Universität zu Köln

Combinatorial Approaches to the EHZ Capacity of Simplices

Inaugural-Dissertation

zur

Erlangung des Doktorgrades der
Mathematisch-Naturwissenschaftlichen Fakultät
der Universität zu Köln

vorgelegt von

Karla Leipold

2026

Berichterstatter (Gutachter):
Professor Frank Vallentin
Professor Alberto Abbondandolo

Tag der mündlichen Prüfung: 08.06.2026

Abstract

This thesis investigates how methods from discrete optimization can be used to study problems in symplectic geometry. Its central object is the Ekeland–Hofer–Zehnder capacity c_{EHZ}, a fundamental symplectic invariant that is monotone under symplectic embeddings and plays an important role in Hamiltonian dynamics and convex symplectic geometry. Although c_{EHZ} is conceptually well understood, explicit computations are notoriously difficult. A central theme of this thesis is that new connections to combinatorial optimization not only clarify the computational complexity of these problems, but also open the door to new and more effective algorithmic approaches.

By specializing a combinatorial formula for the Ekeland–Hofer–Zehnder capacity of polytopes to the case of simplices, we obtain a finite combinatorial formula for c_{EHZ} in this setting. This reveals a discrete structure underlying the symplectic problem and creates a bridge to methods from combinatorial optimization, integer programming, and computational complexity. One of our main results is that computing the Ekeland–Hofer–Zehnder capacity of convex polytopes is NP-hard. This is proved via a reduction from the feedback arc set problem on directed bipartite tournaments. In addition, we relate the problem of deciding affine symplectomorphism of full-dimensional simplices, via polynomial-time reductions, to weighted directed graph isomorphism.

On the algorithmic side, these connections allow us to reformulate the simplex formula as an exact integer linear program for evaluating c_{EHZ} of simplices, providing an effective alternative to brute-force enumeration. They also motivate the algorithmic study of optimization problems that would otherwise be computationally out of reach, such as the search for simplices with large systolic ratio. To this end, the exact evaluation procedure is combined with optimization on the special linear group $\mathrm{SL}(2n, \mathbb{R})$. Since the objective is nonsmooth, we develop a retraction-based manifold optimization approach using Clarke subdifferentials, active branch gradients, and Armijo-type line search. Together, these ideas yield a computational framework that combines exact evaluation, nonsmooth optimization on manifolds, and symmetry detection via weighted graph isomorphism. Computational experiments produce many simplices with systolic ratio exceeding that of the standard simplex in several dimensions. Overall, this thesis shows that discrete optimization provides an effective framework for studying symplectic capacities.

Contents

Introduction 1

I Symplectic Preliminaries and the EHZ Capacity 7

1 Basics of Symplectic Geometry 8

2 The Ekeland–Hofer–Zehnder Capacity 13
- 2.1 The Counterexample to the Viterbo Conjecture 15
- 2.2 Computational Results for the EHZ Capacity of Polytopes 27

3 EHZ Capacity of Simplices 30

II Computational Complexity in Symplectic Geometry 37

4 Computational Complexity 38
- 4.1 Turing Machines 38
- 4.2 Complexity Classes 40

5 NP-Hardness of the Feedback Arc Set Problem on Bipartite Tournaments 44
- 5.1 Notation from Combinatorial Optimization 44
- 5.2 Proof that FAS is NP-Hard on Bipartite Tournaments 45

6 Computing the EHZ Capacity is NP-Hard 50

7 Affine Symplectomorphism of Simplices is GI-Complete 59
- 7.1 Affine Symplectomorphisms of Simplices 61
- 7.2 Reduction from Weighted GI to Anchored Symplectomorphism 63
- 7.3 Reduction from Anchored to Unanchored Affine Symplectomorphism 67

III Searching for Maximizers of the Systolic Ratio 73

8 Problem Statement and Motivation 75

9 ILP Formulation of the EHZ Capacity of Simplices 78
9.1 Linear Programs and Integer Linear Programs 78
9.2 The EHZ Capacity Computed Through a Linear-Ordering ILP 79

10 Optimization on the Special Linear Group $\mathrm{SL}(2n, \mathbb{R})$ 82
10.1 Manifolds, Tangent spaces, and Embedded Submanifolds 83
10.2 Riemannian Metrics and Riemannian Submanifolds 84
10.3 Differentiation on Riemannian Submanifolds 86
10.4 Retractions 90
10.5 Riemannian Gradient Descent (RGD) 93

11 Subgradients and Nonsmoothness in the Objective 97
11.1 The Clarke Subdifferential in Euclidean Space 98
11.2 The Clarke Subdifferential on an Embedded Riemannian Submanifold 102
11.3 Clarke Subdifferential of the Objective $X \mapsto c_{\mathrm{EHZ}}(XK)$ on $\mathrm{SL}(2n, \mathbb{R})$ 106
11.4 The Standard Simplex as a Stationary Point 108

12 Algorithmic Approaches 112
12.1 Retraction-Based Line-Search on Manifolds 112
12.2 Testing Affine Symplectomorphisms 119
12.3 Implementation Details 119
12.4 Results 122
12.5 Discussion 123

Acknowledgements 127

Appendix 129

Notation 132

Bibliography 135

Research Resources and Publications 143

Introduction

The central theme of this thesis is that discrete optimization can be an effective tool for studying questions in symplectic geometry. In particular, symplectic invariants of polyhedral objects often admit a surprisingly explicit combinatorial description, which makes them accessible to methods from discrete optimization and integer programming. This thesis explores this perspective for the Ekeland–Hofer–Zehnder capacity of polytopes, with a particular focus on simplices.

Symplectic geometry comes with a very concrete origin story: It is the geometric language behind Hamilton's formulation of classical mechanics. Building on earlier ideas of Lagrange, Hamilton observed that many mechanical systems can be described in *position–momentum* variables

$$q(t) \in \mathbb{R}^n, \qquad p(t) \in \mathbb{R}^n,$$

together with an energy function (Hamiltonian) $H = H(q, p, t)$, so that the motion is governed by Hamilton's equations (see, McDuff–Salamon [69], Hofer–Zehnder [52]):

$$\dot{p}(t) = \frac{\partial H}{\partial q}\big(q(t), p(t), t\big), \qquad \dot{q}(t) = -\frac{\partial H}{\partial p}\big(q(t), p(t), t\big). \tag{1}$$

These equations are naturally encoded by a nondegenerate skew-symmetric form on the phase space $\mathbb{R}^{2n}$. Writing $z = (q, p) \in \mathbb{R}^{2n}$ and introducing the standard symplectic matrix

$$J \;=\; \begin{pmatrix} 0 & -I_n \\ I_n & 0 \end{pmatrix},$$

Hamilton's equations can be written as $\dot{z}(t) = J\nabla_z H(z(t), t)$. Equivalently, they can be rewritten by the standard symplectic form ω_0 on $\mathbb{R}^{2n}$, defined by

$$\omega_0(u, v) = \langle Ju, v\rangle, \qquad u, v \in \mathbb{R}^{2n}.$$

Symplectic geometry studies manifolds equipped with a symplectic form and transformations that preserve it. This makes symplectic geometry the natural setting for studying structural properties of Hamiltonian dynamics.

A systematic way to quantify symplectic "size" is provided by *symplectic capacities*: numerical invariants that are monotone under symplectic embeddings. Gromov's non-squeezing theorem [35] famously shows that symplectic embeddings are constrained by more than volume, and capacities provide a framework for expressing this additional rigidity. In the convex setting, several a priori different capacities coincide. In particular, the Hofer–Zehnder, first Ekeland–Hofer, Viterbo symplectic homology, and first Gutt–Hutchings capacities agree on convex bodies; see [51, 27, 59, 41, 1, 46]. We denote this common value by c_{EHZ} and refer to it as the *Ekeland–Hofer–Zehnder capacity*. Among its equivalent descriptions, c_{EHZ} can be interpreted as the minimal action of a (generalized) closed characteristic on the boundary, but despite its conceptual importance it is difficult to compute explicitly. A major driving force in the subject is Viterbo's conjectural volume–capacity inequality [94]. For a convex body $K \subset \mathbb{R}^{2n}$ we define the *systolic ratio*

$$\operatorname{sys}_n(K) \;:=\; \frac{c_{\mathrm{EHZ}}(K)}{\big(n!\,\operatorname{Vol}(K)\big)^{1/n}},$$

which is c_{EHZ} normalized by volume in the scale-invariant way suggested by the quadratic homogeneity of capacities. In particular, $\operatorname{sys}_n(B) = 1$ for every Euclidean ball $B \subset \mathbb{R}^{2n}$.

Conjecture (Viterbo [94])**.** Among all convex bodies $K \subset \mathbb{R}^{2n}$ of prescribed volume, the Euclidean ball maximizes every symplectic capacity. For the Ekeland–Hofer–Zehnder capacity one expects $\operatorname{sys}_n(K) \leq 1$ for all full-dimensional convex bodies $K \subset \mathbb{R}^{2n}$.

Since Viterbo formulated this conjecture, a substantial body of work has supported it in many settings. A dimension-independent constant-factor version was proved by Artstein-Avidan, Milman, and Ostrover [7]. Beyond perturbative results, the conjectural inequality has been verified for several concrete families, including certain Lagrangian products [11, 82], monotone toric domains [42], and domains arising from Hamiltonians of classical mechanical type [62].

More recently, Pazit Haim-Kislev and Yaron Ostrover constructed an explicit counterexample to Viterbo's conjecture, based on p-product/Lagrangian-product constructions and billiard-type mechanisms [47, 48], which makes the question of possible maximizers of the systolic ratio even more compelling from a computational perspective.

Organization of the Thesis

Part I: Symplectic background and a simplex formula. Part I introduces the geometric background for the Ekeland–Hofer–Zehnder capacity c_{EHZ}. Building on the finite polytope formula of Haim-Kislev, this part derives a combinatorial formula for c_{EHZ} in the case of simplices, together with an equivalent trace reformulation.

We begin in Section 1 by recalling the linear symplectic setting $(\mathbb{R}^{2n}, \omega_0)$, symplectic maps, and symplectic capacities (following Weinstein [95] and Hofer–Zehnder [52]). In Section 2 we then introduce c_{EHZ} itself and collect some properties of it. Next, in Section 2.1 we go through the counterexample to Viterbo's conjecture constructed by Haim-Kislev and Ostrover in detail, both as motivation and because it is one of the important recent results on the systolic ratio and highlights structural features of c_{EHZ}. Finally, in Section 3, starting from the polytope capacity formula of Haim-Kislev [44], we derive two equivalent formulations for simplices: a purely combinatorial one and a trace reformulation. The combinatorial formulation makes the discrete structure of the problem explicit and serves as the basis for the complexity-theoretic reductions in Part II, while the trace reformulation is better suited to the differentiation arguments used later in Part III.

Part II: Complexity-theoretic hardness and reductions. Part II addresses the question of *intrinsic computational difficulty.* Once c_{EHZ} of polytopes is expressed via a finite permutation-type optimization problem, how hard is it to compute and how hard is it to decide whether two simplices are affinely symplectomorphic?

Our main result in this part is that computing the Ekeland–Hofer–Zehnder capacity of convex polytopes is NP-hard. Unless P = NP, there is no polynomial-time algorithm that computes c_{EHZ} exactly on general polytopes. This provides a complexity-theoretic justification of the qualitative difficulty already emphasized by Hofer and Zehnder [52].

We begin in Section 4 by recalling the standard framework of computational complexity and NP-hardness, following Arora and Barak [4]. Our reduction to c_{EHZ} proceeds through the feedback arc set problem on directed bipartite tournaments. Accordingly, Section 5 introduces the necessary combinatorial notation and reviews the NP-hardness proof for feedback arc set on bipartite tournaments due to Guo, Hüffner, and Moser [39]. The resulting NP-hardness proof for c_{EHZ} is presented in Section 6; this section gives the proof as it appears in joint work of the author with Frank Vallentin [67].

Beyond evaluation complexity, we also study the complexity of deciding symplectic equivalence of simplices. In Section 7 we relate the affine symplectomorphism problem for full simplices to weighted directed graph isomorphism via polynomial-time reductions. This places simplex symplectomorphism in the same complexity landscape as graph isomorphism and supplies a second combinatorial bridge that will later be used from an algorithmic viewpoint.

Part III : Algorithmic approaches. Part III turns the discrete structure identified earlier into a concrete computational framework for studying the systolic ratio within the simplex family. The guiding goal is the optimization problem suggested by Viterbo's conjecture [94]:

Which convex bodies of a fixed volume maximize the systolic ratio? Parts I–II show that for simplices the relevant c_{EHZ} computation admits both a combinatorial description and a trace formulation suitable for exact ILP evaluation and for differentiation-based methods on $\mathrm{SL}(2n, \mathbb{R})$. This makes it possible to carry out exact computations in significantly higher dimensions and to search systematically for simplices with large systolic ratio.

In Section 9 we build on the earlier combinatorial formulation and its connection to the feedback arc set problem to reformulate the permutation problem as a weighted linear ordering / feedback arc set problem and obtain an exact ILP for objective evaluation. The ILP is based on the classical linear-ordering formulations of Grötschel, Jünger, and Reinelt [34]. This yields dramatic speedups over brute-force enumeration of $(2n+1)!$ permutations.

To search for large systolic ratio at fixed volume within the simplex family, we fix the standard simplex S_{2n} and consider its volume-preserving linear images $K_X := XS_{2n}$ with $X \in \mathrm{SL}(2n, \mathbb{R})$. Since $\mathrm{Vol}(K_X) = \mathrm{Vol}(S_{2n})$ for $\det(X) = 1$, maximizing $\mathrm{sys}_n(K_X)$ over this family is equivalent to maximizing $c_{\mathrm{EHZ}}(K_X)$. This turns the geometric problem into a constrained matrix optimization problem on the special linear group:

$$\max\{c_{\mathrm{EHZ}}(XS_{2n}) : X \in \mathrm{SL}(2n, \mathbb{R})\}.$$

Accordingly, Section 10 recalls the general manifold-optimization framework from Boumal [13] and works out its specialization to $\mathrm{SL}(2n, \mathbb{R})$ as an embedded Riemannian submanifold. A key difficulty is that the objective function is nonsmooth. To obtain principled first-order information nonetheless, we use Clarke's generalized derivatives [20]. In Section 11 we combine the explicit smooth branch gradients coming from the trace formula (using standard matrix differential identities, for instance Petersen and Pedersen [77]) with the finite-max structure to represent the Clarke subdifferential as a convex hull of active branch gradients. For the manifold setting we use the Riemannian Clarke calculus and directional derivatives, as developed by Hosseini and Pouryayevali [55]. In Section 11.4 we then use this subdifferential framework to show that the identity matrix $I \in \mathrm{SL}(2n, \mathbb{R})$, corresponding to the standard simplex S_{2n}, is a Clarke-stationary point of the objective $X \mapsto c_{\mathrm{EHZ}}(XS_{2n})$. Finally, in Section 12 we turn these analytic ingredients into an implementable method: a retraction-based *Clarke steepest descent* scheme with Armijo-type backtracking, inspired by the general nonsmooth line-search framework on manifolds developed by Hosseini, Huang, and Yousefpour [54]. Each iteration combines two optimization layers: Solving an ILP, to evaluate c_{EHZ} and extract active maximizers, and a convex quadratic program to compute a minimum-norm element of the resulting convex hull of active gradients. While the available global convergence guarantees require additional assumptions that are not straightforward to verify for our objective, the method provides a robust, systematically derived descent mechanism and in our numerical experiments produces simplices with large systolic ratio.

We also need a practical test for whether two simplices are affinely symplectomorphic. In centred position, this reduces to a weighted graph isomorphism test for the associated skew-symmetric signature matrices. The centred symplectomorphism test and its implementation aspects are developed in Section 12.2, based on the reductions from Section 7.

Combining exact ILP evaluation, explicit Clarke subgradients, nonsmooth manifold descent, and symmetry testing yields, in dimensions $2n \in \{6, 8, 10, 12, 14, 16\}$, simplices whose systolic ratio exceeds that of the standard simplex.

Part I

Symplectic Preliminaries and the EHZ Capacity

To get started and gain some understanding of the mathematical framework of the problems we will be tackling, we recall the basic definitions and facts from symplectic geometry. For background and historical context, we refer to Weinstein's introduction [95]. Our presentation follows the viewpoint and exposition of Hofer and Zehnder [52]. The main objective of this chapter is to introduce the key symplectic invariant studied in this thesis, namely the Ekeland–Hofer–Zehnder capacity, to provide some motivation for its definition, and to work through some known results to gather some understanding of the mathematical framework and open problems. Although symplectic geometry offers powerful methods, our later arguments will use mainly convex/combinatorial/optimization tools rather than symplectic-topological ones; hence we omit those techniques from this chapter. Section 1 recalls a basic background on symplectic geometry and symplectic capacities; readers familiar with these notions may skip it. In Section 2 we introduce the Ekeland–Hofer–Zehnder (EHZ) capacity, which is the symplectic invariant studied throughout this thesis. In Section 2.1 we then discuss in detail the counterexample to the Viterbo conjecture constructed by Haim-Kislev and Ostrover [48]. Next, we describe a combinatorial reformulation of the computation of the EHZ capacity which, for simplices, yields an explicit finite formula in Section 3. We then further rewrite this simplex formula in an equivalent *trace* form , which will later be the convenient interface both for exact computation (via ILP) and for differentiation in our optimization framework. These complementary viewpoints on c_{EHZ} for simplices are developed in this thesis and constitute one of the main contributions of this chapter.

1 Basics of Symplectic Geometry

While symplectic geometry is often developed in the setting of symplectic manifolds, we will work almost exclusively in the linear framework of symplectic vector spaces, since the manifold we are considering is $\mathbb{R}^{2n}$. Let V be a finite-dimensional real vector space and let

$$\omega : V \times V \to \mathbb{R}$$

be a bilinear form with the following properties:

(a) **Skew-symmetry:** $\omega(u, v) = -\omega(v, u)$ for all $u, v \in V$;

(b) **Nondegeneracy:** for every $u \in V \setminus \{0\}$ there exists $v \in V$ such that $\omega(u, v) \neq 0$.

Such a form ω is called a *symplectic form*, and the pair (V, ω) is called a *symplectic vector space*.

The fundamental model example is the pair $(\mathbb{R}^{2n}, \omega_0)$, where ω_0 is defined by

$$\omega_0(u, v) = u^\mathsf{T} J v, \qquad J = \begin{pmatrix} 0 & -I_n \\ I_n & 0 \end{pmatrix},$$

for $u, v \in \mathbb{R}^{2n}$. (Here I_n denotes the $n \times n$ identity matrix.) When the dimension is clear from the context, we simply write ω_0 and J.

The matrix J is invertible (in fact $\det J = 1$), which implies that ω_0 is nondegenerate. Moreover $J^\mathsf{T} = -J$, and hence ω_0 is skew-symmetric:

$$\omega_0(u, v) = u^\mathsf{T} J v = (u^\mathsf{T} J v)^\mathsf{T} = v^\mathsf{T} J^\mathsf{T} u = -\, v^\mathsf{T} J u = -\omega_0(v, u).$$

We refer to $(\mathbb{R}^{2n}, \omega_0)$ as the standard symplectic vector space.

The standard model is not merely an example: in finite dimensions every symplectic vector space is linearly isomorphic to it. Concretely, one can choose a basis in which the symplectic form has the standard matrix J ([52]). It is often convenient to work directly in the standard symplectic vector space $(\mathbb{R}^{2n}, \omega_0)$, which will be our default setting throughout the thesis. In what follows, we view J as the matrix representation of the standard symplectic form ω_0

Next, we consider maps between symplectic vector spaces. Since a symplectic form encodes additional structure, it is natural to single out those maps that preserve it.

Definition 1.1. For $i \in \{1,2\}$ let $U_i \subset \mathbb{R}^{2n}$ be open and let $\omega_i : \mathbb{R}^{2n} \times \mathbb{R}^{2n} \to \mathbb{R}$ be symplectic forms. A differentiable map $\varphi : U_1 \to U_2$ is called *symplectic* if for every $x \in U_1$,

$$\omega_1(u,v) \;=\; \omega_2\big(D\varphi(x)u,\; D\varphi(x)v\big) \qquad \text{for all } \; u, v \in \mathbb{R}^{2n}, \tag{2}$$

where $D\varphi(x)$ is the Jacobian of φ at x. A *symplectic embedding* is an injective smooth symplectic map $\varphi : U_1 \to \mathbb{R}^{2n}$ such that $\varphi(U_1) \subset U_2$ for some target set $U_2 \subset \mathbb{R}^{2n}$, and we write $\varphi : U_1 \overset{s}{\hookrightarrow} U_2$. A *symplectic diffeomorphism* is a symplectic map $\varphi : U_1 \to U_2$ that is a diffeomorphism.

A characteristic feature of symplectic diffeomorphisms is that they preserve volume. In the standard symplectic space $(\mathbb{R}^{2n}, \omega_0)$, condition (2) with ω_1 and ω_2 chosen as ω_0 is equivalent to the matrix identity

$$D\varphi(x)^{\mathsf{T}} J\, D\varphi(x) \;=\; J \qquad \text{for all } x \in \mathbb{R}^{2n}. \tag{3}$$

Indeed, writing $\omega_0(u,v) = u^{\mathsf{T}} J v$ gives

$$u^{\mathsf{T}} J v \;=\; \omega_0(u,v) \;=\; \omega_0(D\varphi(x)u, D\varphi(x)v) \;=\; u^{\mathsf{T}} D\varphi(x)^{\mathsf{T}} J D\varphi(x) v \quad \text{for all } \; u, v,$$

which yields (3). Taking determinants in (3) yields $\det D\varphi(x) \in \{-1, 1\}$. In fact, symplectic matrices always have determinant 1, as proved by Dopico and Johnson [26]. Hence $\det D\varphi(x) = 1$, and therefore φ preserves the Euclidean volume of measurable sets.

In $\mathbb{R}^2$ "symplectic" and "area-preserving" coincide, and total area already distinguishes diffeomorphic domains up to symplectic maps. In higher even dimensions the picture changes drastically: symplectic maps preserve volume, but the symplectic condition is much more rigid than the Jacobian constraint $\det D\varphi = 1$.

To illustrate this, consider the Euclidean ball and the symplectic cylinder (in symplectic coordinates $(x, y) \in \mathbb{R}^n \times \mathbb{R}^n$, with $x = (x_1, \dots, x_n)$ and $y = (y_1, \dots, y_n)$ and r a nonnegative real number):

$$B^{2n}(r) \coloneqq \{(x,y) \in \mathbb{R}^{2n} : \|x\|^2 + \|y\|^2 < r^2\}, \qquad Z^{2n}(R) \coloneqq \{(x,y) \in \mathbb{R}^{2n} : x_1^2 + y_1^2 < R^2\}.$$

The cylinder $Z^{2n}(R)$ has infinite volume when $n > 1$, so a volume comparison cannot possibly prevent putting a large ball into a thin cylinder. Nevertheless, Gromov's celebrated non-squeezing theorem states that symplectically this cannot be done:

Theorem 1.2 ([35])**.** *There exists a symplectic embedding* $B^{2n}(r) \overset{s}{\hookrightarrow} Z^{2n}(R)$ *if and only if* $R \geq r$.

This result already shows that symplectic geometry contains genuine "size" information not captured by volume, it suggests that symplectic maps should preserve additional numerical invariants beyond volume, and that these invariants should obstruct certain embeddings.

A systematic way to quantify symplectic "size" is provided by symplectic capacities. For our purposes it is convenient to restrict to open subsets of $\mathbb{R}^{2n}$ equipped with a symplectic form. Let $\mathcal{M}^{2n}$ denote the collection of open subsets of $\mathbb{R}^{2n}$ and let Ω^{2n} denote the set of symplectic forms on $\mathbb{R}^{2n}$. A symplectic capacity is a map

$$c : \mathcal{M}^{2n} \times \Omega^{2n} \to [0, \infty] ,$$

satisfying the following axioms.

Definition 1.3. A map c is called a *symplectic capacity* if it has the following properties:

(a) ***Monotonicity:*** If $\varphi : (U_1, \omega) \to (U_2, \tau)$ is an injective symplectic map, then $c(U_1, \omega) \leq c(U_2, \tau)$.

(b) ***Conformality:*** For $\alpha \in \mathbb{R} \setminus \{0\}$,

$$c(U_1, \alpha\omega) = |\alpha|\, c(U_1, \omega).$$

(c) ***Normalization and nontriviality:*** In the standard space $(\mathbb{R}^{2n}, \omega_0)$,

$$c(B_1^{2n}, \omega_0) = \pi = c(Z_1^{2n}, \omega_0).$$

Several immediate consequences are worth highlighting. First, if $\varphi : (U_1, \omega) \to (U_2, \tau)$ is a symplectic diffeomorphism, then applying monotonicity to φ and φ^{-1} yields $c(U_1, \omega) = c(U_2, \tau)$; thus capacities are symplectic invariants.

The main reason capacities are useful is that they turn existence questions for symplectic embeddings into inequalities. If the embedding $\varphi : U_1 \overset{s}{\hookrightarrow} U_2$ is symplectic, then by monotonicity

$$c(U_1, \omega_0) \leq c(U_2, \omega_0).$$

In addition to conformality in the symplectic form, capacities are homogeneous of degree 2 under Euclidean scaling of subsets.

Lemma 1.4 ([52])**.** *Let $U \subset \mathbb{R}^{2n}$ be open and let ω be a symplectic form on $\mathbb{R}^{2n}$. Then for every $\alpha \in \mathbb{R} \setminus \{0\}$ and every symplectic capacity c,*

$$c(\alpha U, \omega) \;=\; \alpha^2\, c(U, \omega).$$

Proof. Define $\varphi : \alpha U \to U$ by

$$\varphi(x) = \frac{1}{\alpha}\, x \qquad \text{for all } x \in \alpha U.$$

Clearly, φ is a diffeomorphism and its Jacobian is constant:

$$D\varphi(x) = \frac{1}{\alpha}\, I \qquad \text{for all } x \in \alpha U.$$

Hence for every $x \in \alpha U$ and all $u, v \in \mathbb{R}^{2n}$ we have

$$\omega(u, v) \;=\; \alpha^2\, \omega\big(D\varphi(x)u,\ D\varphi(x)v\big).$$

In other words, φ is a symplectic diffeomorphism from $(\alpha U, \omega)$ onto $(U, \alpha^2\omega)$. Therefore, by monotonicity applied in both directions (i.e. with equality),

$$c(\alpha U, \omega) = c(U, \alpha^2\omega).$$

Finally, conformality gives $c(U, \alpha^2\omega) = |\alpha^2|\, c(U, \omega) = \alpha^2 c(U, \omega)$, which proves the claim. □

Another important note is, that the translation $\varphi : \mathbb{R}^{2n} \to \mathbb{R}^{2n}$ defined by

$$\varphi(x) = x + u,$$

with $u \in \mathbb{R}^{2n}$ is a symplectic diffeomorphism. Indeed, both φ and φ^{-1} are differentiable and their Jacobians are equal to the identity matrix. In particular, φ is symplectic. Consequently, by applying such a translation if necessary, we may assume without loss of generality that $0 \in U_1$ (respectively, that the set under consideration contains the origin).

The existence of capacities implies the "if" direction of Gromov's non-squeezing theorem:

Proof of Gromov's non-Squeezing Theorem (Theorem 1.2): If $R \geq r$, then $B^{2n}(r) \subset Z^{2n}(R)$, hence the inclusion (equivalently, the identity map on $\mathbb{R}^{2n}$ restricted to $B^{2n}(r)$) is a symplectic embedding.

Conversely, assume there exists an injective symplectic map $\varphi : B^{2n}(r) \xhookrightarrow{s} Z^{2n}(R)$. Let c be any symplectic capacity on $(\mathbb{R}^{2n}, \omega_0)$. By monotonicity of c,

$$c\big(B^{2n}(r), \omega_0\big) \ \leq \ c\big(Z^{2n}(R), \omega_0\big).$$

Using normalization and the 2-homogeneity under scaling, we have

$$c\big(B^{2n}(r), \omega_0\big) = \pi r^2 \qquad \text{and} \qquad c\big(Z^{2n}(R), \omega_0\big) = \pi R^2.$$

Therefore $\pi r^2 \leq \pi R^2$, which implies $r \leq R$. □

Two basic examples of symplectic capacities, closely tied to Gromov's non-squeezing theorem, are the *Gromov width*

$$c_G(U) \coloneqq \sup\Big\{\pi r^2 : \ \text{ There exists } \ B^{2n}(r) \ \xhookrightarrow{s} \ U\Big\},$$

and the *cylindrical capacity*

$$c_Z(U) \coloneqq \inf\Big\{\pi R^2 : \ \text{ There exists } \ U \ \xhookrightarrow{s} \ Z^{2n}(R)\Big\},$$

where r is a non-negative real number and U an open subset of $\mathbb{R}^{2n}$.

2 The Ekeland–Hofer–Zehnder Capacity

Up to this point we have discussed the axioms of a symplectic capacity, but not yet a concrete construction. Since Gromov's foundational work [35], many symplectic capacities have been introduced from different perspectives in symplectic geometry and Hamiltonian dynamics. Prominent examples include the Hofer–Zehnder capacity of Hofer and Zehnder [51], the Ekeland–Hofer capacities of Ekeland and Hofer [27], Viterbo's symplectic homology capacity [93], the Gutt–Hutchings capacities [41], Hutchings' embedded contact homology capacities in dimension 4 [57], and capacities arising from rational symplectic field theory as developed by Siegel [88]. For convex bodies $K \subset \mathbb{R}^{2n}$, i.e. compact convex sets with nonempty interior, several of these constructions in fact coincide. More precisely, the Hofer–Zehnder capacity, the first Ekeland–Hofer capacity, Viterbo's symplectic homology capacity, and the first Gutt–Hutchings capacity agree on convex bodies, see for instance, [51, 27, 59, 1, 41, 46, 42]. We refer to this common value as the *Ekeland–Hofer–Zehnder capacity* and denote it by c_{EHZ}.

Since our focus is c_{EHZ}, we omit the general definitions of the other capacities and introduce c_{EHZ} directly. To define c_{EHZ} for a convex body $K \subset \mathbb{R}^{2n}$, we use the (outer) normal cone at boundary points. The brackets $\langle \cdot, \cdot \rangle$ denote the standard Euclidean inner product in the relevant Euclidean space: for $a, b \in \mathbb{R}^n$ it is $\langle a, b \rangle = a^\mathsf{T} b$ and for $A, B \in \mathbb{R}^{n \times n}$ it is $\langle A, B \rangle = \operatorname{tr}(A^\mathsf{T} B)$. The associated norm is always the one induced by the chosen inner product, i.e. $\|x\| := \sqrt{\langle x, x \rangle}$ in the relevant space; in particular, $\|a\| = \sqrt{a^\mathsf{T} a}$ for $a \in \mathbb{R}^n$ and $\|A\| = \|A\|_F = \sqrt{\operatorname{tr}(A^\mathsf{T} A)}$ for $A \in \mathbb{R}^{n \times n}$.

Definition 2.1. Let $K \subset \mathbb{R}^m$ be convex and let $y \in \partial K$. The *(outer) normal cone* of K at y is

$$N_K(y) \ := \ \{\, \nu \in \mathbb{R}^m : \langle \nu, z - y \rangle \leq 0 \text{ for all } z \in K \,\}.$$

When ∂K is smooth, $N_K(y)$ is a ray spanned by the outward unit normal.

From now on we identify the standard symplectic space with

$$\mathbb{R}^{2n} \cong \mathbb{R}^n_q \times \mathbb{R}^n_p,$$

where $q = (q_1, \ldots, q_n)$ denotes the position variables and $p = (p_1, \ldots, p_n)$ the momentum variables. In these coordinates the standard symplectic form is

$$\omega_0 = \sum_{i=1}^n dp_i \wedge dq_i,$$

and the Liouville 1-form is

$$\lambda = \sum_{i=1}^{n} p_i \, dq_i.$$

Besides ω_0 and the induced volume form, there is another quantity that is invariant under symplectic diffeomorphisms: The action. If $\gamma : [0,1] \to \mathbb{R}^{2n}$ is a piecewise C^1 curve, the integral of a 1-form along γ is defined by

$$\int_\gamma \lambda \;\coloneqq\; \int_0^1 \lambda_{\gamma(t)}\big(\dot\gamma(t)\big)\, dt,$$

which is independent of the chosen parametrization. Writing $\gamma(t) = (q(t), p(t))$, this becomes the familiar line integral

$$\int_\gamma \lambda \;=\; \int_0^1 \sum_{i=1}^{n} p_i(t)\, \dot q_i(t)\, dt.$$

Definition 2.2. If γ is a closed curve in $\mathbb{R}^{2n}$, its *action* is

$$\mathcal{A}(\gamma) \;\coloneqq\; \int_\gamma \lambda.$$

Proposition 2.3 ([52])**.** *Let γ be a closed curve in $\mathbb{R}^{2n}$ and let $\varphi : \mathbb{R}^{2n} \to \mathbb{R}^{2n}$ be a symplectic diffeomorphism. Then*

$$\mathcal{A}(\gamma) \;=\; \mathcal{A}(\varphi(\gamma)).$$

Definition 2.4. Let $K \subset \mathbb{R}^{2n}$ be a *convex body.* A *generalized closed characteristic* on ∂K is an absolutely continuous map $\gamma : [0,1] \to \partial K$ such that

$$\gamma(0) = \gamma(1) \qquad \text{and} \qquad J\dot\gamma(t) \in N_K(\gamma(t)) \text{ for almost every } t \in [0,1],$$

where J denotes the standard symplectic matrix on $\mathbb{R}^{2n}$.

In the smooth case this reduces to the usual notion of a closed characteristic on a hypersurface; here we adopt the convention that “closed characteristic” always means the generalized version unless smoothness is explicitly assumed.

Definition 2.5. Let $K \subset \mathbb{R}^{2n}$ be a convex body. The *Ekeland–Hofer–Zehnder capacity* of K is

$$c_{\mathrm{EHZ}}(K) \;\coloneqq\; \min\big\{\mathcal{A}(\gamma) \,:\, \gamma \text{ is a generalized closed characteristic on } \partial K\big\}.$$

By convention, $\inf \emptyset \coloneqq +\infty$. The fact that the minimum is attained whenever $K \neq \emptyset$ is nontrivial and relies on variational methods; see [27, 52].

For smooth convex sets, there is also an equivalent maximization principle—often the more convenient formulation in applications—due to Abbondandolo and Majer [2].

Theorem 2.6 ([2])**.** *Let $K \subset \mathbb{R}^{2n}$ be a convex set with smooth boundary and $0 \in \mathrm{int}(K)$. Let K° denote its polar:*

$$K^\circ := \{\, v \in \mathbb{R}^{2n} : \langle v, z\rangle \leq 1 \textit{ for all } z \in K \,\}.$$

Then

$$\frac{1}{4\, c_{\mathrm{EHZ}}(K)} \;=\; \max\Big\{ \mathcal{A}(\xi) \;\; : \xi : [0,1] \to \mathbb{R}^{2n} \textit{ absolutely continuous,} \\ \xi(0) = \xi(1),\; \dot{\xi}(t) \in K^\circ \textit{ almost everywhere} \Big\}.$$

2.1 The Counterexample to the Viterbo Conjecture

A prominent conjecture on convex bodies is Viterbo's volume–capacity inequality [94], which has driven much of the development of symplectic capacities and their links with classical convex geometry. Let $K \subset \mathbb{R}^{2n}$ be a convex body. We define the *systolic ratio* of K by

$$\mathrm{sys}_n(K) \;:=\; \frac{c_{\mathrm{EHZ}}(K)}{\big(n!\ \mathrm{Vol}(K)\big)^{1/n}},$$

that is, the EHZ capacity normalized by the ω-volume of K. In particular, $\mathrm{sys}_n(B) = 1$ for every Euclidean ball $B \subset \mathbb{R}^{2n}$.

Viterbo Conjecture 2.7 ([94])**.** *Among all convex bodies in $\mathbb{R}^{2n}$ with prescribed volume, the Euclidean ball maximizes every symplectic capacity. Implying, for $K \subset \mathbb{R}^{2n}$ a convex body it is conjectured that*

$$\mathrm{sys}_n(K) \leq 1.$$

A variety of results lend support to Conjecture 2.7: Since Viterbo formulated his conjectural volume–capacity inequality, it has generated a large body of literature, much of which points toward its validity in important cases. The conjecture is known to hold up to a dimension-independent constant by work of Artstein-Avidan, Milman, and Ostrover [7]. Beyond perturbative results, it has been verified for several concrete families, including certain Lagrangian products $A \times B \subset \mathbb{R}^n_q \times \mathbb{R}^n_p \cong \mathbb{R}^{2n}$, where $A \subset \mathbb{R}^n_q$ and $B \subset \mathbb{R}^n_p$, [11, 82], monotone toric domains [42], and domains arising from Hamiltonians of classical mechanical type [62].

A key observation of Artstein-Avidan, Karasev and Ostrover ([5]) is that for centrally symmetric K one has

$$c_{\mathrm{EHZ}}(K \times K^\circ) = 4.$$

If Viterbo's conjectural volume–capacity inequality held for $D = K \times K^\circ \subset \mathbb{R}^{2n}$, then using $c_{\mathrm{EHZ}}(B^{2n}(1)) = \pi$ and $\mathrm{Vol}(B^{2n}(1)) = \pi^n/n!$ we would obtain

$$\left(\frac{4}{\pi}\right)^n \leq \frac{\mathrm{Vol}(K \times K^\circ)}{\mathrm{Vol}(B^{2n}(1))} = \frac{\mathrm{Vol}(K)\,\mathrm{Vol}(K^\circ)}{\pi^n/n!},$$

and therefore

$$\mathrm{Vol}(K)\,\mathrm{Vol}(K^\circ) \;\geq\; \frac{4^n}{n!},$$

which is precisely the symmetric Mahler conjecture. In particular, although Viterbo's conjecture is false in general, this special case remains compatible with the still-open symmetric Mahler conjecture in dimensions $n \geq 4$ (it is known for $n = 2$ and $n = 3$ [60]).

Recently there was a counterexample to the Viterbo conjecture found by Pazit Haim-Kislev and Yaron Ostrover [48] along polytopes, which are constructed as Lagrangian products of regular pentagons in $\mathbb{R}^2$. Using results from their previous paper [47], where they discussed bounds on the systolic ratio on p-products of polytopes, they could extend their counterexample to the Viterbo conjecture in $\mathbb{R}^4$ to counterexamples in $\mathbb{R}^{2n}$ for any $n \in \mathbb{N}_{\geq 2}$. In the following we will introduce the background needed to construct the counterexample of Viterbo's conjecture.

Theorem 2.8 ([48])**.** *Viterbo's conjecture is false in every dimension $2n$ with $n \geq 2$.*

The calculation of the EHZ capacity of the polytope $K \times T$, which yields the counterexample is based on the connection of the EHZ capacity to the computation of Minkowski billiards which are tightly related to the EHZ capacity of convex bodies ([6]). In the following we will briefly introduce the background needed for the proof of Theorem 2.8 and then will go through the proof in more detail.

Definition 2.9. Let $K \subset \mathbb{R}^{2n}$ and $T \subset \mathbb{R}^{2m}$ be convex bodies containing the origin, and let $1 \leq p \leq \infty$. The *p-product* of K and T is the convex body

$$K \times_p T \;:=\; \bigcup_{0 \leq t \leq 1} \Big((1-t)^{1/p} K \times t^{1/p} T \Big) \;\subset\; \mathbb{R}^{2n} \times \mathbb{R}^{2m}.$$

Theorem 2.10 ([47]). *For convex bodies $K \subset \mathbb{R}^{2n}$, $T \subset \mathbb{R}^{2m}$, and $1 \le p \le \infty$,*

$$\mathrm{sys}_{n+m}(K \times_p T)^{m+n} \;\le\; \mathrm{sys}_n(K)^n \, \mathrm{sys}_m(T)^m,$$

where equality holds if and only if $c_{\mathrm{EHZ}}(K) = c_{\mathrm{EHZ}}(T)$ and $p = 2$.

Definition 2.11. Let $T \subset \mathbb{R}^n$ be a convex body. The *support function* of T is

$$h_T(u) \;\coloneqq\; \sup\{\langle x, u\rangle : x \in T\}, \qquad u \in \mathbb{R}^n.$$

Definition 2.12. Let $T \subset \mathbb{R}^n$ be a convex body. We define a (possibly asymmetric) *distance function* $d_{T^\circ} : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ by

$$d_{T^\circ}(x, y) \;\coloneqq\; h_T(x - y).$$

This notation is motivated by the fact that the polar body T° is the unit ball with respect to d_{T°. In particular, if T is the Euclidean unit ball, then d_{T° coincides with the standard Euclidean metric.

Definition 2.13. A closed polygonal curve with vertices $q_1, \dots, q_m$, $m \in \mathbb{N}_{\ge 2}$, on ∂K is a *closed T-billiard trajectory* in K if there are points $p_1, \dots, p_m$ on ∂T such that for every $1 \le j \le m$,

$$\begin{cases} q_{j+1} - q_j \in N_T(p_j), \\ p_{j+1} - p_j \in -\,N_K(q_{j+1}). \end{cases}$$

Definition 2.14. Let $K, T \subset \mathbb{R}^n$ with $n \ge 2$ be convex bodies, and let γ be a closed T-billiard trajectory in K represented by a polygonal curve with vertices $q_1, q_2, \dots, q_m \in \partial K$, where we set $q_{m+1} \coloneqq q_1$. The *T°-length* of γ is defined by

$$\mathrm{Length}_{T^\circ}(\gamma) \;\coloneqq\; \sum_{i=1}^{m} h_T(q_{i+1} - q_i),$$

where h_T denotes the support function of T.

The vertices of the polygonal curve will also be called *bounces*.

We will use the following theorem as a black box and do not revisit its proof here; instead, we focus on how it enables the explicit counterexample in Theorem 2.8.

Theorem 2.15 ([83], [6]). *Denote by $\mathcal{P}(K)$ the set of all polygonal curves which cannot be translated into the interior of K:*

$$\mathcal{P}(K) \coloneqq \{(q_1, \dots, q_m) \ : \ \{q_1, \dots, q_m\} \not\subset \operatorname{int}(K) + t, \quad \textit{for all } t \in \mathbb{R}^n\}.$$

For any two convex bodies $K \subset \mathbb{R}^n_q$ and $T \subset \mathbb{R}^n_p$ one has

$$\begin{aligned} c_{\mathrm{EHZ}}(K \times T) &= \min\{A(\gamma) : \gamma \textit{ is a closed characteristic of } K \times T\} \\ &= \min\{\mathrm{Length}_{T^\circ}(\eta) : \eta \textit{ is a closed } T\textit{-billiard in } K\} \\ &= \min\{\mathrm{Length}_{T^\circ}(\eta) : \eta \in \mathcal{P}(K) \textit{ has at most } n+1 \textit{ vertices}\}. \end{aligned}$$

Now we will go through the proof of Theorem 2.8 in more detail. First we introduce the concrete counterexample which is a polytope constructed as a Lagrangian product $K \times T$. As a concrete geometric setup, consider the decomposition

$$\mathbb{R}^4 = \mathbb{R}^2_q \oplus \mathbb{R}^2_p.$$

Let $K \subset \mathbb{R}^2_q$ be the regular pentagon with vertices

$$\Big\{v_k = \big(\cos(2\pi k/5), \sin(2\pi k/5)\big)\Big\}_{k=0}^{4},$$

and let $T \subset \mathbb{R}^2_p$ be the image of K under rotation by $90°$, i.e. T has vertices

$$\Big\{w_k = \big(\cos(-\pi/2 + 2\pi k/5), \sin(-\pi/2 + 2\pi k/5)\big)\Big\}_{k=0}^{4}.$$

Figure 1 illustrates the two pentagons in the (q_1, q_2)- and (p_1, p_2)-planes.

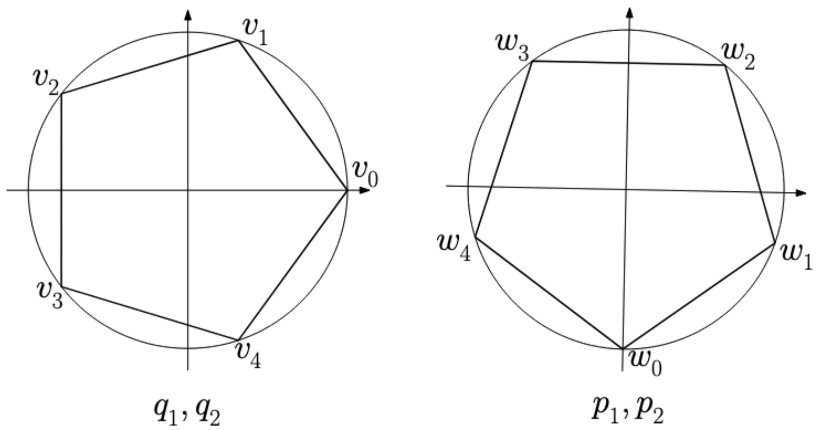


Figure 1: The regular pentagon $K \subset \mathbb{R}^2_q$ and its 90°-rotation $T \subset \mathbb{R}^2_p$. Reproduced from [48, Fig. 1].

Proposition 2.16. *The Ekeland–Hofer–Zehnder capacity of the Lagrangian product $K \times T$ is realized by a 2-bounce T-billiard trajectory running along any diagonal of K, and equals the corresponding T°-length. In particular,*

$$c_{\mathrm{EHZ}}(K \times T) \;=\; 2\cos\Big(\frac{\pi}{10}\Big)\Big(1+\cos\Big(\frac{\pi}{5}\Big)\Big).$$

Proof. As explained above, and in particular in Theorem 2.15, the problem of computing $c_{\mathrm{EHZ}}(K \times T)$ reduces to determining the smallest T°-length among periodic T-billiard trajectories in K. Moreover, it is enough to restrict attention to trajectories with two or three bounce points that cannot be translated into the interior of K.

Recall that, by definition of the support function and our specific choice of T, the T°-length of a vector $v \in \mathbb{R}^2_q$ is computed via

$$h_T(v) \;=\; \max_{0\le k\le 4} \langle v, w_k\rangle.$$

We will, using these facts, compute the EHZ capacity of this specific polygon by working through the possible 2-bounce and 3-bounce trajectories. T is a rotation of K by $\frac{-\pi}{2}$, which we will denote by $T = R_{-\pi/2}K$. We index the vertices of the pentagon cyclically, that is, $v_{k+5} = v_k$ for all $k \in \mathbb{Z}$.

For $a, b \in \partial K$, we write $a \leftrightarrow b$ for the closed polygonal billiard trajectory with bounces alternating between a and b, that is, the periodic 2-bounce orbit

$$a \to b \to a \to b \to \cdots .$$

Thus, for instance, $v_k \leftrightarrow v_{k+2}$ denotes the 2-bounce trajectory between the vertices v_k and v_{k+2}, where the indices are understood modulo 5.

We will dissect the proof into two claims:

Claim 2.17. *The minimal* 2*-bounce trajectory has* T°*-length*

$$2\cos\Big(\frac{\pi}{10}\Big)\Big(1+\cos\Big(\frac{\pi}{5}\Big)\Big)$$

and is realized by a diagonal of K*, namely by any trajectory of the form*

$$v_k \leftrightarrow v_{k+2}, \qquad k \in \{0,1,2,3,4\},$$

with indices understood modulo 5.

Proof of the Claim. An easy computation yields the T°-length of the 2-bounce trajectory of $v_1 \leftrightarrow v_3$. It is left to prove there is no such trajectory with lower T°-length.

A 2-bounce trajectory that cannot be moved to the interior of K needs to connect a vertex v_k and a point $\tilde{q}$ on the opposite edge $[v_{k+2}, v_{k+3}]$ modulo 5:

Suppose a 2-bounce orbit has both bounce points in the relative interiors of two edges E_1, E_2 of K. At an interior point of an edge, the outer normal cone N_K is a single ray, so for such a 2-bounce orbit to be not translatable into $\operatorname{int}(K)$ the two supporting outer normals must be opposite, which forces E_1 and E_2 to be parallel. However, a regular pentagon has no parallel edges. Hence there is no edge-to-edge tight 2-bounce orbit in K. After excluding edge-to-edge tight orbits, one bounce point must be a vertex v_k. If the second bounce point q lies in the relative interior of an edge E with outward normal n_E, then a translation by a vector $t \in \mathbb{R}^2$ moves *both* bounce points into the interior of K whenever t points strictly inward at both v_k and q. At q this means $\langle n_E, t\rangle < 0$; at the vertex v_k it means $\langle n, t\rangle < 0$ for both edge normals n adjacent to v_k. Such a common inward direction exists unless n_E lies in the opposite of the normal cone at v_k, that is, unless $-n_E \in N_K(v_k)$. In a regular pentagon this happens exactly when E is the edge opposite to v_k, namely $[v_{k+2}, v_{k+3}]$ (indices modulo 5). For $\lambda \in [0,1]$, set

$$q_e(\lambda) \coloneqq \lambda v_{k+2} + (1-\lambda) v_{k+3} \in [v_{k+2}, v_{k+3}].$$

Since $T = R_{-\pi/2}K$ and K is regular, the vectors $v_k - v_{k+2}$ and $v_k - v_{k+3}$ are mapped by the rotation $R_{-\pi/2}$ to the outward normals at w_{k+1}. Equivalently, these directions lie in $N_T(w_{k+1})$ and the opposite directions lie in $N_T(w_{k-1})$, so the support-function maximizers stay fixed along the whole edge parametrization. Hence, for all $\lambda \in [0,1]$ one has the support-function identities

$$h_T(v_k - q_e(\lambda)) = \langle v_k - q_e(\lambda), w_{k+1} \rangle, \qquad h_T(q_e(\lambda) - v_k) = \langle q_e(\lambda) - v_k, w_{k-1} \rangle,$$

since $h_T(u) = \max_j \langle u, w_j \rangle$ is linear on each normal cone $N_T(w_j)$ and

$$v_k - q_e(\lambda) = \lambda\,(v_k - v_{k+2}) + (1-\lambda)\,(v_k - v_{k+3}) \in N_T(w_{k+1}), \qquad q_e(\lambda) - v_k \in N_T(w_{k-1})$$

for all $\lambda \in [0,1]$, the maximizers do not switch, as illustrated in Figure 2, hence $h_T(v_k - q_e(\lambda)) = \langle v_k - q_e(\lambda), w_{k+1} \rangle$ and $h_T(q_e(\lambda) - v_k) = \langle q_e(\lambda) - v_k, w_{k-1} \rangle$.

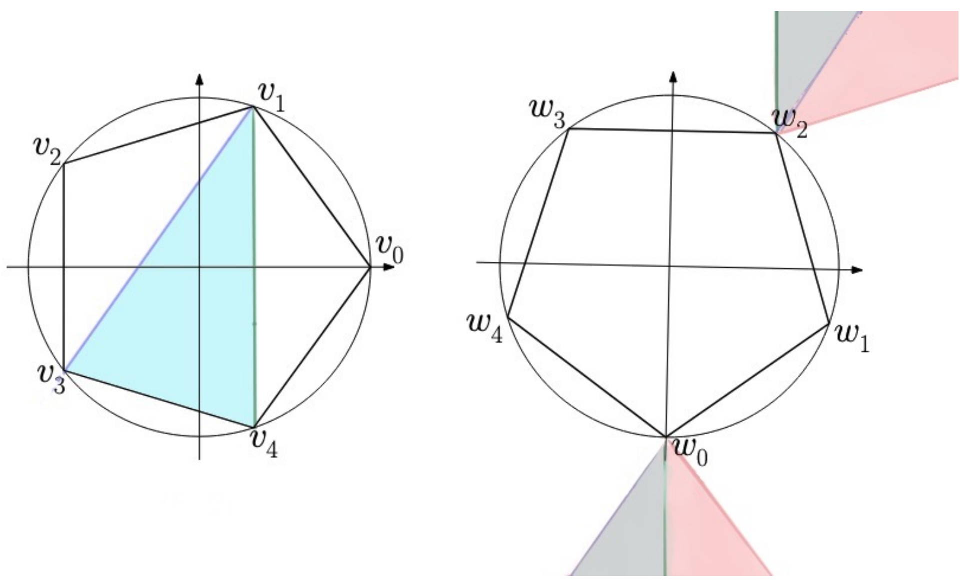


Figure 2: The vector $v_3 - v_1$ always stays in $N_T(w_0)$ and $v_1 - v_3$ in $N_T(w_2)$

Then the T°-length of the 2-bounce orbit $v_k \leftrightarrow q_e(\lambda)$ is

$$L(\lambda) := h_T(v_k - q_e(\lambda)) + h_T(q_e(\lambda) - v_k).$$

Using the assumed identities, we can write

$$L(\lambda) = \langle v_k - q_e(\lambda), w_{k+1}\rangle + \langle q_e(\lambda) - v_k, w_{k-1}\rangle.$$

Regrouping terms gives

$$L(\lambda) = \langle v_k, w_{k+1} - w_{k-1}\rangle + \langle q_e(\lambda), w_{k-1} - w_{k+1}\rangle.$$

Substituting $q_e(\lambda) = \lambda v_{k+2} + (1-\lambda)v_{k+3}$ yields

$$\begin{aligned} L(\lambda) &= \langle v_k, w_{k+1} - w_{k-1}\rangle + \lambda\langle v_{k+2}, w_{k-1} - w_{k+1}\rangle + (1-\lambda)\langle v_{k+3}, w_{k-1} - w_{k+1}\rangle \\ &= C + \lambda \big\langle v_{k+2} - v_{k+3},\, w_{k-1} - w_{k+1}\big\rangle, \end{aligned}$$

where C is a constant independent of λ. Thus $L(\lambda)$ is independent of λ once we show

$$\big\langle v_{k+3} - v_{k+2},\, w_{k+1} - w_{k-1}\big\rangle = 0.$$

Since $w_j = R_{-\pi/2} v_j$, we have

$$w_{k+1} - w_{k-1} = R_{-\pi/2}(v_{k+1} - v_{k-1}).$$

In a regular pentagon, the diagonal direction $v_{k+1} - v_{k-1}$ is parallel to the opposite edge direction $v_{k+3} - v_{k+2}$. Hence $w_{k+1} - w_{k-1}$ is obtained from a vector parallel to $v_{k+3} - v_{k+2}$ by a rotation by $-\pi/2$, and therefore

$$(v_{k+3} - v_{k+2}) \perp (w_{k+1} - w_{k-1}),$$

which implies the dot product above vanishes. This proves that $L(\lambda)$ does not depend on λ and one has that all such 2-bounce trajectories have the same T°-length as the diagonal connecting v_{k+3} and v_k, namely $2\cos\!\left(\frac{\pi}{10}\right)\left(1+\cos\!\left(\frac{\pi}{5}\right)\right)$. $\square$

Claim 2.18. *There is no 3-bounce trajectory in $K \times T$ with a smaller T°-length than $2\cos\!\left(\frac{\pi}{10}\right)\left(1+\cos\!\left(\frac{\pi}{5}\right)\right)$.*

Proof of the Claim. Assume the bounce points of the 3-point bounce trajectory γ to be x_1, x_2, x_3. Since it cannot be translated into the interior, and its three points are distributed along five edges, two of them lie on adjacent edges and the third one on an opposite edge. Since K and T are symmetric under rotation by $\frac{2\pi}{5}$, we may assume without loss of generality that the vertices lie on $[v_1, v_2]$, $[v_4, v_0]$, and $[v_3, v_4]$. Since in general $h_T(t) \neq h_T(-t)$, the clockwise and counterclockwise orientations are a priori distinct.

Fix x_1 and x_3 arbitrarily on their edges and set, for $\lambda \in [0, 1]$,

$$x_2^\lambda := \lambda v_4 + (1 - \lambda)v_0 \in [v_4, v_0].$$

Consider the clockwise family γ_λ with bounce points x_1, x_2^λ, x_3, whose T°-length is

$$\text{Length}_{T^\circ}(\gamma_\lambda) = h_T(x_2^\lambda - x_1) + h_T(x_3 - x_2^\lambda) + h_T(x_1 - x_3).$$

We will show that in the clockwise case a minimizer may be chosen with $x_2 = v_4$. The same conclusion holds for the counterclockwise orientation after the corresponding substitutions. We then conclude in both cases, by subadditivity of the support function, that no 3-point bounce trajectory can have smaller T°-length than the 2-point bounce trajectory running along a diagonal of K.

For a vertex w_k of the convex polygon T, the outer normal cone is

$$N_T(w_k) = \{u \in \mathbb{R}^2 : \ \langle u, w_k\rangle \geq \langle u, w_{k-1}\rangle \text{ and } \langle u, w_k\rangle \geq \langle u, w_{k+1}\rangle\}.$$

Equivalently,

$$u \in N_T(w_k) \iff \langle u, w_k - w_{k-1}\rangle \geq 0 \text{ and } \langle u, w_k - w_{k+1}\rangle \geq 0.$$

For such u, one has $h_T(u) = \langle u, w_k\rangle$.

Set $u(\lambda) := x_3 - x_2^\lambda$. Since $x_2^\lambda = \lambda v_4 + (1 - \lambda)v_0$,

$$u(\lambda) = \lambda(x_3 - v_4) + (1 - \lambda)(x_3 - v_0),$$

so $u(\lambda)$ is a convex combination of $u(1) = x_3 - v_4$ and $u(0) = x_3 - v_0$. A direct computation shows that $u(0), u(1) \in N_T(w_4)$. Since $N_T(w_4)$ is a convex cone, it follows that $u(\lambda) \in N_T(w_4)$ for all $\lambda \in [0, 1]$, and hence

$$h_T(x_3 - x_2^\lambda) = h_T(u(\lambda)) = \langle u(\lambda), w_4\rangle = \langle x_3 - x_2^\lambda, w_4\rangle \qquad \text{for all } \lambda \in [0, 1].$$

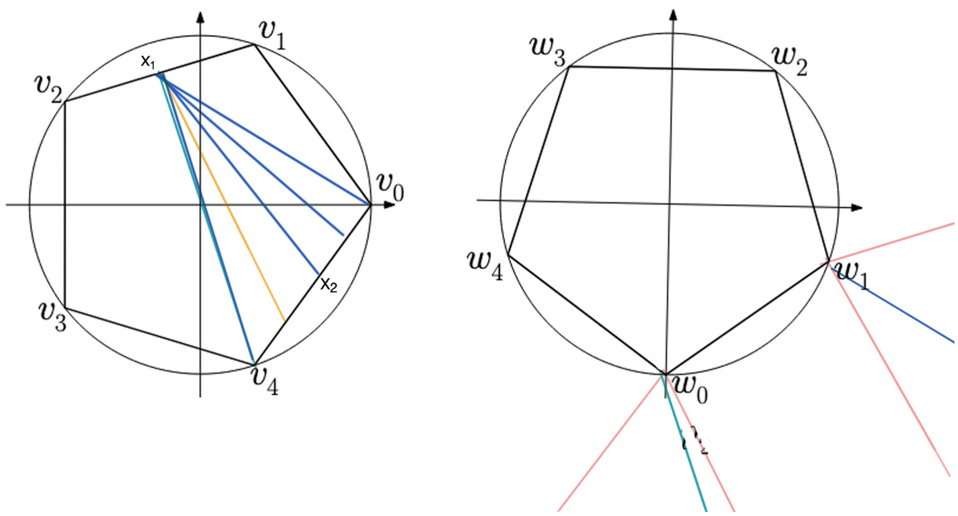


Figure 3: The vectors $x_2 - x_1$ are in $N_T(w_0)$ or $N_T(w_1)$, depending on λ. The orange vector corresponds to the breakpoint Λ.

Along the edge $[v_4, v_0]$ we parametrize $x_2^\lambda = \lambda v_4 + (1-\lambda) v_0$ and set

$$v(\lambda) \coloneqq x_2^\lambda - x_1.$$

Since $v(\lambda)$ depends affinely on λ, each function $\lambda \mapsto \langle v(\lambda), w_k \rangle$ is affine, hence

$$\lambda \longmapsto h_T(v(\lambda)) = \max_{0 \le k \le 4} \langle v(\lambda), w_k \rangle$$

is convex and piecewise affine. In the present configuration ($x_1 \in [v_1, v_2]$ and $x_2^\lambda \in [v_4, v_0]$) the direction $v(\lambda)$ stays inside the union of two adjacent normal cones $N_T(w_0) \cup N_T(w_1)$, as depicted in Figure 3, so only w_0 and w_1 can attain the maximum, and therefore

$$h_T(x_2^\lambda - x_1) = \max\{\langle x_2^\lambda - x_1, w_0 \rangle,\ \langle x_2^\lambda - x_1, w_1 \rangle\}.$$

Consequently there exists a $\Lambda \in [0, 1]$ such that

$$h_T(x_2^\lambda - x_1) = \begin{cases} \langle x_2^\lambda - x_1, w_1 \rangle, & \lambda \le \Lambda, \\ \langle x_2^\lambda - x_1, w_0 \rangle, & \lambda \ge \Lambda, \end{cases}$$

Now write

$$\mathrm{Length}_{T^\circ}(\gamma_\lambda) = h_T(x_2^\lambda - x_1) + h_T(x_3 - x_2^\lambda) + h_T(x_1 - x_3).$$

The last term is constant in λ, and from the previous step we have $h_T(x_3 - x_2^\lambda) = \langle x_3 - x_2^\lambda, w_4\rangle$ for all λ. Thus $\mathrm{Length}_{T^\circ}(\gamma_\lambda)$ is piecewise affine with a single breakpoint at Λ; differentiating and using $\frac{d}{d\lambda}x_2^\lambda = v_4 - v_0$ yields

$$\frac{d}{d\lambda}\mathrm{Length}_{T^\circ}(\gamma_\lambda) = \begin{cases} \langle v_4 - v_0, w_1 - w_4\rangle, & \lambda \le \Lambda, \\ \langle v_4 - v_0, w_0 - w_4\rangle, & \lambda \ge \Lambda. \end{cases}$$

Moreover $\langle v_4 - v_0, w_0 - w_4\rangle = 0$ since $w_0 - w_4 = R_{-\pi/2}(v_0 - v_4)$ is orthogonal to $v_4 - v_0$, while

$$\langle v_4 - v_0, w_1 - w_4\rangle < 0,$$

so $\mathrm{Length}_{T^\circ}(\gamma_\lambda)$ decreases on $[0, \Lambda]$ and is constant on $[\Lambda, 1]$. Hence a minimizer may be taken at $\lambda = 1$, i.e. $x_2 = v_4$, and the corresponding 3-bounce length is

$$\mathrm{Length}_{T^\circ}(\gamma) = h_T(v_4 - x_1) + h_T(x_3 - v_4) + h_T(x_1 - x_3).$$

For the counterclockwise orientation one argues analogously. Consider

$$\widetilde{\gamma}_\lambda := (x_1, x_3, x_2^\lambda), \qquad \text{where} \qquad x_2^\lambda = \lambda v_4 + (1-\lambda)v_0 \in [v_4, v_0].$$

Compared with the clockwise case, the roles of the two λ-dependent terms are interchanged: one checks directly that $x_1 - x_2^\lambda \in N_T(w_3)$ for all $\lambda \in [0,1]$, hence

$$h_T(x_1 - x_2^\lambda) = \langle x_1 - x_2^\lambda, w_3\rangle.$$

Likewise, $x_2^\lambda - x_3$ stays in the union $N_T(w_1) \cup N_T(w_2)$, so there exists $\widetilde{\Lambda} \in [0,1]$ such that

$$h_T(x_2^\lambda - x_3) = \begin{cases} \langle x_2^\lambda - x_3, w_2\rangle, & \lambda \le \widetilde{\Lambda}, \\ \langle x_2^\lambda - x_3, w_1\rangle, & \lambda \ge \widetilde{\Lambda}. \end{cases}$$

Therefore

$$\mathrm{Length}_{T^\circ}(\widetilde{\gamma}_\lambda) = h_T(x_3 - x_1) + h_T(x_2^\lambda - x_3) + h_T(x_1 - x_2^\lambda)$$

is again piecewise affine with a single breakpoint, and differentiating yields

$$\frac{d}{d\lambda}\,\mathrm{Length}_{T^\circ}(\widetilde{\gamma}_\lambda) = \begin{cases} \langle v_4 - v_0,\, w_2 - w_3\rangle, & \lambda \le \widetilde{\Lambda}, \\ \langle v_4 - v_0,\, w_1 - w_3\rangle, & \lambda \ge \widetilde{\Lambda}. \end{cases}$$

As in the clockwise case, the first quantity is negative and the second one is zero. Hence $\mathrm{Length}_{T^\circ}(\widetilde{\gamma}_\lambda)$ decreases on $[0, \widetilde{\Lambda}]$ and is constant on $[\widetilde{\Lambda}, 1]$, so a minimizer may again be taken at $\lambda = 1$, i.e. $x_2 = v_4$. Thus

$$\mathrm{Length}_{T^\circ}(\widetilde{\gamma}) = h_T(x_3 - x_1) + h_T(v_4 - x_3) + h_T(x_1 - v_4).$$

For the clockwise trajectory γ by subadditivity of h_T,

$$h_T(x_3 - v_4) + h_T(x_1 - x_3) \;\ge\; h_T\big((x_3 - v_4) + (x_1 - x_3)\big) = h_T(x_1 - v_4),$$

so $\mathrm{Length}_{T^\circ}(\gamma) \ge h_T(v_4 - x_1) + h_T(x_1 - v_4)$.

In the counterclockwise case with $x_2 = v_4$,

$$\mathrm{Length}_{T^\circ}(\widetilde{\gamma}) = h_T(x_3 - x_1) + h_T(v_4 - x_3) + h_T(x_1 - v_4),$$

and applying subadditivity to $h_T(x_3 - x_1) + h_T(v_4 - x_3) \ge h_T(v_4 - x_1)$ yields the same bound $\mathrm{Length}_{T^\circ}(\widetilde{\gamma}) \ge h_T(v_4 - x_1) + h_T(x_1 - v_4)$. This is exactly the T°-length of the 2-bounce orbit between x_1 and v_4. □

These claims yield the proof of the Proposition. □

Proof of Theorem 2.8. In the case $n = 2$, the areas of both K and T equal

$$A = \frac{5}{2}\sin\Big(\frac{2\pi}{5}\Big).$$

Hence Proposition 2.16 yields

$$\mathrm{sys}(K \times T) = \frac{c_{\mathrm{EHZ}}(K \times T)}{\sqrt{2A}} = \sqrt{\frac{\sqrt{5}+3}{5}} > 1.$$

To obtain counterexamples in higher dimensions, we use the multiplicativity property of the systolic ratio under the symplectic 2-product: if $X \subset \mathbb{R}^{2m}$ and $Y \subset \mathbb{R}^{2n}$ are convex domains with $c_{\mathrm{EHZ}}(X) = c_{\mathrm{EHZ}}(Y)$, then by Theorem 2.10

$$\mathrm{sys}(X \times_2 Y) = \mathrm{sys}(X)\,\mathrm{sys}(Y).$$

The general case follows by taking the symplectic 2-product of $K \times T$ with a Euclidean ball

$$B^{2n-4}\left(\sqrt{\frac{c_{\mathrm{EHZ}}(K \times T)}{\pi}}\right),$$

which has the same EHZ capacity as $K \times T$. □

2.2 Computational Results for the EHZ Capacity of Polytopes

Some notable results have been achieved in computing the EHZ capacity of polytopes. Krupp and Rudolf [65] explored the connections between Minkowski billiards and the EHZ capacity of polytopes in detail and constructed and implemented algorithms calculating the EHZ capacity of Lagrangian products $K \times T$ ([64]). Chaidez and Hutchings introduce a combinatorial model for the Reeb flow on the boundary of a 4-dimensional convex polytope and prove that Reeb orbits on smoothings of the polytope correspond to combinatorial Reeb orbits with the same action, yielding an explicit, algorithmically checkable formula for c_{EHZ} in this setting ([16]). Using this framework, they run computer experiments on k-vertex polytopes in $\mathbb{R}^4$ and report in particular that every 4-simplex satisfies $\mathrm{sys}_4(X) \leq 3/4$, with the apparent maximum $3/4$ attained by the standard simplex, while for $k \geq 6$ they find explicit examples with $\mathrm{sys}_4(X) = 1$, including 6- and 7-vertex families and a rotated 24-cell. In dimension 4 there are exactly four combinatorial types of 4-polytopes with $6 = d+2$ vertices; see Ziegler's book and the exposition by Henk, Richter-Gebert, and Ziegler [98, 49]. For the 6-vertex polytope P that attains systolic ratio 1, one can verify (via the unique affine dependence among its vertices) that P falls into the $(3,3,0)$-type in this classification. As explained by Henk, Richter-Gebert, and Ziegler, this is the unique neighborly type on 6 vertices; consequently P is combinatorially isomorphic to the cyclic polytope $C_4(6)$ [49].

In 2019, Haim-Kislev [44] proved a combinatorial formula for the EHZ capacity of convex polytopes. Since nearly all results in this thesis are based on this approach, we state the formula in detail. In the facet description $P = P(B, c)$, this can be rewritten as follows:

Theorem 2.19 ([44]). *Assume that a 2n-dimensional convex polytope P with k facets is given by linear inequalities*

$$P = P(B,c) = \{x \in \mathbb{R}^{2n} : Bx \le c\} \quad \text{for} \quad B \in \mathbb{R}^{k\times(2n)},\ c \in \mathbb{R}^k.$$

Let $b_1, \dots, b_k \in \mathbb{R}^{2n}$ be the rows of B, written as column vectors, and let $c_1, \dots, c_k$ be the components of c. So the polytope P contains exactly the points $x \in \mathbb{R}^{2n}$ which satisfy the linear inequalities $b_i^\mathsf{T} x \le c_i$, with $i = 1, \dots, k$. Then the EHZ capacity of P is

$$c_{\mathrm{EHZ}}(P) = \frac{1}{2}\left(\max\left\{ \sum_{1\le j<i\le k} \beta_{\sigma(i)}\beta_{\sigma(j)}\,\omega(b_{\sigma(i)}, b_{\sigma(j)}) \ : \ \sigma \in \mathfrak{S}_k,\ \beta \in \mathbb{R}^k_+, \beta^\mathsf{T} c = 1,\ \beta^\mathsf{T} B = 0 \right\}\right)^{-1}, \tag{4}$$

where ω is the standard symplectic form on $\mathbb{R}^{2n}$, where $\mathfrak{S}_k$ denotes the symmetric group of $\{1, \dots, k\}$, and where $\mathbb{R}^k_+$ is the nonnegative orthant.

The formula is combinatorial in the sense that it reduces the computation to a finite collection of quadratic maximization problems, where each maximization is taken over a convex polytope. Thus, rather than optimizing over an infinite-dimensional loop space, one is led to a finite-dimensional optimization problem.

The proof uses Clarke's dual action principle [18] together with rearrangement arguments for dual-action minimizers, and shows that there exists a minimal-action closed characteristic with a particularly simple description: its velocity is piecewise constant, each piece being of the form $C_j J n_i$ for a facet normal n_i and some $C_j > 0$, and each facet direction appears at most once [44].

Krupp used Haim-Kislev's formula to construct a lower bound for the EHZ capacity of polytopes ([64]): Krupp approached the computation by solving the optimization problem of Haim-Kislev in two parts: solving the quadratic program for each permutation separately and then finding the permutation that yields the maximum. The optimization problem for each $\sigma \in \mathfrak{S}_k$ is:

$$p^*_\sigma = \frac{1}{2}\left(\max\left\{ \sum_{1\le j<i\le k} \beta_{\sigma(i)}\beta_{\sigma(j)}\,\omega(b_{\sigma(i)}, b_{\sigma(j)}) \ : \ \beta \in \mathbb{R}^k_+, \beta^\mathsf{T} c = 1,\ \beta^\mathsf{T} B = 0 \right\}\right)^{-1}. \tag{5}$$

Reformulating this problem into a completely positive problem and using the positive semidefinite relaxation to obtain a program $\hat{p}_\sigma$

$$\frac{1}{4\,c_{\mathrm{EHZ}}(P)} \;=\; \max_{\sigma\in\mathrm{Sym}_k} p^*_\sigma \;\leq\; \max_{\sigma\in\mathrm{Sym}_k} \widehat{p}_\sigma, \tag{6}$$

yields an explicit computable upper bound on $1/(4c_{\mathrm{EHZ}}(P))$ and hence a corresponding lower bound on $c_{\mathrm{EHZ}}(P)$.

3 EHZ Capacity of Simplices

We focus on the special case of simplices, for which we can reduce the combinatorial formula from Theorem 2.19. Since translation is a symplectic map, the EHZ capacity is translation invariant, therefore we will assume all the simplices we are considering are centered at the origin.

We will prove that a simplex $S \subset \mathbb{R}^{2n}$, which is centered at the origin can be described by

$$S \coloneqq \{x \in \mathbb{R}^{2n} : Bx \leq \mathbf{e}\},$$

where $\mathbf{e} \in \mathbb{R}^{2n+1}$ is the all-ones vector, and the row vectors $b_1^\mathsf{T}, \ldots, b_{2n+1}^\mathsf{T}$ of $B \in \mathbb{R}^{(2n+1)\times 2n}$, sum to zero. The facet normalization $b_i^\mathsf{T} v_j = 1$ for all $j \neq i$ is standard for simplices: it identifies the facet normals with the vertices of the polar simplex (see, e.g., Grünbaum [38] or Ziegler [98]). In barycentric position, the identity $\sum_{i=0}^n b_i = 0$ then follows by a direct computation, which we include for completeness.

Proposition 3.1. *Let $P \subset \mathbb{R}^n$ be a full-dimensional simplex. Let*

$$c = \frac{1}{n+1} \sum_{i=0}^{n} p_i$$

be its barycenter (where $P = \operatorname{conv}\{p_0, \ldots, p_n\}$), and set $S \coloneqq P - c$. Then S admits a facet hyperplane description

$$S = \{x \in \mathbb{R}^n : Bx \leq \mathbf{e}\},$$

where $B \in \mathbb{R}^{(n+1)\times n}$ has rows $b_0^\mathsf{T}, \ldots, b_n^\mathsf{T}$ satisfying

$$\sum_{i=0}^{n} b_i = 0.$$

Proof. Write $S = \operatorname{conv}\{v_0, \ldots, v_n\}$ with $v_i \coloneqq p_i - c$. Then

$$\sum_{i=0}^{n} v_i = \sum_{i=0}^{n} (p_i - c) = \sum_{i=0}^{n} p_i - (n+1)c = 0.$$

Thus the translated simplex S is centered at the origin (its barycenter is 0).

For each $i \in \{0, \ldots, n\}$ consider the facet

$$F_i = \operatorname{conv}\{v_0, \ldots, v_{i-1}, v_{i+1} \ldots, v_n\}.$$

Since S is a simplex, each facet F_i is contained in a unique supporting hyperplane. We now choose the facet inequality in normalized form by requiring that there is a unique vector $b_i \in \mathbb{R}^n$ such that

$$b_i^\mathsf{T} v_j = 1 \qquad \text{for all } j \neq i. \tag{7}$$

Since $0 \in \operatorname{int}(S)$, the affine hull of $\{v_j\}_{j\neq i}$ does not contain 0, hence these n vectors are linearly independent, so (7) has a unique solution.

With this choice, $F_i \subset \{x \in \mathbb{R}^n : b_i^\mathsf{T} x = 1\}$ and $S \subset \{x \in \mathbb{R}^n : b_i^\mathsf{T} x \leq 1\}$. Stacking the inequalities gives $S = \{x \in \mathbb{R}^n : Bx \leq \mathbf{e}\}$, where B has rows b_i^T. This is exactly the normalization with right-hand side $\mathbf{e}$.

Fix i. Taking the dot product of $\sum_{j=0}^n v_j = 0$ with b_i yields

$$0 = b_i^\mathsf{T} \sum_{j=0}^{n} v_j = \sum_{j\neq i} b_i^\mathsf{T} v_j + b_i^\mathsf{T} v_i = n + b_i^\mathsf{T} v_i,$$

where we used (7) for the n indices $j \neq i$. Thus

$$b_i^\mathsf{T} v_i = -n. \tag{8}$$

Let $s := \sum_{i=0}^n b_i$. For any fixed vertex v_j we compute

$$s^\mathsf{T} v_j = \sum_{i=0}^{n} b_i^\mathsf{T} v_j = b_j^\mathsf{T} v_j + \sum_{i\neq j} b_i^\mathsf{T} v_j = (-n) + n = 0,$$

using (8) for the first term and (7) for the remaining n terms. Therefore s is orthogonal to all vertices $v_0, \ldots, v_n$. Since S is full-dimensional, these vertices span $\mathbb{R}^n$, hence $s = 0$ and

$$\sum_{i=0}^{n} b_i = 0. \qquad \square$$

Since the EHZ capacity is translation invariant, we can assume the simplices we will be working with to have the representation as stated in Proposition 3.1. This leads us to:

Lemma 3.2. *Let the simplex $S = \{x \in \mathbb{R}^{2n} : Bx \leq \mathbf{e}\}$ be of the form where row vectors $b_1, \dots, b_{2n+1}$ of $B \in \mathbb{R}^{(2n+1)\times 2n}$, written as column vectors, sum to zero. Then the computation of $c_{\mathrm{EHZ}}(P(B, \mathbf{e}))$ simplifies to*

$$c_{\mathrm{EHZ}}(S) = \frac{(2n+1)^2}{2} \left(\max \left\{ \sum_{1 \leq j < i \leq 2n+1} \omega(b_{\sigma(i)}, b_{\sigma(j)}) \ : \ \sigma \in \mathfrak{S}_{2n+1} \right\} \right)^{-1}. \tag{9}$$

Proof. Consider the maximization problem in (4) with $k = 2n+1$ and $c = \mathbf{e}$. By assumption $\sum_{i=1}^{2n+1} b_i = 0$, equivalently $\mathbf{e}^{\mathsf{T}} B = 0$, so $\mathbf{e}$ is feasible for the linear constraint $\beta^{\mathsf{T}} B = 0$.

Since S is a full-dimensional simplex in $\mathbb{R}^{2n}$, the facet description matrix $B \in \mathbb{R}^{(2n+1)\times 2n}$ has rank $2n$. Hence $\dim \ker(B^{\mathsf{T}}) = (2n+1) - 2n = 1$, so every β with $\beta^{\mathsf{T}} B = 0$ is a scalar multiple of $\mathbf{e}$. Writing $\beta = \lambda \mathbf{e}$ and using $\beta^{\mathsf{T}} c = 1$ with $c = \mathbf{e}$ gives

$$1 = \beta^{\mathsf{T}} \mathbf{e} = \lambda\, \mathbf{e}^{\mathsf{T}} \mathbf{e} = \lambda\,(2n+1),$$

hence $\lambda = \frac{1}{2n+1}$ and therefore $\beta = \frac{1}{2n+1}\mathbf{e}$. This yields (9). $\square$

We now derive an equivalent trace formulation, which is more convenient when we discuss differentiation later.

Lemma 3.3. *Let $S \subset \mathbb{R}^{2n}$ be a simplex given in facet hyperplane description*

$$S = \{x \in \mathbb{R}^{2n} : Bx \leq \mathbf{e}\}, \qquad B \in \mathbb{R}^{(2n+1)\times 2n}, \ \operatorname{rank}(B) = 2n,$$

where the row vectors of B sum to zero, and let J denote matrix corresponding to the standard symplectic form on $\mathbb{R}^{2n}$. Define

$$W := BJB^{\mathsf{T}} \in \mathbb{R}^{(2n+1)\times(2n+1)}.$$

Let $L \in \{0,1\}^{(2n+1)\times(2n+1)}$ be the strict upper-triangular matrix

$$L_{ij} = \begin{cases} 1, & i < j, \\ 0, & i \geq j, \end{cases}$$

and for $\sigma \in \mathfrak{S}_{2n+1}$ let $P_\sigma \in \{0,1\}^{(2n+1)\times(2n+1)}$ be the permutation matrix, i.e. $(P_\sigma)_{ij} = 1$ iff $i = \sigma(j)$, and

$$L_\sigma := P_\sigma L P_\sigma^{\mathsf{T}}.$$

Then

$$c_{\mathrm{EHZ}}(S) = \frac{(2n+1)^2}{2}\left(\max_{\sigma\in\mathfrak{S}_{2n+1}} \operatorname{tr}(W^{\mathsf{T}}L_\sigma)\right)^{-1} = \frac{(2n+1)^2}{2}\left(-\min_{\sigma\in\mathfrak{S}_{2n+1}} \operatorname{tr}(WL_\sigma)\right)^{-1}. \quad (10)$$

Proof. Fix $\sigma \in \mathfrak{S}_{2n+1}$ and write $W^\sigma := P_\sigma^{\mathsf{T}} W P_\sigma$. Consider the strict upper-triangular matrix L. Since $L_{ij} = 1$ if and only if $i < j$, we have

$$\operatorname{tr}\big((W^\sigma)^{\mathsf{T}} L\big) = \sum_{i,j=1}^{2n+1} (W^\sigma)^{\mathsf{T}}_{ij} L_{ji} = \sum_{i,j=1}^{2n+1} (W^\sigma)_{ji} L_{ji} = \sum_{1\le j<i\le 2n+1} (W^\sigma)_{ij}.$$

Thus,

$$\sum_{1\le j<i\le 2n+1} (P_\sigma^{\mathsf{T}} W P_\sigma)_{ij} = \operatorname{tr}\big((P_\sigma^{\mathsf{T}} W P_\sigma)^{\mathsf{T}} L\big).$$

Using $(P_\sigma^{\mathsf{T}} W P_\sigma)^{\mathsf{T}} = P_\sigma^{\mathsf{T}} W^{\mathsf{T}} P_\sigma$ and cyclicity of trace,

$$\operatorname{tr}\big((P_\sigma^{\mathsf{T}} W P_\sigma)^{\mathsf{T}} L\big) = \operatorname{tr}\big(P_\sigma^{\mathsf{T}} W^{\mathsf{T}} P_\sigma L\big) = \operatorname{tr}\big(W^{\mathsf{T}} P_\sigma L P_\sigma^{\mathsf{T}}\big) = \operatorname{tr}\big(W^{\mathsf{T}} L_\sigma\big).$$

This yields

$$\max_{\sigma\in\mathfrak{S}_{2n+1}} \sum_{1\le j<i\le 2n+1} (P_\sigma^{\mathsf{T}} W P_\sigma)_{ij} = \max_{\sigma\in\mathfrak{S}_{2n+1}} \operatorname{tr}(W^{\mathsf{T}} L_\sigma).$$

Finally, since W is skew-symmetric, $W^{\mathsf{T}} = -W$, we obtain

$$\max_{\sigma\in\mathfrak{S}_{2n+1}} \operatorname{tr}(W^{\mathsf{T}} L_\sigma) = \max_{\sigma\in\mathfrak{S}_{2n+1}} \operatorname{tr}(-W L_\sigma) = -\min_{\sigma\in\mathfrak{S}_{2n+1}} \operatorname{tr}(W L_\sigma),$$

which proves (10).

□

In the following lemma we use the well-known order-reversal symmetry for antisymmetric matrices, which holds in arbitrary dimension; see [25] and Fiorini's description of the full automorphism group in [30].

Lemma 3.4. *Let $W \in \mathbb{R}^{m\times m}$ be skew-symmetric ($W^\top = -W$) and for $\sigma \in \mathfrak{S}_m$ set*

$$S(\sigma) := \sum_{1\le j<i\le m} (P_\sigma^\top W P_\sigma)_{ij}.$$

Then the set of attainable values $\{S(\sigma) : \sigma \in \mathfrak{S}_m\}$ is symmetric about 0: for every σ there exists a permutation τ with

$$S(\tau) = -S(\sigma).$$

In particular,

$$\max_{\sigma \in \mathfrak{S}_m} S(\sigma) = -\min_{\sigma \in \mathfrak{S}_m} S(\sigma) = \max_{\sigma \in \mathfrak{S}_m} |S(\sigma)|.$$

For our purposes this means that, in all subsequent formulas, it is immaterial whether we write a maximum or a minimum, whether we sum over the strictly lower or strictly upper triangle, or whether we replace W by $-W$ (equivalently $W^\top$): all these choices only flip the sign of the objective, and the absolute optimal value is the same.

Proof. Let $\rho \in \mathfrak{S}_m$ be the reversal permutation $\rho(k) \coloneqq m + 1 - k$, and let $\tau \coloneqq \sigma \circ \rho$. Then $P_\tau = P_\sigma P_\rho$ and hence

$$P_\tau^\top W P_\tau = P_\rho^\top (P_\sigma^\top W P_\sigma) P_\rho.$$

Write $Q \coloneqq P_\sigma^\top W P_\sigma$, which is again skew-symmetric. Conjugation by P_ρ simply reverses indices, so for all i, j we have

$$(P_\rho Q P_\rho^\top)_{ij} = Q_{\rho(i),\rho(j)}.$$

Therefore

$$S(\tau) = \sum_{1 \le j < i \le m} (P_\rho Q P_\rho^\top)_{ij} = \sum_{1 \le j < i \le m} Q_{\rho(i),\rho(j)}.$$

Now set $i' \coloneqq \rho(i)$ and $j' \coloneqq \rho(j)$. Since ρ reverses order, the condition $j < i$ is equivalent to $i' < j'$. Hence

$$S(\tau) = \sum_{1 \le i' < j' \le m} Q_{i',j'}.$$

Because Q is skew-symmetric and $Q_{ii} = 0$, we have

$$\sum_{1 \le i' < j' \le m} Q_{i',j'} = -\sum_{1 \le j' < i' \le m} Q_{i',j'} = -S(\sigma),$$

which proves the claim. $\square$

Part II

Computational Complexity in Symplectic Geometry

In this part we show that computing the Ekeland–Hofer–Zehnder (EHZ) capacity of convex polytopes is NP-hard. In particular, unless P = NP, there is no polynomial-time algorithm that computes c_{EHZ} exactly on general polytopes.

We begin in Section 4 by recalling the notion of NP-hardness and computational complexity. Our reduction is from the problem of finding a minimum feedback arc set in a directed bipartite tournament. Accordingly, Section 5 introduces the necessary combinatorial notation and reviews the NP-hardness proof for feedback arc set on bipartite tournaments.

Section 6 reproduces the published proof from Leipold and Vallentin [67]. Readers who are primarily interested in the NP-hardness of computing c_{EHZ} for polytopes may proceed directly to Section 6. Readers seeking additional background on complexity theory and on the feedback arc set reduction should start with Sections 4 and 5; some overlap with Section 6 is therefore unavoidable. In addition, Section 7 studies the computational complexity of deciding affine symplectomorphism between full simplices by relating it to weighted graph isomorphism via polynomial-time reductions.

4 Computational Complexity

In order to make precise what it means for a computational problem to be "hard", we briefly review the standard framework of computational complexity, following the first Chapter of the book of Arora and Barak [4, Chapter 1]. We fix a Turing-machine model, introduce the classes P and NP, and explain polynomial-time reductions and NP-hardness, which are the main tools used in our hardness proofs.

4.1 Turing Machines

In this thesis we fix a standard abstract model of computation: the (deterministic) multi-tape Turing machine (TM). The point of committing to a concrete model is not that TMs are the most convenient to program, but that they provide a clean, machine-independent way to define what it means to be computable and, more importantly for complexity theory, what it means to be efficiently computable. We identify a computational problem with its associated decision problem (language): given an input x, decide whether x belongs to a language $L \subseteq \{0,1\}^*$. Here $\{0,1\}^*$ denotes the set of all finite binary strings (including the empty string), i.e.

$$\{0,1\}^* \coloneqq \bigcup_{n\geq 0} \{0,1\}^n.$$

When convenient, we simply denote the decision problem by L.

Formally, a TM M is described by a tuple (Γ, Q, δ) containing:

- A set Γ of the symbols that M's tapes can contain. We assume that Γ contains a designated blank symbol, denoted $\square$, a designated start symbol, denoted B, and the numbers 0 and 1. We call Γ the alphabet of M.

- A set Q of possible states M's register can be in. We assume that Q contains a designated start state, denoted q_{start}, and a designated halting state, denoted q_{halt}.

- A function
$$\delta : Q \times \Gamma^{k+1} \to Q \times \Gamma^k \times \{L, R\}^{k+1}$$
describing the rule M uses in performing each step. This function is called the transition function of M, where k is the number of work tapes of the machine.

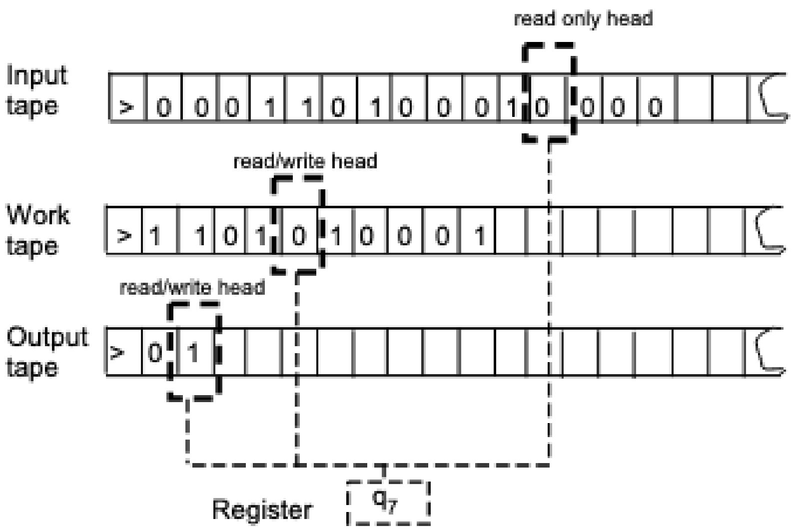


Figure 4: Snapshot of the execution of a multi-tape Turing machine with a read-only input tape and two read/write tapes (work and output). Source: Arora and Barak [4, Fig. 1.1].

In each step, the machine reads the $k+1$ scanned symbols (including the input tape), applies the transition function

$$\delta : Q \times \Gamma^{k+1} \to Q \times \Gamma^{k} \times \{L, R\}^{k+1},$$

updates its state, writes new symbols on the work tapes, and moves each head left (L) or right (R). The machine halts upon entering q_{halt} (and we interpret the output/answer according to a fixed convention, for example an accept/reject state or an output written on a designated tape). For intuition, Figure 4 illustrates a typical snapshot of a multi-tape Turing machine during execution, with a read-only input tape and two read/write tapes (work and output).

By the Church–Turing thesis, every effectively computable procedure can be carried out by a Turing machine; see, for example, Sipser [89, Section 1.2]. We therefore use Turing machines as our basic model of computation.

When we speak of running time, we measure it as the number of elementary TM steps executed until halting. For complexity-theoretic purposes, the particular choice of a standard computational model is largely inessential, since the usual reasonable models of computation simulate one another with at most polynomial overhead. Hence complexity classes such as P and NP are robust under the choice of model.

4.2 Complexity Classes

Complexity theory compares computational problems by the resources required to solve them, most notably running time as a function of the input length. This leads to robust complexity classes, such as P (polynomial-time solvable problems) and NP (problems whose solutions can be verified in polynomial time).

Definition 4.1. Let $T : \mathbb{N} \to \mathbb{N}$ be a function. We let $\mathrm{DTIME}(T(n))$ be the set of all Boolean (one-bit output) functions that are computable in $c \cdot T(n)$ time for some constant $c > 0$.

We often speak of *efficiently computable* when we refer to a problem with polynomial running time, which means it is at most n^c for some constant $c > 0$. We define the following class that captures the efficiently computable functions:

Definition 4.2.

$$\mathrm{P} = \bigcup_{c \geq 1} \mathrm{DTIME}(n^c).$$

Many standard algorithmic tasks are in P. For instance, Gaussian elimination solves linear systems, computes determinants, and, when the matrix is invertible, computes the inverse of an $n \times n$ matrix in at most $c_1 n^3$ arithmetic operations for some constant $c_1 > 0$. Similarly, Gram–Schmidt orthogonalization runs in polynomial time and, for an $n \times n$ input matrix, requires at most $c_2 n^3$ arithmetic operations for some constant $c_2 > 0$. See, for example, Golub and Van Loan [33, Chapters 3–4]. More broadly, sorting, shortest paths, and maximum flows admit polynomial-time algorithms; see, e.g., Cormen–Leiserson–Rivest–Stein [24, Chapters 2–6].

To grasp the idea of problems that are not efficiently computable, one can distinguish between different types of hardness. There are problems that are not decidable at all and there are problems that are solvable in exponential time, which is not considered efficient. A canonical undecidable example is the *Halting Problem*: given a Turing machine M and an input x, decide whether M eventually halts on x. Turing proved that no Turing machine can decide this problem in general [91]; see also Sipser [89, Chapter 4]. Another important phenomenon is that for many problems it is very difficult to find a solution, but it is easy to check whether a given solution is correct. This is captured by the complexity class NP. A common way to define NP is via polynomial-time verification:

Definition 4.3. For every $L \subseteq \{0,1\}^*$, we say that $L \in \mathrm{NP}$ if there exists a polynomial $p : \mathbb{N} \to \mathbb{N}$ and a polynomial-time Turing machine M such that for every $x \in \{0,1\}^*$,

$$x \in L \iff \text{there exists } u \in \{0,1\}^{p(|x|)} \text{ such that } M(x,u) = 1.$$

If $x \in L$ and $u \in \{0,1\}^{p(|x|)}$ satisfies $M(x,u) = 1$, then we say that u is a certificate for x (with respect to the language L and machine M).

Another equivalent definition uses nondeterministic Turing machines. Such a machine, instead of a single transition function δ, has two transition functions δ_0 and δ_1, and a special accept state denoted by q_{accept}. At each computation step M makes an arbitrary (not random) choice of which transition function to apply. For every input x we say $M(x) = 1$ if there exists a sequence of these nondeterministic choices that leads to q_{accept}. If every sequence of choices makes M halt without reaching q_{accept}, then $M(x) = 0$. We say M runs in $T(n)$ time if for every input $x \in \{0,1\}^*$, M reaches either the halting state or q_{accept} within $T(|x|)$ steps.

Definition 4.4. For every function $T : \mathbb{N} \to \mathbb{N}$ and $L \subseteq \{0,1\}^*$, we say that $L \in \mathrm{NTIME}(T(n))$ if there is a constant $c > 0$ and a $cT(n)$-time nondeterministic Turing machine (NDTM) M such that for every $x \in \{0,1\}^*$,

$$x \in L \iff M(x) = 1.$$

The complexity class NP can also be denoted by

$$\mathrm{NP} = \bigcup_{c \geq 1} \mathrm{NTIME}(n^c).$$

There is the easy to prove connection of the complexity classes: letting

$$\mathrm{EXP} \coloneqq \bigcup_{c \geq 1} \mathrm{DTIME}\big(2^{n^c}\big)\,,$$

we have

Theorem 4.5.

$$\mathrm{P} \subseteq \mathrm{NP} \subseteq \mathrm{EXP}.$$

A long-standing open question in complexity theory is whether $\mathrm{P} = \mathrm{NP}$. Informally, it asks whether every problem whose *solutions can be verified* in polynomial time also admits a polynomial-time algorithm to *find* such solutions. If $\mathrm{P} \neq \mathrm{NP}$, then there exist problems in NP that are not in P. It is widely believed that $\mathrm{P} \neq \mathrm{NP}$, this motivates the study of the "hardest" problems in NP, namely NP-complete problems.

Definition 4.6. We say that a language $A \subseteq \{0,1\}^*$ is polynomial-time Karp reducible to a language $B \subseteq \{0,1\}^*$ (sometimes shortened to just "polynomial-time reducible") and denote this by $A \leq_p B$ if there is a polynomial-time computable function

$$f : \{0,1\}^* \to \{0,1\}^*$$

such that for every $x \in \{0,1\}^*$,

$$x \in A \iff f(x) \in B.$$

We say that B is NP-hard if $A \leq_p B$ for every $A \in \mathrm{NP}$. We say that B is NP-complete if B is NP-hard and $B \in \mathrm{NP}$.

Theorem 4.7. *For polynomial-time reductions the following is true:*

(a) If $A \leq_p B$ and $B \leq_p C$, then $A \leq_p C$.

(b) If a language A is NP*-hard and $A \in \mathrm{P}$, then* $\mathrm{P} = \mathrm{NP}$.

(c) If a language A is NP*-complete, then $A \in \mathrm{P}$ if and only if* $\mathrm{P} = \mathrm{NP}$.

Some of the most basic NP-complete problems arise in propositional logic. A *Boolean formula* in variables $u_1, \ldots, u_n$ is built from these variables using the logical connectives AND ($\wedge$), OR ($\vee$), and NOT ($\neg$). For instance,

$$(a \wedge b) \;\vee\; (a \wedge c) \;\vee\; (b \wedge c)$$

evaluates to `True` exactly when a majority of a, b, c are `True`. Given a Boolean formula φ and an assignment $z \in \{0,1\}^n$ (where $1 \equiv$ `True` and $0 \equiv$ `False`), we write $\varphi(z)$ for the value of φ under this assignment. We call φ *satisfiable* if there exists some $z \in \{0,1\}^n$ with $\varphi(z) =$ `True`; otherwise, φ is *unsatisfiable*.

Definition 4.8 (CNF, kCNF)**.** A Boolean formula φ over variables $u_1, \ldots, u_n$ is in *conjunctive normal form* (*CNF*) if it is an AND of clauses, where each clause is an OR of literals. Equivalently,

$$\varphi \;=\; \bigwedge_i \left(\bigvee_j v_{ij} \right),$$

where each v_{ij} is a *literal*, that is, either a variable u_k or its negation $\neg u_k$. The expressions $\left(\bigvee_j v_{ij}\right)$ are called the *clauses* of φ. If every clause contains at most k literals, then φ is called a k*CNF* formula.

For example, the following CNF formula is a 3CNF formula:

$$(u_1 \vee \neg u_2 \vee u_3) \;\wedge\; (u_2 \vee \neg u_3 \vee u_4) \;\wedge\; (\neg u_1 \vee u_3 \vee \neg u_4).$$

Definition 4.9. SAT is the language of all satisfiable CNF formulas, and 3SAT is the language of all satisfiable 3CNF formulas.

Theorem 4.10 (Cook–Levin Theorem ([23])). SAT *is* NP*-complete.*

In the proof of the Cook–Levin theorem one shows that every language in NP can be reduced to SAT. Consequently, by transitivity of polynomial-time reductions, to prove that another language L is NP-complete it suffices to show (i) $L \in$ NP and (ii) SAT $\leq_p L$ (or, equivalently, 3SAT $\leq_p L$). More generally, to prove that a language L' is NP-hard it suffices to reduce to it some language L that is already known to be NP-hard.

The logic behind an NP-hardness proof for a language L' is therefore as follows. Choose a problem L which is known to be NP-hard and give a polynomial-time reduction from L to L'. Concretely, this means constructing a polynomial-time computable map $f : \{0,1\}^* \to \{0,1\}^*$ such that for every instance x,

$$x \in L \iff f(x) \in L'.$$

Equivalently, one describes how to decide L using a hypothetical algorithm for L': given an input x for L, first compute $f(x)$, then run the algorithm for L' on $f(x)$, and finally interpret the output as an answer for x. If L' were solvable efficiently (in polynomial time), then L would be solvable efficiently as well. Thus, unless P = NP, the existence of such a reduction is evidence that L' is NP hard.

5 NP-Hardness of the Feedback Arc Set Problem on Bipartite Tournaments

Since our goal is to obtain a polynomial-time reduction from the feedback arc set problem on bipartite tournaments to the EHZ capacity of polytopes, we dedicate this section to understanding the NP-hardness proof for this problem. We first recall, for reference, some standard notation from combinatorial optimization, which is also stated in Section 6.

5.1 Notation from Combinatorial Optimization

Let $D = (V, A)$ be a directed graph with vertex set V and arc multiset A, allowing for multiple arcs; see Schrijver [87]. Such a directed graph can be represented by its adjacency matrix M of size $|V| \times |V|$, where the entry M_{ij} equals the number of arcs from v_i to v_j.

As usual we define a *walk* in a directed graph to be a sequence of arcs $a_1, \dots, a_k \in A$ such that $a_i = (v_{i-1}, v_i)$ for $i = 1, \dots, k$. A walk is *closed* if the *start vertex* v_0 coincides with the *end vertex* v_k. A *path* is a walk in which the vertices $v_0, \dots, v_k$ are distinct. A *directed circuit* is a closed walk in which all vertices are distinct except that the start vertex v_0 coincides with the end vertex v_k.

A directed graph is called *acyclic* if it does not contain any directed circuits.

For a subset of the vertex set $U \subseteq V$ we define

$$\delta^{\text{out}}(U) = \{(u, v) \in A : u \in U, v \in V \setminus U\}$$

the multiset of arcs of D leaving U, and similarly

$$\delta^{\text{in}}(U) = \{(v, u) \in A : v \in V \setminus U, u \in U\}$$

the arcs entering U. If U consists of only one vertex v we also write $\delta^{\text{out}}(v)$ instead of $\delta^{\text{out}}(\{v\})$, and $\delta^{\text{in}}(v)$ instead of $\delta^{\text{in}}(\{v\})$. The *indegree* of a vertex v is $|\delta^{\text{in}}(v)|$ and its *outdegree* is $|\delta^{\text{out}}(v)|$, where we also count multiple arcs.

A walk in D is called *Eulerian* if each arc of D is traversed exactly once. The directed graph is called *Eulerian* whenever it has a closed Eulerian walk. A directed graph is Eulerian if and only if

(a) for any pair of vertices u, v there is a path with start vertex u and end vertex v, that is, D is *strongly connected*, and

(b) the indegree of every vertex coincides with its outdegree.

A *feedback arc set* of D is a subfamily of arcs which meets all directed circuits of D. That is, removing these arcs from D makes D acyclic. The decision variant of the *feedback arc set problem* asks: given a directed graph $D = (V, A)$ and a natural number δ, does D have a feedback arc set of cardinality at most δ?

The *maximum acyclic subgraph problem* is complementary to the feedback arc set problem. Its optimization variant asks for the maximum cardinality (counting multiplicities) of arcs $A' \subseteq A$ such that the resulting directed subgraph (V, A') of D is acyclic. A subfamily A' of A defines an acyclic subgraph of D if and only if $A \setminus A'$ is a feedback arc set of D.

A *bipartite tournament* is a directed graph $D = (U \cup V, A)$ whose vertex set is the disjoint union $U = \{u_1, \ldots, u_n\}$ and $V = \{v_1, \ldots, v_m\}$ and in which every arc has the property that for each pair (i, j) with $i = 1, \ldots, n$ and $j = 1, \ldots, m$, exactly one of $(u_i, v_j) \in A$ or $(v_j, u_i) \in A$ holds. Thus the underlying undirected graph is a complete bipartite graph without multiple edges.

5.2 Proof that FAS is NP-Hard on Bipartite Tournaments

Since we later reduce from the feedback arc set problem on bipartite tournaments to the EHZ capacity of polytopes, it is worth briefly recalling why this restricted variant is already NP-hard. For general directed graphs, the feedback arc set problem is known to be

NP-complete, as shown by Ausiello, D'Atri, and Protasi [8]. On the other hand, there are graph classes for which the problem becomes polynomial-time solvable: for reducible flow graphs, such an algorithm was given by Ramachandran [79]. Thus it is not immediate how the complexity behaves under structural restrictions. In particular, Guo, Hüffner, and Moser give a polynomial-time reduction from SAT, as defined in Definition 4.9 to the feedback arc set problem on bipartite tournaments [39]. To stay close to the ordering-based viewpoint that also underlies the reduction in Guo–Hüffner–Moser [39], and since this perspective will be useful again later, we briefly recall the equivalent formulation of the feedback arc set problem in terms of linear orders and backward arcs.

Definition 5.1. Let $V = \{v_1, \ldots, v_m\}$. A *linear order* on V is a strict total order $\prec$ on V, that is, a binary relation such that for all distinct $u, v, w \in V$:

(a) exactly one of $u \prec v$ or $v \prec u$ holds;

(b) if $u \prec v$ and $v \prec w$, then $u \prec w$.

For a directed graph $D = (V, A)$, a *topological ordering* is a linear order $\prec$ on V such that every arc points forward:

$$(u, v) \in A \quad \Longrightarrow \quad u \prec v.$$

Given a linear order $\prec$ on V, an arc $(u, v) \in A$ is called *backward* if $v \prec u$.

Remark 5.2. A digraph is acyclic if and only if it admits a topological ordering. Equivalently, asking for a feedback arc set of size at most δ is the same as asking for a linear order $\prec$ of the vertices with at most δ backward arcs, as observed for example by Brandenburg and Hanauer [15].

Theorem 5.3 ([39])**.** *The feedback arc set problem restricted to bipartite tournaments is* NP*-complete.*

Proof sketch (following Guo–Hüffner–Moser [39]). Membership in NP is immediate. For NP-hardness, we reduce from SAT. Let F be a CNF formula with variables $X = \{x_1, \ldots, x_n\}$ and clauses $C = \{C_1, \ldots, C_m\}$.

Construction. We describe the reduction by building the bipartite tournament $D = (U \cup V, A)$ in three layers. First we define the variable gadgets X_i (Figure 5). Second we introduce the clause vertices and their incidence pattern with the gadgets (Figure 6). Finally we add the connector sets H_t, whose role is to enforce global ordering constraints between the gadgets (Figure 7).

We build a bipartite tournament $D = (U \cup V, A)$ whose vertices come in four types: U-vertices, W-vertices, H-vertices, and clause vertices. For each variable x_i we create a gadget X_i consisting of

$$U_i = \{u_i^1, \ldots, u_i^6\} \subseteq U, \qquad W_i = \{w_i^1, \ldots, w_i^{3nm}\} \subseteq V.$$

Inside X_i, all arcs between U_i and W_i are oriented so that $u_i^1, u_i^2, u_i^3, u_i^6$ send arcs to every vertex of W_i, and every vertex of W_i sends arcs to u_i^4, u_i^5. Between different variable gadgets we orient all arcs so that, for $i < j$, every vertex of U_i points to every vertex of W_j, and every vertex of W_i points to every vertex of U_j. Figure 5 summarizes the orientation pattern inside a variable gadget X_i.

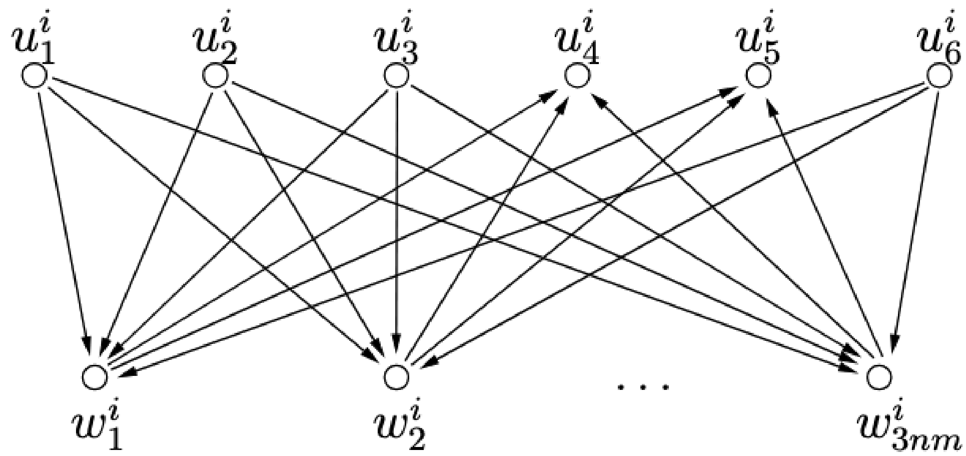


Figure 5: The variable gadget X_i and the orientation of arcs between U_i and W_i (and, schematically, between different gadgets). Reproduced from [39].

To control reachability between gadgets we also add, for each $1 \le t \le n-1$, a set $H_t = \{h_t^1, \dots, h_t^{3nm}\} \subseteq V$ and orient all arcs so that $U_1 \cup \dots \cup U_t$ points to H_t, and H_t points to $U_{t+1} \cup \dots \cup U_n$. The purpose of the sets H_t is to create many directed paths from earlier gadgets to later gadgets; this makes it difficult to mix the gadgets in a low-backward-arc ordering. See Figure 6.

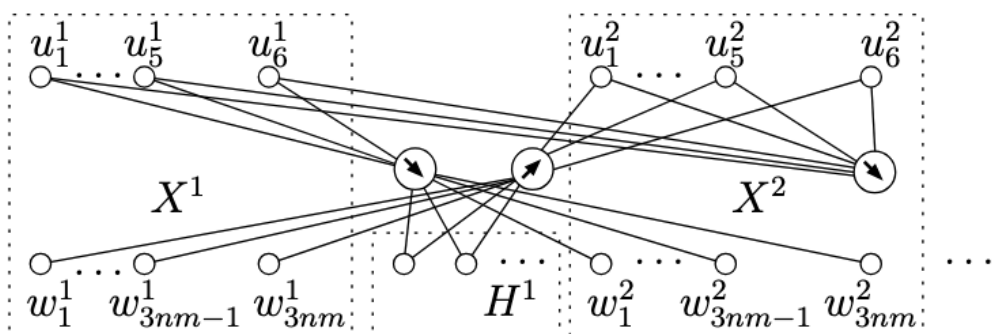


Figure 6: The connector sets H_t and the forced direction of arcs between earlier and later gadgets. Reproduced from [39].

Finally, for each clause C_ℓ we add a clause vertex $c_\ell \in V$, and orient arcs between c_ℓ and the six vertices of each U_j according to whether x_j occurs in C_ℓ positively, negatively, or not at all. The local arc pattern between a clause vertex c_ℓ and a variable gadget X_j depends only on this case; see Figure 7.

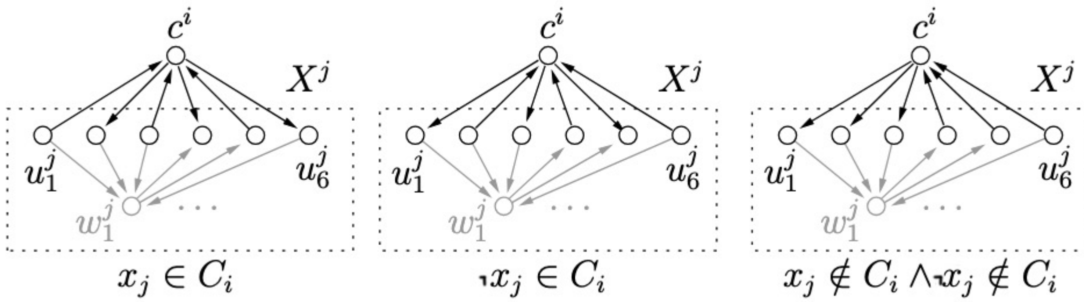


Figure 7: Arc patterns between a clause vertex c_ℓ and a variable gadget X_j, depending on whether x_j appears in C_ℓ positively, negatively, or not at all. Reproduced from [39].

Claim. F is satisfiable if and only if D has an ordering with at most $\delta := 3nm - 2m$ backward arcs (equivalently, a feedback arc set of size at most δ).

($\Longrightarrow$) From a satisfying assignment to a low-backward-arc ordering. If we delete the clause vertices $\{c_1, \ldots, c_m\}$, the remaining digraph is acyclic and admits the topological order

$$s(X_1) \prec r(H_1) \prec s(X_2) \prec \cdots \prec r(H_{n-1}) \prec s(X_n),$$

where $r(V)$ denotes an arbitrary ordering of the vertices in V and each $s(X_i)$ places $\{u_i^1, u_i^2, u_i^3, u_i^6\}$ before W_i and W_i before $\{u_i^4, u_i^5\}$. Given a satisfying assignment T, distribute the clauses among variables that satisfy them. For each variable x_i, insert the clause vertices assigned to x_i into $s(X_i)$ in one of two canonical positions depending on whether $T(x_i) = \texttt{true}$ or $T(x_i) = \texttt{false}$, so the insertion point lies between two pairs of the U_i-vertices. If a clause is satisfied by more than one variable under T, we choose one satisfying variable for the construction. Since the truth assignment satisfies the formula, each clause can be assigned in this way. A counting argument using the fixed arc pattern between c_ℓ and U_j shows that each clause vertex contributes exactly one backward arc with respect to the gadget of the chosen satisfying variable, and exactly three backward arcs with respect to each of the other $n-1$ gadgets. Hence the total number of backward arcs is

$$m \cdot (3(n-1) + 1) = 3nm - 2m = \delta.$$

($\Longleftarrow$) From a feedback arc set to a satisfying assignment. Assume that $A' \subseteq A$ is a feedback arc set with $|A'| \leq \delta$, and let $D_0 := D \setminus A'$ be the resulting acyclic digraph. We first argue that A' cannot delete any arcs except those incident to clause vertices. Indeed, by a counting argument, breaking all directed cycles already forces at least δ deletions among the arcs between clause vertices and variable gadgets. Hence $|A'| = \delta$, and no additional deletions are possible. In particular, every arc of A' is incident to some clause vertex.

Fix a clause vertex c_ℓ. The acyclicity of D_0 imposes a strong restriction on how c_ℓ can remain connected to the variable gadgets. Namely, in D_0 there can be at most one gadget X_j for which c_ℓ is still adjacent (in the underlying undirected sense) to more than three vertices of U_j; otherwise the H-layers provide many directed paths from U_j to $U_{j'}$, which together with the remaining adjacencies of c_ℓ would create a directed cycle through c_ℓ in D_0. Consequently, c_ℓ can keep at most $3(n-1)+5$ of its $6n$ arcs to variable gadgets in D_0.

Summing over all m clauses, this implies that destroying all cycles requires at least

$$6nm - (3nm + 2m) = 3nm - 2m = \delta$$

deletions among the clause–gadget arcs. Since $|A'| \leq \delta$, we must be in the extremal situation: for every clause vertex c_ℓ, the upper bound is attained. Equivalently, for each ℓ there is a *unique* variable gadget X_j such that c_ℓ is adjacent in D_0 to exactly 5 of the 6 vertices in U_j (and to at most 3 vertices in every other U_k). We say that c_ℓ is *linked* to X_j.

Using small collections of arc-disjoint cycles in the local c_ℓ–X_j pattern, one derives two key constraints enforced by the construction:

(a) a clause vertex cannot be linked to a gadget of a variable that does not occur in the clause, and

(b) if multiple clause vertices are linked to the same gadget X_j, then x_j occurs in all the corresponding clauses with a consistent polarity, either always as x_j or always as $\neg x_j$, since otherwise too many arc-disjoint cycles would remain and more than δ deletions would be necessary.

We can therefore define a truth assignment as follows. For each variable x_j that has at least one linked clause c_ℓ, set $T(x_j) = \texttt{true}$ if x_j occurs in C_ℓ, and set $T(x_j) = \texttt{false}$ if $\neg x_j$ occurs in C_ℓ. This is well-defined by (b). Finally, every clause C_ℓ is linked to some variable x_j that appears in C_ℓ with the corresponding polarity, so T satisfies every clause. Hence F is satisfiable. This completes the reduction and proves NP-completeness. $\square$

PROCEEDINGS OF THE
AMERICAN MATHEMATICAL SOCIETY, SERIES B
Volume 11, Pages 603–611 (November 20, 2024)
https://doi.org/10.1090/bproc/250

COMPUTING THE EHZ CAPACITY IS NP-HARD

KARLA LEIPOLD AND FRANK VALLENTIN

(Communicated by Isabella Novik)

ABSTRACT. The Ekeland–Hofer–Zehnder capacity (EHZ capacity) is a fundamental symplectic invariant of convex bodies. We show that computing the EHZ capacity of polytopes is NP-hard. For this we reduce the feedback arc set problem in bipartite tournaments to computing the EHZ capacity of simplices.

1. Introduction

Sympletic capacities are fundamental tools for measuring the "symplectic size" of subsets of the standard symplectic vector space $(\mathbb{R}^{2n}, \omega)$, cf. the book by Hofer and Zehnder [4]. Many symplectic capacities, starting with the Hofer–Zehnder capacity from [4] and the Ekeland–Hofer capacity from [1] have been proved to coincide on convex bodies and there common value is called the Ekeland–Hofer–Zehnder capacity and denoted as c_{EHZ}.

Our starting point is the following combinatorial formulation of the EHZ capacity of convex polytopes due to Haim-Kislev [3]. We assume that a $2n$-dimensional convex polytope P with k facets is given by linear inequalities

$$P = P(B, c) = \{x \in \mathbb{R}^{2n} : Bx \leq c\} \quad \text{for} \quad B \in \mathbb{R}^{k \times (2n)},\ c \in \mathbb{R}^k.$$

Let $b_1, \ldots, b_k \in \mathbb{R}^{2n}$ be the rows of B, written as column vectors, and let $c_1, \ldots, c_k$ be the components of c. So the polytope P contains exactly the points $x \in \mathbb{R}^{2n}$ which satisfy the linear inequalities $b_i^{\mathsf{T}} x \leq c_i$, with $i = 1, \ldots, k$. Then the *EHZ capacity* of P is

$$c_{\mathrm{EHZ}}(P) = \frac{1}{2}\left(\max\left\{ \sum_{1 \leq j < i \leq k} \beta_{\sigma(i)} \beta_{\sigma(j)}\, \omega(b_{\sigma(i)}, b_{\sigma(j)}) \ : \ \sigma \in \mathfrak{S}_k,\ \beta \in \mathbb{R}^k_+,\ \beta^{\mathsf{T}} c = 1,\ \beta^{\mathsf{T}} B = 0 \right\}\right)^{-1}, \tag{1}$$

where ω is the standard symplectic form on $\mathbb{R}^{2n}$,

$$\omega(x, y) = x^{\mathsf{T}} J y \quad \text{for} \quad x, y \in \mathbb{R}^{2n} \quad \text{and} \quad J = \begin{bmatrix} 0 & I_n \\ -I_n & 0 \end{bmatrix},$$

where I_n is the identity matrix with n rows and columns, where $\mathfrak{S}_k$ denotes the symmetric group of $\{1, \ldots, k\}$, and where $\mathbb{R}^k_+$ is the nonnegative orthant.

Hofer and Zehnder state [4, p. 102] "It turns out that it is extremely difficult to compute this capacity." In this paper we give a computational justification of this claim by showing that computing the EHZ capacity is indeed NP-hard, even for the special case of simplices.

Received by the editors February 21, 2024, and, in revised form, September 5, 2024, and October 4, 2024.

Theorem 1.1. *The following decision problem is* NP*-complete: Given the rational inputs $\gamma \in \mathbb{Q}$, a matrix $B \in \mathbb{Q}^{k\times(2n)}$, a vector $c \in \mathbb{Q}^k$. Suppose that $P(B,c)$ is a $2n$-dimensional convex polytope having k facets. Is the EHZ capacity of $P(B,c)$ at most γ? Is $c_{\mathrm{EHZ}}(P(B,z)) \leq \gamma$?*

To prove this theorem we shall reduce the feedback arc set problem in bipartite tournaments to computing the EHZ capacity of a simplex, that is, to $2n$-dimensional polytopes having $k = 2n+1$ facets. The feedback arc set problem in bipartite tournaments is known to be NP-complete; a result by Guo, Hüffner, and Moser [2].

1.1. **Notation from combinatorial optimization.** First we recall, for reference, some standard notation from combinatorial optimization. Let $D = (V, A)$ be a directed graph with vertex set V and with family of arcs A, allowing for multiple arcs; see Schrijver [8]. Such a directed graph can be represented by its adjacency matrix M of size $|V| \times |V|$, where the entry M_{ij} equals the number of arcs from v_i to v_j.

As usual, we define a *walk* in a directed graph to be a sequence of arcs $a_1, \dots, a_k \in A$ so that $a_i = (v_{i-1}, v_i)$ with $i = 1, \dots, k$. A walk is *closed* if *start vertex* v_0 coincides with *end vertex* v_k. A *path* is a walk in which vertices $v_0, \dots, v_k$ are distinct. A *directed circuit* is a closed walk in which all vertices are distinct except that the start vertex v_0 coincides with the end vertex v_k.

A directed graph is called *acyclic* if it does not contain any directed circuits.

For a subset of the vertex set $U \subseteq V$ we define

$$\delta^{\mathrm{out}}(U) = \{(u, v) \in A : u \in U, v \in V \setminus U\}$$

the multiset of arcs of D leaving U, and similarly

$$\delta^{\mathrm{in}}(U) = \{(v, u) \in A : v \in V \setminus U, u \in U\}$$

the arcs entering U. If U consists only out of one vertex v we also write $\delta^{\mathrm{out}}(v)$ instead of $\delta^{\mathrm{out}}(\{v\})$, and $\delta^{\mathrm{in}}(v)$ instead of $\delta^{\mathrm{in}}(\{v\})$. Then, the *indegree* of vertex v is $|\delta^{\mathrm{in}}(v)|$ and its *outdegree* is $|\delta^{\mathrm{out}}(v)|$, where we also count multiple arcs.

A walk in D is called *Eulerian* if each arc of D is traversed exactly once. The directed graph is called *Eulerian* whenever it has a closed Eulerian walk. A directed graph is Eulerian if and only if

(1) for any pair of vertices u, v there is a path with start vertex u and end vertex v; that is, D is *strongly connected*,
(2) the indegree of every vertex coincides with its outdegree.

A *feedback arc set* of D is a subfamily of the arcs which meets all directed circuits of D. That is, removing these arcs from D would make D acyclic. The decision variant of the *feedback arc set problem* asks given a directed graph $D = (V, A)$ and a natural number δ, does D have a feedback arc set of cardinality at most δ?

The *maximum acyclic subgraph problem* is complementary to the feedback arc set problem: Its optimization variant asks for the maximum cardinality (using multiplicities) of arcs $A' \subseteq A$ so that the resulting directed subgraph (V, A') of D is acyclic. A subfamily A' of A defines an acyclic subgraph of D if and only if $A \setminus A'$ is a feedback arc set of D.

A *bipartite tournament* is a directed graph $D = (U \cup V, A)$ whose vertex set is the disjoint union of vertices $U = \{u_1, \dots, u_n\}$ and $V = \{v_1, \dots, v_m\}$ and in which all of its arcs A have the property that for every pair i, j, with $i = 1, \dots, n$,

$j = 1, \dots, m$, either $(u_i, v_j) \in A$ or $(v_j, u_i) \in A$. So the underlying undirected graph is a complete bipartite graph without multiple edges.

1.2. **Outline of the polynomial time reduction: Proof of Theorem 1.1.** Now we outline our polynomial time reduction—the details are provided in the subsequent sections. We describe a polynomial time algorithm to decide if a given bipartite tournament has a feedback arc set of at most δ arcs by calling an oracle to solve the decision variant of the EHZ-capacity of a simplex. We will demonstrate that the feedback arc set of a complete bipartite tournament $D = (U \cup V, A)$ with $n = |U|$ and $m = |V|$, where we assume that $n \geq m$ can be computed by the formula

$$|\tilde{A}| - \frac{1}{2}\left(\left\lfloor \frac{(2n+1)^2}{2c_{\mathrm{EHZ}}(P(\tilde{B}, \mathbf{e}))} + \frac{1}{2} \right\rfloor + \Delta\right) - |\delta^{\mathrm{out}}(x_{2n+1})|. \tag{2}$$

In the following we will explain this formula in detail. In particular, we shall show that all appearing constants can be constructed in polynomial time. Hence, after solving (2) for $c_{\mathrm{EHZ}}(P(\tilde{B}, \mathbf{e}))$, by the NP-hardness of the decision variant of the feedback arc set problem for bipartite tournaments [2], the decision variant of the EHZ capacity of polytopes is also NP-hard.

In Section 2 we transform the complete bipartite tournament D into a simplex $P(\tilde{B}, \mathbf{e})$ through the block matrix $B \in \mathbb{Z}^{(2n+1)\times(2n)}$ defined by

$$B = \begin{bmatrix} I_n & 0 \\ 0 & S \\ -\mathbf{e}^{\mathsf{T}} & -\mathbf{e}^{\mathsf{T}} S \end{bmatrix},$$

where $\mathbf{e}^{\mathsf{T}} = (1 \dots 1)$ is the all-ones (row) vector, and $S \in \{-1, 0, 1\}^{n \times n}$ is defined componentwise by

$$S_{ij} = \begin{cases} 1 & \text{if } (u_i, v_j) \in A, \\ -1 & \text{if } (v_j, u_i) \in A, \\ 0 & \text{otherwise.} \end{cases} \tag{3}$$

For simplicity we suppose in this outline that the matrix S has full rank n. If the matrix S has not full rank, then we have to perform a slight perturbation of it, yielding $\tilde{B}$; we give details about this technical issue in Section 2. It follows that the polytope $P(B, \mathbf{e})$ is a simplex because the row vectors $b_1, \dots, b_{2n+1}$ of B sum up to zero; that is $\sum_{i=1}^{2n+1} b_i = 0$. This implies that the variable $\beta \in \mathbb{R}_+^{2n+1}$ in the maximization problem (1) is determined uniquely: $\beta = \frac{1}{2n+1}\mathbf{e}$. So the computation of $c_{\mathrm{EHZ}}(P(B, \mathbf{e}))$ simplifies to

$$\begin{aligned} & c_{\mathrm{EHZ}}(P(B, \mathbf{e})) \\ & \quad = \frac{(2n+1)^2}{2}\left(\max\left\{ \sum_{1 \leq j < i \leq 2n+1} \omega(b_{\sigma(i)}, b_{\sigma(j)}) \ : \ \sigma \in \mathfrak{S}_{2n+1} \right\}\right)^{-1}. \end{aligned} \tag{4}$$

Since $\omega(b_{\sigma(i)}, b_{\sigma(j)}) = b_{\sigma(i)}^{\mathsf{T}} J b_{\sigma(j)}$, in (4) we are interested in simultaneously permuting rows and columns of the matrix $W = BJB^{\mathsf{T}} \in \mathbb{R}^{(2n+1)\times(2n+1)}$ so that the sum of the entries of its lower triangular part ($i > j$) is maximized. The matrix

W has the following block form

$$W = \begin{bmatrix} I_n & 0 \\ 0 & S \\ -\mathbf{e}^\mathsf{T} & -\mathbf{e}^\mathsf{T} S \end{bmatrix} \begin{bmatrix} 0 & I_n \\ -I_n & 0 \end{bmatrix} \begin{bmatrix} I_n & 0 & -\mathbf{e} \\ 0 & S^\mathsf{T} & -S^\mathsf{T}\mathbf{e} \end{bmatrix} = \begin{bmatrix} 0 & S^\mathsf{T} & -S^\mathsf{T}\mathbf{e} \\ -S & 0 & S\mathbf{e} \\ \mathbf{e}^\mathsf{T} S & -\mathbf{e}^\mathsf{T} S^\mathsf{T} & 0 \end{bmatrix}.$$

In Section 3 we interpret the result of computing $c_{\mathrm{EHZ}}(P(B,\mathbf{e}))$ as given in (4) in terms of the feedback arc set problem. The matrix W is skew-symmetric; $W^\mathsf{T} = -W$. Take the nonnegative part of W to define the matrix M; $M_{ij} = \max\{0, W_{ij}\}$. Then $W = M - M^\mathsf{T}$ and for any permutation matrix P_σ, with $\sigma \in \mathfrak{S}_{2n+1}$, we have

$$P_\sigma^\mathsf{T} W P_\sigma = P_\sigma^\mathsf{T}(M - M^\mathsf{T})P_\sigma = 2P_\sigma^\mathsf{T} M P_\sigma - P_\sigma^\mathsf{T}(M + M^\mathsf{T})P_\sigma. \tag{5}$$

Hence, maximizing the sum of the entries of the lower triangular part of $P_\sigma^\mathsf{T} W P_\sigma$ is equivalent to maximizing the sum of the entries of the lower triangular part of $P_\sigma^\mathsf{T} M P_\sigma$, because $M + M^\mathsf{T}$ is a symmetric matrix and so the sum of the entries of the lower triangular part of $P_\sigma^\mathsf{T}(M + M^\mathsf{T})P_\sigma$ is independent of the permutation σ. We denote this latter sum by Δ, a constant which appears in (2).

The matrix M defines the adjacency matrix of an auxiliary directed graph on $2n+1$ vertices and $|\tilde{A}|$ arcs (appearing in (2)), which is almost the original complete bipartite tournament D. There are two essential differences between the new auxiliary graph and the original graph D: A minor difference occurs when $n > m$, then the new graph has $n - m$ isolated vertices, which simply can be ignored. A major difference is that there is an extra vertex x_{2n+1} which is connected to some vertices in $U \cup V$ by potentially multiple arcs. One feature of the extra vertex is that it makes the auxiliary graph Eulerian.

It is known (the argument is recalled in Section 3) that maximizing the sum of the entries of the lower triangular part of $P_\sigma^\mathsf{T} M P_\sigma$ amounts to solving the maximum acyclic subgraph problem on the auxiliary graph. Using this solution of the maximum acyclic subgraph problem one can determine, by complementarity, an optimal solution of the feedback arc set problem of the auxiliary graph.

Then in Section 4 we relate this optimal solution of the auxiliary graph to the feedback arc set problem of the original graph. For this we shall use the fact (Lemma 4.1) that feedback arc sets of the original graph and of the auxiliary graph differ by exactly $|\delta^{\mathrm{out}}(x_{2n+1})|$ (appearing in (2)) many arcs. □

2. From the feedback arc set problem to the EHZ capacity

Let $D = (U \cup V, A)$ be a complete bipartite tournament with $n = |U|$ and $m = |V|$, where we assume that $n \geq m$. Define the matrix $S \in \{-1, 0, 1\}^{n \times n}$ componentwise as described in (3). If the matrix S is not of full rank n we perform a small perturbation of the row vectors of S. For this let s_i be the ith row of S written as a column vector. Choose a rational $\varepsilon > 0$. Then a desired perturbation can be computed by the following steps, which are all easy, polynomial time, linear algebra computations (for the details we refer to the book by Schrijver [7]). Let $L \subseteq \mathbb{Q}^n$ be the subspace spanned by $s_1, \ldots, s_n$.

(1) Among the vectors $s_1, \ldots, s_n$ choose a basis of L. Without loss of generality we may assume, after reordering, that $s_1, \ldots, s_k$ is such a basis.

(2) Use Gram–Schmidt orthogonalization to compute an orthogonal basis of the orthogonal complement $L^\perp$. Let $t_{k+1}, \ldots, t_n \in \mathbb{Q}^n$ be such an orthogonal basis normalized by $\|t_{k+1}\|_\infty = \ldots = \|t_n\|_\infty = 1$, where $\|\cdot\|_\infty$ denotes the ℓ_∞-norm.

(3) Perturb the vectors $s_{k+1}, \ldots, s_n$ to

$$\tilde{s}_{k+1} = s_{k+1} + \varepsilon t_{k+1},\ \ldots,\ \tilde{s}_n = s_n + \varepsilon t_n.$$

The first k vectors stay unchanged: $\tilde{s}_1 = s_1, \ldots, \tilde{s}_k = s_k$.

In this way we get a perturbed matrix $\tilde{S}$ with rows $\tilde{s}_1, \ldots, \tilde{s}_n$. From this we construct the block matrix

$$\tilde{B} = \begin{bmatrix} I_n & 0 \\ 0 & \tilde{S} \\ -\mathbf{e}^{\mathsf{T}} & -\mathbf{e}^{\mathsf{T}}\tilde{S} \end{bmatrix}$$

The convex hull of $\{\tilde{b}_1, \ldots, \tilde{b}_{2n+1}\}$, where $\tilde{b}_i$ are the row vectors of $\tilde{B}$, is a simplex with the origin in its interior; by construction, the first $2n$ row vectors of $\tilde{B}$ are linearly independent and all $2n+1$ row vectors sum up to zero. The polar of this simplex is equal to $P(\tilde{B}, \mathbf{e})$. Therefore, $P(\tilde{B}, \mathbf{e})$ is a simplex (see, for instance, Ziegler [9, Chapter 2.3]).

In the following we will consider the EHZ capacity of this simplex $P(\tilde{B}, \mathbf{e})$. The principal computational task to determine $c_{\mathrm{EHZ}}(P(\tilde{B}, \mathbf{e}))$ is to determine the maximum

$$\max\left\{ \sum_{1 \le j < i \le 2n+1} (P_\sigma^{\mathsf{T}} \tilde{W} P_\sigma)_{ij} \ :\ \sigma \in \mathfrak{S}_{2n+1} \right\} \tag{6}$$

for $\tilde{W} = \tilde{B} J \tilde{B}^{\mathsf{T}}$ and where P_σ is the permutation matrix corresponding to the permutation $\sigma \in \mathfrak{S}_{2n+1}$. We have

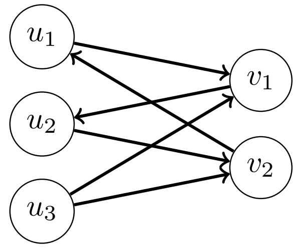
$$\tilde{W} = \begin{bmatrix} 0 & \tilde{S}^{\mathsf{T}} & -\tilde{S}^{\mathsf{T}}\mathbf{e} \\ -\tilde{S} & 0 & \tilde{S}\mathbf{e} \\ \mathbf{e}^{\mathsf{T}}\tilde{S} & -\mathbf{e}^{\mathsf{T}}\tilde{S}^{\mathsf{T}} & 0 \end{bmatrix}.$$

Later on, for interpretation of the EHZ capacity in terms of an (integral) adjacency matrix, it is more convenient to work with W rather than with $\tilde{W}$. Since all entries of $\tilde{W}$ differ from W only by at most $n\varepsilon$, we can choose $\varepsilon = 1/n^4$ so that

$$\left\lfloor \sum_{1 \le j < i \le 2n+1} (P_\sigma^{\mathsf{T}} \tilde{W} P_\sigma)_{ij} + \frac{1}{2} \right\rfloor = \sum_{1 \le j < i \le 2n+1} (P_\sigma^{\mathsf{T}} W P_\sigma)_{ij}$$

for every $\sigma \in \mathfrak{S}_{2n+1}$.

Example 2.1. We illustrate our construction for the following complete bipartite tournament.



Here

$$s_1 = (1, -1, 0),\ s_2 = (-1, 1, 0),\ s_3 = (1, 1, 0),$$

and

$$\tilde{s}_1 = s_1,\ \tilde{s}_2 = s_2,\ \tilde{s}_3 = (1, 1, \varepsilon),$$

therefore,

$$\tilde{W} = \left(\begin{array}{ccc|cccc|c} 0 & 0 & 0 & 1 & -1 & 1 & & -1 \\ 0 & 0 & 0 & -1 & 1 & 1 & & -1 \\ 0 & 0 & 0 & 0 & 0 & \varepsilon & & -\varepsilon \\ \hline -1 & 1 & 0 & 0 & 0 & 0 & & 0 \\ 1 & -1 & 0 & 0 & 0 & 0 & & 0 \\ -1 & -1 & -\varepsilon & 0 & 0 & 0 & & 2+\varepsilon \\ \hline 1 & 1 & \varepsilon & 0 & 0 & -2-\varepsilon & & 0 \end{array}\right).$$

3. Interpretation of the EHZ capacity

In this section we interpret the maximization problem

$$\max\left\{ \sum_{1\le j<i\le 2n+1} (P_\sigma^{\mathsf{T}} W P_\sigma)_{ij} \ : \ \sigma \in \mathfrak{S}_{2n+1} \right\}$$

in terms of the feedback arc set problem.

We define the matrix $M \in \mathbb{R}^{(2n+1)\times(2n+1)}$ by taking the nonnegative part of W so that $M_{ij} = \max\{0, W_{ij}\}$, and then $W = M - M^{\mathsf{T}}$. For any permutation matrix P_σ, with $\sigma \in \mathfrak{S}_{2n+1}$, we have

$$P_\sigma^{\mathsf{T}} W P_\sigma = P_\sigma^{\mathsf{T}}(M - M^{\mathsf{T}})P_\sigma = 2P_\sigma^{\mathsf{T}} M P_\sigma - P_\sigma^{\mathsf{T}}(M + M^{\mathsf{T}})P_\sigma.$$

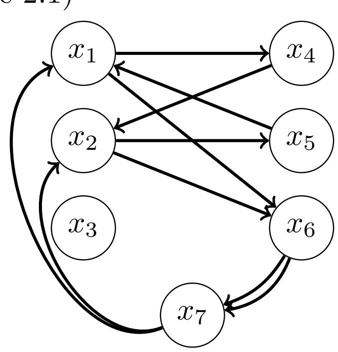
Hence, as noted in the introduction, maximizing the sum of the entries of the lower triangular part of $P_\sigma^{\mathsf{T}} W P_\sigma$ is equivalent to maximizing the sum of the entries of the lower triangular part of $P_\sigma^{\mathsf{T}} M P_\sigma$.

The matrix M defines the adjacency matrix of an auxiliary directed graph $\tilde{D}$ on $2n+1$ vertices, which is, as the example below illustrates, related to D by

(1) identifying $x_1, \dots, x_m$ with $v_1, \dots, v_m$,
(2) adding isolated vertices $x_{m+1}, \dots, x_n$,
(3) identifying $x_{n+1}, \dots, x_{2n}$ with $u_1, \dots, u_n$,
(4) adding the extra vertex x_{2n+1},
(5) reversing all arcs occurring in D.

Example 3.1. (first continuation of Example 2.1)

$$M = \left(\begin{array}{ccc|ccc|c} 0 & 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 2 \\ \hline 1 & 1 & 0 & 0 & 0 & 0 & 0 \end{array}\right)$$



Now we show that the maximization problem

$$\max\left\{ \sum_{1\le j<i\le 2n+1} (P_\sigma^{\mathsf{T}} M P_\sigma)_{ij} \ : \ \sigma \in \mathfrak{S}_{2n+1} \right\} \tag{7}$$

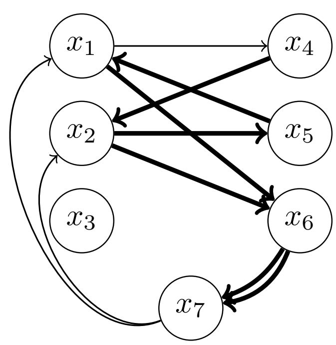
determines a maximum acyclic subgraph in $\tilde{D}$. This is a simple, known fact; the argument uses the concept of topological ordering (see, for instance, Knuth [5]).

Let $D = (V, A)$ be a directed graph with $V = \{v_1, \dots, v_n\}$. A *topological ordering* of D is a permutation $\sigma \in \mathfrak{S}_n$ so that $(v_i, v_j) \in A$ only if $\sigma^{-1}(i) < \sigma^{-1}(j)$, with $i, j = 1, \dots, n$. That is, the permutation σ determines a ranking of the vertices so that vertex v_i becomes ranked at the $\sigma^{-1}(i)$th position. Then there is an arc between v_i and v_j only if vertex v_i is ranked before vertex v_j. There exists a topological ordering of a directed graph if and only if the directed graph is acyclic. Therefore, finding a maximum acyclic subgraph of $\tilde{D}$ is the same as finding a permutation σ for which the number of arcs in $\tilde{D}$ respecting the topological ordering given by σ is maximized, which is

$$\begin{aligned}
&\max\Bigg\{ \sum_{\substack{i,j=1,\dots,n \\ \sigma^{-1}(i)<\sigma^{-1}(j)}} M_{ij} \;:\; \sigma \in \mathfrak{S}_{2n+1} \Bigg\} \\
&\qquad = \max\Bigg\{ \sum_{1\le i<j\le 2n+1} M_{\sigma(i),\sigma(j)} \;:\; \sigma \in \mathfrak{S}_{2n+1} \Bigg\} \\
&\qquad = \max\Bigg\{ \sum_{1\le i<j\le 2n+1} (P_\sigma^{\mathsf{T}} M P_\sigma)_{ij} \;:\; \sigma \in \mathfrak{S}_{2n+1} \Bigg\}.
\end{aligned}$$

To maximize the sum of the lower triangular part instead of the sum of the upper triangular part, that is summing over indices $1 \le j < i \le 2n+1$ instead of $1 \le i < j \le 2n+1$, we simply reverse all the arcs in $\tilde{D}$. This global reversing operation does not change the cardinality of the maximum acyclic subgraph of $\tilde{D}$. Therefore, the maximization problem (7) provides a solution of the maximum acyclic sugraph problem of the auxilary graph $\tilde{D}$.

Example 3.2 (Second continuation of Example 2.1)**.** A maximum acyclic subgraph of the auxiliary graph $\tilde{D}$ is depicted by thick arcs. It has cardinality 7.



The permutation

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 4 & 2 & 7 & 1 & 3 & 5 & 6 \end{pmatrix}$$

gives an optimal topological ordering with

$$P_\sigma^{\mathsf{T}} M P_\sigma = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 2 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

(Note that the permutation matrix P_σ is acting on the rows.)

4. Removing the extra vertex

In the last step we show how the solution of the maximization problem (7), which determines the value of a maximum acyclic subgraph of the auxiliary graph $\tilde{D}$, can be used to solve the feedback arc set problem for the original graph D.

If the extra vertex x_{2n+1} is isolated, then there is nothing to do, otherwise we will make use of the fact, that $\tilde{D}$ is Eulerian (after deleting the isolated vertices). To verify that $\tilde{D}$ is indeed Eulerian, one directly checks, for every vertex x_i, that $|\delta^{\text{in}}(x_i)| = |\delta^{\text{out}}(x_i)|$ holds:

$$
\begin{aligned}
&|\delta^{\text{out}}(x_i)| - |\delta^{\text{in}}(x_i)| \\
&\quad = \sum_{j=1}^{2n+1} M_{ij} - \sum_{j=1}^{2n+1} M_{ji} = e_i^{\mathsf{T}} M \mathbf{e} - \mathbf{e} M e_i^{\mathsf{T}} = e_i^{\mathsf{T}}(M - M^{\mathsf{T}})\mathbf{e} = e_i^{\mathsf{T}} W \mathbf{e} = 0,
\end{aligned}
$$

where e_i is the ith standard basis vector in $\mathbb{R}^{2n+1}$. Moreover, since D is complete, it follows that $\tilde{D}$ is strongly connected.

The following argument is from Perrot and Van-Pham [6, Section 3]. We provide the necessary details to make this paper self-contained.

Lemma 4.1. *The minimum cardinality of a feedback arc set in D equals the minimum cardinality of a feedback arc set in $\tilde{D}$ minus $|\delta^{\text{out}}(x_{2n+1})|$.*

Proof. We show the complementary result for the maximum acyclic subgraph problem.

Let A' be a maximum acyclic subgraph of D. Then we can use these arcs together with the arcs (x_{2n+1}, x_i), with $i = 1, \dots, 2n$, in $\tilde{D}$ leaving the extra vertex x_{2n+1} to determine an acyclic subgraph of $\tilde{D}$ of cardinality $|A'| + |\delta^{\text{out}}(x_{2n+1})|$.

Conversely, let A' be a maximum acyclic subgraph of $\tilde{D}$. Now we describe a procedure to determine a maximum acyclic subgraphs A'' so that the extra vertex x_{2n+1} has no incoming arc. Because A'' is maximum it has to contain all $|\delta^{\text{out}}(x_{2n+1})|$ arcs leaving x_{2n+1}. Removing these arcs determines an acyclic subgraph of the original graph D of cardinality $|A'| - |\delta^{\text{out}}(x_{2n+1})|$.

The procedure works as follows. We consider all vertices R which are reachable by walks from start vertex x_{2n+1} using only in arcs in A'. Define

$$A'' = \big(A' \setminus \delta^{\text{in}}(R)\big) \cup \delta^{\text{out}}(R).$$

Then A'' determines again a maximum acyclic subgraph of $\tilde{D}$ because by removing arcs in $\delta^{\text{in}}(R)$ and by adding arcs from $\delta^{\text{out}}(R)$ we do not introduce directed circuits. Furthermore, we have

$$|A''| \geq |A'| - |\delta^{\text{in}}(R)| + |\delta^{\text{out}}(R)|,$$

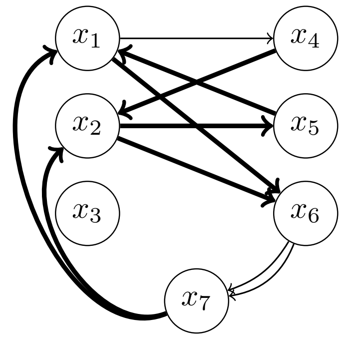
because $\delta^{\text{out}}(R)$, by the definition of R, cannot contain arcs from A'. The graph $\tilde{D}$ is Eulerian (ignoring the isolated vertices). So $|\delta^{\text{in}}(R)| = |\delta^{\text{out}}(R)|$ and hence $|A''| = |A'|$. The subset A'' cannot contain edges from $\delta^{\text{in}}(x_{2n+1})$ because A' is acyclic. □

Example 4.2. (Third continuation of Example 2.1) The additional vertex x_7 has two incoming arcs in the maximum acyclic subgraph A' of $\tilde{D}$.

Now we apply the procedure from the proof of Lemma 4.1 to determine a maximum acyclic subgraph in which x_7 has no incoming arcs. Using only arcs in A', the only vertex reachable from x_7 is x_7 itself. So $R = \{x_7\}$,

$$\delta^{\text{out}}(\{x_7\}) = \{(x_7, x_1), (x_7, x_2)\}, \quad \delta^{\text{in}}(\{x_7\}) = \{(x_6, x_7), (x_6, x_7)\},$$

and $A'' = (A' \setminus \delta^{\text{in}}(\{x_7\})) \cup \delta^{\text{out}}(\{x_7\})$ determines the following maximum acyclic subgraph in which all vertices but the isolated vertex x_3 are reachable from x_7.



Acknowledgments

The authors thank Arne Heimendahl and Andreas Spomer for helpful discussions. The authors also thank Alberto Abbondandolo for careful reading the first version of this paper and for his many valuable suggestions. The authors thank the anonymous referee for the helpful comments, suggestions, and corrections.

The authors are partially supported by the SFB/TRR 191 "Symplectic Structures in Geometry, Algebra and Dynamics" funded by the DFG.

References

[1] I. Ekeland and H. Hofer, *Symplectic topology and Hamiltonian dynamics*, Math. Z. **200** (1989), no. 3, 355–378, DOI 10.1007/BF01215653. MR978597

[2] J. Guo, F. Hüffner, and H. Moser, *Feedback arc set in bipartite tournaments is NP-complete*, Inform. Process. Lett. **102** (2007), no. 2-3, 62–65, DOI 10.1016/j.ipl.2006.11.016. MR2292959

[3] P. Haim-Kislev, *On the symplectic size of convex polytopes*, Geom. Funct. Anal. **29** (2019), no. 2, 440–463, DOI 10.1007/s00039-019-00486-4. MR3945837

[4] H. Hofer and E. Zehnder, *Symplectic invariants and Hamiltonian dynamics*, Birkhäuser Advanced Texts: Basler Lehrbücher. [Birkhäuser Advanced Texts: Basel Textbooks], Birkhäuser Verlag, Basel, 1994, DOI 10.1007/978-3-0348-8540-9. MR1306732

[5] D. E. Knuth, *The art of computer programming. Vol. 1: Fundamental algorithms*, Addison-Wesley Publishing Co., Reading, Mass.-London-Don Mills, Ont., 1969. Second printing. MR0286317

[6] K. Perrot and T. V. Pham, *Feedback arc set problem and NP-hardness of minimum recurrent configuration problem of chip-firing game on directed graphs*, Ann. Comb. **19** (2015), no. 2, 373–396, DOI 10.1007/s00026-015-0266-9. MR3347388

[7] A. Schrijver, *Theory of linear and integer programming*, Wiley-Interscience Series in Discrete Mathematics, John Wiley & Sons, Ltd., Chichester, 1986. A Wiley-Interscience Publication. MR874114

[8] A. Schrijver, *Combinatorial optimization. Polyhedra and efficiency. Vol. A*, Algorithms and Combinatorics, vol. 24,A, Springer-Verlag, Berlin, 2003. Paths, flows, matchings; Chapters 1–38. MR1956924

[9] G. M. Ziegler, *Lectures on polytopes*, Graduate Texts in Mathematics, vol. 152, Springer-Verlag, New York, 1995, DOI 10.1007/978-1-4613-8431-1. MR1311028

Universität zu Köln, Department Mathematik/Informatik, Abteilung Mathematik, Weyertal 86-90, 50931 Köln, Germany
Email address: `karla.leipold@gmail.com`

Universität zu Köln, Department Mathematik/Informatik, Abteilung Mathematik, Weyertal 86-90, 50931 Köln, Germany
Email address: `frank.vallentin@uni-koeln.de`

7 Affine Symplectomorphism of Simplices is GI-Complete

In this section we show that the affine symplectomorphism problem for full-dimensional simplices is **GI**-complete. To formulate this precisely, we first define the three decision problems used in the reductions.

Definition 7.1. Fix $m \in \mathbb{N}$. A *weighted directed graph* on vertex set $[m] \coloneqq \{1, \dots, m\}$ is encoded by a matrix $W \in \mathbb{Q}^{m \times m}$, where W_{ij} is the weight of the directed edge (i, j). Two instances $W_1, W_2 \in \mathbb{Q}^{m \times m}$ are called *isomorphic*, and we write $W_1 \cong W_2$, if there exists a bijection $q : [m] \to [m]$ such that $(W_1)_{ij} = (W_2)_{q(i)\,q(j)}$ for all $i, j \in [m]$. Equivalently, there exists $\sigma \in \mathfrak{S}_m$ such that

$$W_1 = P_\sigma^\mathsf{T} W_2 P_\sigma.$$

The associated decision problem is denoted by **WGI**.

Thus **WGI** asks whether two weighted adjacency matrices differ only by a relabeling of the vertices. It is a weighted variant of the classical graph isomorphism (**GI**) problem. The complexity of **GI** is atypical: it is neither known to be solvable in polynomial time nor known to be NP-complete; on the algorithmic side, Babai proved a quasipolynomial-time algorithm for **GI** [10]. This motivates the standard terminology that a problem is ***GI**-hard* if **GI** reduces to it in polynomial time, and ***GI**-complete* if it is polynomial-time equivalent to **GI** (i.e. both problems reduce to each other); see, e.g., [90]. In this sense, weighted/colored variants such as **WGI** are **GI**-complete, since weights can be encoded within the usual **GI** framework by routine constructions [70, 78].

Definition 7.2. For $n \in \mathbb{N}$, let

$$J \coloneqq \begin{pmatrix} 0 & -I_n \\ I_n & 0 \end{pmatrix}, \qquad \mathrm{Sp}(2n, \mathbb{R}) \coloneqq \{A \in \mathbb{R}^{2n \times 2n} : A^\mathsf{T} J A = J\}.$$

An *affine symplectomorphism* of $\mathbb{R}^{2n}$ is a map $\psi : \mathbb{R}^{2n} \to \mathbb{R}^{2n}$ of the form $\psi(z) = Az + z_0$ with $A \in \mathrm{Sp}(2n, \mathbb{R})$ and $z_0 \in \mathbb{R}^{2n}$.

Let $\Delta(V) = \mathrm{conv}\{v_0, \dots, v_{2n}\}$ and $\Delta(Y) = \mathrm{conv}\{y_0, \dots, y_{2n}\} \subset \mathbb{R}^{2n}$ be full-dimensional simplices. We write

$$\Delta(V) \sim \Delta(Y)$$

if there exists an affine symplectomorphism ψ of $\mathbb{R}^{2n}$ such that $\psi(\Delta(V)) = \Delta(Y)$. The associated decision problem is denoted by **SymplSimp**.

The anchored problem is the version of **SymplSimp** in which one prescribed vertex of the first simplex must be mapped to one prescribed vertex of the second.

Definition 7.3. Let $\Delta(V) = \operatorname{conv}\{v_0, \ldots, v_{2n}\}$ and $\Delta(Y) = \operatorname{conv}\{y_0, \ldots, y_{2n}\} \subset \mathbb{R}^{2n}$ be full-dimensional simplices, and fix vertices v_i of $\Delta(V)$ and y_j of $\Delta(Y)$. We write

$$(\Delta(V), v_i) \sim_{\text{anch}} (\Delta(Y), y_j)$$

if there exists an affine symplectomorphism ψ of $\mathbb{R}^{2n}$ such that
$\psi(\Delta(V)) = \Delta(Y)$ and $\psi(v_i) = y_j$. The associated decision problem is denoted by **AnchSymplSimp**.

The main result of this section is the following.

Theorem 7.4. *Let **WGI** be weighted directed graph isomorphism, and let **SymplSimp** be the affine symplectomorphism problem for full-dimensional simplices. Then **SymplSimp** is **GI**-complete. In particular,*

$$\textbf{WGI} \leq_p \textbf{SymplSimp} \qquad \text{and} \qquad \textbf{SymplSimp} \leq_p \textbf{WGI}.$$

This places the problem studied here into a broader line of geometric isomorphism problems. For instance, Kaibel and Schwartz showed that several isomorphism problems for polytopes are **GI**-hard, and that combinatorial polytope isomorphism is **GI**-complete [61]. In the present section we obtain a related result for a symplectic notion of equivalence: affine symplectomorphism of full-dimensional simplices.

The proof of **WGI** $\leq_p$ **SymplSimp** has two parts and we give the outline of the argumentation here: In Section 7.1 we reformulate anchored affine symplectomorphism in terms of anchored edge matrices and their skew-symmetric signature matrices. This gives a matrix criterion for **AnchSymplSimp**. In Section 7.2 we construct, from a weighted matrix W, an invertible skew-symmetric gadget matrix $K(W)$ whose permutation-conjugacy is equivalent to the isomorphism type of W. We then realize $K(W)$ as the signature matrix of a 0-anchored simplex. This yields a polynomial-time reduction

$$\textbf{WGI} \leq_p \textbf{AnchSymplSimp}.$$

In Section 7.3 we remove the anchor by adding two extra vertices in a distinguished symplectic 2-plane. Their pairings force any affine symplectomorphism of the enlarged simplices to identify the anchor. This gives a polynomial-time reduction

$$\textsc{AnchSymplSimp} \le_p \textsc{SymplSimp}.$$

Combining the two steps gives

$$\textbf{WGI} \le_p \textsc{AnchSymplSimp} \le_p \textsc{SymplSimp}.$$

7.1 Affine Symplectomorphisms of Simplices

We now pass from simplices themselves to their anchored edge matrices. Once an anchor vertex is fixed, the remaining freedom of an affine symplectomorphism is encoded by a permutation of the other vertices. Let $U = \{u_0, \ldots, u_{2n}\} \subset \mathbb{R}^{2n}$ be affinely independent, and write $\Delta(U) \coloneqq \operatorname{conv}(U)$ for the associated full-dimensional simplex. Fix an anchor vertex u_i, and let $\alpha : \{1, \ldots, 2n\} \to \{0, \ldots, 2n\} \setminus \{i\}$ be a bijection, that is, an ordering of the remaining indices. We define the anchored edge matrix by

$$E(U; i, \alpha) \coloneqq \begin{bmatrix} u_{\alpha(1)} - u_i & \cdots & u_{\alpha(2n)} - u_i \end{bmatrix} \in \mathbb{R}^{2n \times 2n}.$$

Its columns are precisely the edge vectors from the anchor u_i to the other vertices. For a second simplex $\Delta(V) = \operatorname{conv}\{v_0, \ldots, v_{2n}\}$, with anchor v_j and a bijection $\beta : \{1, \ldots, 2n\} \to \{0, \ldots, 2n\} \setminus \{j\}$, we use the analogous notation $E(V; j, \beta)$. Affine independence implies that every anchored edge matrix is invertible.

Fix anchors $i, j \in \{0, \ldots, 2n\}$ and orderings α, β of the respective non-anchor index sets, and set

$$E_U \coloneqq E(U; i, \alpha), \qquad E_V \coloneqq E(V; j, \beta).$$

Define the associated *anchored signatures*

$$G_U \coloneqq E_U^{\mathsf{T}} J E_U, \qquad G_V \coloneqq E_V^{\mathsf{T}} J E_V.$$

Theorem 7.5. *Let $\Delta(U) = \operatorname{conv}\{u_0, \ldots, u_{2n}\}$ and $\Delta(V) = \operatorname{conv}\{v_0, \ldots, v_{2n}\}$ be full-dimensional simplices in $\mathbb{R}^{2n}$. Fix anchors i for U and j for V, and fix orderings α, β of the respective non-anchor index sets. Then the following are equivalent:*

(a) $(\Delta(U), u_i) \sim_{\text{anch}} (\Delta(V), v_j)$*;*

(b) there exists a permutation $\sigma \in \mathfrak{S}_{2n}$ such that

$$A \coloneqq E_V P_\sigma E_U^{-1} \in \mathrm{Sp}(2n, \mathbb{R}), \qquad b \coloneqq v_j - Au_i;$$

(c) there exists a permutation $\sigma \in \mathfrak{S}_{2n}$ such that

$$G_U = P_\sigma^\mathsf{T} G_V P_\sigma.$$

Proof. (a)$\Longrightarrow$(b). Let $f(x) = Ax + b$ be an affine symplectomorphism witnessing $(\Delta(U), u_i) \sim_{\text{anch}} (\Delta(V), v_j)$. Then f permutes the vertices, so there is a bijection π with $f(u_k) = v_{\pi(k)}$ and $\pi(i) = j$. Relative to the fixed orderings α, β of the non-anchor vertices, this induces a unique $\sigma \in \mathfrak{S}_{2n}$ such that $\pi(\alpha(a)) = \beta(\sigma(a))$ for all a. Therefore

$$A(u_{\alpha(a)} - u_i) = v_{\beta(\sigma(a))} - v_j \qquad (a = 1, \ldots, 2n),$$

hence $AE_U = E_V P_\sigma$, so $A = E_V P_\sigma E_U^{-1}$. Also $A \in \mathrm{Sp}(2n, \mathbb{R})$ and $b = v_j - Au_i$.

(b)$\Longrightarrow$(a). If $A \coloneqq E_V P_\sigma E_U^{-1} \in \mathrm{Sp}(2n, \mathbb{R})$ and $b \coloneqq v_j - Au_i$, then $f(x) = Ax + b$ is affine symplectic and $AE_U = E_V P_\sigma$ implies $f(u_i) = v_j$ and $f(u_{\alpha(a)}) = v_{\beta(\sigma(a))}$ for all a. Thus f maps the vertex set of $\Delta(U)$ onto that of $\Delta(V)$, hence $f(\Delta(U)) = \Delta(V)$.

(b)$\Longleftrightarrow$(c). Condition (b) is equivalent to $A^\mathsf{T} JA = J$. For $A = E_V P_\sigma E_U^{-1}$ this becomes

$$E_U^{-\mathsf{T}} P_\sigma^\mathsf{T} (E_V^\mathsf{T} J E_V) P_\sigma E_U^{-1} = J,$$

and after multiplying by E_U^T and E_U this is exactly

$$P_\sigma^\mathsf{T} G_V P_\sigma = G_U. \qquad \square$$

Thus, after fixing an anchor, the affine symplectomorphism problem is equivalent to a permutation problem for the anchored signature matrix.

If both anchors are at the origin, then the translation part vanishes, so anchored affine symplectomorphism is equivalent to the existence of a permutation $\sigma \in \mathfrak{S}_{2n}$ such that

$$E_U^\mathsf{T} J E_U = P_\sigma^\mathsf{T} (E_V^\mathsf{T} J E_V) P_\sigma,$$

and in this case the symplectic map is necessarily linear, namely

$$f(x) = Ax, \qquad A = E_V P_\sigma E_U^{-1} \in \mathrm{Sp}(2n, \mathbb{R}).$$

Since translation is an affine symplectomorphism, every anchored instance $(\Delta(U), u_i)$, $(\Delta(V), v_j)$ is equivalent to the 0-anchored instance $(\Delta(U) - u_i, 0)$, $(\Delta(V) - v_j, 0)$. Hence, in the reductions below, we may assume without loss of generality that the anchor is at the origin.

7.2 Reduction from Weighted GI to Anchored Symplectomorphism

Starting from a weighted graph isomorphism instance (W_1, W_2) with $W_1, W_2 \in \mathbb{Q}^{m\times m}$, we construct an invertible skew-symmetric gadget matrix $K(W)$ whose permutation-conjugacy encodes exactly the isomorphism type of W. The gadget consists of the original data block together with a small marker pattern, encoded by the distinguished values β and τ, that forces any weight-preserving permutation to respect the intended block structure. In a second step, we realize $K(W)$ as a symplectic Gram matrix and interpret its columns as the anchored edge vectors of a full-dimensional simplex.

Define the row-sum bound and max-entry bound

$$\Gamma \;:=\; \max\Big\{ \max_i \sum_{j=1}^m |(W_1)_{ij}|,\, \max_i \sum_{j=1}^m |(W_2)_{ij}| \Big\}, \qquad M \;:=\; \max\Big\{ \max_{i,j} |(W_1)_{ij}|,\, \max_{i,j} |(W_2)_{ij}| \Big\}.$$

Set

$$\lambda := 2\Gamma + 1, \qquad \beta := 4\lambda, \qquad \tau := 8\lambda. \tag{11}$$

Then $\lambda > 2M$ and $|\tau| > |\beta|$.

For each $W \in \mathbb{Q}^{m\times m}$ define the upper-triangular marker matrix

$$C(W) \;:=\; \begin{pmatrix} \lambda I_m + W & \beta\mathbf{e} \\ 0 & \tau \end{pmatrix} \in \mathbb{Q}^{(m+1)\times(m+1)}, \tag{12}$$

where $\mathbf{e} = (1, \ldots, 1)^\mathsf{T} \in \mathbb{Q}^m$. Now define the skew-symmetric gadget matrix

$$K(W) \;:=\; \begin{pmatrix} 0 & C(W) \\ -C(W)^\mathsf{T} & 0 \end{pmatrix} \in \mathbb{Q}^{2(m+1)\times 2(m+1)}. \tag{13}$$

Lemma 7.6. *For every $W \in \{W_1, W_2\}$, the matrix $\lambda I_m + W$ is invertible. Hence $C(W)$ is invertible, and therefore $K(W)$ is invertible.*

Proof. Fix $W \in \{W_1, W_2\}$ and write w_{ij} for its entries. By the definition of Γ and $\lambda = 2\Gamma+1$, for every row i we have

$$|\lambda + w_{ii}| \;\geq\; \lambda - |w_{ii}| \;>\; \Gamma \;\geq\; \sum_{j \neq i} |w_{ij}|.$$

Thus $\lambda I_m + W$ is strictly diagonally dominant, hence invertible by the Levy–Desplanques theorem; see, for example, [53, Theorem 6.1.10]. Since $C(W)$ is upper triangular with diagonal blocks $\lambda I_m + W$ and $\tau \neq 0$, it is invertible so $K(W)$ is invertible. □

To analyze the permutation structure of the gadget, it is convenient to regard $K(W)$ as a weighted directed graph with specially marked vertices.

Fix $m \geq 2$ and let

$$\mathcal{V} \coloneqq \{\ell_1, \ldots, \ell_m, \ell_*, r_1, \ldots, r_m, r_*\}.$$

For $K(W) \in \mathbb{R}^{\mathcal{V} \times \mathcal{V}}$ we view $K(W)_{u,v}$ as the weight of the directed edge (u, v).

Write

$$L_d \coloneqq \{\ell_1, \ldots, \ell_m\}, \quad R_d \coloneqq \{r_1, \ldots, r_m\}, \quad L \coloneqq L_d \cup \{\ell_*\}, \quad R \coloneqq R_d \cup \{r_*\}.$$

Then all edges inside $L \times L$ and inside $R \times R$ have weight 0, and for $1 \leq i, j \leq m$,

$$K(W)_{\ell_i, r_j} = (\lambda I_m + W)_{ij}, \qquad K(W)_{\ell_i, r_*} = +\beta, \qquad K(W)_{\ell_*, r_*} = +\tau.$$

Thus the gadget is naturally bipartite: the original matrix data sits on edges from the left data vertices ℓ_i to the right data vertices r_j, while the extra vertices ℓ_*, r_* act as markers that rigidify the whole structure.

Lemma 7.7. *Let $q : \mathcal{V} \to \mathcal{V}$ be a weight-preserving isomorphism from the gadget graph of W_1 to that of W_2, i.e.*

$$K(W_1)_{u,v} = K(W_2)_{q(u),\, q(v)} \qquad \textit{for all } u, v \in \mathcal{V}.$$

Then the following hold:

(i) $q(r_*) = r_*$ *and* $q(\ell_*) = \ell_*$*;*

(ii) $q(L_d) = L_d$ *and* $q(R_d) = R_d$*;*

(iii) there exists a unique permutation $\sigma \in \mathfrak{S}_m$ such that

$$q(\ell_i) = \ell_{\sigma(i)} \qquad \textit{and} \qquad q(r_i) = r_{\sigma(i)} \qquad (i = 1, \ldots, m).$$

Proof. Since $\lambda > 2M$, every entry of $\lambda I_m + W$ has absolute value strictly smaller than

$$\beta = 4\lambda \qquad \text{and} \qquad \tau = 8\lambda.$$

Hence the weights $\pm\beta$ occur only on the edges (ℓ_i, r_*), and the weights $\pm\tau$ occur only on the edge (ℓ_*, r_*). Therefore r_* is the unique vertex incident to exactly m edges of absolute weight β, and ℓ_* is the unique vertex distinct from r_* incident to an edge of absolute weight τ. Since q preserves edge weights, this proves (i).

Now let $x \in \mathcal{V} \setminus \{\ell_*, r_*\}$. Then

$$x \in L_d \quad \Longleftrightarrow \quad |K(W)_{x,r_*}| = \beta.$$

Since q preserves absolute weights and $q(r_*) = r_*$ by part (i), we obtain

$$x \in L_d \quad \Longleftrightarrow \quad q(x) \in L_d \qquad \big(x \in \mathcal{V} \setminus \{\ell_*, r_*\}\big).$$

Hence $q(L_d) = L_d$. Because q is a bijection and fixes ℓ_* and r_*, it follows that $q(R_d) = R_d$ as well. By (ii), $q|_{L_d}$ defines a unique permutation $\sigma \in \mathfrak{S}_m$ via $q(\ell_i) = \ell_{\sigma(i)}$. Moreover, by $\lambda > 2M$, the unique neighbor of ℓ_i in R_d whose edge weight has absolute value greater than M is r_i. Since q preserves absolute weights, this forces

$$q(r_i) = r_{\sigma(i)} \qquad (i = 1, \ldots, m),$$

proving (iii). $\square$

We obtain the central correctness statement of the gadget:

Lemma 7.8. *For $W_1, W_2 \in \mathbb{Q}^{m \times m}$, the following are equivalent:*

(a) there exists $\sigma \in \mathfrak{S}_m$ such that $W_1 = P_\sigma^{\mathsf{T}} W_2 P_\sigma$;

(b) there exists $\tau \in \mathfrak{S}_{2(m+1)}$ such that

$$K(W_1) = P_\tau^{\mathsf{T}} K(W_2) P_\tau.$$

Proof. *(a)⟹(b).* Assume $W_1 = P_\sigma^\mathsf{T} W_2 P_\sigma$ for some $\sigma \in \mathfrak{S}_m$.
Set $P_* := \mathrm{diag}(P_\sigma, 1) \in \mathbb{R}^{(m+1)\times(m+1)}$ and $Q := \mathrm{diag}(P_*, P_*) \in \mathbb{R}^{2(m+1)\times 2(m+1)}$. Then Q is a permutation matrix, say $Q = P_\tau$ for some $\tau \in \mathfrak{S}_{2(m+1)}$. Since $P_\sigma^\mathsf{T}\mathbf{e} = \mathbf{e}$ and $P_\sigma^\mathsf{T}(\lambda I_m + W_2)P_\sigma = \lambda I_m + W_1$, we get $P_*^\mathsf{T} C(W_2) P_* = C(W_1)$, and therefore

$$K(W_1) = Q^\mathsf{T} K(W_2) Q = P_\tau^\mathsf{T} K(W_2) P_\tau.$$

(b)⟹(a). Assume $K(W_1) = P_\tau^\mathsf{T} K(W_2) P_\tau$ for some $\tau \in \mathfrak{S}_{2(m+1)}$, and let $q : \mathcal{V} \to \mathcal{V}$ be the induced relabeling, defined by $P_\tau e_u = e_{q(u)}$. Then $K(W_1)_{u,v} = K(W_2)_{q(u),q(v)}$ for all $u, v \in \mathcal{V}$. By Lemma 7.7, there exists $\sigma \in \mathfrak{S}_m$ such that $q(\ell_i) = \ell_{\sigma(i)}$ and $q(r_j) = r_{\sigma(j)}$. Comparing the (ℓ_i, r_j)-entries gives $(\lambda I_m + W_1)_{ij} = (\lambda I_m + W_2)_{\sigma(i),\sigma(j)}$, and subtracting $\lambda\delta_{ij}$ yields $(W_1)_{ij} = (W_2)_{\sigma(i),\sigma(j)}$, i.e. $W_1 = P_\sigma^\mathsf{T} W_2 P_\sigma$. □

At this point the graph-theoretic part of the construction is complete: the gadget matrix remembers exactly the same isomorphism information as the original weighted graph.

We now pass back from the combinatorial gadget to geometry. Since every invertible skew-symmetric matrix can be realized as a symplectic Gram matrix, we may interpret the columns of a suitable factorization of $K(W)$ as the anchored edge vectors of a simplex. This is standard linear symplectic algebra: every nondegenerate alternating bilinear form admits a symplectic basis, and the usual constructive proofs give an explicit symplectic Gram–Schmidt procedure. Over $\mathbb{Q}$ this can be implemented by Gaussian elimination, hence in polynomial time in the bit-size of the input. See, for instance, [17, Theorem 1.1]; see also [22, Theorem 5.4].

Lemma 7.9 ([17]). *Let $G \in \mathbb{R}^{2N\times 2N}$ be invertible and skew-symmetric, and set*

$$J := \begin{pmatrix} 0 & -I_N \\ I_N & 0 \end{pmatrix}.$$

Then there exists $E \in GL(2N, \mathbb{R})$ such that $G = E^\mathsf{T} J E$.

Set $N := m + 1$ and apply Lemma 7.9 to $G = K(W)$. Choose $E(W) \in GL(2N, \mathbb{Q})$ with $K(W) = E(W)^\mathsf{T} J E(W)$, write $E(W) = [\,\varepsilon_1(W) \ \cdots \ \varepsilon_{2N}(W)\,]$, and define

$$\Delta_W := \mathrm{conv}\{0, \varepsilon_1(W), \ldots, \varepsilon_{2N}(W)\} \subset \mathbb{R}^{2N}. \tag{14}$$

Thus $K(W)$ is precisely the anchored signature matrix of the simplex Δ_W.

Theorem 7.10. *There is a polynomial-time computable transformation, mapping a pair $W_1, W_2 \in \mathbb{Q}^{m\times m}$ to full-dimensional simplices $\Delta_{W_1}, \Delta_{W_2} \subset \mathbb{R}^{2(m+1)}$ such that the following are equivalent:*

(a) $W_1 \cong W_2$;

(b) $(\Delta_{W_1}, 0) \sim_{\text{anch}} (\Delta_{W_2}, 0)$.

Proof. Set $N := m+1$. By Lemma 7.8, condition (a) is equivalent to the existence of $\tau \in \mathfrak{S}_{2N}$ such that

$$K(W_1) = P_\tau^\top K(W_2) P_\tau.$$

For $\nu = 1, 2$, the simplex Δ_{W_ν} has 0-anchored edge matrix $E(W_\nu)$, hence anchored signature

$$E(W_\nu)^\top J E(W_\nu) = K(W_\nu).$$

Therefore, by the anchored-at-zero case of Theorem 7.5, the above condition is equivalent to

$$(\Delta_{W_1}, 0) \sim_{\text{anch}} (\Delta_{W_2}, 0).$$

This proves the claim. □

7.3 Reduction from Anchored to Unanchored Affine Symplectomorphism

Let $\Delta(U), \Delta(V) \subset \mathbb{R}^{2n}$ be full-dimensional simplices. If $\Delta(U) \sim \Delta(V)$, then every anchor vertex of $\Delta(U)$ is mapped to some anchor vertex of $\Delta(V)$ under an affine symplectomorphism; in particular, $\Delta(U) \sim \Delta(V)$ if and only if there exists an anchor pair (u_i, v_j) such that $(\Delta(U), u_i) \sim_{\text{anch}} (\Delta(V), v_j)$. The passage from the unanchored problem to the anchored one introduces only polynomial overhead, however, for a clean many-one reduction, we encode the anchor directly into the instance. Since translation is an affine symplectomorphism, every anchored instance can be translated so that the distinguished vertex becomes the origin. It therefore suffices to work with simplices of the form

$$\Delta = \operatorname{conv}\{0, x_1, \ldots, x_{2n}\} \subset \mathbb{R}^{2n}.$$

We enlarge such a simplex by attaching a distinguished symplectic 2-plane whose geometry forces every affine symplectomorphism of the enlarged simplices to recognize the anchor.

We work on $\mathbb{R}^{2n+2} = \mathbb{R}^{2n} \times \mathbb{R}^2$ with the block-diagonal symplectic matrix

$$J_{2n+2} := \operatorname{diag}(J, J_2), \qquad J_2 := \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix},$$

and the associated symplectic form $\omega(u, v) := u^\mathsf{T} J_{2n+2} v$. Set

$$X := [x_1 \ \cdots \ x_{2n}] \in GL(2n, \mathbb{R}), \qquad K := X^\mathsf{T} J X, \qquad \beta := 1 + \max_{p,q} |K_{pq}|.$$

Embed the old space into $\mathbb{R}^{2n+2}$ by $\iota(x) := (x, 0, 0)$, introduce the two new vertices $u := (0, 1, 0)$ and $w := (0, 0, \beta)$, and define the augmented simplex

$$\Delta^\dagger := \operatorname{conv}\{0, \iota(x_1), \ldots, \iota(x_{2n}), u, w\} \subset \mathbb{R}^{2n+2}.$$

Thus the original simplex lies in the first factor, while u and w span an additional symplectic 2-plane.

Definition 7.11. Let $\Delta \subset \mathbb{R}^{2n+2}$ be a simplex, where $\mathbb{R}^{2n+2}$ is equipped with the symplectic form $\omega(x, y) = x^\mathsf{T} J_{2n+2} y$, and fix $\beta > 0$. For a vertex $p \in \operatorname{Vert}(\Delta)$ we say that p *satisfies* $\mathcal{P}_\beta$, and write $\mathcal{P}_\beta(p)$, if there exist two *distinct* vertices $a, b \in \operatorname{Vert}(\Delta) \setminus \{p\}$ such that

(1) $|\omega(a - p,\, b - p)| = \beta$, and

(2) for every vertex $c \in \operatorname{Vert}(\Delta) \setminus \{p, a, b\}$ one has

$$\omega(a - p,\, c - p) = 0 \qquad \text{and} \qquad \omega(b - p,\, c - p) = 0.$$

Thus $\mathcal{P}_\beta(p)$ says that from p one sees a distinguished pair spanning a large symplectic 2-plane that is orthogonal to everything else. Because affine symplectomorphisms preserve symplectic pairings of difference vectors, the property $\mathcal{P}_\beta$ is invariant under affine symplectomorphisms.

Lemma 7.12. *In the gadget simplex $\Delta^\dagger$, the vertex 0 satisfies $\mathcal{P}_\beta$, and it is the unique vertex with this property.*

Proof. For $p = 0$, choose $a = u$ and $b = w$. Then $|\omega(u, w)| = \beta$, and $\omega(u, \iota(x_k)) = \omega(w, \iota(x_k)) = 0$ for all k, so $P_\beta(0)$ holds.

For uniqueness, let $p \neq 0$ be a vertex of $\Delta^\dagger$ and assume that $P_\beta(p)$ holds. Since all pairings coming from the old simplex have absolute value $< \beta$, condition (1) forces the distinguished pair to be the two special vertices opposite to p.

If $p = \iota(x_i)$, this means $\{a, b\} = \{u, w\}$, and condition (2) gives

$$\omega(u - \iota(x_i),\, \iota(x_j) - \iota(x_i)) = \omega(w - \iota(x_i),\, \iota(x_j) - \iota(x_i)) = 0 \qquad (j \neq i).$$

Since u and w are symplectically orthogonal to the embedded old space, we get

$$\omega(\iota(x_i), \iota(x_j)) = 0 \qquad \text{for all } j = 1, \ldots, 2n.$$

Equivalently, $x_i^\mathsf{T} J x_j = 0$ for all $j = 1, \ldots, 2n$. Thus $x_i^\mathsf{T} J X = 0$, where $X = [x_1 \ \cdots \ x_{2n}]$. Since $X \in \mathrm{GL}(2n, \mathbb{R})$ and J is invertible, also JX is invertible, hence $x_i = 0$, a contradiction. If $p = u$, then necessarily $\{a, b\} = \{0, w\}$; but for any old vertex $\iota(x_k)$ one has

$$\omega(w - u, \iota(x_k) - u) = \beta \neq 0,$$

contradicting condition (2). The case $p = w$ is analogous.

Hence no vertex $p \neq 0$ satisfies P_β. □

Proposition 7.13. *Let $(\Delta_1, 0)$ and $(\Delta_2, 0)$ be anchored full-dimensional simplices in $\mathbb{R}^{2n}$, and construct $\Delta_1^\dagger, \Delta_2^\dagger \subset \mathbb{R}^{2n+2}$ as above using a common β. Then*

$$(\Delta_1, 0) \sim_{\text{anch}} (\Delta_2, 0) \quad \Longleftrightarrow \quad \Delta_1^\dagger \sim \Delta_2^\dagger.$$

Proof. ($\Longrightarrow$). Assume $(\Delta_1, 0) \sim_{\text{anch}} (\Delta_2, 0)$. Then there exists an affine symplectomorphism $f(x) = Ax + b$ with $A \in \mathrm{Sp}(2n, \mathbb{R})$ such that $f(\Delta_1) = \Delta_2$ and $f(0) = 0$. Hence $b = 0$, so f is linear. Its block-diagonal extension

$$\widetilde{A} := \mathrm{diag}(A, I_2)$$

lies in $\mathrm{Sp}(2n + 2, \mathbb{R})$, maps $\iota(x)$ to $\iota(Ax)$, and fixes the two added vertices u, w. Therefore $\widetilde{A}(\Delta_1^\dagger) = \Delta_2^\dagger$, so $\Delta_1^\dagger \sim \Delta_2^\dagger$.

($\Longleftarrow$). Assume $\Delta_1^\dagger \sim \Delta_2^\dagger$, and let $g(z) = Bz + c$ with $B \in \mathrm{Sp}(2n+2, \mathbb{R})$ satisfy $g(\Delta_1^\dagger) = \Delta_2^\dagger$.

Because affine symplectomorphisms preserve symplectic pairings of difference vectors, they preserve the property $\mathcal{P}_\beta$. By Lemma 7.12, the vertex 0 is the unique vertex of each gadget simplex with this property. Hence $g(0) = 0$, so $c = 0$ and $g(z) = Bz$ is linear.

Moreover, the pair $\{u, w\}$ is intrinsically distinguished among the nonzero vertices: it is the unique pair with $|\omega(u, w)| = \beta$ that is symplectically orthogonal to all embedded old vertices. Therefore $B(\{u, w\}) = \{u, w\}$, and the plane $\Lambda := \mathrm{span}\{u, w\}$ is B-invariant.

Now let

$$\Lambda^\omega := \{z \in \mathbb{R}^{2n+2} : \omega(z, \lambda) = 0 \text{ for all } \lambda \in \Lambda\}$$

be the symplectic orthogonal complement of Λ. In our block-diagonal model one has

$$\Lambda^\omega = \{(x, 0, 0) : x \in \mathbb{R}^{2n}\} = \iota(\mathbb{R}^{2n}).$$

Since B is symplectic and preserves Λ, it also preserves Λ^ω. Thus B preserves the embedded copy $\iota(\mathbb{R}^{2n})$ of the old space.

We may therefore define a linear map $A : \mathbb{R}^{2n} \to \mathbb{R}^{2n}$ by $B \circ \iota = \iota \circ A$. Because B preserves the symplectic form and $\omega_{2n+2}(\iota(x), \iota(y)) = \omega_{2n}(x, y)$, it follows that $A \in \mathrm{Sp}(2n, \mathbb{R})$.

Finally, B maps $\Delta_1^\dagger$ onto $\Delta_2^\dagger$, fixes 0, preserves the old subspace $\iota(\mathbb{R}^{2n})$, and therefore maps $\mathrm{conv}\{0, \iota(x_1), \ldots, \iota(x_{2n})\}$ onto the corresponding embedded old simplex. Identifying $\iota(\mathbb{R}^{2n})$ with $\mathbb{R}^{2n}$, we obtain $A(\Delta_1) = \Delta_2$ and $A(0) = 0$, hence $(\Delta_1, 0) \sim_{\text{anch}} (\Delta_2, 0)$. □

Proof of Theorem 7.4. For **WGI** $\leq_p$ **SymplSimp** Theorem 7.10 gives a polynomial-time reduction

$$\textbf{WGI} \leq_p \textbf{AnchSymplSimp}.$$

Proposition 7.13 gives

$$\textbf{AnchSymplSimp} \leq_p \textbf{SymplSimp}.$$

Composing the two reductions proves **WGI** $\leq_p$ **SymplSimp**

Now we prove **SymplSimp** $\leq_p$ **WGI**. Let $\Delta(V) = \mathrm{conv}\{v_0, \ldots, v_{2n}\}$ and $\Delta(W) = \mathrm{conv}\{w_0, \ldots, w_{2n}\}$ be an instance of **SymplSimp**. Since translation is an affine symplectomorphism, we may translate both simplices by their centroids without changing the fact if they are symplectomorphic or not. Thus we assume from now on that both simplices are *centred*,

$$\sum_{k=0}^{2n} v_k = 0, \qquad \sum_{\ell=0}^{2n} w_\ell = 0.$$

Set

$$X := [\, v_0 \ \cdots \ v_{2n} \,] \in \mathbb{R}^{2n \times (2n+1)}, \qquad Y := [\, w_0 \ \cdots \ w_{2n} \,] \in \mathbb{R}^{2n \times (2n+1)},$$

and define the skew-symmetric symplectic pairing matrices

$$W_V := X^\mathsf{T} J X, \qquad W_W := Y^\mathsf{T} J Y.$$

We claim that

$$\Delta(V) \sim \Delta(W) \quad \Longleftrightarrow \quad P_\sigma^{\mathsf{T}} W_W P_\sigma = W_V \text{ for some } \sigma \in \mathfrak{S}_{2n+1}. \tag{15}$$

Indeed, if $f(x) = Ax + b$ is an affine symplectomorphism with $f(v_k) = w_{\sigma(k)}$, then averaging over k and using the centroid condition gives $b = 0$, hence f is linear and $AX = YP_\sigma$. Therefore

$$W_V = X^{\mathsf{T}} JX = (P_\sigma^{\mathsf{T}} Y^{\mathsf{T}} A^{\mathsf{T}}) J (AYP_\sigma) = P_\sigma^{\mathsf{T}} Y^{\mathsf{T}} (A^{\mathsf{T}} JA) YP_\sigma = P_\sigma^{\mathsf{T}} Y^{\mathsf{T}} JYP_\sigma = P_\sigma^{\mathsf{T}} W_W P_\sigma.$$

Conversely, assume $P_\sigma^{\mathsf{T}} W_W P_\sigma = W_V$. Choose $S \subset \{0, \ldots, 2n\}$ with $|S| = 2n$ such that X_S is invertible (and set $T := \sigma(S)$). Restricting to S yields $X_S^{\mathsf{T}} JX_S = Y_T^{\mathsf{T}} JY_T$, so with $A := Y_T X_S^{-1}$ we obtain $A^{\mathsf{T}} JA = J$, hence $A \in \mathrm{Sp}(2n, \mathbb{R})$ and $Av_s = w_{\sigma(s)}$ for all $s \in S$. The remaining vertex follows from the (s, r)-entries, giving $Av_k = w_{\sigma(k)}$ for all k. Thus A maps the vertex set of $\Delta(V)$ onto that of $\Delta(W)$ and hence $A(\Delta(V)) = \Delta(W)$.

Therefore we have constructed a polynomial-time reduction proving **SymplSimp** $\leq_p$ **WGI**.

$\square$

Part III

Searching for Maximizers of the Systolic Ratio

This part of the thesis is focused on the following geometric optimization problem: among convex bodies $K \subset \mathbb{R}^{2n}$ of fixed volume, how large can the systolic ratio

$$\mathrm{sys}_n(K) := \frac{c_{\mathrm{EHZ}}(K)}{(n!\,\mathrm{Vol}(K))^{1/n}}$$

be, and which convex bodies, if any, maximize it? Motivated both by the renewed interest in maximizers of sys_n after recent developments around the Viterbo conjecture, and by combinatorial perspective developed earlier in the thesis, we focus on the family of simplices and on computational approaches to the Ekeland–Hofer–Zehnder capacity c_{EHZ}.

The starting point is the explicit combinatorial formula for the EHZ capacity of a simplex in centred position, derived in Section 3. A central observation of this thesis is that the underlying permutation objective is an instance of the weighted linear-ordering / feedback-arc-set problem. This places the computation of c_{EHZ} for simplices in direct contact with a mature algorithmic toolbox: classical work of Grötschel, Jünger, and Reinelt [34] gives exact integer linear programming formulations for linear ordering, while modern computational studies show that these formulations can perform extremely well in practice.

Section 9 develops this reformulation systematically and uses it as the main computational primitive for objective evaluation.

To study the maximization of the systolic ratio at fixed volume, we fix the standard simplex $S_{2n} \subset \mathbb{R}^{2n}$, which can be described in hyperplane description by

$$S_{2n} = \{x \in \mathbb{R}^{2n} : Bx \leq \mathbf{e}\} \qquad \text{with} \qquad B = (2n+1)\begin{bmatrix} -I_{2n} \\ \mathbf{e}^\mathsf{T} \end{bmatrix}, \tag{16}$$

and consider its volume-preserving linear images XS_{2n} with $X \in \mathrm{SL}(2n, \mathbb{R})$.

This reduces the geometric problem to a constrained matrix optimization problem on the special linear group:

$$\max\{c_{\mathrm{EHZ}}(XS_{2n}) : X \in \mathrm{SL}(2n, \mathbb{R})\}.$$

To treat this problem within a manifold-optimization framework, we develop in Section 10 the basic tools from optimization on manifolds following Boumal [13], and specialize them to the special linear group $\mathrm{SL}(2n, \mathbb{R})$. Equivalently, in the notation introduced in Section 11, we study the matrix objective

$$\Phi : \mathrm{SL}(2n, \mathbb{R}) \to \mathbb{R}, \qquad \Phi(X) \coloneqq c_{\mathrm{EHZ}}(XS_{2n}).$$

A key difficulty is that this objective is nonsmooth: it is the reciprocal of a pointwise maximum of smooth branch functions indexed by permutations. To obtain principled first-order information nonetheless, Section 11 develops a Clarke-subdifferential framework adapted to our setting. On the Euclidean side, the finite-max structure yields an explicit convex-hull representation of the Clarke subdifferential in terms of the active branch gradients; on the manifold side, these ambient Clarke subgradients are projected onto $\mathrm{SL}(2n, \mathbb{R})$ using the Riemannian Clarke calculus defined via retractions, following Clarke [20], Petersen–Pedersen [77], and Hosseini–Pouryayevali [55].

These analytic ingredients are turned into implementable algorithms in the sections that follow. Section 12.1 develops a retraction-based line-search method on $\mathrm{SL}(2n, \mathbb{R})$ tailored to the explicit Clarke-subdifferential structure of our objective. The practical details of this scheme, including the ILP-based extraction of active permutations and the convex quadratic program used to compute minimum-norm Clarke subgradients, are described in Section 12.3. The resulting numerical experiments are reported in Section 12.4. Building on the complexity-theoretic perspective developed earlier in Section 7, Section 12.2 provides an algorithmic test for deciding whether two numerical maximizers are related by an affine symplectomorphism.

8 Problem Statement and Motivation

We now outline the optimization problem studied in this part.

For a convex body $K \subset \mathbb{R}^{2n}$, its systolic ratio is defined by

$$\mathrm{sys}_n(K) \coloneqq \frac{c_{\mathrm{EHZ}}(K)}{(n!\ \mathrm{Vol}(K))^{1/n}}.$$

Our goal is to maximize this quantity over the class of simplices.

To search for simplices with large systolic ratio, we fix the volume and vary only the shape. For this, we choose the standard simplex $S_{2n} \subset \mathbb{R}^{2n}$ as a reference simplex and consider its images under determinant-one linear maps. Writing

$$K_X \coloneqq XS_{2n}, \qquad X \in \mathrm{SL}(2n, \mathbb{R}),$$

we have $\mathrm{Vol}(K_X) = \mathrm{Vol}(S_{2n})$, since $\det(X) = 1$. Hence

$$\mathrm{sys}_n(K_X) = \frac{c_{\mathrm{EHZ}}(K_X)}{(n!\ \mathrm{Vol}(S_{2n}))^{1/n}},$$

so that, within the family $\{K_X : X \in \mathrm{SL}(2n, \mathbb{R})\}$, maximizing the systolic ratio is equivalent to maximizing the Ekeland–Hofer–Zehnder capacity. This leads to the problem

$$\max\{c_{\mathrm{EHZ}}(XS_{2n}) : X \in \mathrm{SL}(2n, \mathbb{R})\}. \tag{17}$$

Accordingly, we introduce the objective function

$$\Phi : \mathrm{SL}(2n, \mathbb{R}) \to \mathbb{R}, \qquad \Phi(X) \coloneqq c_{\mathrm{EHZ}}(XS_{2n}).$$

To derive a more explicit representation of Φ, we recall the formula for the Ekeland–Hofer–Zehnder capacity of a centered simplex established in Section 3. Let

$$S = \{x \in \mathbb{R}^{2n} : Bx \leq \mathbf{e}\}$$

be a simplex in hyperplane description, where the row vectors $b_1, \ldots, b_{2n+1}$ of $B \in \mathbb{R}^{(2n+1)\times 2n}$ satisfy

$$\sum_{i=1}^{2n+1} b_i = 0.$$

If we define

$$W := BJB^{\mathsf{T}} \in \mathbb{R}^{(2n+1)\times(2n+1)},$$

where J denotes the standard symplectic matrix on $\mathbb{R}^{2n}$, and if $L \in \{0,1\}^{(2n+1)\times(2n+1)}$ denotes the strict upper-triangular matrix, then for each $\sigma \in \mathfrak{S}_{2n+1}$ we set

$$L_\sigma := P_\sigma L P_\sigma^{\mathsf{T}}.$$

With this notation,

$$c_{\mathrm{EHZ}}(S) = \frac{(2n+1)^2}{2}\left(\max_{\sigma\in\mathfrak{S}_{2n+1}} \operatorname{tr}(W^{\mathsf{T}} L_\sigma)\right)^{-1} = \frac{(2n+1)^2}{2}\left(-\min_{\sigma\in\mathfrak{S}_{2n+1}} \operatorname{tr}(W L_\sigma)\right)^{-1}. \tag{18}$$

For our purposes it is essential that XS_{2n} remains centered at the origin whenever S_{2n} is centered, since the derivation of (19) applies in this centered setting. This is ensured by the following elementary observation.

Proposition 8.1. *Let $S \subset \mathbb{R}^d$ be a simplex with centroid 0. Then for every invertible linear map $X \in \mathbb{R}^{d\times d}$, the simplex XS also has centroid 0. In particular, if $S \subset \mathbb{R}^{2n}$ is centered at the origin and $X \in \mathrm{SL}(2n,\mathbb{R})$, then XS is centered at the origin.*

Proof. Let $v_1, \dots, v_{d+1}$ be the vertices of the d-simplex S. The vertices of XS are $Xv_1, \dots, Xv_{d+1}$. Since the centroid of S is 0, we have

$$\frac{1}{d+1}\sum_{i=1}^{d+1} v_i = 0.$$

Therefore the centroid of XS equals

$$\frac{1}{d+1}\sum_{i=1}^{d+1} Xv_i = X\left(\frac{1}{d+1}\sum_{i=1}^{d+1} v_i\right) = 0.$$

Hence XS is centered at the origin. □

The EHZ capacity of the simplex XS_{2n} can therefore be written as

$$\Phi(X) = c_{\mathrm{EHZ}}(XS_{2n}) = \frac{(2n+1)^2}{2}\left(\max_{\sigma\in\mathfrak{S}_{2n+1}} f(X,\sigma)\right)^{-1}, \qquad X \in \mathrm{SL}(2n,\mathbb{R}), \tag{19}$$

where, by Lemma 3.3,

$$f(X,\sigma) = \operatorname{tr}\bigl(BX^{-1}J^{\mathsf{T}}(X^{-1})^{\mathsf{T}}B^{\mathsf{T}}L_\sigma\bigr).$$

Thus Φ is obtained by restricting to $\mathrm{SL}(2n,\mathbb{R})$ a function defined in terms of finitely many smooth trace expressions on $\mathrm{GL}(2n,\mathbb{R})$. This representation will be the starting point for the Clarke-type nonsmooth analysis in Section 11.

Since $\mathrm{Vol}(S_{2n}) = \frac{1}{(2n)!}$, the preceding discussion yields the following formulation of the optimization problem.

Theorem 8.2. *Let $S_{2n} = \{x \in \mathbb{R}^{2n} : Bx \leq \mathbf{e}\} \subset \mathbb{R}^{2n}$ be the standard simplex as described in 16. Then the problem of maximizing the systolic ratio over simplices can be formulated as*

$$\max_{X\in\mathrm{SL}(2n,\mathbb{R})} \mathrm{sys}_n(XS_{2n}) = \frac{(2n+1)^2((2n)!)^{1/n}}{2(n!)^{1/n}} \max_{X\in\mathrm{SL}(2n,\mathbb{R})} \left\{ \left(\max_{\sigma\in\mathfrak{S}_{2n+1}} \left\{\operatorname{tr}\bigl(BX^{-1}J^{\mathsf{T}}(X^{-1})^{\mathsf{T}}B^{\mathsf{T}}L_\sigma\bigr)\right\} \right)^{-1} \right\}. \tag{20}$$

9 ILP Formulation of the EHZ Capacity of Simplices

As shown in Section 6, the evaluation of c_{EHZ} for simplices reduces to an optimization problem over linear orders on $2n+1$ vertices. More precisely, the simplex c_{EHZ}-computation can be expressed as a weighted linear ordering problem. In this section we reformulate this problem as an exact integer linear program using the classical linear-ordering formulation from Grötschel–Jünger–Reinelt [34].

9.1 Linear Programs and Integer Linear Programs

Since our exact formulation is an instance of integer linear optimization, we briefly recall the basic LP/ILP framework. We follow Schrijver [86].

A *linear program* (LP) is an optimization problem of the form

$$\min\{c^{\mathsf{T}}x : Ax \le b,\ x \in \mathbb{R}^n\}, \tag{21}$$

where $A \in \mathbb{R}^{m\times n}$, $b \in \mathbb{R}^m$, and $c \in \mathbb{R}^n$. The feasible region

$$P \coloneqq \{x \in \mathbb{R}^n : Ax \le b\}$$

is a (possibly unbounded) polyhedron, and the objective is linear. LPs can be solved efficiently in theory and practice [86].

In many applications, some decision variables are required to be integral, or even binary. This leads to *integer linear programming*. An *integer linear program* (ILP) is of the form

$$\min\{c^{\mathsf{T}}x : Ax \le b,\ x \in \mathbb{Z}^n\}. \tag{22}$$

Unlike LP, ILP is NP-hard in general, and practical solution methods therefore combine continuous relaxations with combinatorial search and inference; see, for example, [96, 21]. A basic tool is the *LP relaxation*, obtained by replacing the integrality condition $x \in \mathbb{Z}^n$ by $x \in \mathbb{R}^n$, that is, by solving (21) with the same linear constraints. For minimization problems, this relaxation yields a lower bound for the ILP optimum.

A wide range of efficient solvers is available, including commercial solvers such as *CPLEX* [58] and open-source solvers such as *SCIP* [84, 85], *HiGHS* [50], and *CBC*. In this thesis, we use the *Gurobi* ILP solver [40].

9.2 The EHZ Capacity Computed Through a Linear-Ordering ILP

We now recall the classical linear-ordering ILP and then specialize it to the matrices arising in the c_{EHZ}-computation for simplices. The key idea is that a linear order of the vertex set can be encoded by an orientation of the complete graph, that is, by a tournament, and that transitivity of this order can be enforced by excluding directed 3-cycles; see, for example, Moon [71], Bang-Jensen–Gutin [12], and Grötschel–Jünger–Reinelt [34]. Recall that a *tournament* is a directed graph $T = (V, A)$ without loops such that for every pair of distinct vertices $u, v \in V$, exactly one of (u, v) and (v, u) belongs to A. Furthermore, let $V = [m]$. A *linear order* on V is a strict total order $\prec$ on V. Equivalently, a linear order can be encoded by a permutation $\sigma \in \mathfrak{S}_m$ in listing form,

$$\sigma(1) \prec \sigma(2) \prec \cdots \prec \sigma(m).$$

Proposition 9.1 ([71, 12]). *A tournament is induced by a unique linear order of its vertices if and only if it contains no directed cycle of length* 3.

Let $C = (c_{ij}) \in \mathbb{R}^{m\times m}$. For $1 \le i < j \le m$, introduce binary variables $y_{ij} \in \{0, 1\}$ by

$$y_{ij} = 0 \iff i \prec j, \qquad y_{ij} = 1 \iff j \prec i.$$

The following ILP is the classical formulation of the linear ordering problem; see Grötschel–Jünger–Reinelt [34], who gave an exact cutting-plane algorithm for solving this problem. For further discussion of exact ILP-based approaches to the feedback arc set problem, see Baharev–Schichl–Neumaier–Achterberg [9]:

$$\begin{aligned}
\min_{y} \quad & \sum_{j=1}^{m} \left(\sum_{k=1}^{j-1} c_{k,j}\, y_{k,j} + \sum_{\ell=j+1}^{m} c_{\ell,j}\,(1 - y_{j,\ell}) \right) & & \\
\text{s.t.} \quad & y_{i,j} + y_{j,k} - y_{i,k} \le 1, & & 1 \le i < j < k \le m, \\
& -y_{i,j} - y_{j,k} + y_{i,k} \le 0, & & 1 \le i < j < k \le m, \\
& y_{i,j} \in \{0, 1\}, & & 1 \le i < j \le m.
\end{aligned} \tag{23}$$

Indeed, the variables y_{ij} determine the relative order of every pair $i < j$, and hence define a tournament on $[m]$. The two families of triangle inequalities exclude directed 3-cycles. Therefore, by Proposition 9.1, every feasible binary solution of (23) corresponds to a unique linear order of the vertices.

To interpret the objective, let $\sigma \in \mathfrak{S}_m$ be the listing permutation associated with such a linear order, so that

$$\sigma(1) \prec \sigma(2) \prec \cdots \prec \sigma(m).$$

For each pair $i < j$, the contribution to the objective is

$$c_{ij}\, y_{ij} + c_{ji}\,(1 - y_{ij}).$$

If $i \prec j$, then $y_{ij} = 0$ and this contribution equals c_{ji}; if $j \prec i$, then $y_{ij} = 1$ and it equals c_{ij}. Hence the objective sums exactly those coefficients that appear below the diagonal after permuting C into the order given by σ. Equivalently,

$$\sum_{j=1}^{m} \left(\sum_{k=1}^{j-1} c_{k,j}\, y_{k,j} + \sum_{\ell=j+1}^{m} c_{\ell,j}\,(1 - y_{j,\ell}) \right) = \sum_{1 \le j < i \le m} (P_\sigma^{\mathsf{T}} C P_\sigma)_{ij}.$$

Thus (23) computes

$$\min_{\sigma \in \mathfrak{S}_m} \sum_{1 \le j < i \le m} (P_\sigma^{\mathsf{T}} C P_\sigma)_{ij}.$$

We now specialize this formulation to the matrices arising in the computation of c_{EHZ} for simplices. Let $n \in \mathbb{N}$. Let $S \subset \mathbb{R}^{2n}$ be a simplex given in facet-hyperplane description by

$$S = \{x \in \mathbb{R}^{2n} : Bx \le \mathbf{e}\}, \qquad B \in \mathbb{R}^{(2n+1)\times 2n},\ \operatorname{rank}(B) = 2n,$$

and let J denote the standard symplectic matrix on $\mathbb{R}^{2n}$. Define

$$W := BJB^{\mathsf{T}} \in \mathbb{R}^{(2n+1)\times(2n+1)}.$$

Then W is skew-symmetric. Recall from Section 6 that

$$c_{\mathrm{EHZ}}(S) = \frac{(2n+1)^2}{2} \left(\max_{\sigma \in \mathfrak{S}_{2n+1}} \sum_{1 \le j < i \le 2n+1} (P_\sigma^{\mathsf{T}} W P_\sigma)_{ij} \right)^{-1}. \tag{24}$$

Applying (23) with coefficients $c_{ij} = W_{ij}$ therefore yields an exact ILP for

$$\min_{\sigma \in \mathfrak{S}_{2n+1}} \sum_{1 \le j < i \le 2n+1} (P_\sigma^{\mathsf{T}} W P_\sigma)_{ij}.$$

Since W is skew-symmetric, reversing the linear order changes the sign of the lower-triangular sum. Consequently,

$$\min_{\sigma \in \mathfrak{S}_{2n+1}} \sum_{1 \leq j < i \leq 2n+1} (P_\sigma^\mathsf{T} W P_\sigma)_{ij} = -\max_{\sigma \in \mathfrak{S}_{2n+1}} \sum_{1 \leq j < i \leq 2n+1} (P_\sigma^\mathsf{T} W P_\sigma)_{ij}.$$

Hence (23) computes exactly the quantity needed in the above formula for $c_{\mathrm{EHZ}}(S)$.

To illustrate the practical advantage of this reformulation, we compared the ILP approach to the brute-force enumeration of all permutations on $2n+1$ vertices. Already for $n=4$ (that is, $2n+1=9$), checking all 9! orderings took about four seconds, whereas Gurobi solved the ILP to proven optimality in about 6×10^{-3} s, a speedup of roughly 700. For $n=5$ (that is, $2n+1=11$), brute-force enumeration became impractical in our setup because of the factorial growth of the search space. In contrast, the ILP remained efficiently solvable in much larger dimensions: in a computational sweep up to $n=70$ (that is, $2n+1=141$), all instances were solved within five minutes. This illustrates that the ILP formulation avoids explicit enumeration of $(2n+1)!$ linear orders and makes large-scale computations possible.

10 Optimization on the Special Linear Group $\mathrm{SL}(2n, \mathbb{R})$

This chapter studies optimization problems with manifold constraints, motivated by the systolic ratio problem stated in Section 8. Since the volume constraint in maximizing the EHZ capacity naturally leads to the special linear group $\mathrm{SL}(2n, \mathbb{R})$, it is helpful to formulate the problem in the language of optimization on manifolds. This viewpoint allows us to bring in tools from smooth manifold optimization, in particular the framework developed by Boumal and Absil–Mahony–Sepulchre [13, 3], and to adapt familiar Euclidean ideas such as gradient-based methods to our setting.

As introduced in Section 8, we work with simplices of the form $K_X = XS_{2n}$, where S_{2n} is the standard simplex (in centered position) and $X \in \mathrm{SL}(2n, \mathbb{R})$.

We develop the manifold-optimization viewpoint needed to address the optimization problem stated in Theorem 8.2. In particular, we discuss how to transfer familiar ideas from Euclidean gradient methods to the special linear group, and we then investigate in more detail how to apply these ideas to the EHZ capacity.

A guiding analogy is the classical gradient descent method in Euclidean space, which is a standard approach to smooth unconstrained optimization problems.

Algorithm 1 Gradient descent ([74])

Require: A differentiable function $f : \mathbb{R}^n \to \mathbb{R}$.

1: **Input:** $x_0 \in \mathbb{R}^n$.
2: **for** $k = 0, 1, 2, \ldots$ **do**
3: Choose a step size $\alpha_k > 0$.
4: Update $x_{k+1} = x_k - \alpha_k \nabla f(x_k)$.
5: **end for**

Our goal is to build an analogous framework on $\mathrm{SL}(2n, \mathbb{R})$. To keep the exposition self-contained, we briefly recall the basic notions from optimization on smooth manifolds, following Boumal [13], and then work out their specialization to $\mathrm{SL}(2n, \mathbb{R})$. The general definitions and standard results are thus taken from the literature, while the concrete formulas and their application to the special linear group are derived here. Some of these notions are stated mainly for later reference. Readers already familiar with the background may skip directly to Lemma 10.22, which summarizes the concrete formulas for $\mathrm{SL}(2n, \mathbb{R})$ used throughout the thesis.

10.1 Manifolds, Tangent spaces, and Embedded Submanifolds

We begin by developing the notions of embedded submanifolds that we will use throughout this chapter following the book of Boumal [13]. We omit the chart-and-atlas viewpoint and focus on a characterization that is well suited for matrix manifolds and is working for our purposes, since we will only consider *smooth* manifolds.

Definition 10.1. Let $U \subset \mathcal{E}$ and $V \subset \mathcal{F}$ be open subsets of finite-dimensional linear spaces. A map $F : U \to V$ is called *smooth* if it is of class C^∞. If F is differentiable at $x \in U$, we denote its differential at x by

$$DF(x) : \mathcal{E} \to \mathcal{F}, \qquad DF(x)[u] = \frac{d}{dt}\bigg|_{t=0} F(x + tu).$$

For a smooth curve $c : I \to \mathcal{E}$, where $I \subset \mathbb{R}$ is an open interval, we write

$$c'(t) = \frac{d}{dt}c(t)$$

for its velocity.

Definition 10.2. Let $\mathcal{E}$ be a d-dimensional linear space. A nonempty subset $\mathcal{M} \subset \mathcal{E}$ is a *(smooth) embedded submanifold* of dimension n if either:

(a) $n = d$ and $\mathcal{M}$ is open in $\mathcal{E}$; or

(b) $n = d - k$ for some $k \geq 1$ and, for each $x \in \mathcal{M}$, there is a neighborhood $U \subset \mathcal{E}$ of x and a smooth map $h : U \to \mathbb{R}^k$ such that

   (a) $y \in U$ satisfies $h(y) = 0$ if and only if $y \in \mathcal{M}$, and

   (b) $\operatorname{rank} Dh(x) = k$ where $Dh(x)$.

Such an h is called a *local defining function.*

A smooth manifold is locally well approximated by a linear space; a convenient way to formalize this is via velocities of smooth curves.

Definition 10.3. Let $\mathcal{M} \subset \mathcal{E}$ and $x \in \mathcal{M}$. The *tangent space* of $\mathcal{M}$ at x is

$$T_x\mathcal{M} \ = \ \{\, c'(0) : c : I \to \mathcal{M} \text{ smooth},\ c(0) = x \,\},$$

where I is any open interval containing 0. The *tangent bundle* is

$$T\mathcal{M} \;=\; \{(x,v) : x \in \mathcal{M},\; v \in T_x\mathcal{M}\}.$$

For embedded submanifolds, the tangent space can equivalently be described in terms of a local defining function.

Theorem 10.4. *Let $\mathcal{M}$ be an embedded submanifold of $\mathcal{E}$, and let $x \in \mathcal{M}$. If $\mathcal{M}$ is an open submanifold, then $T_x\mathcal{M} = \mathcal{E}$. Otherwise, if h is any local defining function at x, then*

$$T_x\mathcal{M} = \ker Dh(x).$$

We now apply this to the special linear group.

Lemma 10.5. *The special linear group* $\mathrm{SL}(n,\mathbb{R})$ *is an embedded submanifold of* $\mathbb{R}^{n\times n}$ *of dimension n^2-1. Moreover, for $X \in \mathrm{SL}(n,\mathbb{R})$ the tangent space is*

$$T_X\mathrm{SL}(n,\mathbb{R}) \;=\; \bigl\{V \in \mathbb{R}^{n\times n} \;\big|\; \mathrm{tr}(X^{-1}V) = 0\bigr\}.$$

Proof. Consider $h : \mathbb{R}^{n\times n} \to \mathbb{R}$ given by $h(X) = \det(X) - 1$. For $X \in \mathrm{SL}(n,\mathbb{R})$ and $H \in \mathbb{R}^{n\times n}$, Jacobi's formula yields

$$Dh_X[H] \;=\; \det(X)\,\mathrm{tr}\bigl(X^{-1}H\bigr) \;=\; \mathrm{tr}\bigl(X^{-1}H\bigr). \tag{25}$$

Hence $\mathrm{rank}\, Dh_X = 1$ for all $X \in \mathrm{SL}(n,\mathbb{R})$, so 1 is a regular value of det, and $\mathrm{SL}(n,\mathbb{R}) = h^{-1}(0)$ is a smooth embedded submanifold of codimension 1, hence dimension n^2-1. By the tangent characterization above,

$$T_X\mathrm{SL}(n,\mathbb{R}) = \ker Dh_X = \{V \in \mathbb{R}^{n\times n} : \mathrm{tr}(X^{-1}V) = 0\}.$$

□

10.2 Riemannian Metrics and Riemannian Submanifolds

The goal of our approach is to obtain gradient-descent-like algorithms on embedded submanifolds of linear spaces. In Euclidean space, the gradient is defined via the inner product: if $U \subset \mathbb{R}^n$ is open and $f : U \to \mathbb{R}$ is differentiable at $x \in U$, then the *(Euclidean) gradient* $\nabla f(x) \in \mathbb{R}^n$ is the unique vector such that

$$Df(x)[p] = \langle \nabla f(x), p\rangle = \nabla f(x)^{\mathsf{T}} p \qquad \text{for all } p \in \mathbb{R}^n,$$

see [74]. To define gradients on manifolds, we therefore need a metric.

Definition 10.6. An *inner product* on $T_x\mathcal{M}$ is a bilinear, symmetric, positive definite map $\langle \cdot, \cdot \rangle_x : T_x\mathcal{M} \times T_x\mathcal{M} \to \mathbb{R}$. It induces a norm $\|u\|_x \coloneqq \sqrt{\langle u, u \rangle_x}$. A *metric* on $\mathcal{M}$ is a choice of inner product $\langle \cdot, \cdot \rangle_x$ for each $x \in \mathcal{M}$.

A *vector field* on a manifold $\mathcal{M}$ is a map $V : \mathcal{M} \to T\mathcal{M}$ such that $V(x) \in T_x\mathcal{M}$ for all $x \in \mathcal{M}$. If V is smooth, we call it a *smooth vector field.* The set of smooth vector fields is denoted $\mathfrak{X}(\mathcal{M})$.

A metric $\langle \cdot, \cdot \rangle_x$ on $\mathcal{M}$ is a *Riemannian metric* if it varies smoothly with x, in the sense that for all $V, W \in \mathfrak{X}(\mathcal{M})$ the function $x \mapsto \langle V(x), W(x) \rangle_x$ is smooth. A *Riemannian manifold* is a manifold equipped with a Riemannian metric.

We will work with an embedded submanifold that inherits their metric from the ambient Euclidean space.

Lemma 10.7. *Let $\mathcal{M}$ be an embedded submanifold of $\mathcal{E}$, and let $\langle \cdot, \cdot \rangle$ be the Euclidean inner product on $\mathcal{E}$. Then the restricted metric $\langle u, v \rangle_x \coloneqq \langle u, v \rangle$ for $u, v \in T_x\mathcal{M}$ defines a Riemannian metric on $\mathcal{M}$. In this case we call $\mathcal{M}$ a* Riemannian submanifold *of $\mathcal{E}$.*

Since $\mathrm{SL}(n, \mathbb{R})$ is an embedded submanifold of $\mathbb{R}^{n \times n}$, and $\mathbb{R}^{n \times n}$ becomes a Euclidean space with the Frobenius inner product $\langle A, B \rangle \coloneqq \mathrm{tr}(A^\mathsf{T} B)$, we equip $\mathrm{SL}(n, \mathbb{R})$ with the restricted metric

$$\langle A, B \rangle_X \;=\; \mathrm{tr}(A^\mathsf{T} B), \qquad \text{for all } A, B \in T_X\mathrm{SL}(n, \mathbb{R}). \tag{26}$$

Hence $\mathrm{SL}(n, \mathbb{R})$ is a Riemannian submanifold of $\mathbb{R}^{n \times n}$.

The normal space at $X \in \mathrm{SL}(n, \mathbb{R})$ is the Frobenius-orthogonal complement of the tangent space:

$$N_X\mathrm{SL}(n, \mathbb{R}) \coloneqq \big(T_X\mathrm{SL}(n, \mathbb{R})\big)^\perp = \{\, Z \in \mathbb{R}^{n \times n} : \langle Z, H \rangle = 0 \text{ for all } H \in T_X\mathrm{SL}(n, \mathbb{R}) \,\}.$$

hence $N_X\mathrm{SL}(n, \mathbb{R})$ is one-dimensional and given by

$$N_X\mathrm{SL}(n, \mathbb{R}) = \mathrm{span}\{(X^{-1})^\mathsf{T}\} = \{\, \alpha\, (X^{-1})^\mathsf{T} : \alpha \in \mathbb{R} \,\}. \tag{27}$$

Definition 10.8. Let $\mathcal{M}$ be a Riemannian submanifold with Riemannianinduced metric $\langle \cdot, \cdot \rangle$ and induced norm $\|v\| \coloneqq \sqrt{\langle v, v \rangle}$ on $T_x\mathcal{M}$. Let $\gamma : [a, b] \to \mathcal{M}$ be a piecewise C^1 curve. Its *(Riemannian) length* is

$$L(\gamma) \coloneqq \int_a^b \|\dot{\gamma}(t)\|\, dt.$$

The *Riemannian distance* induced by the metric is

$$d_{\mathcal{M}}(p,q) \coloneqq \inf\big\{\, L(\gamma) \ : \ \gamma \text{ piecewise } C^1,\ \gamma(a)=p,\ \gamma(b)=q \,\big\}, \qquad p,q\in\mathcal{M}.$$

10.3 Differentiation on Riemannian Submanifolds

To formulate gradient-based methods on $\mathrm{SL}(n,\mathbb{R})$, we need a concrete way to differentiate functions on Riemannian submanifolds. The key point is that, in the embedded setting, derivatives can be computed via smooth extensions to the ambient Euclidean space. This leads in particular to explicit formulas for the Riemannian gradient, which in our situation turns out to be the orthogonal projection of the ambient Euclidean gradient onto the tangent space. In the following, we first discuss smooth extensions and differentials, and then use this framework to derive the gradient and Hessian formulas needed later for optimization on $\mathrm{SL}(n,\mathbb{R})$.

Definition 10.9. Let $\mathcal{M}$ and $\mathcal{M}'$ be embedded submanifolds of $\mathcal{E}$ and $\mathcal{E}'$, respectively. A map $F:\mathcal{M}\to\mathcal{M}'$ is *smooth at* $x\in\mathcal{M}$ if there exists an open neighborhood $U\subset\mathcal{E}$ of x and a smooth map $\overline{F}:U\to\mathcal{E}'$ such that $F(y)=\overline{F}(y)$ for all $y\in\mathcal{M}\cap U$. Such $\overline{F}$ is called a *(local) smooth extension* of F around x. The map F is *smooth* if it is smooth at every $x\in\mathcal{M}$.

In our applications we mostly consider real-valued maps. In this case, a local smooth extension of $f:\mathcal{M}\to\mathbb{R}$ near $x\in\mathcal{M}$ is a smooth function $\bar f:U\to\mathbb{R}$ (with $U\subset\mathcal{E}$ open) such that $\bar f|_{\mathcal{M}\cap U}=f$.

Proposition 10.10. *Let $\mathcal{M}$ and $\mathcal{M}'$ be embedded submanifolds of $\mathcal{E}$ and $\mathcal{E}'$. A map $F:\mathcal{M}\to\mathcal{M}'$ is smooth if and only if $F=\overline{F}\big|_{\mathcal{M}}$ for some smooth $\overline{F}$ defined on a neighborhood of $\mathcal{M}$ in $\mathcal{E}$. In particular, $f:\mathcal{M}\to\mathbb{R}$ is smooth if and only if it admits a smooth extension $\bar f$ to a neighborhood of $\mathcal{M}$ in $\mathcal{E}$.*

Definition 10.11. The *differential* of $F:\mathcal{M}\to\mathcal{M}'$ at $x\in\mathcal{M}$ is the linear map

$$DF(x): T_x\mathcal{M}\to T_{F(x)}\mathcal{M}'$$

defined by

$$DF(x)[v] \ = \ \frac{d}{dt}F(c(t))\bigg|_{t=0}, \tag{28}$$

where c is any smooth curve in $\mathcal{M}$ with $c(0)=x$ and $c'(0)=v$.

If F admits a smooth extension $\overline{F}$ near x, then $F \circ c = \overline{F} \circ c$, and the usual chain rule in $\mathcal{E}$ yields

$$DF(x)[v] = D\overline{F}(x)[v] \qquad \text{for all } v \in T_x\mathcal{M}. \tag{29}$$

Proposition 10.12. *With notation as above,*

$$DF(x) \;=\; D\overline{F}(x)\big|_{T_x\mathcal{M}}.$$

Definition 10.13 ([13])**.** Let $f : \mathcal{M} \to \mathbb{R}$ be smooth on a Riemannian manifold $(\mathcal{M}, \langle \cdot, \cdot \rangle_x)$. The *Riemannian gradient* of f is the vector field $\operatorname{grad} f$ uniquely defined by

$$Df(x)[v] \;=\; \langle v, \operatorname{grad} f(x) \rangle_x \qquad \text{for all } x \in \mathcal{M},\; v \in T_x\mathcal{M}. \tag{30}$$

Let $\langle \cdot, \cdot \rangle$ be the ambient Euclidean inner product on $\mathcal{E}$, let $f : \mathcal{M} \to \mathbb{R}$ be smooth, and let $\bar{f}$ be a smooth extension to $\mathcal{E}$. Then for all $v \in T_x\mathcal{M}$,

$$\langle v, \operatorname{grad} f(x) \rangle_x \;=\; Df(x)[v] \;=\; D\bar{f}(x)[v] \;=\; \langle v, \nabla \bar{f}(x) \rangle, \tag{31}$$

Decompose the ambient gradient as

$$\nabla \bar{f}(x) = \big(\nabla \bar{f}(x)\big)^{\parallel} + \big(\nabla \bar{f}(x)\big)^{\perp},$$

where $\big(\nabla \bar{f}(x)\big)^{\parallel} \in T_x\mathcal{M}$ denotes the tangential component and $\big(\nabla \bar{f}(x)\big)^{\perp} \in (T_x\mathcal{M})^{\perp}$ the normal component with respect to the Euclidean inner product. Since $v \in T_x\mathcal{M}$ is orthogonal to the normal component,

$$\big\langle v, \operatorname{grad} f(x) \big\rangle_x \;=\; \big\langle v, \big(\nabla \bar{f}(x)\big)^{\parallel} \big\rangle. \tag{32}$$

Therefore, for an embedded Riemannian submanifold, the gradient of a smooth function on the manifold is obtained by orthogonally projecting the ambient Euclidean gradient of a smooth extension onto the tangent space, where $\nabla \bar{f}(x) \in \mathcal{E}$ denotes the ambient Euclidean gradient.

Definition 10.14. The *tangent-space orthogonal projector* $\operatorname{Proj}_x : \mathcal{E} \to \mathcal{E}$ is characterized by:

(a) $\operatorname{im}(\operatorname{Proj}_x) = T_x\mathcal{M}$;

(b) $\operatorname{Proj}_x \circ \operatorname{Proj}_x = \operatorname{Proj}_x$;

(c) $\langle u - \mathrm{Proj}_x(u), v\rangle = 0$ for all $v \in T_x\mathcal{M}$ and all $u \in \mathcal{E}$.

Proposition 10.15. *Let $X \in \mathrm{SL}(n,\mathbb{R})$ and $V \in \mathbb{R}^{n\times n}$. Define*

$$\mathrm{Proj}_X(V) \;=\; V \;-\; \frac{\mathrm{tr}(X^{-1}V)}{\mathrm{tr}\big(X^{-1}(X^{-1})^\mathsf{T}\big)}\,(X^{-1})^\mathsf{T}. \tag{33}$$

Then $\mathrm{Proj}_X(V) \in T_X\mathrm{SL}(n,\mathbb{R})$, and Proj_X is the Frobenius-orthogonal projection $\mathbb{R}^{n\times n} \to T_X\mathrm{SL}(n,\mathbb{R})$.

Proof. Let $\mathcal{M} = \mathrm{SL}(n,\mathbb{R})$ with Frobenius inner product $\langle U, V\rangle = \mathrm{tr}(U^\top V)$ and

$$T_X\mathcal{M} = \{B \in \mathbb{R}^{n\times n} : \mathrm{tr}(X^{-1}B) = 0\}.$$

Define

$$P_X(V) \coloneqq V - \frac{\mathrm{tr}(X^{-1}V)}{\mathrm{tr}\big(X^{-1}(X^{-1})^\mathsf{T}\big)}(X^{-1})^\mathsf{T}.$$

(i) Tangency. Compute

$$\mathrm{tr}(X^{-1}P_X(V)) = \mathrm{tr}(X^{-1}V) - \frac{\mathrm{tr}(X^{-1}V)}{\mathrm{tr}\big(X^{-1}(X^{-1})^\mathsf{T}\big)}\;\mathrm{tr}\big(X^{-1}(X^{-1})^\mathsf{T}\big) = 0,$$

so $P_X(V) \in T_X\mathcal{M}$.

(ii) Idempotence. Let $W \coloneqq P_X(V)$. By (i), $W \in T_X\mathcal{M}$, so $\mathrm{tr}(X^{-1}W) = 0$ and therefore

$$P_X(W) = W - \frac{\mathrm{tr}(X^{-1}W)}{\mathrm{tr}(X^{-1}(X^{-1})^\mathsf{T})}(X^{-1})^\mathsf{T} = W.$$

Hence $P_X(P_X(V)) = P_X(V)$.

(iii) Orthogonality. For any $B \in T_X\mathcal{M}$ we have

$$V - P_X(V) = \frac{\mathrm{tr}(X^{-1}V)}{\mathrm{tr}\big(X^{-1}(X^{-1})^\mathsf{T}\big)}(X^{-1})^\mathsf{T},$$

hence

$$\begin{aligned}\langle V - P_X(V), B\rangle &= \mathrm{tr}\left(\left(\frac{\mathrm{tr}(X^{-1}V)}{\mathrm{tr}(X^{-1}(X^{-1})^\mathsf{T})}(X^{-1})^\mathsf{T}\right)^\mathsf{T} B\right)\\ &= \frac{\mathrm{tr}(X^{-1}V)}{\mathrm{tr}(X^{-1}(X^{-1})^\mathsf{T})}\;\mathrm{tr}(X^{-1}B) = 0,\end{aligned}$$

since $\mathrm{tr}(X^{-1}B) = 0$ for $B \in T_X\mathcal{M}$. Thus $V - P_X(V) \perp T_X\mathcal{M}$. $\square$

Thus, to compute the gradient of a smooth function $f : \mathrm{SL}(n, \mathbb{R}) \to \mathbb{R}$, it suffices to compute the ambient gradient $\nabla \bar{f}(X)$ of a smooth extension $\bar{f} : \mathbb{R}^{n\times n} \to \mathbb{R}$ and then project it to the tangent space:

$$\mathrm{grad}\, f(X) = \mathrm{Proj}_X\big(\nabla \bar{f}(X)\big) = \nabla \bar{f}(X) - \frac{\mathrm{tr}\big(X^{-1}\nabla \bar{f}(X)\big)}{\mathrm{tr}\big(X^{-1}(X^{-1})^{\mathsf{T}}\big)}\,(X^{-1})^{\mathsf{T}}.$$

For second-order approaches, we now define the Hessian on Riemannian submanifolds, similarly to the gradient, the Hessian on a Riemannian manifold can also be obtained by using the projection map.

Definition 10.16. Let $U \subset \mathbb{R}^n$ be open and let $f : U \to \mathbb{R}$ be C^2. The *(Euclidean) Hessian* of f at $x \in U$ is the linear map

$$\nabla^2 f(x) : \mathbb{R}^n \to \mathbb{R}^n$$

defined by

$$\nabla^2 f(x)[p] \;:=\; D(\nabla f)(x)[p], \qquad p \in \mathbb{R}^n.$$

Equivalently, in coordinates $\nabla^2 f(x)$ is represented by the matrix $\big(\partial_i \partial_j f(x)\big)_{i,j=1}^{n}$.

Lemma 10.17. *Let $\mathcal{M}$ be a Riemannian manifold and consider a smooth function $f : \mathcal{M} \to \mathbb{R}$. Let $\bar{G}$ be a smooth extension of* $\mathrm{grad}\, f$ *—that is, $\bar{G}$ is any smooth vector field defined on a neighborhood of $\mathcal{M}$ in the embedding space such that $\bar{G}(x) = \mathrm{grad}\, f(x)$ for all $x \in \mathcal{M}$. Then, The* Riemannian Hessian *of the smooth function $f : \mathcal{M} \to \mathbb{R}$ at $x \in \mathcal{M}$ is the linear map*

$$\mathrm{Hess}\, f(x) : T_x\mathcal{M} \to T_x\mathcal{M}$$

can be derived by

$$\mathrm{Hess}\, f(x)[u] = \mathrm{Proj}_x\left(D\bar{G}(x)[u]\right).$$

So for our circumstances a gradient on the manifold is obtained after projecting it, hence

$$\begin{aligned}
\bar{G}(X) &= \mathrm{grad}\, f(X) && (34)\\
&= \mathrm{Proj}_X\big(\nabla \bar{f}(X)\big) && (35)\\
&= \nabla \bar{f}(X) - \frac{\mathrm{tr}\big(X^{-1}\nabla \bar{f}(X)\big)}{\mathrm{tr}\big(X^{-1}(X^{-1})^{\mathsf{T}}\big)}(X^{-1})^{\mathsf{T}}. && (36)
\end{aligned}$$

For the Hessian it follows

$$\begin{aligned}\operatorname{Hess} f(X)[V] &= \operatorname{Proj}_X\left(D\bar{G}(X)[V]\right)\\ &= \operatorname{Proj}_X\big(\nabla^2 \bar{f}(X)[V]\big) - \operatorname{Proj}_X\big(DK(X)[V]\big), \qquad \text{with}\\ K(X) &:= \frac{\operatorname{tr}\big(X^{-1}\nabla \bar{f}(X)\big)}{\operatorname{tr}\big(X^{-1}(X^{-1})^\top\big)}(X^{-1})^\top.\end{aligned}$$

We compute $DK(X)[V]$ by using the product rule and the identities mentioned before. We define

$$a(X) := \operatorname{tr}\big(X^{-1}\nabla \bar{f}(X)\big), \qquad b(X) := \operatorname{tr}\big(X^{-1}(X^{-1})^\top\big).$$

Then

$$\begin{aligned}\operatorname{Proj}_X\big(DK(X)[V]\big) &= \operatorname{Proj}_X\left(\frac{Da(X)[V]\,b(X) - a(X)\,Db(X)[V]}{b(X)^2}(X^{-1})^\top - \frac{a(X)}{b(X)}(X^{-1})^\top V^\top (X^{-1})^\top\right)\\ &= -\operatorname{Proj}_X\left(\frac{a(X)}{b(X)}(X^{-1})^\top V^\top (X^{-1})^\top\right)\\ &= -\frac{\operatorname{tr}\big(X^{-1}\nabla \bar{f}(X)\big)}{\operatorname{tr}\big(X^{-1}(X^{-1})^\top\big)}\operatorname{Proj}_X\left((X^{-1})^\top V^\top (X^{-1})^\top\right).\end{aligned}$$

Putting these things together, we again obtain

$$\operatorname{Hess} f(X)[V] = \operatorname{Proj}_X\big(\nabla^2 \bar{f}(X)[V]\big) + \frac{\operatorname{tr}\big(X^{-1}\nabla \bar{f}(X)\big)}{\operatorname{tr}\big(X^{-1}(X^{-1})^\top\big)}\operatorname{Proj}_X\left((X^{-1})^\top V^\top (X^{-1})^\top\right). \tag{37}$$

10.4 Retractions

Optimization methods on manifolds follow the Euclidean template: given an iterate $x \in \mathcal{M}$ and a search direction $v \in T_x\mathcal{M}$, we would like to move from x along v while staying on the manifold. In Euclidean space the update is simply $x \leftarrow x - \alpha v$ for a positive scalar α. On a curved manifold, a canonical way to move is to follow the *geodesic* starting at x with initial velocity v, that is, to use the Riemannian exponential map $\operatorname{Exp}_x(\alpha v)$. However, computing Exp exactly can be expensive or unavailable in closed form. Retractions provide a practical substitute: they produce a curve on $\mathcal{M}$ that matches the geodesic to first order at the base point. This is precisely what is needed to transfer local (first-order) optimization ideas from Euclidean space to manifolds.

Definition 10.18. A retraction on a manifold $\mathcal{M}$ is a smooth map

$$R : T\mathcal{M} \to \mathcal{M}, \qquad (x, v) \mapsto R_x(v),$$

such that for each $(x, v) \in T\mathcal{M}$ the curve $c(t) = R_x(tv)$ satisfies

$$c(0) = x \qquad \text{and} \qquad c'(0) = v.$$

A retraction maps tangent vectors back to the manifold and locally models the exponential map. In particular, if $x_+ = R_x(\alpha v)$, then for small $\alpha > 0$ the update stays on $\mathcal{M}$ and satisfies

$$R_x(\alpha v) = x + \alpha v + o(\alpha) \quad \text{in local coordinates,}$$

so it agrees with the Euclidean step to first order. This makes retractions the natural mechanism for manifold-compatible updates in gradient-type algorithms.

We view $\mathrm{SL}(n, \mathbb{R})$ as an embedded submanifold of the Euclidean space $\mathcal{E} = \mathbb{R}^{n\times n}$ equipped with the Frobenius inner product. Recall that

$$T_X\mathrm{SL}(n, \mathbb{R}) = \{\, V \in \mathcal{E} : \ \mathrm{tr}(X^{-1}V) = 0 \,\}.$$

A particularly convenient choice is the exponential-type retraction, hence recall:

Definition 10.19. Let $A \in \mathbb{R}^{n\times n}$. The (matrix) exponential of A is defined by the absolutely convergent power series

$$\exp(A) \;:=\; \sum_{k=0}^{\infty} \frac{1}{k!}\, A^k \;=\; I + A + \frac{1}{2!}A^2 + \frac{1}{3!}A^3 + \cdots .$$

It satisfies $\exp(0) = I$ and, for all $s, t \in \mathbb{R}$, the map $t \mapsto \exp(tA)$ is a smooth curve with

$$\frac{d}{dt}\exp(tA) = A\, \exp(tA) = \exp(tA)\, A.$$

Proposition 10.20. *Let $\mathcal{M} = \mathrm{SL}(n, \mathbb{R})$ and let*
$T\mathcal{M} = \{(X, V) : X \in \mathrm{SL}(n, \mathbb{R}),\ V \in T_X\mathrm{SL}(n, \mathbb{R})\}$. *Define*

$$R : T\mathcal{M} \to \mathcal{M}, \qquad R_X(V) := X \exp\bigl(X^{-1}V\bigr).$$

Then R is a retraction on $\mathcal{M}$.

Proof. Fix $(X, V) \in T\mathcal{M}$ and consider the curve

$$c(t) := R_X(tV) = X \exp\big(tX^{-1}V\big).$$

Then

$$c(0) = X \exp(0) = X.$$

To compute the derivative at $t = 0$, set $A := X^{-1}V$ and note that

$$\frac{d}{dt} \exp(tA)\Big|_{t=0} = A.$$

Hence, by the matrix chain rule,

$$c'(0) = X \left(\frac{d}{dt} \exp(tA)\Big|_{t=0} \right) = XA = X(X^{-1}V) = V.$$

This proves that R satisfies the defining properties of a retraction. Smoothness follows since $(X, V) \mapsto X^{-1}V$ is smooth on $T\mathcal{M}$ and the matrix exponential is smooth. □

Remark 10.21. There are two notions of exponential that appear in manifold optimization, and it is important not to confuse them. The symbol $\exp(\cdot)$ denotes the *matrix exponential* acting on elements of the linear space $\mathbb{R}^{n\times n}$, defined by the power series above. The symbol $\mathrm{Exp}_x(\cdot)$ denotes the *Riemannian exponential map* at a point $x \in \mathcal{M}$, which maps a tangent vector $v \in T_x\mathcal{M}$ to the point reached at time 1 along the (unique) geodesic starting at x with initial velocity v.

These maps live in different settings: exp is defined on a linear space of matrices, while Exp_x depends on the Riemannian metric and the geometry of $\mathcal{M}$.

Now we state all the information about $\mathrm{SL}(n, \mathbb{R})$ we gathered and will be needed in the upcoming approaches:

Lemma 10.22. *Let $n \geq 2$ and consider the special linear group*

$$\mathrm{SL}(n, \mathbb{R}) := \{X \in \mathbb{R}^{n\times n} : \det(X) = 1\},$$

viewed as an embedded submanifold of $\mathbb{R}^{n\times n}$ equipped with the Frobenius inner product $\langle A, B\rangle := \mathrm{tr}(A^\top B)$. Then the following statements hold (proved throughout this chapter and collected here for reference).

(a) ***Manifold structure and dimension.*** *$\mathrm{SL}(n,\mathbb{R})$ is a smooth embedded submanifold of $\mathbb{R}^{n\times n}$ of dimension n^2-1.*

(b) ***Tangent space.*** *For $X \in \mathrm{SL}(n,\mathbb{R})$,*

$$T_X\mathrm{SL}(n,\mathbb{R}) = \{V \in \mathbb{R}^{n\times n} : \mathrm{tr}(X^{-1}V) = 0\}$$

(c) ***Orthogonal projection onto the tangent space.*** *For any $V \in \mathbb{R}^{n\times n}$, the Frobenius-orthogonal projection onto $T_X\mathrm{SL}(n,\mathbb{R})$ is*

$$\mathrm{Proj}_X(V) = V - \frac{\mathrm{tr}(X^{-1}V)}{\mathrm{tr}\big(X^{-1}(X^{-1})^{\mathsf{T}}\big)}\,(X^{-1})^{\mathsf{T}}.$$

In particular, if $\bar{f}$ is a smooth extension of $f : \mathrm{SL}(n,\mathbb{R}) \to \mathbb{R}$ to an open subset of $\mathbb{R}^{n\times n}$, then the Riemannian gradient under the induced metric is

$$\mathrm{grad}\, f(X) = \mathrm{Proj}_X\big(\nabla \bar{f}(X)\big).$$

(d) ***Exponential retraction.*** *The map*

$$R_X : T_X\mathrm{SL}(n,\mathbb{R}) \to \mathrm{SL}(n,\mathbb{R}), \qquad R_X(V) := X\exp\big(X^{-1}V\big),$$

is a retraction on $\mathrm{SL}(n,\mathbb{R})$.

10.5 Riemannian Gradient Descent (RGD)

All of the differential-geometric ingredients collected above (tangent spaces, orthogonal projection, and a retraction such as the exponential update) fit together into the standard *Riemannian gradient descent* scheme on $\mathrm{SL}(n,\mathbb{R})$. We follow the presentation in [13] and recall only the high-level idea and the key assumptions. Although our main algorithms later are nonsmooth (Clarke-type) methods, it is useful to briefly review the smooth Riemannian framework: it provides the guiding analogy for our manifold subgradient descent and clarifies which parts of the analysis rely on smoothness. We first state the smooth Riemannian gradient method in abstract form. The subsequent convergence statement is derived from two assumptions: a uniform lower bound on the objective and a sufficient decrease estimate along the iterates. We then introduce a pullback regularity assumption under which the backtracking line search yields step sizes satisfying this sufficient decrease condition.

Algorithm 2 RGD: Riemannian gradient descent

Require: A Riemannian manifold $\mathcal{M}$, a differentiable function $f : \mathcal{M} \to \mathbb{R}$, and a retraction $R : T\mathcal{M} \to \mathcal{M}$.

1: **Input:** $x_0 \in \mathcal{M}$.
2: **for** $k = 0, 1, 2, \ldots$ **do**
3: Choose a step size $\alpha_k > 0$.
4: Set $s_k := -\alpha_k \operatorname{grad} f(x_k) \in T_{x_k}\mathcal{M}$.
5: Update $x_{k+1} = R_{x_k}(s_k)$.
6: **end for**

Conceptually, each iteration consists of two steps: first compute a tangent-space search direction by projecting the ambient Euclidean gradient onto the tangent space (thus obtaining the Riemannian gradient); then move along this tangent direction and map back to the manifold via the retraction. Concretely,

$$x_{k+1} = R_{x_k}\big(-\alpha_k \operatorname{grad} f(x_k)\big), \qquad \operatorname{grad} f(x_k) = \operatorname{Proj}_{x_k}\big(\nabla \bar{f}(x_k)\big),$$

where $\bar{f}$ is a smooth extension of f to an open subset of the ambient Euclidean space.

The convergence analysis of RGD in [13] is based on two high-level assumptions: a uniform lower bound on the objective and a *sufficient decrease* property along the iterates. Together, these imply that the algorithm produces points with vanishing Riemannian gradient norm.

Assumption 10.23. There exists $f_{\text{low}} \in \mathbb{R}$ such that $f(x) \geq f_{\text{low}}$ for all $x \in \mathcal{M}$.

Assumption 10.24. At each iteration, the algorithm achieves sufficient decrease: there exists a constant $c > 0$ such that, for all k,

$$f(x_k) - f(x_{k+1}) \ \geq \ c \,\|\operatorname{grad} f(x_k)\|^2. \tag{38}$$

Proposition 10.25. *Let f be a smooth function on a Riemannian manifold $\mathcal{M}$ satisfying Assumption 10.23. Let $x_0, x_1, x_2, \ldots$ be iterates satisfying Assumption 10.24 with constant c. Then*

$$\lim_{k\to\infty} \|\operatorname{grad} f(x_k)\| = 0.$$

In particular, every accumulation point (if any) is a critical point. Moreover, for all $K \geq 1$, there exists $k \in \{0, \ldots, K-1\}$ such that

$$\|\operatorname{grad} f(x_k)\| \;\leq\; \sqrt{\frac{f(x_0) - f_{\text{low}}}{c}}\;\frac{1}{\sqrt{K}}.$$

Importantly, this does not assert that the iterates (x_k) converge to a single critical point. It guarantees only that every accumulation point—if there are any—is critical. Thus, the practical question is how to choose step sizes so that Assumption 10.24 holds.

To guarantee sufficient decrease in practice, we choose the step sizes α_k adaptively by a backtracking line search. The relevant pullback regularity assumption is the following.

Assumption 10.26. For a given subset $S \subset T\mathcal{M}$, there exists a constant $L > 0$ such that for all $(x, s) \in S$,

$$f\big(R_x(s)\big) \;\leq\; f(x) + \langle \operatorname{grad} f(x),\, s\rangle + \frac{L}{2}\,\|s\|^2. \tag{39}$$

We use the following sufficient decrease condition.

Definition 10.27. Let $(\mathcal{M}, \langle\cdot,\cdot\rangle)$ be a Riemannian manifold, $f : \mathcal{M} \to \mathbb{R}$ smooth, and $R : T\mathcal{M} \to \mathcal{M}$ a retraction. Fix $r \in (0, 1)$. A step size $\alpha > 0$ satisfies the *Armijo–Goldstein condition* at $x \in \mathcal{M}$ if

$$f(x) - f\big(R_x(-\alpha \operatorname{grad} f(x))\big) \;\geq\; r\,\alpha\,\|\operatorname{grad} f(x)\|^2. \tag{40}$$

The backtracking procedure below returns, by construction, a step size satisfying Definition 10.27.

Algorithm 3 Backtracking line search (Armijo–Goldstein)

Require: Parameters $\tau, r \in (0, 1)$ (e.g. $\tau = \frac{1}{2}$ and $r = 10^{-4}$), trial step $\bar{\alpha} > 0$.

1: **Input:** $x \in \mathcal{M}$.
2: Set $\alpha \leftarrow \bar{\alpha}$.
3: **while** $f(x) - f\big(R_x\big(-\alpha \operatorname{grad} f(x)\big)\big) < r\,\alpha\,\|\operatorname{grad} f(x)\|^2$ **do**
4: $\quad\alpha \leftarrow \tau\,\alpha$.
5: **end while**
6: **Output:** α.

Under Assumption 10.26, the Armijo–Goldstein condition yields the sufficient decrease estimate required in Assumption 10.24. Combined with Assumption 10.23, this gives the following convergence result for Riemannian gradient descent with backtracking.

Lemma 10.28. *Let $f : \mathcal{M} \to \mathbb{R}$ be smooth and satisfy Assumption 10.23 on a Riemannian manifold $(\mathcal{M}, \langle \cdot, \cdot \rangle)$. Let R be a retraction and assume that the pullback $f \circ R$ satisfies Assumption 10.26 on a set $S \subset T\mathcal{M}$ with constant $L > 0$. Let $x_0, x_1, x_2, \ldots$ be the iterates generated by Algorithm 2 with step sizes chosen by Algorithm 3, using fixed parameters $\tau, r \in (0, 1)$ and trial steps $\bar{\alpha}_0, \bar{\alpha}_1, \ldots$.*

If for every k,

$$\big\{(x_k, -\alpha \operatorname{grad} f(x_k)) : \alpha \in [0, \bar{\alpha}_k]\big\} \subset S \qquad \text{and} \qquad \liminf_{k\to\infty} \bar{\alpha}_k > 0,$$

then

$$\lim_{k\to\infty} \|\operatorname{grad} f(x_k)\| = 0.$$

Thus, under the above assumptions, the backtracking Riemannian gradient method drives the norm of the Riemannian gradient to zero, so any accumulation point is first-order critical. In this sense, the smooth Riemannian theory provides the model for the manifold-based descent methods we consider later.

11 Subgradients and Nonsmoothness in the Objective

In this section we study the nonsmoothness of the objective function associated with the fixed standardsimplex $S_{2n} \subset \mathbb{R}^{2n}$. Rather than differentiating the Ekeland–Hofer–Zehnder capacity as a function on convex bodies directly, we consider

$$\Phi : \mathrm{SL}(2n, \mathbb{R}) \to \mathbb{R}, \qquad X \mapsto c_{\mathrm{EHZ}}(XS_{2n}).$$

where X is an invertible linear map on $\mathbb{R}^{2n}$. Since our optimization problem is carried out on $\mathrm{SL}(2n, \mathbb{R})$, we distinguish between an ambient objective on the open set

$$U := \mathrm{GL}(2n, \mathbb{R}) \subset \mathcal{E} := \mathbb{R}^{2n \times 2n}$$

and its restriction to the embedded submanifold

$$\mathcal{M} := \mathrm{SL}(2n, \mathbb{R}) \subset U.$$

More precisely, we write

$$\bar{\Phi} : U \to \mathbb{R}, \qquad \Phi := \bar{\Phi}|_{\mathcal{M}}.$$

Thus the bar notation has the same meaning as in the previous sections: it denotes an ambient function whose restriction to $\mathcal{M}$ is the function of interest. The difference from the smooth setting considered earlier is that $\bar{\Phi}$ is no longer smooth in general, but only locally Lipschitz.

In our application we take $\mathcal{E} = \mathbb{R}^{2n \times 2n}$ equipped with the trace inner product

$$\langle A, B \rangle := \mathrm{tr}(A^{\mathsf{T}} B).$$

Although $\mathcal{E}$ is not literally $\mathbb{R}^N$, it is a finite-dimensional Hilbert space, so Clarke's theory applies verbatim; see Clarke [20]. Concretely, after choosing an orthonormal basis of $\mathcal{E}$, the resulting linear isometry $\mathcal{E} \simeq \mathbb{R}^N$ shows that all definitions and results can be formulated in a coordinate-free way on $\mathcal{E}$.

The objective $\bar{\Phi}$ is built from finitely many smooth branch functions. For each $\sigma \in \mathfrak{S}_{2n+1}$ where B is the facet matrix describing S_{2n}, we obtain

$$\bar{f}_\sigma(X) := \mathrm{tr}\big(BX^{-1} J^{\mathsf{T}} (X^{-1})^{\mathsf{T}} B^{\mathsf{T}} L_\sigma\big), \qquad X \in U.$$

Each branch $\bar{f}_\sigma$ is C^∞ on U, since inversion is smooth on $\mathrm{GL}(2n,\mathbb{R})$. The nonsmoothness enters only through the finite maximum

$$\bar{F}(X) \coloneqq \max_{\sigma\in\mathfrak{S}_{2n+1}} \bar{f}_\sigma(X),$$

and the objective can be written as

$$\bar{\Phi}(X) = \frac{(2n+1)^2}{2}\,\bar{F}(X)^{-1}. \tag{41}$$

By the sign-symmetry property proved in Remark 3.4 for skew-symmetric weights, the values $\{\bar{f}_\sigma(X)\}_{\sigma\in\mathfrak{S}_{2n+1}}$ occur in $\pm$-pairs. Hence

$$\max_\sigma \bar{f}_\sigma(X) = \max_\sigma |\bar{f}_\sigma(X)| = -\min_\sigma \bar{f}_\sigma(X) \geq 0.$$

In particular, for full-dimensional simplices we may (and will) assume that $\bar{F}(X) > 0$, so that the reciprocal in (41) is well defined and the absolute-value normalization does not change the objective. On the manifold $\mathcal{M} = \mathrm{SL}(2n,\mathbb{R})$ we then consider the restricted objective

$$\Phi = \bar{\Phi}|_{\mathcal{M}}.$$

We now develop the Clarke subdifferential calculus needed for this nonsmooth objective. We begin in the Euclidean setting in Section 11.1, where we recall the basic notions of Clarke directional derivatives, subdifferentials, and regularity, and specialize them to finite max-functions and their reciprocals, which is precisely the structure appearing in (41). We then pass to embedded Riemannian submanifolds in Section 11.2, explaining how ambient Clarke subdifferentials project onto tangent spaces. Finally, in Section 11.3, we apply this framework to the restriction of $\bar{\Phi}$ to $\mathrm{SL}(2n,\mathbb{R})$ and obtain an explicit description of $\partial_C\Phi(X)$ in terms of the active trace branches and their projected gradients. Later, in Section 11.4, we use this subdifferential framework to show that the identity matrix $I \in \mathrm{SL}(2n,\mathbb{R})$, corresponding to the standard simplex S_{2n}, is a Clarke-stationary point of the objective $X \mapsto c_{\mathrm{EHZ}}(XS_{2n})$ in every $\mathrm{SL}(2n,\mathbb{R})$ with $n \geq 1$.

11.1 The Clarke Subdifferential in Euclidean Space

In this section we collect the basic tools from Clarke's nonsmooth analysis [20] that we will use later to differentiate the objective arising from our trace-based formulation of the Ekeland–

Hofer–Zehnder capacity. More precisely, we recall local Lipschitz continuity, the Clarke directional derivative and subdifferential, regularity, and a chain rule, and then specialize these notions to finite max-functions and their reciprocals.

Definition 11.1. Let $A \subset \mathcal{E}$ and let $f : A \to \mathbb{R}^m$. We say that f is *locally Lipschitz at* $x_0 \in A$ if there exist a neighborhood W of x_0 in $\mathcal{E}$ and a constant $L > 0$ such that

$$\|f(x) - f(y)\| \le L\|x - y\| \qquad \text{for all } x, y \in A \cap W.$$

We say that f is *locally Lipschitz on* A if it is locally Lipschitz at every point $x_0 \in A$.

Definition 11.2 ([20])**.** Let $f : \mathcal{E} \to \mathbb{R}$ be locally Lipschitz. For $x \in \mathcal{E}$ and $d \in \mathcal{E}$ define the *Clarke directional derivative* by

$$f^\circ(x; d) := \limsup_{\substack{x' \to x \\ t \downarrow 0}} \frac{f(x' + td) - f(x')}{t}.$$

The *Clarke subdifferential* of f at x is

$$\partial_C f(x) := \big\{\, s \in \mathcal{E} :\ \langle s, d \rangle \le f^\circ(x; d) \text{ for all } d \in \mathcal{E} \big\}.$$

If the one-sided *directional derivative*

$$f'(x; d) := \lim_{t \downarrow 0} \frac{f(x + td) - f(x)}{t}$$

exists, then $f'(x; d) \le f^\circ(x; d)$.

We also record the notion of *(Clarke-)regularity*, since regular functions admit a particularly convenient calculus and will allow us to apply chain rules in the form needed later.

Definition 11.3 ([20])**.** A locally Lipschitz function $f : \mathcal{E} \to \mathbb{R}$ is *(Clarke-)regular* at x if $f'(x; d)$ exists for every $d \in \mathcal{E}$ and

$$f'(x; d) = f^\circ(x; d) \qquad \text{for all } d \in \mathcal{E}.$$

The main calculus tool we will use is the *Clarke chain rule*, which describes the subdifferential of a composition $f = g \circ h$ under mild hypotheses.

Theorem 11.4 ([20]). *Let $U \subset \mathcal{E}$ and let $h : U \to \mathbb{R}^n$ be given by $h = (h_1, \dots, h_n)$. Let $g : \mathbb{R}^n \to \mathbb{R}$, and define $f \coloneqq g \circ h$. Assume that each h_i is locally Lipschitz near $x \in U$ and that g is locally Lipschitz near $h(x)$. Then*

$$\partial_C f(x) \subset \overline{\mathrm{conv}}\Big\{ \sum_{i=1}^{n} \alpha_i \, \xi_i \;:\; \xi_i \in \partial_C h_i(x),\ \alpha \in \partial_C g(h(x)) \Big\}.$$

Moreover, equality holds under the following additional hypotheses, that g is strictly differentiable at $h(x)$ and $n = 1$ (then the closure is unnecessary);

Proposition 11.5. *Let $f_1, \dots, f_N : \mathcal{E} \to \mathbb{R}$ be C^1, and define*

$$F(x) \coloneqq \max_{1 \le i \le N} f_i(x), \qquad G(x) \coloneqq \frac{1}{F(x)}.$$

Assume that $F(x) > 0$ for all $x \in \mathcal{E}$. Then G is locally Lipschitz on $\mathcal{E}$.

Proof. Fix $x_0 \in \mathcal{E}$ and set

$$F_0 \coloneqq F(x_0) > 0.$$

Choose $r > 0$. Since each f_i is C^1, its gradient is continuous and hence bounded on the compact ball $\overline{B}(x_0, r)$. Define

$$L_i \coloneqq \sup_{x \in \overline{B}(x_0, r)} \|\nabla f_i(x)\| < \infty, \qquad L \coloneqq \max_{1 \le i \le N} L_i.$$

By the mean value inequality, for all $x, y \in \overline{B}(x_0, r)$ and every i,

$$|f_i(x) - f_i(y)| \le L\|x - y\|.$$

Using

$$\Big| \max_i a_i - \max_i b_i \Big| \le \max_i |a_i - b_i|,$$

we obtain, for all $x, y \in \overline{B}(x_0, r)$,

$$|F(x) - F(y)| \le L\|x - y\|.$$

Thus F is Lipschitz on $\overline{B}(x_0, r)$.

If $L = 0$, then each f_i is constant on $\overline{B}(x_0, r)$, hence F and G are constant on $B(x_0, r)$, and the claim follows. Assume now that $L > 0$, and set

$$\delta \coloneqq \min\left\{r, \frac{F_0}{2L}\right\}.$$

Then for every $z \in B(x_0, \delta)$,

$$F(z) \geq F(x_0) - |F(z) - F(x_0)| \geq F_0 - L\|z - x_0\| > F_0 - L\delta \geq \frac{F_0}{2}.$$

Hence F is bounded away from zero on $B(x_0, \delta)$.

Now let $x, y \in B(x_0, \delta)$. Then $F(x), F(y) \geq F_0/2$, and therefore

$$\begin{aligned} |G(x) - G(y)| &= \left|\frac{1}{F(x)} - \frac{1}{F(y)}\right| \\ &= \frac{|F(x) - F(y)|}{F(x)F(y)} \\ &\leq \frac{L\|x - y\|}{(F_0/2)^2} \\ &= \frac{4L}{F_0^2}\|x - y\|. \end{aligned}$$

Thus G is Lipschitz on $B(x_0, \delta)$. Since x_0 was arbitrary, G is locally Lipschitz on $\mathcal{E}$. □

A key fact in Clarke calculus is that the subdifferential of a finite maximum has an explicit description in terms of the gradients of the *active* functions.

Proposition 11.6 ([20]). *Let $\{f_\sigma\}_{\sigma \in \Sigma}$ be a finite family of C^1 functions $f_\sigma : \mathcal{E} \to \mathbb{R}$, and define the* max-function

$$F(x) \coloneqq \max_{\sigma \in \Sigma} f_\sigma(x).$$

Then F is (Clarke-)regular and

$$\partial_C F(x) = \operatorname{conv}\{\nabla f_\sigma(x) : \ \sigma \in \Sigma, \ f_\sigma(x) = F(x)\}.$$

Combining this max-formula with the chain rule yields a clean expression for the Clarke subdifferential of the reciprocal $G = 1/F$ at points where $F > 0$.

Proposition 11.7. *Let $F : \mathcal{E} \to \mathbb{R}$ be locally Lipschitz near x and (Clarke-)regular at x, and assume $F(x) > 0$. Define the* reciprocal function *G by*

$$G := \frac{1}{F}$$

on a neighborhood of x where $F > 0$. Then G is locally Lipschitz near x and

$$\partial_C G(x) = -\frac{1}{F(x)^2}\,\partial_C F(x) = \operatorname{conv}\left\{-\frac{1}{F(x)^2}\,\xi : \ \xi \in \partial_C F(x)\right\}. \tag{42}$$

Proof. Let $\phi : (0, \infty) \to \mathbb{R}$ be given by $\phi(t) = 1/t$. Then ϕ is C^1 on $(0, \infty)$; in particular it is strictly differentiable at $F(x)$:

Since F is locally Lipschitz near x, it is continuous near x. As $F(x) > 0$, there exist $\delta > 0$ and $a > 0$ such that $F(y) \geq a$ for all $y \in B(x, \delta)$. Hence $G = \phi \circ F$ is well-defined and locally Lipschitz on $B(x, \delta)$, as a composition of locally Lipschitz maps.

Now apply Theorem 11.4 with $n = 1$, $h = F$, and outer function ϕ to obtain

$$\partial_C(\phi \circ F)(x) = \phi'(F(x))\,\partial_C F(x).$$

Since $\phi'(t) = -1/t^2$, this yields (42). □

11.2 The Clarke Subdifferential on an Embedded Riemannian Submanifold

Let $\mathcal{M} \subset \mathcal{E}$ be an embedded Riemannian submanifold equipped with the induced Riemannian metric and the associated Riemannian distance $d_{\mathcal{M}}$. Our goal in this section is to obtain a convenient way to work with Clarke subgradients on $\mathcal{M}$ while performing computations in the ambient Euclidean space $\mathcal{E}$. Our presentation follows closely the Riemannian Clarke calculus developed by Hosseini and Pouryayevali [55] and the subsequent algorithmic framework of Hosseini, Huang, and Yousefpour [54], and we also draw on the embedded-submanifold perspective summarized in of Zhang, Yang, and Song [97].

Throughout this section we use the following convention. Let $U \subset \mathcal{E}$ be open, let $\mathcal{M} \subset \mathcal{E}$ be an embedded submanifold, and let

$$\bar{f} : U \to \mathbb{R}$$

be an ambient function. Its restriction to the submanifold is denoted by

$$f \coloneqq \bar{f}|_{\mathcal{M}\cap U}.$$

In the application later on, $U = \mathrm{GL}(2n, \mathbb{R})$, $\mathcal{M} = \mathrm{SL}(2n, \mathbb{R})$, and $\bar{f}$ will be instantiated by the ambient max-function $\bar{F}$ or the ambient objective $\bar{\Phi}$.

We begin by recalling local Lipschitz continuity in the intrinsic metric sense on a Riemannian manifold. This is the minimal regularity assumption needed to define Clarke's generalized directional derivative and subdifferential.

Definition 11.8. Let $(\mathcal{M}, g)$ be a finite-dimensional Riemannian manifold with Riemannian distance $d_{\mathcal{M}}$, and let $f : \mathcal{M} \to \mathbb{R}$. We say that f is locally Lipschitz at $x \in \mathcal{M}$ if there exist a neighborhood $V \ni x$ and a constant $L > 0$ such that

$$|f(p) - f(q)| \le L\, d_{\mathcal{M}}(p, q) \qquad \text{for all } p, q \in V.$$

In the embedded setting, ambient Lipschitz continuity implies intrinsic Lipschitz continuity after restriction, since the Riemannian distance dominates the ambient Euclidean distance locally.

Once local Lipschitz continuity is available, Clarke's construction extends to manifolds. A convenient way to define the generalized directional derivative intrinsically is via a retraction pullback, as in [55, 54].

Definition 11.9 ([55])**.** Let $f : \mathcal{M} \to \mathbb{R}$ be locally Lipschitz and let $R : T\mathcal{M} \to \mathcal{M}$ be a retraction. Write $R_x \coloneqq R|_{T_x\mathcal{M}}$ and define the pullback $\hat{f}_x : T_x\mathcal{M} \to \mathbb{R}$ by $\hat{f}_x(\eta) \coloneqq f(R_x(\eta))$. The Clarke generalized directional derivative of f at x in direction $v \in T_x\mathcal{M}$ is

$$f^\circ(x; v) \coloneqq \hat{f}_x^\circ(0_x; v),$$

where $0_x \in T_x\mathcal{M}$ denotes the zero vector in $T_x\mathcal{M}$. The Clarke subdifferential of f at $x \in \mathcal{M}$ is

$$\partial_C f(x) \coloneqq \big\{\xi \in T_x\mathcal{M} :\ \langle \xi, v\rangle \le f^\circ(x; v) \text{ for all } v \in T_x\mathcal{M}\big\}.$$

A key identity for locally Lipschitz functions is the support-function formula, which shows that $f^\circ(x; v)$ is the support function of the compact convex set $\partial_C f(x)$:

$$f^\circ(x; v) = \max\{\langle \xi, v\rangle :\ \xi \in \partial_C f(x)\}, \qquad v \in T_x\mathcal{M}. \tag{43}$$

To obtain clean calculus rules, we will often require a regularity assumption. Intuitively, regularity rules out pathological discrepancies between one-sided directional derivatives and generalized directional derivatives.

Definition 11.10. Let $f : \mathcal{M} \to \mathbb{R}$ be locally Lipschitz and let $R : T\mathcal{M} \to \mathcal{M}$ be the retraction fixed in Definition 11.9. For $x \in \mathcal{M}$, define the pullback

$$\hat{f}_x : T_x\mathcal{M} \to \mathbb{R}, \qquad \hat{f}_x(\eta) \coloneqq f(R_x(\eta)).$$

We say that f is *regular at x along $T_x\mathcal{M}$* if for every $v \in T_x\mathcal{M}$ the one-sided directional derivative of the pullback

$$\hat{f}'_x(0_x; v) \coloneqq \lim_{t \downarrow 0} \frac{\hat{f}_x(tv) - \hat{f}_x(0_x)}{t}$$

exists and satisfies

$$\hat{f}'_x(0_x; v) = \hat{f}^\circ_x(0_x; v) = f^\circ(x; v).$$

Equivalently,

$$\lim_{t \downarrow 0} \frac{f(R_x(tv)) - f(x)}{t} = f^\circ(x; v) \qquad \text{for all } v \in T_x\mathcal{M}.$$

The next result explains how ambient and intrinsic Clarke objects compare in tangent directions. In particular, it justifies the idea that, under appropriate regularity, Clarke subgradients on $\mathcal{M}$ can be obtained by projecting ambient Clarke subgradients onto $T_x\mathcal{M}$.

Theorem 11.11 ([97])**.** *Let $\mathcal{M}$ be an embedded submanifold of a Euclidean space $\mathcal{E}$. Let $U \subset \mathcal{E}$ be open, let $\bar{f} : U \to \mathbb{R}$ be Lipschitz at $x \in \mathcal{M} \cap U$, and set $f \coloneqq \bar{f}|_{\mathcal{M}\cap U}$. Then:*

(i) $\bar{f}^\circ(x; d) \geq f^\circ(x; d)$ for every $d \in T_x\mathcal{M}$.

(ii) $\partial_C f(x) \subset \mathrm{Proj}_x\big(\partial_C \bar{f}(x)\big)$.

(iii) If f is regular at x along $T_x\mathcal{M}$, then $\partial_C f(x) = \mathrm{Proj}_x\big(\partial_C \bar{f}(x)\big)$.

In the application later on, the ambient max-function $\bar{F}$ will be (Clarke-)regular at the points of interest, because $\bar{F}$ is a finite maximum of smooth functions. The following lemma shows that ambient regularity implies the regularity along $T_x\mathcal{M}$ needed to invoke projection-based formulae on the submanifold.

Lemma 11.12. *Let $\mathcal{M} \subset \mathcal{E}$ be a C^1 embedded submanifold with the induced metric. Let $U \subset \mathcal{E}$ be open, let $\bar{f} : U \to \mathbb{R}$ be locally Lipschitz near $x \in \mathcal{M} \cap U$, and set $f \coloneqq \bar{f}|_{\mathcal{M}\cap U}$. Assume that $\bar{f}$ is (Clarke-)regular at x in U. Then f is regular at x along $T_x\mathcal{M}$.*

Proof. Fix $v \in T_x\mathcal{M}$ and define the retraction curve

$$\gamma(t) := R_x(tv).$$

Then $\gamma(0) = x$ and $\gamma'(0) = v$ (since $DR_x(0_x) = \mathrm{Id}_{T_x\mathcal{M}}$). In particular,

$$\gamma(t) = x + tv + o(t) \qquad (t \to 0),$$

meaning

$$\frac{\|\gamma(t) - (x + tv)\|}{|t|} \to 0 \qquad (t \to 0).$$

Since $\bar{f}$ is locally Lipschitz near x, there exist $L > 0$ and a neighborhood $W \subset U$ of x such that

$$|\bar{f}(a) - \bar{f}(b)| \le L\|a - b\| \qquad \text{for all } a, b \in W.$$

For t small we have $\gamma(t)$, $x + tv \in W$, hence

$$\left| \frac{f(\gamma(t)) - f(x)}{t} - \frac{\bar{f}(x + tv) - \bar{f}(x)}{t} \right| = \left| \frac{\bar{f}(\gamma(t)) - \bar{f}(x + tv)}{t} \right| \le L\, \frac{\|\gamma(t) - (x + tv)\|}{|t|} \xrightarrow[t\downarrow 0]{} 0.$$

Therefore, whenever the Euclidean directional derivative $\bar{f}'(x; v)$ exists, the pullback directional quotient along $R_x(tv)$ has the same limit. Since $\bar{f}$ is regular at x, $\bar{f}'(x; v)$ exists for every $v \in \mathcal{E}$ and

$$\bar{f}'(x; v) = \bar{f}^\circ(x; v).$$

Hence, for our fixed $v \in T_x\mathcal{M}$,

$$\hat{f}_x'(0_x; v) = \lim_{t\downarrow 0} \frac{f(R_x(tv)) - f(x)}{t} = \bar{f}'(x; v) = \bar{f}^\circ(x; v).$$

Since $\bar{f}$ is locally Lipschitz near x, $f = \bar{f}|_{\mathcal{M}\cap U}$ is locally Lipschitz at x on $\mathcal{M}$. Consequently, the pullback $\hat{f}_x = f \circ R_x$ is locally Lipschitz near 0_x in $T_x\mathcal{M}$. By the Euclidean inequality for Clarke directional derivatives (applied to $\hat{f}_x : T_x\mathcal{M} \to \mathbb{R}$),

$$\hat{f}_x'(0_x; v) \le \hat{f}_x^\circ(0_x; v) = f^\circ(x; v).$$

Moreover, Theorem 11.11(i) gives

$$f^\circ(x; v) \le \bar{f}^\circ(x; v) \qquad \text{for all } v \in T_x\mathcal{M}.$$

Combining these identities and inequalities yields

$$f^\circ(x;v) \;\le\; \bar{f}^\circ(x;v) \;=\; \hat{f}_x'(0_x;v) \;\le\; f^\circ(x;v),$$

hence

$$\hat{f}_x'(0_x;v) = f^\circ(x;v) \qquad \text{for all } v \in T_x\mathcal{M}.$$

Therefore f is regular at x along $T_x\mathcal{M}$ in the sense of Definition 11.10. □

Finally, we record the basic first-order necessary optimality condition in Clarke calculus: local extrema are necessarily Clarke-stationary. This is precisely why our algorithms later aim to drive generalized gradients to zero, i.e. to approach points x with $0 \in \partial_C f(x)$.

Proposition 11.13 ([20])**.** *Let $f : \mathcal{M} \to \mathbb{R}$ be locally Lipschitz near x. If f attains a local minimum or a local maximum at x, then*

$$0 \in \partial_C f(x).$$

If $0 \in \partial_C f(x)$, then we say that x is a (Clarke-)stationary point.

11.3 Clarke Subdifferential of the Objective $X \mapsto c_{\mathrm{EHZ}}(XK)$ on $\mathrm{SL}(2n,\mathbb{R})$

We now apply the preceding Euclidean and manifold Clarke calculus to the function $\Phi(X) = c_{\mathrm{EHZ}}(XS_{2n})$. Let

$$\mathcal{E} = \mathbb{R}^{2n\times 2n}, \qquad U := \mathrm{GL}(2n,\mathbb{R}) \subset \mathcal{E}, \qquad \mathcal{M} := \mathrm{SL}(2n,\mathbb{R}) \subset U,$$

where $\mathcal{E}$ is equipped with the Frobenius metric

$$\langle A, B\rangle = \mathrm{tr}(A^\mathsf{T} B).$$

For each $\sigma \in \mathfrak{S}_{2n+1}$, let $\bar{f}_\sigma : U \to \mathbb{R}$ be the smooth branch function

$$\bar{f}_\sigma(X) := \mathrm{tr}\big(BX^{-1}J^\mathsf{T}(X^{-1})^\mathsf{T}B^\mathsf{T}L_\sigma\big).$$

Define the ambient finite maximum

$$\bar{F}(X) := \max_{\sigma\in\mathfrak{S}_{2n+1}} \bar{f}_\sigma(X), \qquad X \in U.$$

Since $\bar{F}$ is the maximum of finitely many smooth functions, it is locally Lipschitz and (Clarke-)regular on U, and Proposition 11.6 yields

$$\partial_C \bar{F}(X) = \operatorname{conv}\{\nabla \bar{f}_\sigma(X) : \ \sigma \in \mathcal{A}_X\}, \qquad \mathcal{A}_X := \{\sigma \in \mathfrak{S}_{2n+1} : \ \bar{f}_\sigma(X) = \bar{F}(X)\}. \tag{44}$$

Since $\bar{F}(X) > 0$ on the relevant domain, we define the ambient objective

$$\bar{\Phi}(X) := \frac{(2n+1)^2}{2} \frac{1}{\bar{F}(X)}, \qquad X \in U.$$

Its restriction to $\mathcal{M} = \mathrm{SL}(2n, \mathbb{R})$ is precisely

$$\Phi := \bar{\Phi}|_{\mathcal{M}}, \qquad \Phi(X) = c_{\mathrm{EHZ}}(XS_{2n}).$$

A convenient closed form for the Euclidean gradient of each branch is

$$\nabla \bar{f}_\sigma(X) = -\,(X^{-1})^\mathsf{T} \Big(B^\mathsf{T} L_\sigma^\mathsf{T} B\, X^{-1} J \ + \ B^\mathsf{T} L_\sigma B\, X^{-1} J^\mathsf{T}\Big)(X^{-1})^\mathsf{T}.$$

Since the scalar map $t \mapsto 1/t$ is smooth on $(0, \infty)$, Proposition 11.7 gives

$$\partial_C \left(\frac{1}{\bar{F}}\right)(X) = -\frac{1}{\bar{F}(X)^2}\, \partial_C \bar{F}(X).$$

Combining this with (44) yields, for every $X \in U$,

$$\partial_C \bar{\Phi}(X) = \operatorname{conv}\left\{-\frac{(2n+1)^2}{2} \frac{1}{\bar{F}(X)^2}\, \nabla \bar{f}_\sigma(X) : \ \sigma \in \mathcal{A}_X\right\}. \tag{45}$$

Since $\bar{\Phi}$ is locally Lipschitz on U, and $\mathcal{M} = \mathrm{SL}(2n, \mathbb{R})$ is an embedded submanifold of $\mathcal{E}$, Lemma 11.12 and Theorem 11.11 imply that for $X \in \mathrm{SL}(2n, \mathbb{R})$,

$$\partial_C \Phi(X) = \operatorname{Proj}_X\big(\partial_C \bar{\Phi}(X)\big).$$

Therefore we obtain:

$$\partial_C \Phi(X) = \operatorname{conv}\left\{-\frac{(2n+1)^2}{2} \frac{1}{\bar{F}(X)^2}\, \operatorname{Proj}_X\big(\nabla \bar{f}_\sigma(X)\big) : \ \sigma \in \mathcal{A}_X\right\}. \tag{46}$$

11.4 The Standard Simplex as a Stationary Point

Motivated by the expectation that the standard simplex $S_{2n} \subset \mathbb{R}^{2n}$ is a local maximizer of the systolic ratio among simplices, we now apply the Clarke-subdifferential framework to the objective

$$\Phi : \mathrm{SL}(2n,\mathbb{R}) \to \mathbb{R}, \qquad \Phi(X) \coloneqq c_{\mathrm{EHZ}}(XS_{2n}).$$

Under this parametrization, the standard simplex S_{2n} corresponds to the identity matrix $I \in \mathrm{SL}(2n,\mathbb{R})$. We will show that I is a Clarke-stationary point of Φ, that is,

$$0 \in \partial_C \Phi(I).$$

Equivalently, the standard simplex is Clarke-stationary for the optimization problem under consideration. Let $S_{2n} \subset \mathbb{R}^{2n}$ be the standard simplex from (16), with facet matrix B. For the associated objective

$$\Phi : \mathrm{SL}(2n,\mathbb{R}) \to \mathbb{R}, \qquad X \mapsto c_{\mathrm{EHZ}}(XS_{2n}),$$

Lemma **??** yields

$$\Phi(X) = \frac{(2n+1)^2}{2} \left(\max_{\sigma \in \mathfrak{S}_{2n+1}} \operatorname{tr}\!\big(BX^{-1}J^{\mathsf{T}}(X^{-1})^{\mathsf{T}}B^{\mathsf{T}}L_\sigma\big) \right)^{-1}, \qquad X \in \mathrm{SL}(2n,\mathbb{R}).$$

In particular,

$$\Phi(I) = c_{\mathrm{EHZ}}(S_{2n}) = \frac{(2n+1)^2}{2} \left(\max_{\sigma \in \mathfrak{S}_{2n+1}} \operatorname{tr}\!\big(BJ^{\mathsf{T}}B^{\mathsf{T}}L_\sigma\big) \right)^{-1}.$$

For every $n \in \mathbb{N}$, the maximum in

$$\max_{\sigma \in \mathfrak{S}_{2n+1}} \bar{f}_\sigma(I)$$

is attained by the following two linear orders on the vertex set $[2n+1]$:

$$\pi_1 = (1, 3, 5, \dots, 2n+1, 2, 4, \dots, 2n)$$

and

$$\pi_2 = (1, 2n, 2n-2, \dots, 4, 2n+1, 2n-1, 2n-3, \dots, 3, 2).$$

Let $\sigma_1, \sigma_2 \in \mathfrak{S}_{2n+1}$ denote the associated permutations, defined by

$$\sigma_r(k) = \pi_r(k), \qquad k = 1, \ldots, 2n+1, \quad r = 1, 2.$$

Applying $\sigma \mapsto L_\sigma = P_\sigma L P_\sigma^\mathsf{T}$, and writing the resulting $(2n+1) \times (2n+1)$ matrices in a 5×5 block form, yields

$$L_{\sigma_1} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ \mathbf{e} & C^\mathsf{T} & 0 & D^\mathsf{T} & 0 \\ 1 & \mathbf{e}^\mathsf{T} & 0 & \mathbf{e}^\mathsf{T} & 1 \\ \mathbf{e} & C^\mathsf{T} & 0 & C^\mathsf{T} & 0 \\ 1 & \mathbf{e}^\mathsf{T} & 0 & \mathbf{e}^\mathsf{T} & 0 \end{bmatrix},$$

$$L_{\sigma_2} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ \mathbf{e} & C & 0 & D & \mathbf{e} \\ 1 & \mathbf{e}^\mathsf{T} & 0 & \mathbf{e}^\mathsf{T} & 1 \\ \mathbf{e} & C & 0 & C & \mathbf{e} \\ 1 & 0 & 0 & 0 & 0 \end{bmatrix},$$

where $\mathbf{e} \in \mathbb{R}^{n-1}$ is the column vector of all ones, and $C, D \in \mathbb{R}^{(n-1)\times(n-1)}$ are given by

$$C = \begin{bmatrix} 0 & 1 & 1 & \cdots & 1 \\ 0 & 0 & 1 & \cdots & 1 \\ 0 & 0 & 0 & \cdots & 1 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 0 \end{bmatrix}, \qquad D = \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ 0 & 1 & 1 & \cdots & 1 \\ 0 & 0 & 1 & \cdots & 1 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix}.$$

To see that these permutations are indeed active at $X = I$, we compute $\operatorname{tr}(W_0 L_{\sigma_i})$ for $i \in \{1, 2\}$, where

$$W_0 \coloneqq \frac{1}{(2n+1)^2} BJB^\mathsf{T} = \begin{bmatrix} 0 & -I & \mathbf{e} \\ I & 0 & -\mathbf{e} \\ -\mathbf{e}^\mathsf{T} & \mathbf{e}^\mathsf{T} & 0 \end{bmatrix}.$$

We obtain

$$W_0 L_{\sigma_1} = \begin{bmatrix} 0 & -\mathbf{e}^\mathsf{T} & 0 & -\mathbf{e}^\mathsf{T} & -1 \\ 0 & -C & 0 & -C & -\mathbf{e} \\ -1 & 0 & 0 & 0 & 0 \\ 0 & C & 0 & D & \mathbf{e} \\ 1 & \mathbf{e}^\mathsf{T} & 0 & 0 & 1 \end{bmatrix},$$

$$W_0 L_{\sigma_2} = \begin{bmatrix} 0 & 0 & 0 & 0 & -1 \\ 0 & D & 0 & D & 0 \\ -1 & -\mathbf{e}^\mathsf{T} & 0 & -\mathbf{e}^\mathsf{T} & 0 \\ 0 & -D & 0 & -C & 0 \\ 0 & \mathbf{e}^\mathsf{T} & 0 & 0 & 1 \end{bmatrix}.$$

Therefore,

$$\operatorname{tr}(W_0 L_{\sigma_1}) = \operatorname{tr}(W_0 L_{\sigma_2}) = n.$$

It follows that

$$\bar{F}(I) = n \qquad \text{and hence} \qquad \Phi(I) = c_{\mathrm{EHZ}}(S_{2n}) = \frac{1}{2n}.$$

Recall that P_I denotes the Frobenius-orthogonal projection onto the tangent space

$$T_I \mathrm{SL}(2n, \mathbb{R}) = \{V \in \mathbb{R}^{2n \times 2n} : \operatorname{tr}(V) = 0\},$$

and is given explicitly by

$$P_I(V) = V - \frac{\operatorname{tr}(V)}{2n} I_{2n}.$$

At $X = I$, the branch-gradient formula from the previous subsection simplifies to

$$\nabla \bar{f}_\sigma(I) = -(B^\mathsf{T} L_\sigma^\mathsf{T} B J + B^\mathsf{T} L_\sigma B J^\mathsf{T}).$$

Thus the corresponding projected gradients satisfy

$$P_I\big(B^\mathsf{T} L_{\sigma_1}^\mathsf{T} B J + B^\mathsf{T} L_{\sigma_1} B J^\mathsf{T}\big) = -\frac{1}{n} \begin{bmatrix} C^\mathsf{T} - C & C - C^\mathsf{T} \\ C^\mathsf{T} - C & C - C^\mathsf{T} \end{bmatrix} = -\,P_I\big(B^\mathsf{T} L_{\sigma_2}^\mathsf{T} B J + B^\mathsf{T} L_{\sigma_2} B J^\mathsf{T}\big).$$

Now define

$$\xi_i := -\frac{(2n+1)^2}{2} \, \frac{1}{\bar{F}(I)^2} \, \mathrm{Proj}_I\big(\nabla \bar{f}_{\sigma_i}(I)\big), \qquad i = 1, 2.$$

Since σ_1 and σ_2 are active at I, formula (46) shows that $\xi_1, \xi_2 \in \partial_C \Phi(I)$. Moreover, the above computation shows that $\operatorname{Proj}_I(\nabla \bar{f}_{\sigma_2}(I)) = -\operatorname{Proj}_I(\nabla \bar{f}_{\sigma_1}(I))$, and hence $\xi_2 = -\xi_1$. Therefore $0 = \frac{1}{2}(\xi_1 + \xi_2) \in \partial_C \Phi(I)$. Thus I is a Clarke-stationary point of the optimization problem

$$X \longmapsto \Phi(X) = c_{\mathrm{EHZ}}(XS_{2n}) \qquad \text{on } \mathrm{SL}(2n, \mathbb{R}).$$

With these computations we have shown that the necessary first-order conditions for the standard simplex to be a local maximizer in the Clarke sense are fulfilled. To investigate whether it is actually an extreme point, one needs second-order methods.

12 Algorithmic Approaches

In this section we turn the structural results developed so far into computational procedures. Our goal is to search for simplices with high systolic ratio, in particular among centred simplices. The objective functions arising from the EHZ capacity admit convenient reformulations (Section 9), but remain *nonsmooth* because they are given by maxima over finitely many trace expressions. This naturally suggests subgradient-type methods on manifolds. We then exploit the special structure of our objective, namely that its Clarke subdifferential can be computed explicitly as the convex hull of finitely many active branch gradients, and use this to construct a descent method driven by the minimum-norm Clarke subgradient. The practical implementation of this scheme, including the extraction of active permutations and the convex QP used to compute minimum-norm subgradients, is described in Section 12.3.

Since large-scale experiments produce many candidate simplices, we also need an algorithmic way to decide when two outputs represent the same simplex up to affine symplectomorphisms. This is the topic of Section 12.2, where we use the graph-isomorphism viewpoint developed earlier to test affine symplectomorphism efficiently in the centred setting.

The numerical outcomes of these computations are collected in Section 12.4, where we report the best values found in dimensions $n = 3, \ldots, 8$. Finally, Section 12.5 discusses some of the questions suggested by these experiments, including the motivation for second-order tools near multi-active stationary points and possible alternative optimization approaches on related matrix groups such as $\mathrm{SO}(n)$.

12.1 Retraction-Based Line-Search on Manifolds

Let $(\mathcal{M}, \langle\cdot,\cdot\rangle)$ be a (complete) Riemannian manifold and let $f : \mathcal{M} \to \mathbb{R}$ be locally Lipschitz. Many first-order schemes on manifolds follow the Euclidean blueprint: at $x_k \in \mathcal{M}$ compute a search direction $p_k \in T_{x_k}\mathcal{M}$, choose a step size $\alpha_k > 0$ by an inexact line search, and update via a retraction

$$x_{k+1} = R_{x_k}(\alpha_k p_k),$$

see, for instance, Absil, Mahony and Sepulchre [3] and Gabay [32]. For nonsmooth locally Lipschitz objectives, one replaces gradients by generalized derivatives; a systematic Riemannian Clarke calculus is developed by Hosseini and Pouryayevali [55]. Practical methods include ε-subgradient schemes [36], gradient sampling [56], trust-region methods [37], and retraction-based line-search schemes with nonsmooth Armijo/Wolfe conditions developed by Hosseini–Huang–Yousefpour [54].

In our setting, a natural starting point is the conceptual Riemannian subgradient scheme of Ferreira–Oliveira [29], as implemented for manifold optimization in `Manopt` [14]. Its convergence theory, however, relies on a positive sectional-curvature assumption. For $\mathrm{SL}(n,\mathbb{R})$ equipped with the Frobenius-induced metric, this assumption fails, as follows from the sectional-curvature computation of Rosevear, Sottile, and Wong [81]. Thus, in our setting, the Ferreira–Oliveira framework does not provide descent guarantees for an arbitrary choice of subgradient direction, and this motivates the more selective retraction-based line-search method developed below. The line-search framework of Hosseini–Huang–Yousefpour [54] provides such an alternative in a general nonsmooth manifold setting: it uses an ε-subdifferential as an algorithmic outer approximation of $\partial_C f(x)$, built by sampling/transporting subgradients from a neighborhood. In our setting, however, $\partial_C f(x)$ is available in closed form at every iterate, typically as the convex hull of finitely many smooth Riemannian gradients from a finite active set. Hence we can adapt this line-search perspective to our problem while working directly with $\partial_C f(x)$ and omitting the ε-approximation machinery.

Definition 12.1. Let $f : \mathcal{M} \to \mathbb{R}$ and let $R : T\mathcal{M} \to \mathcal{M}$ be a retraction. Fix $x \in \mathcal{M}$, and let $0_x \in T_x\mathcal{M}$ denote the zero vector. Define the pullback

$$\hat{f}_x : T_x\mathcal{M} \to \mathbb{R}, \qquad \hat{f}_x(\eta) := f(R_x(\eta)).$$

A direction $p \in T_x\mathcal{M}$ is called a *descent direction* for f at x if there exists $\alpha > 0$ such that for every $t \in (0, \alpha)$,

$$f(R_x(tp)) - f(x) = \hat{f}_x(tp) - \hat{f}_x(0_x) < 0.$$

Lemma 12.2. *If $f^\circ(x; p) < 0$, then p is a descent direction at x.*

Proof. Let

$$\hat{f}_x(\eta) := f(R_x(\eta)).$$

Assume that p is not a descent direction at x. Then, by Definition 12.1, for every $\alpha > 0$ there exists $t \in (0, \alpha)$ such that

$$f(R_x(tp)) - f(x) \geq 0.$$

Equivalently,

$$\hat{f}_x(tp) - \hat{f}_x(0_x) \geq 0.$$

Hence we can choose a sequence $t_k \downarrow 0$ such that

$$\hat{f}_x(t_k p) - \hat{f}_x(0_x) \geq 0 \qquad \text{for all } k.$$

Since $t_k > 0$, it follows that

$$\frac{\hat{f}_x(t_k p) - \hat{f}_x(0_x)}{t_k} \geq 0 \qquad \text{for all } k,$$

and therefore

$$\limsup_{k\to\infty} \frac{\hat{f}_x(t_k p) - \hat{f}_x(0_x)}{t_k} \geq 0.$$

By Definition 11.9,

$$f^\circ(x;p) = \hat{f}_x^\circ(0_x;p),$$

and by definition of the Clarke directional derivative,

$$\hat{f}_x^\circ(0_x;p) = \limsup_{\substack{\eta\to 0_x\\ t\downarrow 0}} \frac{\hat{f}_x(\eta + tp) - \hat{f}_x(\eta)}{t}.$$

In particular, this limsup includes the special choice $\eta = 0_x$ and $t = t_k$, so

$$\hat{f}_x^\circ(0_x;p) \geq \limsup_{k\to\infty} \frac{\hat{f}_x(t_k p) - \hat{f}_x(0_x)}{t_k} \geq 0.$$

Thus

$$f^\circ(x;p) \geq 0.$$

We have shown that if p is not a descent direction at x, then $f^\circ(x;p) \geq 0$. The contrapositive yields the claim. $\square$

The preceding implication is consistent with the discussion in the literature: Hosseini and Uschmajew [56] note that negativity of the Clarke generalized directional derivative yields a descent direction in the Riemannian setting, referring to the Euclidean result of Mäkelä–Neittaanmäki [68, Theorem 5.2.5].

A particularly natural choice of search direction for locally Lipschitz objectives is obtained by taking the element of smallest norm in the Clarke subdifferential. This is the analog of the steepest-descent direction in smooth optimization and is also the direction used (in an approximate form of ε) in [54].

Definition 12.3. Let $x \in \mathcal{M}$ and let $\partial_C f(x) \subset T_x\mathcal{M}$ be the Clarke subdifferential. Since $\partial_C f(x)$ is nonempty, compact, and convex, there exists at least one element of minimal norm, i.e. there exists $w(x) \in \partial_C f(x)$ such that

$$\|w(x)\| = \min\{\|\xi\| : \ \xi \in \partial_C f(x)\}.$$

We call any such $w(x)$ a *minimum-norm Clarke subgradient* at x and define the associated search direction

$$p(x) := -\, w(x) \in T_x\mathcal{M}.$$

Lemma 12.4. *Let $x \in \mathcal{M}$ and let $w = w(x)$ be a minimum-norm Clarke subgradient at x in the sense of Definition 12.3. Then*

$$f^\circ(x; -w) = -\|w\|^2.$$

In particular, if $w \neq 0$ (equivalently $0 \notin \partial_C f(x)$), then $p = -w$ is a descent direction *at x.*

Proof. Set $C := \partial_C f(x) \subset T_x\mathcal{M}$. By construction, $w \in C$ minimizes the convex function $\xi \mapsto \frac{1}{2}\|\xi\|^2$ over the nonempty, closed, convex set C.

Minimizing $\xi \mapsto \frac{1}{2}\|\xi\|^2$ over C is equivalent to finding the point of C closest to the origin, i.e. $w = \arg\min_{\xi \in C} \|\xi\|$. Since C is nonempty, closed, and convex, this minimizer exists and is unique, and it is precisely the metric projection of 0 onto C, denoted $\mathrm{proj}_C(0)$. A standard characterization of metric projections onto closed convex sets is the variational inequality

$$\langle \xi - w,\ 0 - w\rangle \le 0 \qquad \text{for all } \xi \in C,$$

see, e.g. by Rockafellar, [80]. Therefore

$$\langle \xi,\, -w\rangle \le -\|w\|^2 \qquad \text{for all } \ \xi \in C. \tag{47}$$

Taking $\xi = w$ in (47) shows that equality is attained, hence

$$\max_{\xi \in C} \langle \xi,\, -w\rangle = \langle w,\, -w\rangle = -\|w\|^2.$$

Using the support-function identity (43) gives

$$f^\circ(x; -w) = \max_{\xi \in \partial_C f(x)} \langle \xi,\, -w\rangle = -\|w\|^2.$$

If $w \neq 0$, then $f^\circ(x; -w) < 0$, and by Lemma 12.2 the direction $-w$ is a descent direction at x. □

The above construction is the Clarke-subdifferential analogue of the minimum-norm ε-subgradient directions used in [37, Theorem 3.12] and [54, Theorem 2.13]. Since $\partial_C f(x)$ is nonempty, compact, and convex, the minimum-norm element $w(x)$ is well defined; the identity $f^\circ(x; -w(x)) = -\|w(x)\|^2$ is obtained by the same projection argument as in [54, Theorem 2.13].

Definition 12.5 ([54])**.** Fix $c_1 \in (0, 1)$. A step size $\alpha > 0$ satisfies the *(nonsmooth) Armijo condition* if

$$f(R_x(\alpha p)) - f(x) \leq c_1 \, \alpha \, f^\circ(x; p). \tag{48}$$

This is the manifold analogue of the usual sufficient decrease test; cf. [54] and the nonsmooth equivalent of the Armijo condition for smooth functions in manifolds.

Proposition 12.6. *Let $x \in \mathcal{M}$ be such that $0 \notin \partial_C f(x)$, and let $p = -w(x)$ where $w(x)$ is any minimum-norm Clarke subgradient.*

Then $f^\circ(x; p) = -\|w(x)\|^2 < 0$ and the Armijo condition (48) is equivalent to

$$f(R_x(\alpha p)) \leq f(x) - c_1 \, \alpha \, \|w(x)\|^2. \tag{49}$$

When f is C^1, the Clarke directional derivative coincides with the classical one, $f^\circ(x; p) = \langle \operatorname{grad} f(x), p \rangle$. In particular, for the gradient step $p = -\operatorname{grad} f(x)$ we have $f^\circ(x; p) = -\|\operatorname{grad} f(x)\|^2$, and our nonsmooth Armijo condition $f(R_x(\alpha p)) - f(x) \leq c_1 \alpha f^\circ(x; p)$ reduces exactly to the standard smooth Armijo (sufficient decrease) inequality

$$f(x) - f\big(R_x(-\alpha \operatorname{grad} f(x))\big) \geq c_1 \, \alpha \, \|\operatorname{grad} f(x)\|^2.$$

On the basis of the subgradient-descent algorithm discussed in [29] and the algorithms discussed in this section, we developed a line-search algorithm which is a special case of the algorithm in [54]:

Algorithm 4 Exact Clarke steepest descent with Armijo backtracking

Input: Riemannian manifold $\mathcal{M}$, locally Lipschitz and regular $f : \mathcal{M} \to \mathbb{R}$, retraction R, parameters $c_1 \in (0, 1)$, $\rho \in (0, 1)$, initial step $\alpha_0 > 0$, tolerance $\delta > 0$, start $x_1 \in \mathcal{M}$.

Assumption: for each $x \in \mathcal{M}$ the Clarke subdifferential is available in the form $\partial_C f(x) = \operatorname{conv}\{g_1(x), \dots, g_m(x)\} \subset T_x\mathcal{M}$.

for $k = 1, 2, 3, \dots$ **do**
  Compute the minimum-norm Clarke subgradient: find $w_k \in \partial_C f(x_k)$ with

$$\|w_k\| = \min\{\|\xi\| : \ \xi \in \partial_C f(x_k)\}.$$

  if $\|w_k\| \le \delta$ **then**
    return x_k ▷ approximately *Clarke-stationary*
  end if
  Descent direction: $p_k \leftarrow -w_k$
  Backtracking: $\alpha \leftarrow \alpha_0$
  while $f(R_{x_k}(\alpha p_k)) > f(x_k) - c_1\, \alpha\, \|w_k\|^2$ **do**
    $\alpha \leftarrow \rho\, \alpha$
  end while
  Update: $x_{k+1} \leftarrow R_{x_k}(\alpha p_k)$
end for

If

$$\partial_C f(x) = \operatorname{conv}\{g_1(x), \dots, g_m(x)\} \subset T_x\mathcal{M},$$

then $w(x) = \sum_{i=1}^{m} \lambda_i g_i(x)$ where λ solves the convex quadratic program

$$\min_{\lambda \in \mathbb{R}^m} \left\| \sum_{i=1}^{m} \lambda_i g_i(x) \right\|^2 \quad \text{s.t.} \quad \lambda_i \ge 0,\ \sum_{i=1}^{m} \lambda_i = 1. \tag{50}$$

Thus, in the *explicit active-set* case, one replaces the sampling/transport loop of $\partial_\varepsilon f(x)$-based methods by the QP (50).

In [54] there has been developed a more advanced version of the algorithm we are using. In our setting, we apply their framework to the special case where the Clarke subdifferential is available explicitly and, moreover, we choose the positive semidefinite matrix P (used there as a preconditioning/generalization option) to be the identity I. Hosseini, Huang and

Yousefour show in [54], that all accumulation points are Clarke-stationary. Reframing the corresponding result to our setting yields:

Lemma 12.7. *Assume $\mathcal{M}$ is complete and f is locally Lipschitz, the initial level set $N := \{x \in \mathcal{M} : \; f(x) \leq f(x_1)\}$ is bounded, each step length α_k is chosen by Armijo backtracking, i.e. satisfies (49) with $p_k = -w_k$ and w_k the minimum-norm Clarke subgradient at x_k. Under the assumptions above, either the algorithm terminates in finitely many steps with $0 \in \partial_C f(x_k)$, or every accumulation point $\bar{x}$ of $\{x_k\}$ is* Clarke-stationary*.*

We have nearly all the assumptions fulfilled: In our setting, let $\mathcal{E} = \mathbb{R}^{2n\times 2n}$ be equipped with the Frobenius inner product $\langle A, B\rangle = \operatorname{tr}(A^\mathsf{T} B)$, so $\mathcal{E}$ is (isometric to) Euclidean space and hence complete. The determinant map $\det : \mathcal{E} \to \mathbb{R}$ is smooth, and for every $X \in \mathrm{SL}(2n, \mathbb{R})$ its differential $D(\det)_X$ is nonzero, so 1 is a regular value; therefore $\mathrm{SL}(2n, \mathbb{R}) = \det^{-1}(1)$ is an embedded submanifold of $\mathcal{E}$ by the regular level set theorem as stated by Lee [66, Cor. A.26]. Moreover, since det is continuous and $\{1\}$ is closed, the set $\mathrm{SL}(2n, \mathbb{R})$ is closed in $\mathcal{E}$. Since closed embedded submanifolds of complete Riemannian manifolds are complete (with the induced metric), as recorded for example in Florit's notes [31, Corollary 37], it follows that $\mathrm{SL}(2n, \mathbb{R})$ is a complete Riemannian manifold with the metric induced by the Frobenius inner product.

The function we are working with being locally Lipschitz has been discussed already. Sadly we cannot guarantee that the initial level set $N := \{x \in \mathcal{M} : \; f(x) \leq f(x_1)\}$ is bounded. Hence, with this approach for this function we cannot prove that we reach accumulation points which are Clarke-stationary. But since we can guarantee descent we implemented this algorithm anyway and achieved interesting results.

12.2 Testing Affine Symplectomorphisms

By the centred reduction from the proof of Theorem 7.4, two full-dimensional simplices $\Delta(V) = \operatorname{conv}\{v_0, \dots, v_{2n}\}$ and $\Delta(W) = \operatorname{conv}\{w_0, \dots, w_{2n}\} \subset \mathbb{R}^{2n}$ are affinely symplectomorphic if and only if, after translating both simplices to centroid 0, their skew-symmetric pairing matrices

$$W_V := X^\mathsf{T} J X, \qquad W_W := Y^\mathsf{T} J Y \qquad \big(X = [\, v_0 \ \cdots \ v_{2n} \,],\ Y = [\, w_0 \ \cdots \ w_{2n} \,]\big)$$

are permutation–conjugate:

$$\text{There exists a } \sigma \in \mathfrak{S}_{2n+1} \text{ such that } \quad P_\sigma^\mathsf{T} W_W P_\sigma = W_V.$$

Thus detecting symplectic duplicates reduces to a single instance of weighted (edge-labeled) graph isomorphism on $2n+1$ vertices, rather than enumerating $(2n+1)!$ permutations. This brings in the full algorithmic toolbox for graph isomorphism, such as canonical labeling and refinement–individualization methods (e.g. `nauty`/`Traces` [70]). In our implementation we use `networkx` [43] to search for a weight-preserving isomorphism between the complete graphs defined by W_V and W_W, together with simple invariants (row/column weight multisets) to prune the permutation search.

12.3 Implementation Details

In this section we summarize the main implementation choices behind our retraction-based descent method on $\mathrm{SL}(2n, \mathbb{R})$. Concretely, we implemented Algorithm 4 from the previous section and now explain how its abstract ingredients (tangent space, projection, retraction, objective evaluation, and Clarke steepest-descent direction) are realized in our setting.

Manifold model: $\mathrm{SL}(2n, \mathbb{R})$ as an embedded submanifold.
We work on $\mathrm{SL}(2n, \mathbb{R})$ as an embedded submanifold of the Euclidean space $\mathcal{E} = \mathbb{R}^{2n \times 2n}$ equipped with the Frobenius inner product $\langle A, B \rangle = \operatorname{tr}(A^\mathsf{T} B)$. Using the manifold material from Section 10, the tangent space at $X \in \mathrm{SL}(2n, \mathbb{R})$ is

$$T_X \mathrm{SL}(2n, \mathbb{R}) = \{\, V \in \mathcal{E} : \ \operatorname{tr}(X^{-1} V) = 0 \,\}.$$

Given an ambient direction $V \in \mathcal{E}$, we use the explicit Frobenius-orthogonal projection $\mathrm{Proj}_X : \mathcal{E} \to T_X\mathrm{SL}(2n, \mathbb{R})$ defined by

$$\mathrm{Proj}_X(V) = V - \lambda\,(X^{-1})^{\mathsf{T}}, \qquad \lambda = \frac{\mathrm{tr}(X^{-1}V)}{\mathrm{tr}(X^{-1}(X^{-1})^{\mathsf{T}})}.$$

Equivalently, Proj_X removes the component of V in the normal direction $(X^{-1})^{\mathsf{T}}$, which spans the Frobenius-orthogonal complement of $T_X\mathrm{SL}(2n, \mathbb{R})$.

Retraction / update rule.
To update along a tangent direction we require a retraction $R_X : T_X\mathrm{SL}(2n, \mathbb{R}) \to \mathrm{SL}(2n, \mathbb{R})$, as developed in Section 10. For the actual iteration steps we use the exponential update

$$\mathrm{Exp}_X(\xi) = X \exp(X^{-1}\xi), \qquad \xi \in T_X\mathrm{SL}(2n, \mathbb{R}),$$

which is more expensive than elementary retractions but empirically very robust.

Objective Evaluation via the ILP formulation.
As explained in Section 9, evaluating c_{EHZ} reduces to solving a linear ordering ILP, which we solve with Gurobi [40]. At each iterate X we compute:

- The objective value by solving the ILP instance induced by the current matrix $W(X)$,
- a set of optimal solutions from the solver's solution pool, which represent active maximizers (L-matrices corresponding to active permutations),
- for each active solution L, an ambient "branch gradient" $G_L(X) \in \mathcal{E}$ obtained from the smooth trace expression, and its tangent projection $\mathrm{Proj}_X(G_L(X)) \in T_X\mathrm{SL}(2n, \mathbb{R})$.

This turns the abstract assumption in Algorithm 4,

$$\partial_C f(X) = \mathrm{conv}\{g_1(X), \dots, g_m(X)\} \subset T_X\mathcal{M},$$

into a concrete representation of the Clarke subdifferential as a finite convex hull, as derived in Section 11.

Clarke steepest-descent direction via a convex QP.
Given the projected active gradients

$$g_1(X), \dots, g_m(X) \in T_X\mathrm{SL}(2n, \mathbb{R}), \qquad \partial_C f(X) = \mathrm{conv}\{g_1(X), \dots, g_m(X)\},$$

Algorithm 4 chooses the search direction by the minimum-norm element of this convex set. Equivalently, we solve the convex quadratic program

$$\min_{\lambda \in \mathbb{R}^m} \left\| \sum_{i=1}^{m} \lambda_i g_i(X) \right\|^2 \quad \text{s.t.} \quad \lambda_i \geq 0,\ \sum_{i=1}^{m} \lambda_i = 1, \tag{51}$$

and set

$$w(X) = \sum_{i=1}^{m} \lambda_i g_i(X), \qquad p(X) = -w(X).$$

We solve (51) with a standard convex QP routine (implemented via `cvxpy`, typically using OSQP).

Step sizes.
A full Armijo backtracking line search is robust, but in our setting it is costly: each trial step requires a fresh evaluation of the objective, hence another ILP solve. For this reason, we use a phased strategy:

- In early phases we use fixed step sizes ("loose", "middle", "tight", etc.) to move quickly toward a promising region with relatively few ILP evaluations.
- We allow small tolerated increases in the objective value in these phases (within prescribed tolerances) to avoid excessive backtracking and to prevent stalling due to overly conservative step sizes.
- Only in the final phase we enforce an Armijo-type acceptance test, so that the method behaves like a genuine descent scheme close to a stationary region.

Computational bottlenecks.
Two costs dominate the runtime:

(a) *ILP solves.* Evaluating the objective and extracting active solutions requires solving an ILP and querying the solution pool. This is the dominant cost when the line search performs many trial steps.

(b) *Minimum-norm QP.* As n increases, the number of active solutions (and hence the number of active gradients) can grow quickly, and the dimension of the gradients grows as well. Consequently, the QP (51) can become the new bottleneck.

To keep the method feasible we cap the number of extracted optimal ILP solutions (solution-pool cut-off, in our experiments typically at 5000). This preserves the same mathematical structure, but in practice it replaces the full active set by a truncated active set and may therefore reduce accuracy at higher dimensions. Within these computational constraints, the resulting routine is a practical realization of Algorithm 4, specialized to $\mathrm{SL}(2n,\mathbb{R})$ and to the explicit Clarke subdifferential structure derived from the ILP-based trace formulation.

12.4 Results

We report the best values of the objective

$$\Phi(X) = c_{\mathrm{EHZ}}(XS_{2n})$$

obtained by our retraction-based Clarke steepest-descent scheme (Algorithm 4) for dimensions $2n$ with $n = 2, \ldots, 8$. As a stationarity diagnostic we record the norm of the minimum-norm Clarke subgradient

$$w(X^\star) \in \partial_C \Phi(X^\star)$$

returned by the convex QP (51); small values of $\|w(X^\star)\|$ indicate that the output is approximately Clarke-stationary.

As a baseline we recall that the standard simplex S_{2n} satisfies $c_{\mathrm{EHZ}}(S_{2n}) = \frac{1}{2n}$ (see the computation in Section 11), so it is natural to also report the ratio

$$\frac{\Phi(X^\star)}{\Phi(I)} = \frac{c_{\mathrm{EHZ}}(X^\star S_{2n})}{c_{\mathrm{EHZ}}(S_{2n})} \; = \; 2n\, c_{\mathrm{EHZ}}(X^\star S_{2n}),$$

which measures the improvement over the standard simplex in each dimension. Our computations empirically corroborate the observations of Chaidez and Hutchings [16], that in dimension 4 ($n = 2$) the standard simplex appears to maximize the systolic ratio among simplices. In dimension $n = 3$, our best run attains a value that is slightly larger than the value 0.8333555549629893 reported by Haim-Kislev [45]. Across all tested dimensions, the computed upper bounds to $\|w(X^\star)\|$ are typically small, indicating that the returned points are close to Clarke-stationary for the truncated active sets used in the QP. Nevertheless, the experiments also suggest that slight further improvements may still be possible, though only at the scale of very small steps and at significant additional computational cost.

Empirically, one recurring outcome of the algorithm is the the value of the standard simplex $c_{\mathrm{EHZ}}(S_{2n}) = \frac{1}{2n}$. This behavior is frequent in the lower dimensions, but the hit rate decreases as n increases: in higher dimensions, the algorithm less often returns the standard simplex as a maximizer encountered during the runs, and more often converges to

n	$m = 2n+1$	max $c_{\mathrm{EHZ}}(XS_{2n})$	$2n\, c_{\mathrm{EHZ}}(X^\star S_{2n})$	$\mathrm{sys}_n(X^\star S_{2n})$	$\mathrm{sys}_n(S_{2n})$
2	5	**0.2500000000**	**1.0000000000**	**0.8660254037**	**0.8660254037**
3	7	**0.1689718761**	**1.0138312566**	**0.8334409621**	**0.8220706914**
4	9	**0.1267600379**	**1.0140803031**	**0.8115395331**	**0.8002714682**
5	11	**0.1016619511**	**1.0166195113**	**0.8003405068**	**0.7872566854**
6	13	**0.0846278014**	**1.0155336162**	**0.7907037574**	**0.7786091418**
7	15	**0.0720974936**	**1.0093649110**	**0.7796810846**	**0.7724471855**
8	17	**0.0625503327**	**1.0008053229**	**0.7684524643**	**0.7678341100**

Table 1: Best values observed for $\Phi(X) = c_{\mathrm{EHZ}}(XS_{2n})$ for $n = 2, \dots, 8$, together with the corresponding Viterbo-normalized systolic ratios $\mathrm{sys}_n(K) := c_{\mathrm{EHZ}}(K)/(n!\,\mathrm{Vol}(K))^{1/n}$. Since $\mathrm{Vol}(XS_{2n}) = \mathrm{Vol}(S_{2n})$, the improvement factor over the standard simplex is $c_{\mathrm{EHZ}}(X^\star S_{2n})/c_{\mathrm{EHZ}}(S_{2n}) = 2n\, c_{\mathrm{EHZ}}(X^\star S_{2n})$.

other Clarke-stationary candidates. The simplices corresponding to the reported values are documented in the appendix. A link to the code used to generate these results is included there as well.

12.5 Discussion

We conclude with a brief discussion of second-order questions suggested by our results.

Second-Order Considerations.

For smooth functions on $\mathrm{SL}(n, \mathbb{R})$, second-order information is naturally encoded by the Riemannian Hessian, and in Section 10.3 we derived explicit formulas for this object. For the nonsmooth max-type objectives studied in this thesis, however, there is no comparably satisfactory intrinsic second-order theory currently available on $\mathrm{SL}(n, \mathbb{R})$. In particular, at points where several smooth branches are active, one should not expect branchwise curvature alone to capture the full local behavior.

A natural ambient point of reference for such questions is the variational second-order theory of Mordukhovich. In the Euclidean setting, the second-order chain rule of Mordukhovich–Rockafellar [72] and the max-formula of Emich–Henrion [28] show that for finite maxima of smooth functions, the relevant second-order behavior is governed not only by the Hessians of the active branches, but also by additional terms that encode how the active set changes under perturbation. Since our objective has exactly this max-structure, these results indicate why branchwise Hessians alone are not expected to be sufficient, and they suggest the form that a more satisfactory second-order theory in our setting should take.

What we did compute in our numerical experiments are the Riemannian Hessians of the active smooth branches. These serve as a natural second-order diagnostic. Empirically, at multi-active stationary points, the Hessians of the active branches are typically neither positive nor negative semidefinite. Thus, branchwise second-order information by itself does not explain the observed local stability of the numerically detected stationary points. This strongly suggests that the relevant local maxima occur at nonsmooth switching points of the active branches, where several branches meet and remain competitive under perturbation.

We now return to $\mathcal{M} = \mathrm{SL}(n, \mathbb{R}) \subset \mathcal{E}$. As in the smooth case, our guiding principle is: *differentiate in the ambient space, restrict to tangent directions, and project back to the tangent space.* Recall that the Frobenius-orthogonal projection onto $T_X\mathcal{M}$ is

$$\mathrm{Proj}_X(V) = V - \frac{\mathrm{tr}\big(X^{-1}V\big)}{\mathrm{tr}\big(X^{-1}(X^{-1})^\mathsf{T}\big)}(X^{-1})^\mathsf{T}, \qquad V \in \mathcal{E}. \tag{52}$$

For each $\sigma \in \Sigma$, let $\bar{f}_\sigma$ be a C^2 ambient extension of $f_\sigma\big|_{\mathcal{M}}$. Then the smooth branch gradient and Hessian on $\mathcal{M}$ are obtained exactly as in Section 10.3:

$$\mathrm{grad}\, f_\sigma(X) = \mathrm{Proj}_X\big(\nabla \bar{f}_\sigma(X)\big), \qquad \mathrm{Hess}\, f_\sigma(X)[V] = \mathrm{Proj}_X\big(D\bar{G}_\sigma(X)[V]\big),$$

where $\bar{G}_\sigma$ is any smooth ambient extension of $X \mapsto \mathrm{grad}\, f_\sigma(X)$.

In the applications studied in this thesis, the ambient gradients of the branch functions are of the form

$$\nabla_X f_\sigma(X) = (X^{-1})^\mathsf{T} \left(BL_\sigma^\mathsf{T} B^\mathsf{T} X^{-1} J^\mathsf{T} - BL_\sigma B^\mathsf{T} X^{-1} J^\mathsf{T}\right) (X^{-1})^\mathsf{T}.$$

Consequently, to obtain $\nabla_X^2 \bar{f}_\sigma(X)[V]$ we compute

$$D\big((X^{-1})^\mathsf{T} \left(BL_\sigma^\mathsf{T} B^\mathsf{T} X^{-1} J^\mathsf{T} - BL_\sigma B^\mathsf{T} X^{-1} J^\mathsf{T}\right) (X^{-1})^\mathsf{T}\big)\, [V].$$

After applying the product rule this yields:

$$\begin{aligned}
\mathrm{Hess}\, f_\sigma(X)[V] = \mathrm{Proj}_X\Big(&-(X^{-1})^\mathsf{T} V^\mathsf{T} (X^{-1})^\mathsf{T} \big(BL_\sigma^\mathsf{T} B^\mathsf{T} X^{-1} J^\mathsf{T} - BL_\sigma B^\mathsf{T} X^{-1} J^\mathsf{T}\big)(X^{-1})^\mathsf{T} \\
&-(X^{-1})^\mathsf{T} \big(BL_\sigma^\mathsf{T} B^\mathsf{T} X^{-1} V X^{-1} J^\mathsf{T} - BL_\sigma B^\mathsf{T} X^{-1} V X^{-1} J^\mathsf{T}\big)(X^{-1})^\mathsf{T} \\
&-(X^{-1})^\mathsf{T} \big(BL_\sigma^\mathsf{T} B^\mathsf{T} X^{-1} J^\mathsf{T} - BL_\sigma B^\mathsf{T} X^{-1} J^\mathsf{T}\big)(X^{-1})^\mathsf{T} V^\mathsf{T} (X^{-1})^\mathsf{T}\Big) \\
&+ \frac{\mathrm{tr}\big(X^{-1} \nabla \bar{f}_\sigma(X)\big)}{\mathrm{tr}\big(X^{-1}(X^{-1})^\mathsf{T}\big)} \mathrm{Proj}_X\big((X^{-1})^\mathsf{T} V^\mathsf{T} (X^{-1})^\mathsf{T}\big)\,.
\end{aligned}$$

This is the branchwise second-order formula that we actually implemented. At a stationary point $\bar{X}$, we compute the Hessians of the active smooth branches and test them for positive or negative semidefiniteness. As noted above, these branch Hessians are typically indefinite at multi-active points, which reinforces the view that branchwise curvature alone does not explain the local maximality mechanism.

Motivated by the ambient second-order theory for finite maxima, one may nevertheless regard the following projected object as a plausible second-order model on $T_{\bar{X}}\mathcal{M}$. For $V \in T_{\bar{X}}\mathcal{M}$, define

$$d_\sigma(V) \coloneqq Df_\sigma(\bar{X})[V] = \langle \operatorname{grad} f_\sigma(\bar{X}), V \rangle, \qquad d(V) \coloneqq (d_\sigma(V))_{\sigma\in\Sigma} \in \mathbb{R}^{|\Sigma|},$$

and set

$$\mathcal{H}^{\max}_{\bar{X},\lambda}(V) \coloneqq \left\{ \sum_{\sigma\in\Sigma} \lambda_\sigma \operatorname{Hess} f_\sigma(\bar{X})[V] + \sum_{\sigma\in\Sigma} q_\sigma \operatorname{grad} f_\sigma(\bar{X}) \;\middle|\; q \in \partial^2\theta\big(H(\bar{X}), \lambda\big)\big(d(V)\big) \right\} \subset T_{\bar{X}}\mathcal{M}. \tag{53}$$

We do not claim a full rigorous manifold theory for (53) in the present thesis. Rather, this formula should be understood as a conceptual model that indicates the direction in which a genuinely nonsmooth second-order theory on $\mathrm{SL}(n,\mathbb{R})$ would have to proceed: combining branch curvature with the geometry of switching between active branches.

A Rotationally Constrained Variant: Optimization on $\mathrm{SO}(2n)$

In addition to searching over the entire special linear group $\mathrm{SL}(2n,\mathbb{R})$, one can also study the *rotation-only* variant obtained by restricting the shape parameter to the special orthogonal group $\mathrm{SO}(2n) \coloneqq \{X \in \mathbb{R}^{2n\times 2n} : X^\mathsf{T} X = I,\ \det(X) = 1\} \subset \mathrm{SL}(2n,\mathbb{R})$.

Geometrically, this means that we only rotate the reference simplex, rather than allowing general volume-preserving deformations. This restriction is still nontrivial in our setting: symplectic capacities are not invariant under arbitrary Euclidean rotations unless the rotation is symplectic (i.e. lies in $\mathrm{Sp}(2n,\mathbb{R}) \cap \mathrm{SO}(2n)$); hence $c_{\mathrm{EHZ}}(XS)$ typically changes with $X \in \mathrm{SO}(2n)$. From the manifold-optimization viewpoint, $\mathrm{SO}(2n)$ is a smooth embedded submanifold of $\mathbb{R}^{2n\times 2n}$. Equipped with the Frobenius inner product $\langle A, B\rangle = \operatorname{tr}(A^\mathsf{T} B)$, it becomes a Riemannian submanifold, and the basic ingredients (tangent space, projection, retraction) have standard closed forms; see Boumal [13]. Restricting the search space from $\mathrm{SL}(2n,\mathbb{R})$ to $\mathrm{SO}(2n)$ yields a compact feasible set and a conceptually clean "rotation-only" benchmark problem. While we do not pursue this variant further here, it provides a useful sanity check and an alternative experiment.

Acknowledgements

I gratefully acknowledge financial support by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) through the Collaborative Research Centre / Transregio 191 *Symplectic Structures in Geometry, Algebra and Dynamics* and the support through the GreenEPS project, funded by the Federal Ministry for Research, Technology and Space (BMFTR).

I am grateful to my supervisor, Professor Frank Vallentin, for his guidance and for the inspiration, trust, and freedom he gave me in developing my work.

I also thank Professor Alberto Abbondandolo for kindly agreeing to serve as second reviewer of this thesis and for being a motivating conversation partner.

I thank Andreas Spomer for going through the PhD years alongside me. Sharing this time, with its challenges and successes, meant a great deal. I also thank Arne Heimendahl for many valuable discussions and for the time we spent sharing an office, even if that time was unfortunately too short.

I also thank Marc Zimmermann, Annette Koenen and the entire group for the supportive atmosphere, the many helpful conversations, and the enjoyable breaks that made the trip to the office worthwhile.

I am grateful to my family for their constant support. I especially thank my grandfather, who first got me excited about mathematics.

I also thank Lukas Winkel, Stefan Kupp, Arne Heimendahl, Jorin Dornemann, and Andreas Spomer for proofreading parts of this thesis.

Finally, I thank my friends for listening, for their encouragement, and for sharing life and adventures beyond mathematics.

Appendix

In this appendix we record the simplices $K_X = \{z \in \mathbb{R}^{2n} : X^{-1}Bz \leq \mathbf{e}\}$ recovered from the gradient descent method of Section 12.1 that produced the largest systolic ratio in the available dimensions. Their exact c_{EHZ}-values and systolic ratios are also listed in Table 1, furthermore they are contained in the GitHub respository.

The exact computation of c_{EHZ} and of the systolic ratio is implemented in `ehzofsimplex.py`, based on the trace reformulation and ILP procedure from Section 9. The gradient descent search is implemented in `gradientdescent.py`. The repository also contains code for testing affine symplectomorphism of simplices, corresponding to Section 12.2 and its theoretical background in Section 7; see `Symplectomorphismtest.py`. These programs, together with brief instructions for their use, are available at

https://github.com/kleipold/CombinatorialApproachestoehzcapacityofsimplex.

The material contained in this repository in the version available up to the `20.03.2026` was developed by the author in connection with the research presented in this dissertation.

The numerical experiments in this thesis were carried out in an exploratory manner rather than as a systematic benchmarking study. Still, the computational effort was substantial: for each dimension $n = 3, \dots, 6$, at least 10,000 runs of the gradient-based method were carried out. Individual runs were typically allowed up to about 10,000 iterations, a deliberately conservative choice, since in lower dimensions convergence was often observed earlier. For these lower-dimensional cases, the computations were feasible on a standard computer and typically terminated within one day, whereas the results for the higher-dimensional cases $n = 7$ and $n = 8$ were obtained using HPC resources.

Dimension $2n = 6$.

$$(X^\star)^{-1}B = \begin{bmatrix} -3.4594173020e-01 & 4.6063569115e-01 & 4.6665298109e-01 & -3.1225645494e-01 & -2.6488464076e-01 & -2.6461648451e-01 & 2.6041063816e-01 \\ 1.0017392276e+00 & 1.0578346299e+00 & -4.1207834098e-01 & 8.4586896646e-02 & -5.8910004912e-01 & 3.7618475531e-01 & -1.5191671193e+00 \\ -2.3601180876e-02 & -9.3534343866e-03 & -5.5533418872e-01 & -1.0269637896e+00 & 1.0096498069e+00 & 5.9378474481e-01 & 1.1818041880e-02 \\ -2.7402504637e-01 & -5.3950783972e-01 & -3.3944579047e-01 & 1.7942989392e-01 & -3.4654252047e-01 & -8.5855137084e-01 & 2.1786426739e+00 \\ -7.4757744364e-02 & 2.8824798679e-01 & -4.6089755369e-01 & 7.5189181289e-01 & -5.2585792551e-01 & 4.5514071909e-01 & -4.3376729520e-01 \\ -8.1219556203e-02 & -5.6938478629e-01 & 8.0963565800e-02 & -3.8207289078e-01 & -1.2213686682e-01 & 1.6898450382e-01 & 9.0486603047e-01 \end{bmatrix}$$

$$\begin{array}{ll} c_{\mathrm{EHZ}}(X^\star S_{2n}) & 0.1689718761 \\ \mathrm{sys}_n(X^\star S_{2n}) & 0.8334409621 \end{array}$$

Dimension $2n = 8$.

$$(X^\star)^{-1}B = \begin{bmatrix}
-7.5875612133e-01 & -8.3340317343e-01 & -2.9763767170e-01 & 4.7534506196e-01 & -7.9263650418e-01 & 6.7261791642e-01 & -4.2436534172e-01 & -1.3164009777e-02 & 1.9719998438e+00 \\
-6.6459583621e-01 & -2.3344052023e-02 & 1.2773133077e+00 & 3.7784932088e-01 & -3.1498578021e-01 & 5.3934695994e-01 & -5.4288893283e-01 & 2.7272047410e-02 & -6.7596703466e-01 \\
9.4218461176e-01 & 6.3851224200e-01 & -6.6114035778e-01 & 4.0685169190e-01 & -4.8705548644e-01 & 3.6733848927e-02 & 6.8284538880e-02 & 2.0013802665e-02 & -9.6438489192e-01 \\
9.1357014765e-01 & 6.4599789634e-02 & -8.3527689036e-01 & -1.3192003423e+00 & 1.6205550772e-02 & -7.7881331765e-01 & 1.0470588563e-01 & 6.1955599612e-02 & 1.7722535770e+00 \\
-3.3403032382e-01 & -1.1178547464e-02 & 9.5838771402e-01 & -9.3017357967e-02 & 4.0537163063e-01 & 4.6471852852e-01 & 2.7753385849e-01 & -6.6195457003e-01 & -1.0058309324e+00 \\
1.2404735402e+00 & 1.0559833389e+00 & -3.2818602432e-01 & -3.3560878631e-01 & -1.3172223329e-01 & 6.4295644819e-02 & 8.8864221856e-01 & 1.8691799972e-01 & -2.6407956982e+00 \\
-5.7325497947e-01 & -4.6438255794e-01 & -6.8165631284e-01 & -1.4550688185e-01 & 8.2541148701e-02 & 1.3900138236e-01 & -5.1623262207e-01 & 1.8466321350e-01 & 1.9748276096e+00 \\
-6.0124807248e-02 & 2.3764787120e-01 & -8.3010611534e-01 & -3.3635288398e-01 & -3.0188563811e-01 & -6.6010760287e-01 & 1.9531001645e-01 & -3.2999122795e-01 & 2.0856103879e+00
\end{bmatrix}$$

$$\begin{array}{ll}
c_{\mathrm{EHZ}}(X^\star S_{2n}) & 0.1267600379 \\
\mathrm{sys}_n(X^\star S_{2n}) & 0.8115395331
\end{array}$$

Dimension $2n = 10$.

$$(X^\star)^{-1}B = \begin{bmatrix}
2.6061577684e-01 & -1.1508043299e-02 & 3.6252641225e-01 & 3.8752798182e-01 & -3.3115442292e-01 & -5.4290388569e-01 & -7.7224312436e-01 & -4.4819688046e-01 & 5.1744974093e-01 & -7.6118671445e-02 & 6.5400511633e-01 \\
-4.9580296423e-01 & 1.7592612015e-01 & 2.8289285049e-02 & 5.3818530989e-01 & -2.6158894035e-01 & 5.1912369350e-01 & -4.6031120857e-01 & 4.1114096648e-02 & -1.8140508763e+00 & -8.5844027388e-02 & 1.8149595116e+00 \\
1.6035354492e-01 & -1.4690932502e-01 & 2.9751890889e-02 & -2.4663179019e-01 & 5.6998705465e-01 & -4.7576458884e-01 & 1.4827721880e+00 & -2.8860688517e-01 & -3.1557225799e-02 & -3.0187494967e-01 & -7.5151991373e-01 \\
8.9078540156e-02 & -5.3781998855e-02 & 3.7487538979e-01 & -1.6567411973e-01 & -3.1433822889e-01 & -2.4236129053e-01 & 5.3689982564e-01 & 4.7614349806e-01 & 7.4633517586e-01 & 8.2502592548e-01 & -2.2722027170e+00 \\
1.4637454430e-01 & -1.5836667863e-01 & 8.5008462585e-02 & -1.6390290160e-01 & -6.0890451811e-01 & -6.1492498227e-01 & -6.8324509772e-01 & -4.0441400419e-01 & 1.7445258206e-01 & 4.7795932871e-01 & 1.7499632649e+00 \\
3.4756077012e-01 & 5.7174869386e-01 & -4.7873665990e-01 & -3.0927407232e-01 & 8.6430089416e-01 & 6.4137162797e-01 & -5.4462191881e-02 & 8.1624922800e-01 & -5.0524552480e-01 & -8.6775536517e-02 & -1.8067372287e+00 \\
-1.0806309044e-01 & 3.9332114787e-03 & 5.4571344297e-01 & -3.7679691117e-01 & -6.8012636822e-01 & -7.1511664016e-01 & 8.2044469710e-01 & -7.2494638673e-01 & 5.5227012566e-01 & 2.6439488411e-01 & 4.1829303541e-01 \\
7.4296655908e-01 & -3.4013453700e-01 & -1.3082730365e+00 & 3.0140566602e-01 & 1.7400295770e-01 & 6.1889473442e-01 & -4.4028713335e-01 & 3.7279877161e-01 & 3.1850641711e-01 & -4.9359059912e-01 & 5.3710200051e-02 \\
1.8810967337e-01 & 7.8204303779e-02 & 7.2952557367e-01 & 6.0667798022e-03 & 2.5208684141e-03 & -2.3800207014e-01 & -7.4012695796e-02 & -9.7645805961e-01 & 2.5616773484e-01 & 5.1530188864e-01 & -4.8742399697e-01 \\
7.6932909143e-02 & 1.0197774576e-01 & -2.5873253329e-01 & -1.6873498085e-01 & -7.9468475774e-01 & -7.0157484025e-01 & -3.2603821009e-01 & -1.0692168526e-01 & 1.2142140366e+00 & 8.5780304200e-01 & 1.0575927401e-01
\end{bmatrix}$$

$$\begin{array}{ll}
c_{\mathrm{EHZ}}(X^\star S_{2n}) & 0.1016619511 \\
\mathrm{sys}_n(X^\star S_{2n}) & 0.8003405068
\end{array}$$

Dimension $2n = 12$.

$$(X^\star)^{-1}B = \begin{bmatrix}
-4.1942947907e-01 & -4.5140508773e-01 & 6.4789354726e-01 & -9.9693996520e-01 & -4.9720713606e-01 & -9.0134132917e-02 & -1.4136828108e-03 & -6.4340314256e-01 & -5.6921042421e-02 & 5.7768765980e-01 & 9.4419997268e-01 & 2.5348967938e-01 & 7.3358280965e-01 \\
-6.7882174292e-01 & -5.9104183900e-01 & 2.1372926621e-01 & -1.3996562361e+00 & 1.3587434827e-01 & -1.5006334795e+00 & -3.7319269947e-03 & 2.5204716260e-01 & -1.4490296008e-01 & 4.9678853112e-01 & 5.9887916594e-02 & -4.3870563143e-01 & 3.5991665912e+00 \\
7.9843185375e-01 & -3.8400860082e-01 & 1.4822240778e-01 & 2.5337729929e-01 & -5.4560021109e-01 & -3.1210864948e-02 & -1.7726687053e-02 & -3.0494693549e-01 & -6.8807344798e-02 & 7.4751207236e-01 & -8.9877382202e-01 & -6.5070074966e-01 & 9.5423158271e-01 \\
1.1178647629e+00 & 6.1295662035e-02 & 2.0360345183e-01 & 1.2088491407e+00 & 1.6311937851e-01 & 6.7325523261e-01 & -9.4339017860e-02 & -4.4898007210e-01 & -8.4381820062e-02 & -4.9202291390e-01 & -5.4971120682e-01 & -2.1082729880e-01 & -1.5477252990e+00 \\
3.0289957157e-01 & -3.5412230477e-01 & 4.2266847738e-01 & 1.6801746560e-01 & -1.5189430667e-01 & 5.5534243767e-01 & 1.1339553792e-01 & 1.4223033539e-01 & -2.0046014258e-01 & -1.1425667974e-01 & -4.5650561377e-01 & 1.4321565762e-01 & -5.7053043563e-01 \\
-1.0040977696e-01 & -3.1502378665e-01 & 2.4744073672e-01 & 2.6352747778e-01 & 6.1288915432e-01 & 2.3786747533e-03 & -2.1456147871e-01 & 2.4837146725e-01 & -8.7316600303e-02 & -2.3301960439e-01 & 3.2490520643e-01 & 1.9910595288e-01 & -9.4828742312e-01 \\
4.7611084556e-01 & 9.0300419155e-01 & -6.3608963098e-02 & 2.9072597075e-01 & 5.4105144593e-02 & 6.0647906119e-02 & 1.8859412473e-01 & 2.5192234196e-02 & -2.0267346508e-01 & -2.0186323509e-01 & 3.5678165913e-01 & 4.8891083863e-01 & -2.3759272520e+00 \\
-3.5018962063e-01 & 2.2776970978e-01 & -3.1073079152e-02 & 1.1371996447e-01 & 2.4953637849e-01 & 2.6485312533e-01 & 1.1879216270e-01 & 6.1893580630e-01 & 7.3821645525e-01 & -2.2105027196e-01 & 5.2968656581e-01 & -1.8241613829e-01 & -2.0767810581e+00 \\
5.2670164325e-01 & -1.2999344892e-01 & 5.8447126205e-01 & 5.4209669543e-01 & -5.0349178781e-01 & 5.4997029719e-01 & -3.6745930500e-01 & 1.9875151838e-01 & -9.2295072198e-02 & -2.6253261634e-01 & -1.5861638444e-02 & 3.9922166982e-01 & -1.4295792174e+00 \\
1.0109600495e+00 & 2.8707824408e-01 & -8.1880613330e-01 & -7.5163131274e-02 & -5.9685016629e-01 & -1.8542691175e-01 & 4.5671601451e-01 & -2.6405352112e-01 & -9.5798626608e-01 & 2.7635411361e-01 & -2.5997976539e-01 & -5.0794636712e-01 & 1.6351038406e+00 \\
-1.0649642015e+00 & 2.5022032478e-01 & -4.8104172126e-01 & -1.7331929935e+00 & 3.7739332691e-01 & -3.8745748290e-01 & -5.9767665957e-02 & 4.2740531649e-02 & 1.7328824709e-01 & 4.9259647688e-01 & 1.0062390033e+00 & 2.0862212455e-02 & 1.3630839421e+00 \\
-4.5252217575e-01 & -2.4228186723e-01 & -3.8763783334e-02 & 7.5115470057e-02 & 9.3503991535e-01 & -7.5845283842e-02 & -2.9567389752e-01 & 6.1468475373e-01 & 1.0710171297e+00 & -7.1686638918e-01 & 2.4030452128e-01 & 7.7524700196e-01 & -1.8894553952e+00
\end{bmatrix}$$

$$\begin{array}{ll}
c_{\mathrm{EHZ}}(X^\star S_{2n}) & 0.0846278014 \\
\mathrm{sys}_n(X^\star S_{2n}) & 0.7907037574
\end{array}$$

Dimension $2n = 14$.

$$
(X^\star)^{-1}B = \begin{bmatrix}
1.8429701305e-01 & 3.6418471811e-01 & 5.6434586024e-01 & 2.0835801809e-01 & -4.6472622600e-02 & 5.6702180891e-01 & 6.8623293702e-01 & 1.7434164077e-01 & 1.5566090609e-01 & 1.4436798405e-01 \\
2.2792062784e-02 & -1.3204917927e-01 & 1.1045512776e-01 & 5.3807185238e-01 & -3.5416081274e+00 \\
1.2500111379e-01 & 6.3094797950e-01 & 9.4932482833e-02 & -1.6108337445e-01 & 2.4470564466e-01 & 6.5244694486e-01 & -3.7916416936e-03 & -4.4689301389e-01 & 4.6193174826e-01 & -4.4398138007e-01 \\
2.0877098239e-01 & -1.2396527463e-01 & 2.3431884261e-01 & 2.1907481530e-01 & -1.6924158695e+00 \\
7.3975311947e-01 & 1.0094857014e+00 & 3.5975327396e-01 & 3.6408119608e-01 & -2.8005984503e-01 & -2.7050443862e-01 & -2.1021199741e-01 & 5.7387753894e-01 & 4.8343366318e-02 & 2.4110923627e-01 \\
5.5034305938e-01 & -2.8797469612e-01 & 1.5995264230e-02 & -2.2767575852e-01 & -2.6263150204e+00 \\
-1.3102715349e-01 & -2.7592880981e-01 & 1.4357765046e-01 & -2.2792014401e-01 & 7.4990964978e-01 & 8.7862980966e-02 & -4.7195166275e-01 & 4.4482507424e-01 & -8.5518928983e-02 & 2.0297721376e-01 \\
2.4080687610e-01 & -2.5151122869e-01 & -4.2715212322e-01 & -5.5166341174e-01 & 5.5271401740e-01 \\
5.2551630034e-01 & -6.6371822816e-01 & 2.6970697345e-01 & 1.9388533251e-01 & 6.0456415200e-01 & -2.2146252145e-02 & 4.6229719788e-01 & -1.7511078510e-01 & 3.7261450588e-02 & 3.4879772405e-01 \\
1.7651794936e-01 & -4.1664147433e-01 & 6.6180672343e-01 & -5.9955696629e-01 & -1.4031800976e+00 \\
-6.3376612284e-01 & -8.6043975931e-02 & 8.7030645891e-01 & 4.7231803307e-01 & -6.9374011602e-02 & 1.0099608494e-01 & 5.1594825833e-01 & -6.8953846488e-02 & -2.6110597387e-02 & -4.0224793127e-01 \\
8.5335466761e-01 & 2.2006418553e-01 & 3.8139956484e-01 & 2.5623231386e-01 & -2.3841230816e+00 \\
2.3296364992e-01 & -6.4503286757e-01 & -2.8664465716e-01 & 9.0808782506e-03 & 6.0859730932e-01 & 4.8481357917e-02 & -2.2391603899e-01 & 4.6732557992e-02 & -2.4191232446e-02 & -6.7702071181e-02 \\
-3.5019049878e-01 & -4.1911065930e-02 & -2.0735551293e-01 & 5.3646770662e-02 & 8.4744142094e-01 \\
5.9046946535e-02 & 3.7466351137e-01 & 4.6158543081e-01 & 1.8692908277e-01 & 1.1380311163e+00 & 1.2738870628e-01 & 2.9471793312e-02 & -3.3124299218e-01 & 1.4370734607e-01 & -1.8977999137e-02 \\
1.1649189339e-01 & 1.1684754298e-02 & -6.1900765225e-02 & 7.2799532565e-02 & -2.3096783571e+00 \\
-3.1031329176e-01 & -2.3062801170e-01 & -6.7844726383e-01 & 2.2658318617e-01 & -5.8353060343e-02 & -8.8405420497e-02 & -4.7042546056e-02 & -4.3735706764e-01 & 1.2172742100e-01 & 5.1400632771e-01 \\
-1.0020856648e-01 & -1.7521553551e-01 & 4.8111421158e-01 & 3.2691260635e-01 & 4.5562701100e-01 \\
1.0119878272e-01 & 1.2367237340e-01 & 2.2531364138e-02 & 7.5789046846e-02 & -7.2339397062e-01 & -1.3569132025e-01 & -6.2074080253e-02 & -2.1539271137e-02 & 2.2194377060e-01 & -3.3552590227e-02 \\
3.4971128807e-02 & 2.2001980576e-01 & -3.6182730685e-01 & -2.1612879905e-01 & 7.5408106611e-01 \\
6.7755118933e-01 & -4.4449925374e-02 & -7.7448309955e-01 & -5.0490486740e-01 & 5.0786319905e-03 & -2.2709340366e-01 & -2.1710600819e-02 & 6.0575923746e-02 & -4.3409712392e-01 & -1.0344748337e-02 \\
-9.9981534456e-02 & 7.8727182281e-02 & -1.3027042291e+00 & -1.1947219319e-01 & 2.7173087985e+00 \\
-4.6931804029e-01 & -9.8742458793e-01 & 2.4291387824e-01 & -9.9261882778e-01 & -1.6061610590e-01 & 3.0743900351e-01 & -5.2834815334e-01 & -3.3504829515e-01 & -6.8006103982e-01 & -1.6350046797e-01 \\
-2.9241582834e-01 & 2.4125744321e-01 & -3.1329084544e-01 & 1.3554559233e-01 & 3.9954862747e+00 \\
9.3637505727e-02 & 6.3641251465e-01 & -7.6771585372e-01 & 1.0545991199e-01 & -1.2611207943e+00 & -2.1014705299e-01 & -1.2798596862e-01 & 1.6777511533e-01 & 4.8466911915e-01 & 1.0247628936e-01 \\
5.2010855660e-02 & 6.5270269132e-01 & 3.1643206307e-01 & 6.3698704062e-02 & -3.0830510067e-01 \\
-9.3811246657e-02 & -7.5003903806e-02 & -7.8829500808e-01 & 6.8169219535e-01 & -5.1506487858e-01 & -3.6119125340e-02 & 1.5089467895e-01 & 7.6210525253e-02 & -1.0658747588e-01 & 3.0207371590e-01 \\
-4.2194154425e-01 & 5.1088669105e-02 & 3.8980579549e-01 & -3.7526367883e-01 & 7.6032128137e-01
\end{bmatrix}
$$

$c_{\mathrm{EHZ}}(X^\star S_{2n})$	0.0720974936
$\mathrm{sys}_n(X^\star S_{2n})$	0.7796810846

Dimension $2n = 16$.

$$
(X^\star)^{-1}B = \begin{bmatrix}
-6.1780246107e-01 & -3.7356396514e-01 & 1.0776935342e+00 & -1.9150690332e-01 & -6.4923872024e-01 & 1.2916464249e-01 & -2.2830739703e-01 & 9.3829325517e-02 & -1.1124201173e-01 & -2.9342452007e-01 \\
-3.3751491339e-01 & 1.8978721350e-01 & 3.1959590486e-01 & 4.5251243346e-01 & -4.7285628568e-01 & 5.0166437192e-01 & 5.1120975173e-01 \\
-3.6446914027e-02 & 4.6561605596e-02 & -3.2031306628e-01 & -2.2649787937e-01 & -1.0255854833e+00 & -8.9149161393e-01 & -2.0730465484e-01 & -2.3792600521e-01 & -2.2451605269e-01 & 1.5313308507e+00 \\
-1.1655639909e-01 & 3.6683624758e-01 & -4.9858770655e-01 & -5.2359324947e-02 & -1.6599809238e-01 & -3.2628550090e-01 & 2.3851399896e+00 \\
2.0086082948e-01 & -7.5425660612e-02 & 7.2765871385e-02 & 3.5469652023e-01 & 7.5845391333e-02 & 4.1014603002e-01 & -5.7642313281e-01 & 1.2724201672e-01 & -8.4555962333e-04 & -2.3715119698e-02 \\
3.3255358250e-01 & -1.1389889375e-01 & -4.8839880482e-01 & 1.2947268237e-01 & 6.2061051604e-01 & 7.5532596532e-02 & -1.1210188653e+00 \\
-2.5395059935e-01 & -3.1562028019e-02 & 9.8163392877e-02 & 3.1024087213e-01 & -8.7756999204e-03 & 3.1812511611e-01 & 3.3644983504e-01 & 3.3990298075e-01 & 1.6788619088e-01 & -4.6640549769e-01 \\
9.3584625133e-02 & 4.0482233457e-01 & -9.4072706996e-02 & 1.6391900151e-01 & -1.0033802092e-01 & -4.6379681953e-01 & -8.1419297658e-01 \\
5.7437501683e-01 & 1.9078478719e-01 & -2.0423991635e-02 & 8.6974528233e-02 & 1.5856698632e-02 & 3.2023470812e-01 & -9.3206168094e-02 & 8.1336872968e-02 & 2.9799760418e-01 & -2.8282828778e-02 \\
1.1889470380e-02 & 2.2982117012e-01 & -5.7735027002e-01 & -4.7365663118e-01 & 8.9822631945e-02 & 5.3278212063e-01 & -1.2389557195e+00 \\
3.3793724145e-01 & 3.5055726961e-01 & -4.1535763521e-01 & -4.8157384995e-02 & -1.0874678657e-02 & 4.9539546760e-01 & 1.9310835931e-02 & 5.0393702846e-01 & 5.1242262790e-01 & -8.3454974806e-01 \\
2.1027894128e-01 & -4.6197628471e-01 & -2.1366542963e-01 & 6.4505503281e-02 & 1.8020717666e-01 & 1.5037879587e-01 & -8.4034972679e-01 \\
-1.6019295180e-01 & 1.3046938780e-01 & 8.5641042576e-01 & -2.7905548559e-01 & 2.0387522067e-01 & 2.0951985598e-02 & 6.7712612512e-01 & -1.2827229507e-01 & -4.8984601446e-01 & 8.2490520122e-01 \\
-2.6159177621e-01 & -2.7578288371e-01 & 6.5984782813e-01 & -4.4212980501e-01 & -7.8194288504e-02 & 3.1889276469e-01 & -1.5774134386e+00 \\
1.7873510580e-01 & -1.2987825220e-01 & -2.5086681136e-01 & 4.9837895358e-01 & 3.6292119481e-01 & -3.0338942168e-01 & -1.8676850028e-01 & -2.6337770590e-01 & 1.1403457449e-01 & -7.3140467928e-02 \\
4.3390151248e-01 & 4.5283688651e-02 & 3.3461107236e-01 & -3.9606343601e-01 & 3.7625901037e-01 & 2.1571745109e-01 & -9.5635796828e-01 \\
-4.7113607065e-01 & -4.6988505467e-01 & -6.1010405411e-01 & -3.9601443035e-02 & -7.3716548272e-01 & -5.3596689889e-01 & 5.0335313083e-02 & 4.1214512823e-01 & 5.0218429964e-01 & 4.4726417076e-01 \\
-1.6100854902e-01 & 3.4946159489e-01 & -6.4709549199e-01 & 6.8691616691e-02 & -7.8257772764e-02 & -2.9400042123e-01 & 2.2141391158e+00 \\
4.5620434434e-01 & 1.7969541072e-01 & -2.6100736285e-03 & -1.1402995573e-02 & 2.0370392119e-01 & -4.3225612531e-01 & -3.8125672752e-01 & -5.6696440807e-01 & 2.0501312283e-01 & -4.5617871007e-01 \\
-2.7537552070e-01 & -3.1090299145e-01 & -6.3319183533e-03 & -2.3604480878e-01 & 4.6836955435e-01 & -1.3442672911e-01 & 1.3007646551e+00 \\
-1.5908849043e-01 & 2.2221030174e-01 & -4.8703862776e-02 & -4.8962662256e-01 & 1.4363398773e-01 & 2.4572492448e-01 & 3.3041166428e-01 & -4.6507604803e-01 & 7.6451181996e-01 & -4.9897033344e-01 \\
-9.0972071894e-03 & -1.2534337633e-01 & 1.0448168208e-01 & 3.5309138362e-01 & 1.8002633535e-01 & -2.2826566954e-01 & -3.1992048896e-01 \\
3.9687408106e-01 & 2.8268573644e-01 & -5.7101082319e-01 & 6.1841539731e-02 & 1.0515875668e+00 & 3.0392312230e-01 & 5.4172821828e-03 & 2.6717791509e-01 & -6.8528967471e-01 & 6.9404731386e-01 \\
1.8494908612e-01 & 7.2270153809e-02 & 6.5622084064e-01 & -6.1304431486e-01 & 5.3460629829e-01 & 4.7577318109e-02 & -2.6898334416e+00 \\
-8.1294980098e-02 & -1.3482522744e-01 & -8.3887304548e-02 & -4.7173865785e-02 & -1.0168073926e-01 & -5.2946826443e-01 & -2.7919641842e-01 & 4.9199914355e-01 & 5.3034273270e-01 & -3.2228464990e-01 \\
7.9745933925e-03 & -3.7811505075e-02 & -2.4635363772e-01 & 6.1525093804e-01 & -6.8973765106e-01 & 2.1154074334e-01 & 6.9660609273e-01 \\
-2.0786396986e-01 & -3.8926542117e-01 & 2.9713604299e-01 & -5.5877859936e-01 & 5.4382626157e-01 & -4.6863828746e-01 & 2.1958878372e-01 & -1.7189450761e-01 & -5.3421685676e-01 & 2.0897559216e-01 \\
1.8690041998e-01 & 1.6029698280e-01 & 1.0345753120e+00 & -2.0400720980e-01 & -2.1913438090e-01 & -5.0442825051e-01 & 6.0692808826e-01 \\
-3.0188349192e-01 & 1.7888293460e-01 & 6.9222147945e-01 & 3.3653183376e-01 & -2.2480800999e-01 & 6.0270757888e-01 & -2.4827773667e-01 & 3.5668527705e-01 & -8.2621230982e-02 & 3.6276222760e-01 \\
2.1306758486e-01 & -1.2851958402e-01 & -7.2250033891e-02 & -2.8852428282e-01 & -7.1005800297e-01 & 2.9184572442e-01 & -9.7776226734e-01 \\
-1.8344298093e-01 & 2.6597368975e-01 & 1.5314425816e-02 & -6.7520040211e-02 & -1.7462651277e-01 & 5.6608881760e-02 & -1.3463738178e-01 & -1.2521299661e-01 & 1.3072427473e-01 & -1.0324494967e+00 \\
2.4052809372e-01 & 1.3994295097e-01 & -5.6448729223e-01 & 2.2490693879e-01 & 3.8625524788e-01 & 6.4270736376e-02 & 7.5785146143e-01
\end{bmatrix}
$$

$c_{\mathrm{EHZ}}(X^\star S_{2n})$	0.0625503327
$\mathrm{sys}_n(X^\star S_{2n})$	0.7684524643

Notation

Symbol	Meaning
General / linear algebra	
$\mathbb{R}, \mathbb{C}, \mathbb{Z}, \mathbb{N}$	Real/complex numbers, integers, naturals.
$[m]$	The set $\{1, 2, \ldots, m\}$.
$\langle \cdot, \cdot \rangle$	Standard inner product in the ambient Euclidean space.
$\\| \cdot \\|$	Norm induced by $\langle \cdot, \cdot \rangle$ (Euclidean/Frobenius): $\\|u\\| = \sqrt{\langle u, u \rangle}$.
$\mathrm{tr}(\cdot)$, $\det(\cdot)$	Trace and determinant.
$A^\top$	Transpose of a matrix A.
I_n	Identity matrix in $\mathbb{R}^{n \times n}$.
I	Identity matrix (dimension clear from context).
$\mathbf{e}$	All-ones vector (dimension clear from context).
e_i	i-th standard basis vector (Einheitsvektor) in $\mathbb{R}^n$.
$\mathrm{rank}(A)$	Rank of a matrix A.
$\mathrm{conv}(S)$	Convex hull of a set S.
Symplectic geometry	
$(\mathbb{R}^{2n}, \omega_0)$	Standard symplectic vector space.
ω_0	Standard symplectic form on $\mathbb{R}^{2n}$.
J	Matrix corresponding to the standard symplectic form (as fixed in the thesis).
$\mathrm{Sp}(2n, \mathbb{R})$	Symplectic group.
$c_{\mathrm{EHZ}}(K)$	Ekeland–Hofer–Zehnder capacity of a convex body K.
$\mathrm{sys}_n(K)$	Systolic ratio of a convex body K.
Convex geometry / polytopes	
$K \subset \mathbb{R}^d$	Convex body.
$h_K(u)$	Support function of K.
$N_K(x)$	Normal cone of K at a boundary point x.
S_{2n}	The (centered) standard simplex in $\mathbb{R}^{2n}$ as defined in (16).
Permutations, orders, and graphs	
$\mathfrak{S}_m$	Symmetric group on $[m] = \{1, \ldots, m\}$.
$\sigma \in \mathfrak{S}_m$	A permutation of $[m]$; when used to encode a linear order, $\sigma(1) \prec \sigma(2) \prec \cdots \prec \sigma(m)$, so σ lists the elements in order.

Symbol	Meaning
P_σ	Permutation matrix associated with σ, defined by $P_\sigma e_i = e_{\sigma(i)}$; equivalently, $(P_\sigma)_{ij} = 1$ iff $i = \sigma(j)$.
L	Strict upper-triangular 0–1 matrix of appropriate size; for $\sigma \in \mathfrak{S}_m$, we set $L_\sigma := P_\sigma L P_\sigma^\top$.
linear order on $[m]$	Total order on $[m]$; equivalently, an ordered list of its elements.
$D = (V, A)$	Directed graph (digraph) with vertex set V and arc set A.
topological order	Linear order on V consistent with all arcs;
backward arc	Arc directed opposite to a chosen linear order.
FAS	Feedback arc set problem: remove a minimum number or weight of arcs to make a digraph acyclic.
Optimization on manifolds	
$\mathrm{SL}(n, \mathbb{R})$	Special linear group $\{X \in \mathbb{R}^{n\times n} : \det X = 1\}$.
$T_X\mathrm{SL}(n, \mathbb{R})$	Tangent space at X (as characterized in the thesis).
Proj_X	Frobenius-orthogonal projection $\mathbb{R}^{n\times n} \to T_X\mathrm{SL}(n, \mathbb{R})$.
$R_X(\cdot)$	Retraction at X.
$D\bar{f}(X)$	Differential of a smooth ambient extension $\bar{f}$ at X.
$\nabla\bar{f}(X)$	Ambient (Euclidean) gradient of a smooth extension $\bar{f}$.
$\operatorname{grad} f(X)$	Riemannian gradient (projection of the ambient gradient).
$\partial^C f(x)$	Clarke subdifferential (Euclidean or Riemannian, as specified).
$\mathcal{A}_X$	Active set at X (set of maximizers/active branches, as defined in the thesis).

Bibliography

[1] A. Abbondandolo, J. Kang, Symplectic homology of convex domains and Clarke's duality, Duke Math. J. 171 (2022), 739–830.

[2] A. Abbondandolo, P. Majer, A non-squeezing theorem for convex symplectic images of the Hilbert ball, Calc. Var. Partial Differential Equations 54 (2015), 1469–1506.

[3] P.-A. Absil, R. Mahony, R. Sepulchre, Optimization Algorithms on Matrix Manifolds, Princeton University Press, Princeton, NJ (2008).

[4] S. Arora, B. Barak, Computational complexity: A modern approach, Cambridge University Press, Cambridge (2009).

[5] S. Artstein-Avidan, R. Karasev, Y. Ostrover, From symplectic measurements to the Mahler conjecture, Duke Math. J. 163 (2014), 2003–2022.

[6] S. Artstein-Avidan, Y. Ostrover, Bounds for Minkowski billiard trajectories in convex bodies, Int. Math. Res. Not. IMRN (2014), no. 1, 165–193.

[7] S. Artstein-Avidan, V. Milman, Y. Ostrover, The M-ellipsoid, symplectic capacities and volume, Comment. Math. Helv. 83 (2008), 359–369.

[8] G. Ausiello, A. D'Atri, M. Protasi, Structure preserving reductions among convex optimization problems, J. Comput. Syst. Sci. 21 (1980), 136–153.

[9] A. Baharev, H. Schichl, A. Neumaier, T. Achterberg, An exact method for the minimum feedback arc set problem, J. Exp. Algorithmics 26 (2021), 1–28.

[10] L. Babai, Graph isomorphism in quasipolynomial time, Proc. 48th ACM Sympos. Theory Comput. (STOC) (2016), 684–697; full version J. ACM 63 (2016), no. 2, Article 8.

[11] A. Balitskiy, Equality cases in Viterbo's conjecture and isoperimetric billiard inequalities, Int. Math. Res. Not. IMRN (2020), 1957–1978.

[12] J. Bang-Jensen, G. Z. Gutin, Digraphs: Theory, Algorithms and Applications, 2nd ed., Springer, London (2009).

[13] N. Boumal, An Introduction to Optimization on Smooth Manifolds, Cambridge University Press, Cambridge (2023).

[14] N. Boumal, B. Mishra, P.-A. Absil, R. Sepulchre, Manopt, a Matlab toolbox for optimization on manifolds, J. Mach. Learn. Res. 15 (2014), no. 42, 1455–1459.

[15] F. J. Brandenburg, K. Hanauer, Sorting heuristics for the feedback arc set problem, Technical Report MIP-1104, University of Passau (2011).

[16] J. Chaidez, M. Hutchings, Computing Reeb dynamics on four-dimensional convex polytopes, J. Comput. Dyn. 8 (2021), 403–445.

[17] A. Cannas da Silva, Lectures on Symplectic Geometry, Lecture Notes in Mathematics 1764, Springer, Berlin–Heidelberg (2001).

[18] F. H. Clarke, A classical variational principle for periodic Hamiltonian trajectories, Proc. Amer. Math. Soc. 76 (1979), 186–188.

[19] F. H. Clarke, Optimization and Nonsmooth Analysis, John Wiley & Sons, New York (1983).

[20] F. H. Clarke, Optimization and Nonsmooth Analysis, Classics in Appl. Math. 5, SIAM, Philadelphia, PA (1990).

[21] M. Conforti, G. Cornuéjols, G. Zambelli, Integer Programming, Graduate Texts in Mathematics 271, Springer, Cham (2014).

[22] K. Conrad, Bilinear forms, Expository notes.

[23] S. A. Cook, The complexity of theorem-proving procedures, in Proc. 3rd Annual ACM Sympos. Theory Comput., ACM, New York (1971), 151–158.

[24] T. H. Cormen, C. E. Leiserson, R. L. Rivest, C. Stein, Introduction to Algorithms, 3rd ed., MIT Press, Cambridge, MA (2009).

[25] J.-P. Doignon, S. Fiorini, G. Joret, P. Maillon, Facets of the linear ordering polytope, Preprint (2007).

[26] F. M. Dopico, C. R. Johnson, Complementary bases in symplectic matrices and a proof that their determinant is one, Linear Algebra Appl. 419 (2006), 772–778.

[27] I. Ekeland, H. Hofer, Symplectic topology and Hamiltonian dynamics, Math. Z. 200 (1989), 355–378.

[28] K. Emich, R. Henrion, A simple formula for the second-order subdifferential of maximum functions, Vietnam J. Math. 42 (2014), 467–478.

[29] O. P. Ferreira, P. R. Oliveira, Subgradient algorithm on Riemannian manifolds, J. Optim. Theory Appl. 97 (1998), 93–104.

[30] S. Fiorini, Determining the automorphism group of the linear ordering polytope, Discrete Appl. Math. 112 (2001), 121–128.

[31] L. A. Florit, Riemannian Geometry: A class guide, Lecture notes, IMPA.

[32] D. Gabay, Minimizing a differentiable function over a differentiable manifold, J. Optim. Theory Appl. 37 (1982), 177–219.

[33] G. H. Golub, C. F. Van Loan, Matrix Computations, 4th ed., Johns Hopkins University Press, Baltimore, MD (2013).

[34] M. Grötschel, M. Jünger, G. Reinelt, A cutting plane algorithm for the linear ordering problem, Oper. Res. 32 (1984), 1195–1220.

[35] M. Gromov, Pseudo holomorphic curves in symplectic manifolds, Invent. Math. 82 (1985), 307–347.

[36] P. Grohs, S. Hosseini, ε-subgradient algorithms for locally Lipschitz functions on Riemannian manifolds, Adv. Comput. Math. 42 (2016), no. 2, 333–360.

[37] P. Grohs, S. Hosseini, Nonsmooth trust region algorithms for locally Lipschitz functions on Riemannian manifolds, IMA J. Numer. Anal. 36 (2016), no. 3, 1167–1192.

[38] B. Grünbaum, Convex Polytopes, 2nd ed., Graduate Texts in Mathematics 221, Springer, New York (2003).

[39] J. Guo, F. Hüffner, H. Moser, Feedback arc set in bipartite tournaments is NP-complete, Inform. Process. Lett. 102 (2007), 62–65.

[40] Gurobi Optimization, LLC, Gurobi Optimizer Reference Manual, (2026), https://www.gurobi.com.

[41] J. Gutt, M. Hutchings, Symplectic capacities from positive S^1-equivariant symplectic homology, Algebr. Geom. Topol. 18 (2018), 3537–3600.

[42] J. Gutt, M. Hutchings, V. G. B. Ramos, Examples around the strong Viterbo conjecture, J. Fixed Point Theory Appl. 24 (2022), no. 2, Paper No. 41, 22 pp.

[43] A. A. Hagberg, D. A. Schult, P. J. Swart, Exploring network structure, dynamics, and function using NetworkX, in G. Varoquaux, T. Vaught, J. Millman (eds.), Proc. 7th Python Sci. Conf. (SciPy 2008), Pasadena, CA (2008), 11–15.

[44] P. Haim-Kislev, On the symplectic size of convex polytopes, Geom. Funct. Anal. 29 (2019), 440–463.

[45] P. Haim-Kislev, The EHZ capacity of cubes and simplices, Master's thesis, School of Mathematical Sciences, Tel Aviv University (2019).

[46] P. Haim-Kislev, R. Hind, Y. Ostrover, On the existence of symplectic barriers, To appear in Selecta Math.; preprint, arXiv:2301.01822 (2023).

[47] P. Haim-Kislev, Y. Ostrover, Remarks on symplectic capacities of p-products, Internat. J. Math. 34 (2023), Paper No. 2350021, 21 pp.

[48] P. Haim-Kislev, Y. Ostrover, A counterexample to Viterbo's conjecture, To appear in Ann. of Math.; preprint, arXiv:2405.16513 (2024).

[49] M. Henk, J. Richter-Gebert, G. M. Ziegler, Basic properties of convex polytopes, in Handbook of Discrete and Computational Geometry, 3rd ed., CRC Press, Boca Raton, FL (2017), 383–413.

[50] HiGHS Developers, HiGHS – High Performance Optimization Software, https://highs.dev/, accessed Jan. 31, 2026.

[51] H. Hofer, E. Zehnder, A new capacity for symplectic manifolds, in Analysis, et cetera, Academic Press, Boston, MA (1990), 405–427.

[52] H. Hofer, E. Zehnder, Symplectic Invariants and Hamiltonian Dynamics, Modern Birkhäuser Classics, Springer, Basel (2011).

[53] R. A. Horn, C. R. Johnson, Matrix Analysis, 2nd ed., Cambridge University Press, Cambridge (2013).

[54] S. Hosseini, W. Huang, R. Yousefpour, Line search algorithms for locally Lipschitz functions on Riemannian manifolds, SIAM J. Optim. 28 (2018), 596–619.

[55] S. Hosseini, M. R. Pouryayevali, Generalized gradients and characterization of epi-Lipschitz sets in Riemannian manifolds, Nonlinear Anal. 74 (2011), 3884–3895.

[56] S. Hosseini, A. Uschmajew, A Riemannian gradient sampling algorithm for nonsmooth optimization on manifolds, SIAM J. Optim. 27 (2017), 173–189.

[57] M. Hutchings, Quantitative embedded contact homology, J. Differential Geom. 88 (2011), 231–266.

[58] IBM Documentation, Branch-and-cut in CPLEX, IBM ILOG CPLEX Optimization Studio Docs, https://www.ibm.com/docs/en/icos/22.1.0?topic=concepts-branch-cut-in-cplex, accessed Jan. 31, 2026.

[59] K. Irie, Symplectic homology of fiberwise convex sets and homology of loop spaces, J. Symplectic Geom. 20 (2022), 417–470.

[60] H. Iriyeh, M. Shibata, Symmetric Mahler's conjecture for the volume product in the 3-dimensional case, Duke Math. J. 169 (2020), 1077–1134.

[61] V. Kaibel, A. Schwartz, On the complexity of polytope isomorphism problems, Graphs Combin. 19 (2003), 215–230.

[62] R. Karasev, A. Sharipova, Viterbo's conjecture for certain Hamiltonians in classical mechanics, Arnold Math. J. 5 (2019), 483–500.

[63] P. D. Khanh, V. V. H. Khoa, B. S. Mordukhovich, V. T. Phat, Second-order subdifferential optimality conditions in nonsmooth optimization, SIAM J. Optim. 35 (2025), no. 2, 678–711.

[64] S. Krupp, Calculating the EHZ capacity of polytopes, Ph.D. thesis, Universität zu Köln (2020).

[65] S. Krupp, D. Rudolf, A regularity result for shortest generalized billiard trajectories in convex bodies in $\mathbb{R}^n$, Geom. Dedicata 216 (2022), Article 55.

[66] J. M. Lee, Introduction to Riemannian Manifolds, 2nd ed., Graduate Texts in Mathematics 176, Springer, Cham (2018).

[67] K. Leipold, F. Vallentin, Computing the EHZ capacity is NP-hard, Proc. Amer. Math. Soc. Ser. B 11 (2024), 603–611.

[68] M. M. Mäkelä, P. Neittaanmäki, Nonsmooth Optimization: Analysis and Algorithms with Applications to Optimal Control, World Scientific, River Edge, NJ (1992).

[69] D. McDuff, D. Salamon, Introduction to Symplectic Topology, 3rd ed., Oxford University Press, Oxford (2017).

[70] B. D. McKay, A. Piperno, Practical graph isomorphism, II, J. Symbolic Comput. 60 (2014), 94–112.

[71] J. W. Moon, Topics on Tournaments, Holt, Rinehart and Winston, New York (1968).

[72] B. S. Mordukhovich, R. T. Rockafellar, Second-order subdifferential calculus with applications to tilt stability in optimization, SIAM J. Optim. 22 (2012), 953–986.

[73] Y. Nir, On closed characteristics and billiards in convex bodies, Master's thesis, Tel Aviv University (2013).

[74] J. Nocedal, S. J. Wright, Numerical Optimization, 2nd ed., Springer, New York (2006).

[75] Y. Ostrover, When symplectic topology meets Banach space geometry, in Proc. Int. Congr. Math. 2014, Vol. 2, Kyoto (2014), 959–981.

[76] K. Perrot, T. Van-Pham, Feedback arc set problem and NP-hardness of minimum recurrent configuration problem of chip-firing game on directed graphs, Ann. Comb. 19 (2015), 373–396.

[77] K. B. Petersen, M. S. Pedersen, The Matrix Cookbook, Technical University of Denmark (2012), Version 20121115.

[78] A. Piperno, Isomorphism test for digraphs with weighted edges, in 18th International Symposium on Experimental Algorithms (SEA 2018), LIPIcs 103, Schloss Dagstuhl–Leibniz-Zentrum für Informatik (2018), Art. 30.

[79] V. Ramachandran, Finding a minimum feedback arc set in reducible flow graphs, Journal of Algorithms 9 (1988), 299–313.

[80] R. T. Rockafellar, Convex Analysis, Princeton University Press, Princeton, NJ (1970).

[81] A. Rosevear, S. Sottile, W. W. Y. Wong, Geodesic motion on $\mathsf{SL}(n)$ with the Hilbert–Schmidt metric, La Matematica 1 (2022), 343–386.

[82] D. Rudolf, Viterbo's conjecture for Lagrangian products in $\mathbb{R}^4$ and symplectomorphisms to the Euclidean ball, Preprint, arXiv:2203.02294 (2022).

[83] D. Rudolf, The Minkowski billiard characterization of the EHZ-capacity of convex Lagrangian products, J. Dyn. Differ. Equ. 36 (2024), 2773–2791.

[84] SCIP Optimization Suite, SCIP Optimization Suite, https://www.scipopt.org/, accessed Jan. 31, 2026.

[85] C. Hojny et al., The SCIP Optimization Suite 10.0, Optimization Online, technical report (2025), https://optimization-online.org/wp-content/uploads/2025/11/scipopt-100.pdf, accessed Jan. 31, 2026.

[86] A. Schrijver, Theory of Linear and Integer Programming, Wiley-Interscience Series in Discrete Mathematics and Optimization, John Wiley & Sons, Chichester (1986).

[87] A. Schrijver, Combinatorial Optimization, Springer, Berlin (2003).

[88] K. Siegel, Higher symplectic capacities, Preprint, arXiv:1902.01490 (2019).

[89] M. Sipser, Introduction to the Theory of Computation, 3rd ed., Cengage Learning, Boston, MA (2012).

[90] J. Torán, Reductions to graph isomorphism, Lecture notes (2008).

[91] A. M. Turing, On computable numbers, with an application to the Entscheidungsproblem, Proc. London Math. Soc. (2) 42 (1936), 230–265.

[92] C. Viterbo, Capacités symplectiques et applications, in Séminaire Bourbaki, Vol. 1988/89, Exp. 714, Astérisque 177–178 (1989), 345–362.

[93] C. Viterbo, Functors and computations in Floer homology with applications. I, Geom. Funct. Anal. 9 (1999), 985–1033.

[94] C. Viterbo, Metric and isoperimetric problems in symplectic geometry, J. Amer. Math. Soc. 13 (2000), 411–431.

[95] A. Weinstein, Symplectic geometry, Bull. Amer. Math. Soc. (N.S.) 5 (1981), 1–13.

[96] L. A. Wolsey, Integer Programming, 2nd ed., John Wiley & Sons, Hoboken, NJ (2020).

[97] L.-H. Zhang, W.-H. Yang, R. Song, Optimality conditions for the nonlinear programming problems on Riemannian manifolds, Pacific J. Optim. 10 (2014), 415–434.

[98] G. M. Ziegler, Lectures on Polytopes, Graduate Texts in Mathematics 152, Springer-Verlag, New York (1995).

Research Resources and Publications

Publication List

Published

- K. Leipold, F. Vallentin, *Computing the EHZ capacity is* NP-*hard*, Proc. Amer. Math. Soc. Ser. B 11 (2024), 603–611.

In preparation

- K. Leipold, F. Vallentin, *Affine symplectomorphism of simplices is graph-isomorphism complete*.

Used Tools and Resources

This dissertation made use of computational resources and digital tools. Numerical experiments and optimization computations were carried out using the DFG-funded HPC system RAMSES (Research Accelerator for Modeling and Simulation with Enhanced Security) provided by the ITCC (IT Center University of Cologne). We furthermore thank the ITCC (IT Center University of Cologne) for providing compute resources on the DFG-funded HPC (High Performance Computing) system RAMSES (Research Accelerator for Modeling and Simulation with Enhanced Security) as well as support (DFG funding number: INST 216/512-1 FUGG).

Digital tools, including ChatGPT, were used in a supportive manner for language editing of parts of the text. All mathematical content, proofs, computations, interpretations, and conclusions were developed, checked, and are the sole responsibility of the author.