i \otimes \mathbb{C} \cong \mathbb{W}_i^0 \oplus \mathbb{W}_i^0$, $\chi_i = 2\chi_i^0$, and $d_i^0 = \frac{d_i}{2}$. In this case, the projection onto $\mathbb{W}_i$ and $\mathbb{W}_i^0$ is essentially the same, and $\chi_i^0$ is real-valued. Consequently,

$$\Pi_i = \frac{d_i^0}{|G|} \sum_{g \in G} \chi_i^0(g^{-1})\rho_g = \frac{d_i}{4|G|} \sum_{g \in G} \chi_i(g^{-1})\rho_g. \tag{3.14}$$

□

Lemma 3.3.1 can be extended to the algebra of equivariant maps by the universal property of direct sum. (3.5) results in

$$\mathrm{End}_G(\mathbb{V}) = \mathrm{End}_G\Big(\bigoplus_{i=1}^{m} \mathbb{W}_i^{n_i}\Big) = \mathrm{Hom}_G\left(\bigoplus_{i=1}^{m} \mathbb{W}_i^{n_i}, \bigoplus_{j=1}^{m} \mathbb{W}_j^{n_j}\right) \tag{3.15}$$

$$\cong \bigoplus_{i,j} \mathrm{Hom}_G(\mathbb{W}_i^{n_i}, \mathbb{W}_j^{n_j}). \tag{3.16}$$

Additionally, by Schur's lemma, there is no nontrivial $G$-homomorphism between non isomorphic simple $G$-modules. Extending it to isotypic representations, $\mathrm{Hom}_G(\mathbb{W}_i^{n_i}, \mathbb{W}_j^{n_j})$ is trivial when $i \neq j$. This results in the following *-algebra isomorphism:

$$\mathrm{End}_G(\mathbb{V}) \cong \bigoplus_{i=1}^{m} \mathrm{End}_G(\mathbb{W}_i^{n_i}). \tag{3.17}$$

From this point, it is straightforward to show $\tau \mapsto \Pi_i \tau \Pi_i^*$ is the projection of $\mathrm{End}_G(\mathbb{V})$ onto $\mathrm{End}_G(\mathbb{W}_i^{n_i})$ associated with (3.17). Finally, combining this with lemma 3.2.2 yields the extension of lemma 3.2.2 to the general case.

**Lemma 3.3.2.** *There exists the *-algebra embedding*

$$End_G(\mathbb{V}) \hookrightarrow Mat_N(\mathbb{R}), \quad \tau \mapsto \bigoplus_{i=1}^{m} A_i \otimes I_{d_i/k_i}, \tag{3.18}$$

*where* $A_i = \iota_{\mathbb{W}_i} \circ \theta(\Pi_i \tau \Pi_i^*)$ *is a* $k_i n_i \times k_i n_i$ *matrix, and* $N$, $d_i$, *and* $k_i$ *are the dimension of* $\mathbb{V}$, $\mathbb{W}_i$, *and* $End_G(\mathbb{W}_i)$, *respectively. Additionally, If all* $\mathbb{W}_i$ *are real-type, then the embedding is an isomorphism.*

A basis $\epsilon$ for a representation $\mathbb{V}$ is called symmetry-adapted if, for every $\tau \in \mathrm{End}_G(\mathbb{V})$, the matrix $[\tau]_\epsilon$ has the form of the image of (3.18) for some $k_i n_i \times k_i n_i$ matrices $A_i$. When $\mathbb{V}$ is identified with $\mathbb{R}^N$, a matrix $U$ is called a symmetry-adapted basis if, for every $A \in$

$\mathrm{End}_G(\mathbb{R}^N)$, the conjugate $U^{-1}AU$ has the form of the image of (3.18) for some $k_i n_i \times k_i n_i$ matrices $A_i$. The following lemma is the extension of lemma 3.2.3.

**Lemma 3.3.3.** *Each orthonormal symmetry-adapted basis for $\mathrm{End}_G(\mathbb{V})$ corresponds to a *-algebra embedding in* (3.18). *Moreover,*

$$\mathfrak{E}(\mathbb{V}) = \left\{ \bigcup_{i=1}^{m} \epsilon_i \;\middle|\; \epsilon_i \in \mathfrak{E}(\mathbb{W}_i^{n_i}) \right\}. \tag{3.19}$$

The proof of this lemma is similar to that of lemma 3.2.3, so we omit it. We close this section with two corollaries. Corollary 3.3.4, which is needed for our results in chapter 4, states the implication of lemma 3.3.2 and lemma 3.3.3 for SDD matrices. Corollary 3.3.5 follows directly from Lemma lemma 3.3.3 and shows an orthonormal symmetry-adapted basis associated with the $G$-equivariant maps of an orthogonal (resp. unitary) subrepresentation of $\mathbb{V}$ can be constructed with any orthonormal symmetry-adapted basis associated with $\mathrm{End}_G(\mathbb{V})$.

**Corollary 3.3.4.** *Let $\mathbb{V}$ be an orthogonal representation. Given* (3.11) *where $k_i n_i \leq 2$, for any symmetry-adapted basis $\epsilon \in \mathfrak{U}(\mathbb{V})$ and any positive semidefinite operator $\tau \in End_G(\mathbb{V})$, $[\tau]_\epsilon$ is an SDD matrix.*

*Proof.* By lemma 3.3.2 and lemma 3.3.3, any symmetry-adapted basis maps $\tau$ to a matrix with blocks of size $k_i n_i$. Since the embedding (3.18) preserves involution, any PSD operator $\tau$ is mapped to a PSD matrix, with all the blocks being PSD. Since the size of each block is smaller than 2, all the blocks are also SDD by remark 2.3.2 and so is the whole matrix. □

**Corollary 3.3.5.** *Let $\mathbb{S}$ be an orthogonal (resp. unitary) subrepresentation of an orthogonal (resp. unitary) representation $\mathbb{V}$, and define*

$$End_G(\mathbb{V}|\mathbb{S}) := \left\{ \tau \in End_G(\mathbb{V}) \;\middle|\; \tau|_{\mathbb{S}^\perp} = 0 \right\}. \tag{3.20}$$

*Then, every symmetry adapted-basis associated with* $End_G(\mathbb{S})$ *extends to one associated with* $End_G(\mathbb{V}|\mathbb{S})$.

## 3.4 Construction of blocks

This section aims to construct the block matrices $A_i$ in (3.18) for the orthogonal representation $(\mathbb{R}^N, \rho)$. In the previous section, $\Pi_i$ is shown to be the orthogonal projection of $\mathbb{V}$ onto $\mathbb{W}_i^{n_i}$ associated with the decomposition (3.11). In this case where $\mathbb{V} = \mathbb{R}^N$, $\Pi_i$ is a matrix with rank $d_i n_i$. To construct $A_i$, we require an orthogonal projection from $\mathbb{R}^N$ onto a subspace of dimension $k_i n_i$. If $d_i = k_i$, then $\Pi_i$ does the job since then the rank of $\Pi_i$ will be $d_i n_i = k_i n_i$. Let $U_i$ be the $N \times k_i n_i$ matrix whose columns are formed by any orthonormal collection of columns of $p_i$; for brevity, we denote this with $U_i := \text{orth}(\Pi_i)$. Then, $A_i$ corresponding to the image of $A \in \text{End}_G(\mathbb{R}^N)$ under (3.18) can be constructed via $A_i = U_i^T A U_i$.

Groups with quaternion-type representations $W_i$ whose representations can have $d_i \geq k_i$ are quite rare; the smallest group with this property has a degree 48 and is called binary Octahedral group or conformal special linear group on the field $\mathbb{F}_3^2$. However, this is not the case when $\mathbb{W}_i$ is real-type. The next lemma presents an orthogonal projection from $\mathbb{R}^N$ onto a subspace of dimension $k_i n_i$ for real-type irreducible representations.

**Lemma 3.4.1.** *Let* $R_i(g)$ *be any orthogonal realization of* $\mathbb{W}_i$ *whose* $(\alpha, \beta)$ *element is denoted by* $R_i(g)(\alpha, \beta)$. *Here,* $R_i(g)$ *are* $d_i \times d_i$ *real matrices. Let*

$$\Pi_i(\alpha, \beta) := \frac{d_i}{k_i|G|} \sum_{g \in G} R_i(g^{-1})(\alpha, \beta)\rho_g. \tag{3.21}$$

*If* $k_i = 1$, *then the maps* $k_i \Pi_i(\alpha, \alpha)$ *are mutally orthogonal projections from* $\mathbb{V}$ *onto its subspace isomorphic to* $Hom_G(\mathbb{W}_i, \mathbb{W}_i^{n_i})$ *with dimension* $k_i n_i$.

The proof of this lemma (when each $k_i = 1$) is based on the isomorphism $\mathbb{W}_i^{n_i} \cong \mathbb{W}_i \otimes \mathrm{Hom}_G(\mathbb{W}_i, \mathbb{W}_i^{n_i})$ and can be found in [Serre, 1977]. The author conjectures that this lemma could also be true when $k_i = 2$ and $k_i = 4$ by establishing an isomorphism $\mathbb{W}_i^{n_i} \otimes \mathrm{End}_G(\mathbb{W}_i) \cong \mathbb{W}_i \otimes \mathrm{Hom}_G(\mathbb{W}_i, \mathbb{W}_i^{n_i})$.

**Remark 3.4.2.** *Let $\boldsymbol{U}_i^{\alpha} = orth(\Pi_i(\alpha, \alpha))$ be $N \times k_i n_i$ matrix whose columns are formed by any orthonormal collection of columns of $\Pi_i(\alpha, \alpha)$. Then, When $k_i = 1$, $A_i$ in lemma 3.3.2 can be obtained via $A_i = (\boldsymbol{U}_i^{\alpha})^T A \boldsymbol{U}_i^{\alpha}$ for any $\alpha$.*

# Chapter 4

# Characterization of invariant SDDPs

In this chapter, we identify some symmetry groups and representations that ensure either $f_U = f$ or $f_U \leq f$ for some symmetry-adapted bases $U$ in $G$-invariant SDDPs, where the underlying representation is induced via (2.7). We focus on signed-permutation representations since these are the representations that keep the SDD cone invariant, according to theorem 2.3.3. In section 4.1, we determine groups and representations that guarantee the existence of $U \in \mathfrak{U}(\rho)$ such that $f_U = f$. To achieve this, we determine for which groups and representations there exists a symmetry-adapted basis $U$ satisfying (2.12) for $\mathcal{K} = \mathcal{S}(N)_{2,+}$. In sections 4.2 and 4.3, we identify groups such that $f_U \leq f$ for any symmetry-adapted basis $U$ independent of the $G$-representation $\rho$. We also identify groups such that, for each of its permutation representations, there exists at least a symmetry-adapted basis $U$ guaranteeing $f_U \leq f$. For a group $G$ that is not in this category, we derive algebraic conditions that identify those $G$-representations $\rho$ that guarantee $f_U \leq f$ for any symmetry-adapted basis. We rely on the algebraic conditions that guarantee (2.13) for a symmetry-adapted basis $U$. Since the SDD elements of $\mathrm{End}_G(\mathbb{R}^N)$ are self-adjoint, (2.13) is equivalent to

$$U^T \left(\mathcal{S}(N)_{2,+} \cap \mathrm{End}_G(\mathbb{R}^N)\right) U \subset \mathcal{S}(N)_{2,+}. \qquad (2.13')$$

We interchange the inclusion (2.13') and (2.13) as needed. We denote the set of matrices that satisfy these inclusions for representation $\rho$ with size $N$ as $\mathfrak{A}(\rho, N)$. While $N$ is the size of $\rho$, we do not omit it for clarity.

## 4.1 SDD-invariant symmetry-adapted bases

Although $\mathcal{S}(N)_{2,+}$ is preserved under conjugation by signed permutation representations, it is not necessarily preserved under conjugation by the symmetry-adapted bases $U$ associated with these representations, as $U$ is not always a signed permutation matrix. The following lemma characterizes the representations whose associated symmetry-adapted bases do preserve $\mathcal{S}(N)_{2,+}$.

**Theorem 4.1.1.** *Given a $G$-invariant SDDP, there exists $U \in \mathfrak{U}(\rho)$ satisfying* (2.12) *with $\mathcal{K} = \mathcal{S}(N)_{2,+}$ if and only if $\rho$ is a signed diagonal representation (without any permutation).*

*Proof.* Assume $\rho : G \to O(N)$ is a signed diagonal representation, representing a cyclic group of order two, $G = \{e, g\}$, where $\rho_e$ is $N \times N$ identity matrix and $\rho_g$ is an $N \times N$ diagonal matrix whose diagonal entries are $\pm 1$. Columns of any symmetry-adapted basis are formed from any orthonormal collection of columns of $\Pi_1 = \frac{1}{2}(\rho_e + \rho_g)$ and that of $\Pi_2 = \frac{1}{2}(\rho_e - \rho_g)$ by remark 3.4.2. Since these matrices are diagonal, any orthogonal symmetry-adapted basis is a signed permutation matrix. Therefore, they preserve $\mathcal{S}(N)_{2,+}$ by Theorem 2.3.3.

For the reverse direction, assume $\rho$ involves permutation. It is sufficient to show that any symmetry-adapted basis $U$ has at least one column with more than one nonzero element because then Theorem 2.3.3 would imply $U^T \mathcal{S}(N)_{2,+} U \neq \mathcal{S}(N)_{2,+}$ for $N \geq 3$. Let $\mathbb{W}_1$ denote the trivial representation of $G$ that occurs $n_1 > 0$ times in $\mathbb{R}^N$. Some columns of $U$ are constructed by choosing an orthonormal basis for $W_1^{n_1}$ in an irreducible-representation decomposition of $\mathbb{R}^N$. These bases are constructed by choosing any collection of independent columns of

$$\frac{1}{|G|}\sum_{g\in G}\chi_1(g^{-1})\rho_g = \frac{1}{|G|}\sum_{g\in G}\rho_g, \tag{4.1}$$

where $\chi_1$ denotes the character of the trivial representation. Since not all $\rho_g$ are diagonal matrices, Matrix 4.1 contains at least one column with more than one nonzero element and so $U$ as well. □

The proof of this theorem illustrates that the relation (2.12) holds for any $U \in \mathfrak{U}_\rho$ if $\rho$ is a signed diagonal representation. This yields the following corollary.

**Corollary 4.1.2.** *When $\rho$ is a signed diagonal representation, $f_U = f$ holds for all orthogonal symmetry-adapted bases.*

## 4.2 Group characterization

Theorem 4.1.1 asserts that if $\rho$ is not a signed diagonal representation, then $f_U = f$ is not guaranteed. This implies that when the underlying representation involves permutation, the symmetry-adapted bases are no longer signed permutation matrices. Therefore, we aim to identify situations that guarantee $f_U \leq f$ in the presence of permutation. We focus on (unsigned) permutation representations. The existence of an orthogonal symmetry-adapted basis that satisfies (2.13) is equivalent to saying that $\mathfrak{A}(\rho, N) \cap \mathfrak{U}(\rho)$ is non-empty. On the other hand, all $U \in \mathfrak{U}(\rho)$ satisfying (2.13) is equivalent to $\mathfrak{U}(\rho) \subset \mathfrak{A}(\rho, N)$. We categorize representations into three types: fully adaptable, partially adaptable, and unadaptable, all in relation to the SDD cone. A representation $\rho$ is called fully adaptable if $\mathfrak{U}(\rho) \subset \mathfrak{A}(\rho, N)$, unadaptable if $\mathfrak{U}(\rho) \cap \mathfrak{A}(\rho, N)$ is empty set, and partially adaptable otherwise. To make use of theorem 2.4.2, we state its stronger version in the context of equivariant maps as theorem 4.2.2 below. First, the following lemma states the relationship among the symmetry-adapted bases associated with an algebra and those associated with its subalgebras.

**Lemma 4.2.1.** *Let* $(\mathbb{R}^{N'}, \rho')$ *be an orthogonal subrepresentation of* $(\mathbb{R}^{N}, \rho)$.

1. *For every* $U' \in \mathfrak{U}(\mathbb{R}^{N'})$ *that satisfies (resp. does not satisfy)* (2.13) *for* $N'$ *and* $\rho'$*, there exists a corresponding* $U \in \mathfrak{U}(\mathbb{R}^{N})$ *that satisfies (resp. does not satisfy)*

$$U^T \left(\mathcal{S}(N)_{2,+} \cap End_G(\mathbb{R}^N \,|\, \mathbb{R}^{N'})\right) U \subset \mathcal{S}(N)_{2,+}. \tag{4.2}$$

2. *If* $\rho$ *is partially (resp. fully) adaptable, then* $\rho'$ *is also partially (resp. fully) adaptable.*

*Proof.* Let $P'$ be an $N' \times N$ matrix whose columns are formed by elements of $\mathcal{B}$ that span $\mathbb{R}^{N'}$. Then, $P' : \mathbb{R}^N \to \mathbb{R}^{N'}$ is a $G$-equivariant epimorphism, implying that $P'AP'^T \in \mathrm{End}_G(\mathbb{R}^{N'})$ for every $A \in \mathrm{End}_G(\mathbb{R}^N)$. Additionally, $P'\mathcal{S}(N)_{2,+}P'^T = \mathcal{S}(N')_{2,+}$ by remark 2.3.4.

For the proof of the first part, the map $U' \mapsto P^T U' P$ sends the elements of $\mathfrak{U}(\mathbb{R}^{N'})$ satisfying (resp. not satisfying) (2.13) for $N'$ and $\rho'$ to the elements of $\mathfrak{U}(\mathbb{R}^{N})$ satisfying (resp. not satisfying) (4.2). This is a stronger version of corollary 3.3.5.

For the proof of the second part, let $\rho$ be partially adaptable, and suppose $U \in \mathfrak{U}(\rho) \cap \mathfrak{A}(\rho, N)$. For every $A' \in \mathcal{S}(N')_{2,+} \cap \mathcal{S}(N')^G$, $P^T A' P \in \mathcal{S}(N)_{2,+} \cap \mathcal{S}(N)^G$, and so

$$U^T P^T A' P U \subset \mathcal{S}(N)_{2,+} \Rightarrow P U^T P^T A' P U P^T \subset \mathcal{S}(N')_{2,+}. \tag{4.3}$$

Consequently, $PUP^T \in \mathfrak{U}(\rho') \cap \mathfrak{A}(\rho', N')$ and $\rho'$ is partially adaptable. Now, let $\rho$ be fully adaptable. Every $U' \in \mathfrak{U}(\rho')$ satisfies (2.13) because otherwise it contradicts the fact that $\mathfrak{U}(\rho) \subset \mathfrak{A}(\rho, N)$. Consequently, $\rho'$ is fully adaptable. □

We would like to conclude if $\rho'$ and $\rho''$ were partially (resp. fully) adaptable, then $\rho' \oplus \rho''$ would be also partially (resp. fully) adaptable. Although this is not generally true, we identify situations that guarantee the statement. Let $(\mathbb{R}^{\mathbb{N}}, \rho)$ be a permutation representation. As stated in chapter 2, this representation is induced by an action of $G$ on $\mathfrak{B}$, resulting in $L$ disjoint orbits $\mathfrak{O}_j$. Defining $\mathbb{S}_l := \text{span}(\mathfrak{O}_l)$, the next theorem is followed.

**Theorem 4.2.2.** *A permutation representation $\mathbb{R}^N$ is partially (resp. fully) adaptable if and only if each $\mathbb{S}_p \oplus \mathbb{S}_q$ is partially (resp. fully) adaptable.*

*Proof.* This theorem is another version of theorem 2.4.2. The $\Rightarrow$ direction is a direct consequence of Part 2 of lemma 4.2.1. For the reverse direction assume each $\mathbb{S}_p \oplus \mathbb{S}_q$ is partially (resp. fully) adaptable, then, by Part 1 of lemma 4.2.1, for some (resp. all) $U \in \mathfrak{U}(\mathbb{R}^N)$,

$$U^T \left( \mathcal{S}(N)_{2,+} \cap \text{End}_G(\mathbb{R}^N | S_p \oplus S_q) \right) U \subset \mathcal{S}(N)_{2,+} \tag{4.4}$$

holds. This is exactly equivalent to say, for some (resp. all) $U \in \mathfrak{U}(\mathbb{R}^N)$, (2.29) holds. Finally, theorem 2.4.2 implies for some (resp. all) $U \in \mathfrak{U}(\mathbb{R}^N)$, (2.13) holds, or equivalently $\mathfrak{U}(\mathbb{R}^N) \cap \mathfrak{A}(\rho, N) \neq \emptyset$ (resp. $\mathfrak{U}(\mathbb{R}^N) \subset \mathfrak{A}(\rho, N)$). □

This theorem is interesting because it involves algebra while fundamentally utilizing a set containing the extreme rays of the SDD cone, which is a geometric object. It is also useful because, if $\mathbb{R}^N$ is fully adaptable, one can freely choose a symmetry-adapted basis and ensure $f_U \leq f$. If $\mathbb{R}^N$ is partially adaptable, one can find a symmetry-adapted basis that guarantees $f_U \leq f$. We simplify this theorem so that it can be applied to determine full adaptability using the orbital structure of $\rho$.

**Definition 4.2.3** (Mono semisimple representation)**.** We call a representation $(\mathbb{V}, \rho)$ a mono semisimple (MSS) representation if each irreducible representation $\mathbb{W}_i$ occurs at most once in the decomposition of $\mathbb{V}$ into the direct sum of isotypic components. In another world, $\mathbb{V}$ is an MSS representation if $n_i$ are either zero or one in (3.11).

This definition helps us to state the simplified version of theorem 4.2.2 more concisely.

**Theorem 4.2.4.** *Let $\mathbb{R}^N$ be an orthogonal representation.*

1. *Suppose $\mathbb{R}^N$ is absolutely irreducible (i.e only real-type $\mathbb{W}_i$ appears in $\mathbb{R}^N$). If each $\mathbb{S}_l$ $(1 \leq l \leq L)$ is MSS, then $\mathbb{R}^N$ is fully adaptable.*

2. *Suppose only real-type and complex-type $W_i$ appear in $\mathbb{R}^N$. If each $\mathbb{S}_l$ $(1 \leq l \leq L)$ is MSS and each complex-type $\mathbb{W}_i$ appears once in $\mathbb{R}^N$, then $\mathbb{R}^N$ is fully adaptable.*

*Proof.* For the proof of the first part, each $\mathbb{S}_l$ is MSS, implying

$$\mathbb{S}_p \oplus \mathbb{S}_q \cong \bigoplus_i \mathbb{W}_i^{n_i}, \tag{4.5}$$

where $n_i \leq 2$. Since each $\mathbb{W}_i$ is real-type, $k_i = 1$ and thus $k_i n_i \leq 2$. Corollary 3.3.4 implies that $\mathbb{S}_p \oplus \mathbb{S}_q$ is fully adaptable, and by theorem 4.2.2, so is $\mathbb{R}^N$. The second part follows a similar proof. In each $\mathbb{S}_p \oplus \mathbb{S}_q$, a real-type $\mathbb{W}_i$ appears at most twice and a complex-type $\mathbb{W}_i$ appears at most once. In either case, $k_i n_i \leq 2$, and the result follows from corollary 3.3.4. □

We conjecture the following stronger version of theorem 4.2.4. [Boman et al., 2005] showed that the cone of SDD matrices is equivalent to the cone of symmetric factor-width-two matrices. The sparsity of a matrix has a close connection with the possibility of having a smaller factor width. In much the same way a change of basis can make a sparse matrix dense, we conjecture that there exists a symmetry-adapted basis that increases the factor width of sparse SDD matrices.

**Conjecture 4.2.5.** *A permutation $G$-representation $\mathbb{R}^N$ is fully adaptable if and only if each real-type irreducible representation appears at most twice, and each complex-type irreducible representation appears at most once, in every $\mathbb{S}_p \bigoplus \mathbb{S}_q$.*

This conjecture implies that fully adaptable representations would not contain any quaternion-type irreducible representation. Indeed, if some $\mathbb{S}_p \bigoplus \mathbb{S}_q$ contains $n_i$ times $W_i$ such that

$k_i n_i \geq 3$, then there would exist $U \in \mathfrak{U}(\mathbb{S}_p \bigoplus \mathbb{S}_q)$ such that $U^T AU$ would not be SDD for some SDD matrix $A$ in $\mathrm{End}_G(\mathbb{S}_p \bigoplus \mathbb{S}_q)$.

One can check whether each $\mathbb{S}_l$ is MSS. The number of times each $\mathbb{W}_i$ appears in $\mathbb{S}_l$ is calculated based on the character of $\mathbb{S}_l$ and $\mathbb{W}_i$; see [Serre, 1977, Fulton and Harris, 2004] for details. It is also possible to find a stronger condition guaranteeing that each $\mathbb{S}_l$ is MSS based on the orbital structure of $\rho$. This requires establishing a relation between each $\mathbb{S}_l$ and the real regular representation $r_G : G \to O(|G|)$ of $G$, which the following lemma provides.

**Lemma 4.2.6.** *Each $\mathbb{S}_l$ is isomorphic to a subrepresentation of the real regular representation $r_G$.*

*Proof.* Since $G$ acts on $\mathfrak{O}_l$ transitively, there is an intertwining from $G$ onto $\mathfrak{O}_l$. This makes $\mathfrak{O}_l$ a homomorphic image of $G$ as a $G$-set. Consequently, $\mathbb{S}_l$ is a homomorphic image of the regular representation, and the result follows. □

The next lemma identifies a wide range of fully adaptable representations for small groups and a narrow range for large groups, based on their orbital structure. This lemma is independent of the size of the representation, rather it depends on the size of the group.

**Lemma 4.2.7.** *If $|\mathfrak{O}_l| = \dim(S_l) < 4$, then $\mathbb{S}_l$ is MSS.*

*Proof.* Let each $G$-irreducible representation $\mathbb{W}_i$ appear $n_i$ times in the regular representation of $G$. If the ground field is the complex numbers, then $n_i = d_i$ for each $i$. If the ground field is the real numbers, then using the proof of lemma 3.3.1 it is straightforward to show that $n_i$ is either $d_i$, $\frac{d_i}{2}$, or $\frac{d_i}{4}$. Here, $\mathbb{S}_l$ is a subrepresentation of irreducible representation by lemma 4.2.6. Therefore, the number of times that $\mathbb{W}_i$ appears in $\mathbb{S}_l$ is at most equal to the dimension of $\mathbb{W}_i$. Since $\dim(S_l) < 4$, no irreducible representations appear more than once in $\mathbb{S}_l$. □

This lemma can be used along with the fact that $|\mathfrak{O}_l|$ divides $|G|$ (by the orbit-stabilizer theorem) to identify fully adaptable representations. This is illustrated by the following example.

**Example 4.2.8.** Let $G = \mathfrak{S}_3$ be the symmetric group of order 3, and let $(\mathbb{R}^N, \rho)$ be any permutation representation of $G$. By the orbit-stabilizer theorem, $|\mathfrak{O}_l|$ is either 1, 2, 3, or 6. If $\rho$ does not contain any orbit of size 6, then by lemma 4.2.7 all the $\mathbb{S}_l$ are MMS. Additionally, irreducible representations of $G$ are all real-type. Therefore, $\mathbb{R}^N$ is fully adaptable if $\rho$ does not contain any orbit of size 6, by Part 1 of theorem 4.2.4. This asserts that, if an invariant SDDP with any size has the symmetry of group $\mathfrak{S}_3$, then $f_U \leq f$ for any symmetry-adapted basis, provided that $\rho$ does not contain any orbit of size 6. Additionally, if $\rho$ contains one orbit of size 6, $\mathbb{S}_p \oplus \mathbb{S}_q$ can be shown to be fully adaptable. We also claim that, for cases where $\rho$ contains more than one orbit of size 6, we can construct a counterexample demonstrating no symmetry-adapted basis with the desired property exists. This means any permutation representation of $\mathfrak{S}_3$ with more than two orbits of size 6 is unadaptable; otherwise, they are fully adaptable.

While it is possible to find criteria guaranteeing full adaptability based on the size of $\rho$, we are not interested in such criteria since it would not be useful. This is because conic optimization problems are often very large in practice. Instead, our results apply to practical problems, which is a significant advantage of our results. The conclusion of example 4.2.8 is independent of the size of the representation. It depends solely on the size of each orbit, which in turn depends on the size of the group. The drawback of this criteria, however, is that it does not identify any partially adaptable $\rho$. To deal with this restriction, we present how one can check the full and partial adaptability of any permutation representation of $G$ by analyzing only one special representation. This new criterion depends not on the representation but merely on the group itself. Therefore, we classify groups into representation-independent and representation-dependent. If some of the permutation representations of $G$ are partially adaptable, some fully adaptable, and some unadaptable regardless of the size of $\rho$, then we call $G$ representation-dependent. Otherwise, we call $G$ representation-independent. The following theorem identifies three sub-cases of the representation-independent case, providing equivalent conditions for two sub-cases and a sufficient condition for the third, all in terms of $r_G \oplus r_G$. It is the main result of the thesis.

**Theorem 4.2.9.**
1. *Every permutation representation of G is fully adaptable if and only if $r_G \oplus r_G$ is fully adaptable.*
2. *Every permutation representation of G is either fully adaptable or partially adaptable if and only if $r_G \oplus r_G$ is partially adaptable.*
3. *If $r_G \oplus r_G$ is unadaptable, then every permutation representation of G with size equal or greater than $2|G|$ that contains at least two orbits of size $|G|$ is unadaptable.*

*Proof.* For statements 1 and 2, the second part of each equivalence is a special case of the first part, so we only need to prove the 'if' directions. Fix arbitrary permutation representation $\mathbb{R}^N$. By lemma 4.2.6, $\mathbb{R}^N$ is a direct sum of $\mathbb{S}_l$ where each $\mathbb{S}_l$ isomorphic to a subrepresentation of the regular representation. If $r_G \oplus r_G$ is fully (resp. partially) adaptable, Part 2 of lemma 4.2.1 implies that for every subrepresentation $\mathbb{S}_p$ and $\mathbb{S}_q$ of $\mathbb{R}^{2|G|}$, $\mathbb{S}_p \oplus \mathbb{S}_q$ is fully (resp. partially) adaptable. This implies that $\mathbb{R}^N$ is fully (resp. partially) adaptable by theorem 4.2.2, which completes the proof of the first and second statements.

To prove the third statement, fix arbitrary permutation representation $\mathbb{R}^N$ that contains at least two orbits of size $|G|$, and assume $r_G \oplus r_G$ is unadaptable. The only transitive permutation representation with size $|G|$ is the regular representation. Therefore, each orbit of size $|G|$ corresponds to $r_G$. Since $\mathbb{R}^N$ contains at least two orbits of size $|G|$, it contains $r_G \oplus r_G$. The result then follows from theorem 4.2.2. □

We call group $G$ representation-independent of type 1 if all of its permutation representations are partially adaptable and representation-independent of type 2 if all of its permutation representations are either fully or partially adaptable. It is representation-independent of type 3 if all of its representations with a size equal to or greater than $2|G|$ are unadaptable. Given $G$-invariant SDDP with any arbitrary permutation representation $\rho$ of size $N$, one seeks to find whether every $U \in \mathfrak{U}(\mathbb{R}^N)$ satisfies $f_U \leq f$, or whether there exists $U \in \mathfrak{U}(\mathbb{R}^N)$ satisfying $f_U \leq f$. Theorem 4.2.9 offers that it is enough to check whether $G$ is representation-independent of type 1 or 2 without analyzing the underlying representation. This simplifies the computation considerably since the size of the representation $\rho$ in practical problems is very big, for instance, a matrix of size $10^4 \times 10^4$ in combinatorial problems from [de Klerk et al., 2011], while the size of $r_G \oplus r_G$ is $2|G|$. Moreover, if one finds the desired $U$ for $\rho = r_G \oplus r_G$, then the desired $U$ can be constructed for any $\rho$. Theorem 4.2.9 also can be further simplified by combining it with theorem 2.3.3 to obtain the following corollary.

**Corollary 4.2.10.** *Given a $G$-invariant SDDP with permutation representation $\rho$ of size $N$, there exists $U \in \mathfrak{U}(\mathbb{R}^N)$ ensuring $f_U \leq f$ if there exists $U \in \mathfrak{U}(r_G \oplus r_G)$ satisfying*

$$\begin{aligned} U^T \Sigma^{2|G|}_{r_G \oplus r_G}(1,1) U &\subset \mathcal{S}(2|G|)_{2,+} \\ U^T \Sigma^{2|G|}_{r_G \oplus r_G}(1,2) U &\subset \mathcal{S}(2|G|)_{2,+}. \end{aligned} \tag{4.6}$$

Moreover, $f_U \leq f$ for all $U \in \mathfrak{U}(\mathbb{R}^N)$ if (4.6) holds for all $U \in \mathfrak{U}(r_G \oplus r_G)$.

*Proof.* When $\rho = r_G \oplus r_G$, there are only two orbits, So $1 \leq p \leq q \leq 2$ in (2.29). Additionally, since the orbits are the same type $\Sigma^{2|G|}_{r_G \oplus r_G}(1,1) = \Sigma^{2|G|}_{r_G \oplus r_G}(2,2)$, and the result follows. □

Corollary 4.2.10 provides an algorithm to check if a $U \in \mathfrak{U}(\mathbb{R}^{2|G|})$ satisfies (4.6) and thus (2.29). If it is satisfied, then one can construct $U \in \mathfrak{U}(\mathbb{R}^N)$ with the desired property using lemma 4.2.1. Given $U \in \mathfrak{U}(r_G \oplus r_G)$ and $\rho = r_G \oplus r_G$ , algorithm 1 illustrates whether (4.6) is satisfied or not.

**Algorithm 1** Existence of symmetry-adapted basis for the direct sum of two regular representation

$A_0 = [x\ y; y\ z]$
**for** p=1:1:$|G|$ **do**
  **for** q=p:1:2$|G|$ **do**
    $A = \sum_{g \in G} \rho_g [e_p\ \ e_q] A_0 [e_p\ \ e_q]^T \rho_g^T$
    Check each blocks of $U^T AU$ is SDD for all $xy \geq z^2$
  **end for**
**end for**

This algorithm checks whether $\frac{(|G|)(3|G|+1)}{2}$ block diagonal matrices with $m$ blocks of size at most $2k_i d_i$ are SDD or not. To illustrate how much our calculation is simplified, consider $G$ being $\mathfrak{S}^3$, the symmetric group of order three. In this case, the algorithm checks whether 57 block diagonal matrices with 3 blocks of size at most 4 are SDD or not.

## 4.3 Combined analysis

Combining theorem 4.2.4 with theorem 4.2.9 can identify representation-independent groups of types 1 based on the orbital structure of their regular representations. This is illustrated in the next lemma for totally orthogonal groups. A group is called totally orthogonal if all of its real irreducible representations are real-type.

**Lemma 4.3.1.** *For a totally orthogonal group $G$, if its regular representation is MSS, then $G$ is representation-independent of type 1.*

*Proof.* When $G$ is totally orthogonal, $r_G$ is absolutely irreducible. Additionally, the representation $r_G \oplus r_G$ has only 2 orbits each corresponding to $r_G$. So, if $r_G$ is MSS, then $r_G \oplus r_G$ is fully adaptable by Part 1 of theorem 4.2.4. The fact that $G$ is representation-independent of type 1 is now concluded from theorem 4.2.9. □

Based on this lemma, we identify some representation-independent groups of type 1 in the following example.

**Example 4.3.2.** Let $G = Z_2$ be the cyclic group of order 2. Then $G$ has only two real-type irreducible representations of dimension 1, namely the trivial and sign representations. Hence, $G$ is totally orthogonal and $r_G$ is MSS implying $Z_2$ is a representation-independent group of type 1. This can be extended to groups that are a direct product of $Z_2$. Let $G$ be $Z_2^t := Z_2 \times \cdots \times Z_2$ be a direct product of $t$ copies of $Z_2$. Then $G$ has $2^t$ real-type irreducible representation of dimension 1. All of these irreducible representations appear once in $r_G$. Therefore, $Z_2^t$ is totally orthogonal, and $r_G$ is an absolutely irreducible MSS representation, implying $Z_2^t$ is a representation-independent group of type 1. This asserts that whenever an SDDP has a symmetry of group $G = Z_2^t$ for any $t$, $f_U \leq f$ for any symmetry-adapted basis.

We used algorithm 1 combined with theorem 4.2.9 to identify some representation-independent groups and their types. We further derive conditions on full or partial adaptability of the representations of some groups that we identify as representation-independent groups of type 2 or as representation-dependent groups. Our result is stated in table 4.1. Any representation with a small enough size is fully adaptable. As mentioned, characterizing representations based on their size is not practical. For this reason, representation-independent groups of type 1 aside, the results in table 4.1 is only considered for the representation with a size equal to or greater than $2|G|$.

Table 4.1: A non-empty square in the columns RI1, RI2, or RI3 determines $G$ is a representation-independent group of type 1, 2, or 3, respectively. Similarly, RD corresponds to representation-dependent groups. In the columns corresponding to RI2 and RD, we further determine which orbital structure makes the $G$-representations fully or partially adaptable.

| Group | RI1 | RI2 | | RD | | | RI3 |
|---|---|---|---|---|---|---|---|
| | | Fully adaptable | Partially adaptable | Fully adaptable | Partially adaptable | Unadaptable | |
| $Z_2$ | x | | | | | | |
| $Z_4$ | | never | always | | | | |
| $Z_n$ for $n \neq 1, 2, 4$ | | | | | | | x |
| $Z_2^t$ for any $t$ | x | | | | | | |
| $\mathfrak{S}_3$ | | | | no orbit of size 6 | never | orbit of size 6 | |
| $\mathfrak{D}_4$ | | never | always | | | | |

# Chapter 5

# Nonnegativity and polynomial optimization

This chapter presents the application of symmetry reduction in polynomial optimization problems, specifically those that result in invariant SDDPs. First, we review some well-known results in real algebraic geometry and polynomial optimization. That can be found in books such as [Powers, 2021, Theobald, 2024]. We consider real polynomials in $n$ variables $p(x) := p(x_1, \cdots, x_n)$. Nonnegative polynomials in $n$ variables form a convex cone in the infinite-dimensional vector space of $n$-variate polynomials. We denote this cone by $\mathcal{P}_n$, and by $\mathcal{P}_{n,d}$ when we restrict to maximum polynomial degree $d$. A subset of $\mathbb{R}^n$ is basic semialgebraic if it is formed by a finite union of polynomial inequalities, i.e.,

$$L(h_1, \cdots, h_k) := \{x \in \mathbb{R}^n \mid h_i(x) \in \mathcal{P}_{n,d} \ \ \forall i \in [1, k]\}. \tag{5.1}$$

Equalities are also allowed in basic semialgebraic sets since $h(x) = 0$ equivalents to $h(x) \geq 0$ and $-h(x) \geq 0$. For the rest of the discussion, we consider $L := L(h_1, \cdots, h_k)$ a basic semialgebraic set unless otherwise specified. A set of nonnegative polynomials on $L$ is denoted by $\mathcal{P}(L)$. A common simple polynomial optimization problem has the form

$$p^* := \inf \quad p(x) \qquad \text{subject to} \quad x \in L, \tag{5.2}$$

or equivalently,

$$p^* = \inf \quad -\lambda \qquad \text{subject to} \quad p(x) + \lambda \in \mathcal{P}(L). \tag{5.3}$$

This problem is of NP-hard computational complexity. Solving such problems requires a certificate of nonnegativity for real multivariate polynomials. Positivstellensatz or certificate of nonegativity in Dutch is a fundamental problem in real algebraic geometry. If $p(x)$ can be written as a sum of squares of polynomials, i.e.,

$$p(x) = \sum_i p_i(x)^2, \tag{5.4}$$

then $p(x) \in \mathcal{P}_{n,d}$ and called a sum-of-squares (SOS) polynomial. However, not all nonnegative polynomials are SOS. We let $\mathcal{SOS}_{n,d}$ denote the proper cone of of $n$-variate SOS polynomials of degree at most $d$, Hilbert showed that $\mathcal{SOS}_{n,d} = \mathcal{P}_{n,d}$ if and only if either $n = 1$, $d = 2$, or $n = 4$ and $d = 2$ [Hilbert, 1888]. Let

$$\text{QM}(L) := \text{QM}(h_1, \cdots, h_k) = \{s_0 + \sum_{i=1}^{k} s_i(x)h_i(x) \mid s_i(x) \in \mathcal{SOS}_n \ \ \forall i \in [0,k]\} \tag{5.5}$$

define the quadratic module of $L$. A quadratic module QM is said to be Archimedean if, for any polynomial $p(x)$, there exists a natural number $N$ such that $N \pm p \in \text{QM}$. Putinar showed if $\text{QM}(L)$ is Archimedean and $p(x)$ is positive on $L$, then $p(x) \in \text{QM}(L)$. This is known as the Putinar Positivstellensatz. [Lasserre, 2001] relaxed problem (5.3) to a series of relatively tractable polynomial optimization problems

$$p^{K,d} := \min \quad -\lambda \qquad \text{subject to} \quad \begin{aligned} & p(x) + \lambda = s_0(x) + \textstyle\sum_{i=1}^{k} s_i(x) h_i(x) \\ & s_i(x) \in \mathcal{K} \end{aligned} \tag{5.6}$$

where $\mathcal{K} = \mathcal{SOS}_{n,d}$, and he showed that $k^{K,d}$ converges to the optimal value of (5.3) when $d$ increases. In other words,

$$p^{sos} := \lim_{d\to\infty} p^{sos,d} = p^{*}. \tag{5.7}$$

This is known as the Lasserre hierarchy, and problem (5.6) with $\mathcal{K}$ being an SOS cone is an instance of SOS optimization. If the quadratic module of $L$ is not Archimedean, $p^{sos}$ may not be $p^{*}$; for a convergence on the cases where $L$ is not compact see [Mai et al., 2022]. However, SOS optimization is still widely used in such cases due to its practical effectiveness.

## 5.1 SOS optimization and its relaxations

In contrast to (5.3), SOS optimizations are practically solvable when $n$ and $d$ are not too large. Indeed, [Parrilo, 2003] showed this problem can be solved using SDP.

**Lemma 5.1.1** ([Parrilo, 2003]). *Let the monomial vector basis $m(x)$ be the $N \times 1$ vector whose elements are monomials $x_1^{d_1} \cdots x_n^{d_n}$ such that $0 \le d_1 + \cdots d_m \le d$ ($N = \binom{n+d-1}{d}$). There exists a matrix $Q \in \mathcal{S}(N)_+$, known as Gram matrix, such that $s(x) = m(x)^T Q m(x)$ if and only if $s(x) \in \mathcal{SOS}_{n,d}$.*

This demonstrates that the task of finding a SOS polynomial in $n$ variables of degree at most $d$ is equivalent to the task of finding a PSD matrix of size $\binom{n+d-1}{d}$. One can show problem (5.6) with $\mathcal{K} = \mathcal{SOS}_{n,d}$ is equivalent to (1.1) with $\mathcal{K} = \mathcal{S}(N)_+$, $N = \binom{n+d-1}{d}$. The first constraint of (5.6) determines the affine subspace $\mathcal{L}$ and its second constraint determines the cone $\mathcal{K}$ in (1.1).

To find $p^{sos}$ in practice, $d$ does not need to go to infinity, instead, there are upper bounds for the value $d$. However, based on $N = \binom{n+d-1}{d}$, $N$ grows very fast with the increase of $d$. Solving SDP with large $N$ is not fast nor precise. This challenge has led mathematicians and optimizers to approximate large SDPs with more computationally efficient optimizations. [Ahmadi and Majumdar, 2019] suggested instead choosing $\mathcal{K}$ to be $\mathcal{SBS}_{n,d}$ the cone of sum-of-binomial-squares (SBS) polynomials, in which case (5.3) turns into an SDDP rather than an SDP.

**Lemma 5.1.2** ([Ahmadi and Majumdar, 2019]). *There exists matrix $Q \in \mathcal{S}(N)_{2,+}$ such that $s(x) = m(x)^T Q m(x)$ if and only if $s(x) \in \mathcal{SBS}_{n,d}$.*

Since $\mathcal{SBS}_{n,d} \subset \mathcal{SOS}_{n,d}$, problem (5.3) with the cone $\mathcal{SBS}_{n,d}$ might arise on its own but arises naturally as a strengthening of an SOS condition, where this cone acts as an inner approximation to the SOS cone. When viewed as a relaxed version of SOS problems, their current formulation lacks a convergence guarantee even if the quadratic module of $L$ is Archimedean, unlike SOS problems. Indeed, [Josz, 2017] constructs an example in which (5.3) with $\mathcal{K} = \mathcal{SBS}_{n,d}$ gives strictly bigger value, i.e.,

$$p^{sos} < p^{sbs} := \lim_{d \to \infty} p^{sbs,d}. \tag{5.8}$$

Recent papers have addressed this issue by either reformulating (5.3) [Roebers et al., 2021] or modifying interior point methods to achieve similar convergence properties [Roig-Solvas and Sznaier, 2022]. Although their practical applicability is not yet clear, these developments suggest an increasing interest in replacing the PSD (resp. SOS) cone with the SDD (resp. SBS) cone. Thus, extending symmetry reduction techniques to problems involving the SDD and SBS cones is a logical step.

## 5.2 Invariant polynomial optimization

We say problem (5.3) is $G$-invariant if $p(x)$ and $L$ are $G$-invariant with respect to $\rho^0 : G \to O(n)$. That is, for all $g \in G$, $p(\rho_g^0 x) = p(x)$ and $\rho_g^0(L) \subset L$. This representation induces another representation $\rho : G \to O(N)$ such that, for each $g \in G$, $\rho_g m(x) = m(\rho_g^0 x)$, where $m(x)$ in the monomial vector basis defined in lemma 5.1.1 and lemma 5.1.2. The analogous program (1.1) to invariant (5.3) is $G$-invariant with respect to $\sigma$, which is induced by $\rho$ via (2.5).

From now on, we treat (5.3) with $\mathcal{K} = \mathcal{SOS}_{n,d}$ with the optimal value $p^{sos,d}$ as (1.1) with $\mathcal{K} = \mathcal{S}(N)_+$ with optimal value $f^{sdp}$ and (5.3) with $\mathcal{K} = \mathcal{SBS}_{n,d}$ with the optimal value $p^{sbs,d}$ as (1.1) with $\mathcal{K} = \mathcal{S}(N)_{2,+}$ with optimal value $f^{sdd}$. As discussed, while $f_U^{sdp} = f^{sdp}$, generally $f_U^{sdd} \neq f^{sdd}$. We also illustrate that $f_U^{sdd}$ can provide a better bound for $f^{sdd}$ as a result of symmetry reduction.

The numerical experiment in this section is solved via MOSEK [ApS, 2024] in Julia. The optimization problems are formulated with the Julia package "Sumsofsquares" [Legat et al., 2017] and the symmetry reduction method is used with "Symbolicwedderburn" [Kaluba et al., 2019].

**Example 5.2.1** ([Magron and Wang, 2023], Example 25)**.** To minimize $\lambda$ such that $P = 1 + x_1^4 + x_2^4 - x_1 x_2^2 - x_1^2 x_2 + 5x_1 x_2 - \lambda$ lies within the SOS cone, an SDP must be solved. Let $f^{sdp}$ represent the optimal value of this SDP. Approximating this SDP by its $\mathcal{S}(N)_{2,+}$-relaxation leads to an SDDP with an optimal value denoted as $f^{sdd}$. Given the symmetry in $P$, where swapping $x_1$ and $x_2$ does not alter the polynomial, this property extends to the corresponding SDP and SDDP, which share a cyclic group of order two symmetry. While symmetry reduction applied to the SDP maintains its optimal value, applying it to the SDDP using a symmetry-adapted basis $U$ yields program (1.3) with an optimal value $f_U^{sdd}$. As shown in Table 5.1, $f_U^{sdd}$ serves as a significantly better upper bound for $f^{sdp}$ compared to $f^{sdd}$.

Table 5.1: Optimal values derived from different optimization

| $f^{sdp}$ | $f^{sdd}$ | $f_U^{sdd}$ |
|---|---|---|
| -2.2033 | -6.9165 | -2.2894 |

# Chapter 6

# Conclusion and future works

In this thesis, we illustrated how extending symmetry reduction techniques to cone programming with basis-dependent cones entails several difficulties not present in the case of the SDP cone. A general geometric perspective is given on how one can approach such challenges. This perspective shows how one can take advantage of algebraic structure to find geometric properties of proper cones. We focus on the geometry of the SDD cone which is equivalent to the cone of factor-width-two matrices. However, we have not extended our results to other proper cones. Exploring the geometry of factor-width-k matrices or diagonally dominant matrices in the presence of symmetry are natural next steps in this research program.

The cone of SBS polynomials is isomorphic to the cone of SDD matrices. While we investigate the decomposition of invariant SDD matrices, one can find an analogous result for the decomposition of invariant SBS polynomials. Instances of the decomposition of SBS polynomials with special structure can be found in [Powers, 2021, Chapter 12] due to Reznick's work. In most of these examples, the polynomials have symmetry of symmetric groups $\mathfrak{S}_n$, while our result can be considered for polynomials having symmetry of a broader range of groups. What we did in this thesis has a counterpart in real algebraic geometry. One can seek the algebraic structures of SOS polynomials that guarantee they are SBS as well.

In real algebraic geometry and optimization, there is a common question of which spectrahedron (or the feasible region of an SDP) is second-order cone representable. For instance, [Scheiderer, 2021] answered this question for convex sets in a plane. While generalizing this result to higher dimensions is a complicated task, one can attempt this generalization for spectahedra with group symmetries. There is strong motivation to investigate this question since, in some instances of invariant SDDPs, while $f^{sdd} < f^{psd}$, it is observed that $f_U^{sdd} = f^{psd}$ for some symmetry-adapted basis $U$. Ultimately, a broad direction for future work involves characterizing symmetry groups and the regions (not necessarily convex) exhibiting such symmetries, with the aim of lifting these regions into the feasible regions of semidefinite programming, second-order cone programming, or linear programming.

# References


Amir Ali Ahmadi and Anirudha Majumdar. DSOS and SDSOS optimization: more tractable alternatives to sum of squares and semidefinite optimization. *SIAM Journal on Applied Algebra and Geometry*, 3(2):193–230, 2019.

MOSEK ApS. *MOSEK optimizer API for Julia. Version 10.2.3.*, 2024. URL https://docs.mosek.com/10.2/juliaapi.pdf.

Christine Bachoc, Dion C Gijswijt, Alexander Schrijver, and Frank Vallentin. Invariant semidefinite programs. *Handbook on Semidefinite, Conic and Polynomial Optimization*, pages 219–269, 2012.

Mattew Badger. Division algebras over the real numbers. *University of Pittsburgh, Pittsburgh*, 2006.

Jeff Bezanson, Alan Edelman, Stefan Karpinski, and Viral B Shah. Julia: a fresh approach to numerical computing. *SIAM Review*, 59(1):65–98, 2017. URL https://doi.org/10.1137/141000671.

Erik G Boman, Doron Chen, Ojas Parekh, and Sivan Toledo. On factor width and symmetric H-matrices. *Linear Algebra and Its Applications*, 405:239–248, 2005.

Stephen P Boyd and Lieven Vandenberghe. *Convex optimization*. Cambridge University Press, 2004.

Etienne De Klerk and Renata Sotirov. Exploiting group symmetry in semidefinite programming relaxations of the quadratic assignment problem. *Mathematical Programming*, 122:225–246, 2010.

Etienne de Klerk, Cristian Dobre, and Dmitrii V asechnik. Numerical block diagonalization of matrix *-algebras with application to semidefinite programming. *Mathematical Programming*, 129:91–111, 2011.

William Fulton and Joe Harris. *Representation theory: a first course*, volume 129. Springer Science & Business Media, 2004.

Terry Gannon. *Moonshine beyond the Monster: The bridge connecting algebra, modular forms and physics*. Cambridge University Press, 2007.

Karin Gatermann and Pablo A Parrilo. Symmetry groups, semidefinite programs, and sums of squares. *Journal of Pure and Applied Algebra*, 192(1-3):95–128, 2004.

Philippe Gille and Tamás Szamuely. *Central simple algebras and Galois cohomology*, volume 165. Cambridge University Press, 2017.

Alexander Heaton, Serkan Hosten, and Isabelle Shankar. Symmetry adapted gram spectrahedra. *SIAM Journal on Applied Algebra and Geometry*, 5(1):140–164, 2021.

David Hilbert. Über die darstellung definiter formen als summe von formenquadraten. *Mathematische Annalen*, 32(3):342–350, 1888.

Nathan Jacobson. *Finite-dimensional division algebras over fields*. Springer Science & Business Media, 2009.

Cédric Josz. Counterexample to global convergence of DSOS and SDSOS hierarchies. *arXiv preprint arXiv:1707.02964*, 2017.

Marek Kaluba, Piotr W Nowak, and Narutaka Ozawa. Aut $(\mathbb{F}_5)$ has property (T). *Mathematische Annalen*, 375:1169–1191, 2019.

Max A Knus, Alexander Merkurjev, Markus Rost, and Jean P Tignol. *The book of involutions*, volume 44. American Mathematical Society, 1998.

Tsit Y Lam. *A First Course in Noncommutative Rings*, volume 131. Springer, 2001.

Serge Lang. *Algebra*, volume 211. Springer Science & Business Media, 2002.

Jean B Lasserre. Global optimization with polynomials and the problem of moments. *SIAM Journal on Optimization*, 11(3):796–817, 2001.

Benoît Legat, Chris Coey, Robin Deits, Joey Huchette, and Amelia Perry. Sum-of-squares optimization in Julia. In *The First Annual JuMP-dev Workshop*, 2017.

Takanori Maehara and Kazuo Murota. A numerical algorithm for block-diagonal decomposition of matrix *-algebras with general irreducible components. *Japan Journal of Industrial and Applied Mathematics*, 27:263–293, 2010.

Victor Magron and Jie Wang. SONC optimization and exact nonnegativity certificates via second-order cone programming. *Journal of Symbolic Computation*, 115:346–370, 2023.

Ngoc Hoang Anh Mai, Jean-Bernard Lasserre, and Victor Magron. Positivity certificates and polynomial optimization on non-compact semialgebraic sets. *Mathematical Programming*, pages 1–43, 2022.

Philippe Moustrou, Helen Naumann, Cordian Riener, Thorsten Theobald, and Hugues Verdure. Symmetry reduction in AM/GM-based optimization. *SIAM Journal on Optimization*, 32(2):765–785, 2022.

Philippe Moustrou, Cordian Riener, and Hugues Verdure. Symmetries in polynomial optimization. In *Polynomial Optimization, Moments, and Applications*, pages 53–111. Springer, 2023.

Kazuo Murota, Yoshihiro Kanno, Masakazu Kojima, and Sadayoshi Kojima. A numerical algorithm for block-diagonal decomposition of matrix-algebras with application to semidefinite programming. *Japan Journal of Industrial and Applied Mathematics*, 27 (1):125–160, 2010.

Pablo A Parrilo. Semidefinite programming relaxations for semialgebraic problems. *Mathematical Programming*, 96:293–320, 2003.

Donald S Passman. *A course in ring theory*. American Mathematical Society, 2004.

Sven C Polak. Symmetry reduction to optimize a graph-based polynomial from queueing theory. *SIAM Journal on Applied Algebra and Geometry*, 6(2):243–266, 2022.

Victoria Powers. *Certificates of positivity for real polynomials*. Springer, 2021.

Cordian Riener, Thorsten Theobald, Lina Jansson Andrén, and Jean B Lasserre. Exploiting symmetries in SDP-relaxations for polynomial optimization. *Mathematics of Operations Research*, 38(1):122–141, 2013.

Lorenz M Roebers, Juan C Vera, and Luis F Zuluaga. Sparse non-SOS Putinar-type positivstellensatze. *arXiv preprint arXiv:2110.10079*, 2021.

Biel Roig-Solvas and Mario Sznaier. A globally convergent LP and SOCP-based algorithm for semidefinite programming. *arXiv preprint arXiv:2202.12374*, 2022.

Claus Scheiderer. Second-order cone representation for convex sets in the plane. *SIAM Journal on Applied Algebra and Geometry*, 5(1):114–139, 2021.

Alexander Schrijver. A comparison of the delsarte and lovász bounds. *IEEE Transactions on Information Theory*, 25(4):425–429, 1979.

Jean P Serre. *Linear representations of finite groups*, volume 42. Springer, 1977.

Thorsten Theobald. *Real algebraic geometry and optimization*, volume 241. American Mathematical Society, 2024.