RELATIVE ENTROPY METHODS IN CONSTRAINED POLYNOMIAL AND SIGNOMIAL OPTIMIZATION

THORSTEN THEOBALD

Abstract. Relative entropy programs belong to the class of convex optimization problems. Within techniques based on the arithmetic-geometric mean inequality, they facilitate to compute nonnegativity certificates of polynomials and of signomials.
While the initial focus was mostly on unconstrained certificates and unconstrained optimization, recently, Murray, Chandrasekaran and Wierman developed conditional techniques, which provide a natural extension to the case of convex constrained sets. This expository article gives an introduction into these concepts and explains the geometry of the resulting conditional SAGE cone. To this end, we deal with the sublinear circuits of a finite point set in $\mathbb{R}^n$, which generalize the simplicial circuits of the affine-linear matroid induced by a finite point set to a constrained setting.

1. Introduction

Relative entropy programs provide a class of convex optimization problems [5]. They are concerned with optimizing linear functions over affine sections of the relative entropy cone

$$K_{\mathrm{rel}}^n \;=\; \mathrm{cl}\left\{(x,y,\tau) \in \mathbb{R}^n_{>0} \times \mathbb{R}^n_{>0} \times \mathbb{R}^n \;:\; x_i \ln \frac{x_i}{y_i} \le \tau_i \;\text{ for all } i\right\},$$

where cl denotes the topological closure. Relative entropy programming contains as a subclass geometric programing and the special case $n = 1$ of the relative entropy cone can be viewed as a reparametrization of the exponential cone. Beside applications in fields such as engineering and information theory, the last years have shown an exciting and powerful application of relative entropy programming for optimization of polynomials and of signomials (i.e., exponential sums). Namely, within techniques based on the *arithmetic-geometric mean inequality* (AM/GM inequality), relative entropy programs facilitate to compute nonnegativity certificates of polynomials and of signomials. These techniques can also be combined with other nonnegativity certificates, such as sums of squares.

A *signomial*, also known as *exponential sum* or *exponential polynomial*, is a sum of the form

$$f(x) \;=\; \sum_{\alpha \in \mathcal{T}} c_\alpha \exp(\langle \alpha, x \rangle)$$

with real coefficients c_α and a finite ground support set $\mathcal{T} \subset \mathbb{R}^n$. Here, $\langle \cdot, \cdot \rangle$ is the usual scalar product and we point out explicitly that the definition of a signomial also allows for negative or non-integer entries in the elements of $\mathcal{T}$. Exponential sums can be seen as a generalization of polynomials: when $\mathcal{T} \subset \mathbb{N}^n$, the transformation $x_i = \ln y_i$ gives polynomial functions $y \mapsto \sum_{\alpha \in \mathcal{T}} c_\alpha y^\alpha$ on $\mathbb{R}^n_{>0}$. For example, we have $f = 5\exp(2x_1 + 3x_2) - 3\exp(4x_2 + x_3)$

versus $p = 5y_1^2y_2^3 - 3y_2^4y_3$. When $\mathcal{T} \subset \mathbb{N}^n$, a signomial f is nonnegative on $\mathbb{R}^n$ if and only if its associated polynomial p is nonnegative on $\mathbb{R}_+^n$, where $\mathbb{R}_+$ denotes the set of nonnegative real numbers.

Signomial optimization has additional modeling power compared to polynomial optimization. For example, the non-integer exponent $\frac{1}{2}$ is possible in signomial optimization, which corresponds to square roots. This leads to additional applications, for example, in chemical reaction networks [22], aircraft design optimization [38] or epidemiological models [31].

The following idea connects global nonnegativity certificates for polynomials and for signomials to the AM/GM inequality. This basic insight goes back to Reznick [34] and was further developed by Pantea, Koeppl, Craciun [32], Iliman and de Wolff [14] as well as Chandrasekaran and Shah [4]. For support points $\alpha_0, \ldots, \alpha_m \in \mathbb{R}^n$ and $\lambda = (\lambda_1, \ldots, \lambda_m) \in \mathbb{R}_+^m$ with $\sum_{i=1}^m \lambda_i = 1$ and $\sum_{i=1}^m \lambda_i \alpha_i = \alpha_0$, the signomial

$$\sum_{i=1}^m \lambda_i \exp(\langle \alpha_i, x \rangle) - \exp(\langle \alpha_0, x \rangle)$$

is nonnegative on $\mathbb{R}^n$. This is a consequence of the weighted AM/GM inequality, see Section 2.3.

During the early developments of AM/GM-based optimization of polynomials and signomials, most work concentrated on unconstrained certificates and unconstrained optimization. In the recent work [25], Murray, Chandrasekaran and Wierman presented an extension of the relative entropy methods to a conditional setting with a convex constrained set. In this situation, the constrained approach provides a much more suitable framework than earlier initial approaches of addressing constraints in AM/GM-based optimization based on a mix with a Krivine-type Positivstellensatz [4, 9].

The goal of this expository article is to offer an introduction into the relative entropy methods for unconstrained and for constrained polynomial and signomial optimization. The resulting cones are called the *SAGE cone* and the *conditional SAGE cone*, where SAGE is the acronym for *Sums of Arithmetic-Geometric Exponentials*. The geometry of the SAGE cone is governed by the simplicial circuits of the affine-linear matroid induced by the support $\mathcal{T}$ [12, 24, 36]. In order to exhibit the geometry of the conditional SAGE cone, we spell out and study the sublinear circuits of a finite point set in $\mathbb{R}^n$, which generalize the simplicial circuits to a constrained setting [26].

Sublinear circuits of polyhedral sets have specifically been studied in [28]. Meanwhile, the conditional SAGE approaches has also been extended towards hierarchies and Positivstellensätze for conditional SAGE [35] and to additional non-convex constraints [10].

2. From relative entropy programming to the SAGE cone

We provide some background on relative entropy programming and explain the main concepts of the SAGE cone.

2.1. **Cones and optimization.** Conic optimization is concerned with optimization problems of the form $\inf\{c^T x : Ax = b, x \in K\}$ for a convex cone $K \subset \mathbb{R}^n$, where $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$ and $c \in \mathbb{R}^n$. Usually, we assume that the cone K is closed, full-dimensional and pointed, where

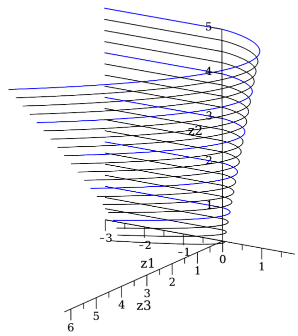

FIGURE 1. The exponential cone.

pointed means that K contains no lines. A closed, full-dimensional and pointed cone is called a *proper* cone. If a self-concordant barrier function for K is known, then conic optimization problems over K can be approached efficiently through interior point methods (see, e.g., [30]).

Prominent special cases of conic programming are linear programming and semidefinite programming. Linear programming (LP) can be viewed as conic programming over the nonnegative cone $K = \mathbb{R}^n_+$. Linear programming arises in polynomial optimization, for example, in LP-relaxations via Handelman's Theorem [13] or in the DSOS approach (diagonally dominant sums of squares [1]).

Semidefinite programming (SDP) can be viewed as conic programming over the cone $\mathcal{S}_n^+$ of positive semidefinite symmetric $n \times n$-matrices. Semidefinite programming is ubiquitous in polynomial optimization through sums of squares. In particular, with respect to constrained optimization, semidefinite programming is tightly connected to Lasserre's hierarchical relaxation and thus to Putinar's Positivstellensatz and to moments (see, e.g., [17]). Recent developments on the use of semidefinite programming in polynomial optimization include the improved exploitation of sparsity, see [19, 37].

2.2. **The exponential cone and the relative entropy cone.** The exponential cone and the relative entropy cone are rather young cones within the development of convex optimization. The *exponential cone* is defined as the three-dimensional cone

$$K_{\exp} \;=\; \mathrm{cl}\left\{ z \in \mathbb{R} \times \mathbb{R}_+ \times \mathbb{R}_{>0} \;:\; \exp\left(\frac{z_1}{z_3}\right) \le \frac{z_2}{z_3} \right\}.$$

In 2006, Nesterov gave a self-concordant barrier function [29], see also [6]. This enables efficient interior-point algorithms for approximating optimization problems over $K_{\exp}$. The exponential cone is depicted in Figure 1.

Remark 1. For any convex function $\varphi : \mathbb{R} \to \mathbb{R}$, the *perspective function* $\tilde{\varphi} : \mathbb{R} \times \mathbb{R}_{>0} \to \mathbb{R}$ is defined as $\tilde{\varphi}(x, y) \;=\; y\varphi\left(\frac{x}{y}\right)$. The closure of the epigraph of the perspective function,

$$\mathrm{cl}\left\{ (t, x, y) \in \mathbb{R} \times \mathbb{R} \times \mathbb{R}_{>0} \;:\; t \ge y\varphi\left(\frac{x}{y}\right) \right\},$$

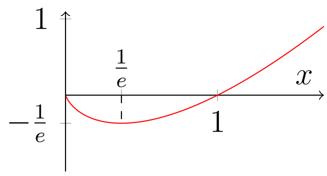

FIGURE 2. The negative entropy function.

is known to be a closed convex cone. The exponential cone is exactly the cone which arises from this construction for the exponential function $\varphi : x \mapsto \exp(x)$.

The exponential cone is a non-symmetric cone, where 'symmetric' means that a cone is homogeneous and self-dual. Optimization over a Cartesian product of exponential cones contains as a special case the class of geometric programming, which has many applications in engineering [3]. Moreover, the exponential cone has applications, for example, in maximum likelihood estimation or logistic regression [20]. The dual of the exponential cone is

$$K_{\exp}^* \ = \ \mathrm{cl} \left\{ s \in \mathbb{R}_{<0} \times \mathbb{R}_+ \times \mathbb{R} \ : \ \exp\left(\frac{s_3}{s_1}\right) \le -\frac{e \cdot s_2}{s_1} \right\},$$

where e denotes Euler's number. Meanwhile, the exponential cone is implemented in software tools, such as `CVXPY` [7], `ECOS` [8] and `MOSEK` [20].

The relative entropy cone can be seen as a reparametrized version of the exponential cone. For a formal definition, we consider the *negative entropy function*

$$f : \mathbb{R}_{>0} \to \mathbb{R}, \ f(x) = x \ln x,$$

see Figure 2.

The *relative entropy function* is defined as $D : \ \mathbb{R}_{>0} \times \mathbb{R}_{>0} \to \mathbb{R}$, $D(x,y) = x \ln \frac{x}{y}$. It is a convex function in $z = (x,y)$, also called a *jointly convex* function in x and y. To see the joint convexity, observe that the Hessian of D evaluates to

$$\nabla^2 D(x,y) \ = \ \begin{pmatrix} \frac{1}{x} & -\frac{1}{y} \\ -\frac{1}{y} & \frac{x}{y^2} \end{pmatrix}.$$

The relative entropy function can be extended to vectors $x, y \in \mathbb{R}^n_{>0}$ by setting $D(x,y) = \sum_{i=1}^n x_i \ln \frac{x_i}{y_i}$. The *relative entropy cone* is defined as

$$K_{\mathrm{rel}}^1 \ := \ \mathrm{cl} \left\{ (x, y, \tau) \in \mathbb{R}_{>0} \times \mathbb{R}_{>0} \times \mathbb{R} \ : \ D(x,y) \le \tau \right\}.$$

Here, the upper index "1" indicates that x, y and τ are scalars. The relative entropy cone can be viewed as a reparametrization of the exponential cone, because of the equivalences

$$x \ln \frac{x}{y} \le \tau \quad \Longleftrightarrow \quad \exp\left(-\frac{\tau}{x}\right) \le \frac{y}{x} \quad \Longleftrightarrow \quad (-\tau, y, x) \in K_{\exp}.$$

More generally, the relative entropy cone can be extended to triples of vectors by defining $K^n_{\text{rel}} \subset \mathbb{R}^{3n}$ as

$$K^n_{\text{rel}} \ = \ \text{cl}\left\{(x,y,\tau) \in \mathbb{R}^n_{>0} \times \mathbb{R}^n_{>0} \times \mathbb{R}^n \ : \ x_i \ln \frac{x_i}{y_i} \le \tau_i \ \text{ for all } i\right\}.$$

This allows to model the n-variate relative entropy condition

$$D((x_1,\ldots,x_n),(y_1,\ldots,y_n)) \ := \ \sum_{i=1}^n x_i \ln \frac{x_i}{y_i} \le t$$

as

$$\exists \tau \in \mathbb{R}^n \ \text{ with } (x,y,\tau) \in K^n_{\text{rel}} \ \text{ and } \ \mathbf{1}^T\tau = t,$$

where $\mathbf{1}$ denotes the all-ones vector.

2.3. **The basic AM/GM idea.** The following idea, going back to Reznick [34] and further developed in [4, 14, 32], connects global nonnegativity certificates for polynomials and for signomials to the AM/GM inequality. Consider support points $\alpha_0,\ldots,\alpha_m \in \mathbb{R}^n$ such that α_0 is a convex combination of $\alpha_1,\ldots,\alpha_m$, that is, $\alpha_0 = \sum_{i=1}^m \lambda_i\alpha_i = \alpha_0$ with $\sum_{i=1}^m \lambda_i = 1$ and $\lambda \in \mathbb{R}^m_+$. Then the signomial

$$\sum_{i=1}^m \lambda_i \exp(\langle \alpha_i, x\rangle) - \exp(\langle \alpha_0, x\rangle) \tag{2.1}$$

is nonnegative on $\mathbb{R}^n$. Namely, we can use the following weighted AM/GM inequality, which can easily be derived from the strict convexity of the univariate function $x \mapsto -\ln x$ on the domain $(0,\infty)$.

Theorem 2 (Weighted arithmetic-geometric mean inequality)**.** *For each $z \in \mathbb{R}^n_+$ and $\lambda \in \mathbb{R}^n_+$ with $\sum_{i=1}^n \lambda_i = 1$, we have*

$$\sum_{i=1}^n \lambda_i z_i \ \ge \ \prod_{i=1}^n z_i^{\lambda_i}.$$

The nonnegativity of (2.1) follows from the weighted AM/GM inequality through

$$\sum_{i=1}^m \lambda_i \exp(\langle \alpha_i, x\rangle) \ \ge \ \prod_{i=1}^m (\exp(\langle \alpha_i, x\rangle))^{\lambda_i} \ = \ \exp(\langle \alpha_0, x\rangle) \tag{2.2}$$

for all $x \in \mathbb{R}^n$. Clearly, sums of such exponential sums of the form (2.1) are nonnegative as well. We will see later how to generalize this core idea from the unconstrained setting to the constrained case with respect to a set X.

For the class of signomials, we assume that an underlying finite ground support set $\mathcal{T} \subset \mathbb{R}^n$ is given. When considering subsets of the ground support, we usually employ the convention that $\mathcal{A}$ refers to terms with positive coefficients and $\mathcal{B}$ (or β in case of single elements) refers to terms with possibly negative coefficients.

Let f be a general signomial whose coefficients except at most one are positive,

$$f(x) = \sum_{\alpha\in\mathcal{A}} c_\alpha \exp(\langle \alpha, x\rangle) + d\exp(\langle \beta, x\rangle) \text{ with } c_\alpha > 0 \text{ and } d \in \mathbb{R}. \tag{2.3}$$

Chandrasekaran and Shah [4] have given the following exact characterization of the nonnegativity of f in terms of the coefficients $(c_\alpha)_{\alpha\in\mathcal{A}}$ and d. This insight establishes a fundamental connection from the nonnegativity of signomials (or polynomials) with at most one negative coefficient to the relative entropy function.

Theorem 3. *The signomial f in (2.3) is nonnegative if and only if there exists $\nu \in \mathbb{R}_+^{\mathcal{A}}$ with $\sum_{\alpha\in\mathcal{A}} \nu_\alpha \alpha = (\sum_{\alpha\in\mathcal{A}} \nu_\alpha)\beta$ and $D(\nu, ec) \le d$, where D denotes the relative entropy function and e is Euler's number.*

As revealed by the proof, a natural way to see why the relative entropy function occurs is through duality theory. For a function $g : \mathbb{R}^n \to \mathbb{R}$, the conjugate $g^* : \mathbb{R}^n \to \mathbb{R}$ is defined by

$$g^*(y) \ = \ \sup_x \left(y^T x - g(x) \right).$$

The conjugate function of $g(x) = e^x$ is $g^*(y) = y \ln y - y$ on the domain $\mathbb{R}_+$, where we use the convention $0 \cdot \ln 0 := 0$. Using standard computation rules from convex optimization, the conjugate function of

$$g(x) \ = \ \sum_{i=1}^n c_i e^{x_i} \quad \text{with } c_1, \ldots, c_n > 0$$

is $D(y, ec)$, where $c := (c_1, \ldots, c_n)^T$.

Proof. The nonnegativity of the signomial f is equivalent to the nonnegativity of $f(x)\exp(\langle -\beta, x\rangle)$ and thus also equivalent to $\sum_{\alpha\in\mathcal{A}} c_\alpha \exp(\langle \alpha, x\rangle - \langle \beta, x\rangle) \ge -d$ for all $x \in \mathbb{R}^n$. The function $\sum_{\alpha\in\mathcal{A}} c_\alpha \exp(\langle \alpha, x\rangle - \langle \beta, x\rangle)$ is, as a sum of convex functions, convex as well. Its infimum can be formulated as the convex optimization problem

$$\inf_{x\in\mathbb{R}^n,\ t\in\mathbb{R}^{\mathcal{A}}} \sum_{\alpha\in\mathcal{A}} c_\alpha t_\alpha \ \text{ s.t. } \ \exp(\langle \alpha - \beta, x\rangle) \le t_\alpha \quad \forall \alpha \in \mathcal{A},$$

where the inequality constraints can also be written as $\langle \alpha - \beta, x\rangle \le \ln t_\alpha$. The primal problem satisfies Slater's condition from convex optimization, because a Slater point can be constructed by considering an arbitrary point $x \in \mathbb{R}^n$ and choosing all t_α sufficiently large. For the Lagrange dual, we obtain, by reducing to the conjugate function together with the computation rules above,

$$\sup_{\nu\in\mathbb{R}_+^{\mathcal{A}}} -D(\nu, ec) \ \text{ s.t. } \ \sum_{\alpha\in\mathcal{A}} \nu_\alpha \alpha \ = \ (\sum_{\alpha\in\mathcal{A}} \nu_\alpha)\beta.$$

Due to Slater's condition, we have strong duality and the dual optimum is attained. This shows the theorem. □

Example 4. We consider the Motzkin-type signomial $f_\delta = e^{4x+2y} + e^{2x+4y} + 1 + \delta e^{2x+2y}$ with some parameter $\delta \in \mathbb{R}$. In order to determine the smallest δ such that f_δ is nonnegative, we can consider the signomial

$$g_\delta(x) := f_\delta(x)\, e^{-2x-2y} \ = \ e^{2x} + e^{2y} + e^{-2x-2y} + \delta.$$

Since g_δ has at most one negative term, we can formulate the nonnegativity of g_δ in terms of the relative entropy condition

$$\begin{array}{rcl} \inf \delta & & \\ \nu_1 \begin{pmatrix} 2 \\ 0 \end{pmatrix} + \nu_2 \begin{pmatrix} 0 \\ 2 \end{pmatrix} + \nu_3 \begin{pmatrix} -2 \\ -2 \end{pmatrix} & = & 0, \\ \nu_1 \ln \frac{\nu_1}{e \cdot 1} + \nu_2 \ln \frac{\nu_2}{e \cdot 1} + \nu_3 \ln \frac{\nu_3}{e \cdot 1} & \leq & \delta, \\ \nu & \in & \mathbb{R}^3_+, \ \delta \in \mathbb{R}. \end{array}$$

The minimal δ satisfying this condition is $\delta = -3$ and $\min_{x,y} f_{\delta=-3} = 0$.

In the algorithmic access via conic optimization as described above, it is essential that the vector ν in Theorem 3 is not normalized to, say, $\sum_{\alpha \in \mathcal{A}} \nu_\alpha = 1$. Indeed, normalizing the vector ν in that theorem gives a formulation of a nonnegativity condition which can be viewed as a slight generalization of the AM/GM consideration (2.2).

Proposition 5. *The signomial f in (2.3) is nonnegative if and only if there exists $\lambda \in \mathbb{R}^{\mathcal{A}}_+$ with $\sum\limits_{\alpha \in \mathcal{A}} \lambda_\alpha \alpha = \beta$, $\sum_{\alpha \in \mathcal{A}} \lambda_\alpha = 1$ and*

$$\prod_{\alpha \in \mathcal{A} \text{ with } \lambda_\alpha > 0} \left(\frac{c_\alpha}{\lambda_\alpha} \right)^{\lambda_\alpha} \geq -d\,. \tag{2.4}$$

Proof. If there exists $\lambda \in \mathbb{R}^{\mathcal{A}}_+$ satisfying (2.4), using the weighted AM/GM-inequality with weights $(\lambda_\alpha)_{\alpha \in \mathcal{A}}$ gives

$$\begin{array}{rcl} \sum\limits_{\alpha \in \mathcal{A}} c_\alpha \exp(\langle \alpha, x \rangle) & \geq & \prod\limits_{\lambda_\alpha > 0} \left(\frac{1}{\lambda_\alpha} c_\alpha \exp(\langle \alpha, x \rangle \right)^{\lambda_\alpha} \ = \ \prod\limits_{\lambda_\alpha > 0} \left(\frac{c_\alpha}{\lambda_\alpha} \right)^{\lambda_\alpha} \exp(\langle \beta, x \rangle) \\ & \geq & -d \exp(\langle \beta, x \rangle) \end{array}$$

for all $x \in \mathbb{R}$. Hence, f is nonnegative.

Conversely, if f is nonnegative, then, by Theorem 3, there exists $\nu \in \mathbb{R}^{\mathcal{A}}_+$ with $\sum_{\alpha \in \mathcal{A}} \nu_\alpha \alpha = (\sum_{\alpha \in \mathcal{A}} \nu_\alpha) \beta$ and $D(\nu, ec) \leq d$. Set $\lambda_\alpha = \frac{\nu_\alpha}{\mathbf{1}^T \nu}$ for all $\alpha \in \mathcal{A}$ and we can assume that $c_\alpha > 0$ for all $\alpha \in \mathcal{A}$. The convex univariate function $h : \mathbb{R}_{>0} \to \mathbb{R}$, $s \mapsto D(s\lambda, ec)$ has the derivative $h'(s) = \ln s + D(\lambda, c)$ and thus takes its minimum at $s^* = e^{-D(\lambda,c)}$ with minimal value

$$\begin{array}{rcl} h(s^*) & = & e^{-D(\lambda,c)} \sum\limits_{\alpha \in \mathcal{A}} \lambda_\alpha \ln \frac{e^{-D(\lambda,c)} \lambda_\alpha}{e \cdot c_\alpha} \ = \ e^{-D(\lambda,c)} (D(\lambda, c) - D(\lambda, c) - 1) \\ & = & -e^{-D(\lambda,c)} \ = \ - \prod\limits_{\alpha \in \mathcal{A} \text{ with } \lambda_\alpha > 0} \left(\frac{c_\alpha}{\lambda_\alpha} \right)^{\lambda_\alpha} . \end{array}$$

Hence,

$$d \ \geq \ D(\nu, ec) \ \geq \ h(s^*) \ = \ - \prod_{\alpha \in \mathcal{A} \text{ with } \lambda_\alpha > 0} \left(\frac{c_\alpha}{\lambda_\alpha} \right)^{\lambda_\alpha} .$$

□

2.4. **The SAGE cone (Sums of Arithmetic-Geometric Exponentials).** Building upon the AM/GM idea, Chandrasekaran and Shah [4] have introduced the following cones of nonnegative signomials. The elements of these cones admit a nonnegativity certificate based on the AM/GM inequality. For given $\mathcal{A}$ and $\beta \notin \mathcal{A}$, the *AGE cone* $C_{\mathrm{AGE}}(\mathcal{A}, \beta)$ is defined as

$$C_{\mathrm{AGE}}(\mathcal{A}, \beta) = \left\{ f \,:\, f = \sum_{\alpha \in \mathcal{A}} c_\alpha e^{\langle \alpha, x \rangle} + d e^{\langle \beta, x \rangle} \text{ is nonnegative, } c \in \mathbb{R}_+^{\mathcal{A}} \right\}.$$

Given a finite set $\mathcal{T} \subset \mathbb{R}^n$, the *SAGE cone* $C(\mathcal{T})$ is then defined as

$$C(\mathcal{T}) := \sum_{\beta \in \mathcal{T}} C_{\mathrm{AGE}}(\mathcal{T} \setminus \{\beta\}, \beta).$$

It consists of signomials which admit a decomposition as a sum of AGE signomials. The SAGE cone supports handling sparse signomials. A crucial property is that it allows cancellation-free representations. This was shown by Wang [36] in the polynomial setting and by Murray, Chandrasekaran and Wierman [24] in the signomial setting.

Theorem 6. *Let f be a signomial with support $\mathcal{T}$. If $f \in C(\mathcal{T}')$ for some $\mathcal{T}' \supseteq \mathcal{T}$, then $f \in C(\mathcal{T})$. In words, if a signomial f supported on $\mathcal{T}$ has a SAGE certificate with respect to some larger support set $\mathcal{T}'$, then the SAGE certificate also exists on the support set $\mathcal{T}$ itself.*

Membership of a signomial to the SAGE cone can be formulated in terms of a relative entropy program. For disjoint $\emptyset \neq \mathcal{A} \subset \mathbb{R}^n$ and $\mathcal{B} \subset \mathbb{R}^n$, write

$$C(\mathcal{A}, \mathcal{B}) \ := \ \sum_{\beta \in \mathcal{B}} C_{\mathrm{AGE}}(\mathcal{A} \cup \mathcal{B} \setminus \{\beta\}, \beta).$$

Hence, signomials in $C(\mathcal{A}, \mathcal{B})$ can only have negative coefficients within the subset $\mathcal{B}$. It holds $C(\mathcal{A}, \mathcal{B}) \ = \ \{f = \sum_{\alpha \in \mathcal{A}} c_\alpha e^{\langle \alpha, x \rangle} + \sum_{\beta \in \mathcal{B}} c_\beta e^{\langle \beta, x \rangle} \in C(\mathcal{A} \cup \mathcal{B}) \,:\, c_\alpha \geq 0 \text{ for } \alpha \in \mathcal{A}\}$. This allows the following relative entropy formulation to decide whether a given signomial

$$f \ = \ \sum_{\alpha \in \mathcal{A}} c_\alpha \exp(\langle \alpha, x \rangle) + \sum_{\beta \in \mathcal{B}} c_\beta \exp(\langle \beta, x \rangle) \tag{2.5}$$

with $c_\alpha \geq 0$ for $\alpha \in \mathcal{A}$ and $c_\beta < 0$ for $\beta \in \mathcal{B}$ is contained in the SAGE cone.

Theorem 7. [24] *The signomial f in (2.5) is contained in $C(\mathcal{A}, \mathcal{B})$ if and only for every $\beta \in \mathcal{B}$ there exist $c^{(\beta)} \in \mathbb{R}_+^{\mathcal{A}}$ and $\nu^{(\beta)} \in \mathbb{R}_+^{\mathcal{A}}$ such that*

$$\begin{array}{rcll} \sum\limits_{\alpha \in \mathcal{A}} \nu_\alpha^{(\beta)} \alpha & = & \big(\sum\limits_{\alpha \in \mathcal{A}} \nu_\alpha^{(\beta)}\big) \beta & \text{for } \beta \in \mathcal{B}, \\ D(\nu^{(\beta)}, e \cdot c^{(\beta)}) & \leq & c_\beta & \text{for } \beta \in \mathcal{B}, \\ \sum\limits_{\beta \in \mathcal{B}} c_\alpha^{(\beta)} & \leq & c_\alpha & \text{for } \alpha \in \mathcal{A}. \end{array}$$

For combining AM/GM-techniques for polynomials on $\mathbb{R}^n$ and $\mathbb{R}_+^n$, Katthän, Naumann and the current author have developed the $\mathcal{S}$-cone [16]. AM/GM techniques can also be combined with sums of squares to hybrid methods, see [15]. Moreover, an implementation of the SAGE cone, called `Sageopt` was provided by Murray [23].

3. Conditional nonnegativity over convex sets

Murray, Chandrasekaran and Wierman [25] generalized AM/GM optimization from the unconstrained setting to the constrained setting over a convex set $X \subset \mathbb{R}^n$. Denote by $\sigma_X(y) = \sup\{y^T x : x \in X\}$ the *support function* of X from classical convex geometry. σ_X is a convex function $\mathbb{R}^n \to \mathbb{R}_+ \cup \{\infty\}$. If X is polyhedral, then σ_X is linear on every normal cone of X. The support function σ_X arises naturally in optimization as the conjugate function of the indicator function

$$\mathbb{1}_X(x) \;=\; \begin{cases} 0 & x \in X, \\ \infty & \text{otherwise} \end{cases}$$

of a convex set X.

We begin with the crucial insight that for a signomial with at most one negative term, the nonnegativity on X ("*conditional nonnegativity*") can be formulated in terms of a relative entropy program involving also the support function of X. Let $\mathcal{T} := \mathcal{A} \cup \{\beta\}$ and $f(x) = \sum_{\alpha \in \mathcal{T}} c_\alpha \exp(\langle \alpha, x \rangle)$ with $c_\alpha \geq 0$ for $\alpha \in \mathcal{A}$. As short notation, for a given $\nu \in \mathbb{R}^{\mathcal{T}}$ we write $\mathcal{T}\nu := \sum_{\alpha \in \mathcal{T}} \alpha \nu_\alpha$.

Theorem 8. [25] *The signomial f in (2.3) is nonnegative* on X *if and only if there exists $\nu \in \mathbb{R}^{\mathcal{T}} \setminus \{0\}$ with $\sum_{\alpha \in \mathcal{T}} \nu_\alpha = 0$, $\nu_{|\mathcal{A}} \geq 0$ and $\sigma_X(-\mathcal{T}\nu) + D(\nu_{|\mathcal{A}}, ec_{|\mathcal{A}}) \leq c_\beta$. Here, $\nu_{|\mathcal{A}}$ denotes to the restriction of the vector ν to the coordinates of $\mathcal{A}$.*

Proof. Generalizing the idea of the proof of Theorem 3, we now observe that f is nonnegative on X if and only if $\sum_{\alpha \in \mathcal{A}} c_\alpha \exp(\langle \alpha, x \rangle - \langle \beta, x \rangle) \geq -c_\beta$ for all $x \in X$. The infimum of the left convex function can be formulated as the convex program

$$\inf_{x \in X,\ t \in \mathbb{R}^{\mathcal{A}}} \sum_{\alpha \in \mathcal{A}} c_\alpha t_\alpha \quad \text{s.t.} \quad \exp(\langle \alpha - \beta, x \rangle) \leq t_\alpha \quad \forall \alpha \in \mathcal{A}.$$

Strong duality holds, and the dual is

$$\max_{\nu \in \mathbb{R}^{\mathcal{T}} \setminus \{0\}} -(\sigma_X(-\mathcal{T}\nu) + D(\nu_{|\mathcal{A}}, ec_{|\mathcal{A}})) \quad \text{s.t.} \quad \sum_{\alpha \in \mathcal{T}} \nu_\alpha = 0,\ \nu_{|\mathcal{A}} \geq 0.$$

Hence, f is nonnegative if and only if this maximum is larger than or equal to $-c_\beta$. □

It is useful to consider also the following alternative formulation of the characterization in Theorem 8. For $\beta \in \mathcal{T}$, set

$$N_\beta \;=\; \left\{ \nu \in \mathbb{R}^{\mathcal{T}} \,:\, \nu_{\setminus\beta} \geq \mathbf{0},\ \sum_{\alpha \in \mathcal{T}} \nu_\alpha = 0 \right\},$$

where $\nu_{\setminus\beta}$ refers to the vector ν in which the component β has been removed. The choice of the name N reflects that the coordinate β is the only coordinate which may be negative.

Corollary 9. *The signomial f in (2.3) is nonnegative on X if and only if there exists $\nu \in N_\beta \setminus \{\mathbf{0}\}$ with $\sigma_X(-\mathcal{T}\nu) + D(\nu_{|\mathcal{A}}, ec_{|\mathcal{A}}) \leq c_\beta$.*

The following theorem characterizes nonnegativity in terms of a normalized dual variable and thus generalizes Proposition 5.

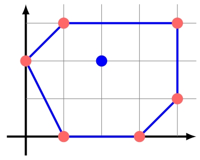

FIGURE 3. In the case $X = \mathbb{R}^n$, the exponent vectors of terms with negative coefficients lie in the convex hull of the exponent vectors of terms with positive coefficients.

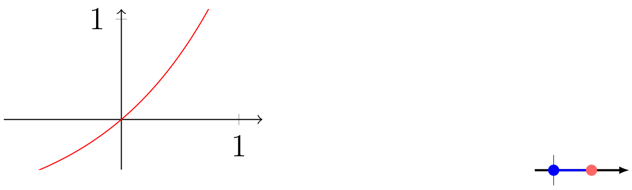

FIGURE 4. Left: The graph of the function $f(x) = \exp(x) - \exp(0)$. Right: The support point of the negative coefficient (visualized in blue) is not contained in the convex hull of the set of support points with positive coefficient (visualized in red).

Theorem 10. [26] *The signomial f in (2.3) is nonnegative* on X *if and only if there exists $\lambda \in N_\beta$ with $\lambda_\beta = -1$ and*

$$\prod_{\alpha \in \mathcal{A} \text{ with } \lambda_\alpha > 0} \left(\frac{c_\alpha}{\lambda_\alpha} \right)^{\lambda_\alpha} \geq -c_\beta \exp(\sigma_X(-\mathcal{T}\lambda)).$$

For given $\mathcal{A}$ and $\beta \notin \mathcal{A}$, define the *conditional AGE cone* $C_{X,\mathrm{AGE}}(\mathcal{A}, \beta)$ as

$$\left\{ f : f = \sum_{\alpha \in \mathcal{A}} c_\alpha e^{\langle \alpha, x \rangle} + d e^{\langle \beta, x \rangle} \text{ nonnegative } on\ X,\ c \in \mathbb{R}_+^{\mathcal{A}} \right\}. \tag{3.1}$$

The *conditional SAGE cone* is defined as

$$C_X(\mathcal{T}) = \sum_{\beta \in \mathcal{T}} C_{X,\mathrm{AGE}}(\mathcal{T} \setminus \{\beta\}, \beta). \tag{3.2}$$

These cones are abbreviated as the *X-AGE cone* and the *X-SAGE cone*. The result on cancellation-free representation from Theorem 6 also holds for the constrained situation.

In the transition from the unconstrained optimization to the constrained optimization, the following key change in geometry and combinatorics takes place. Let f be a signomial of the form $f = \sum_{\alpha \in \mathcal{A}} c_\alpha e^{\langle \alpha, x \rangle} + d e^{\langle \beta, x \rangle}$ with $c_\alpha \geq 0$ for all $\alpha \in \mathcal{A}$. If f is nonnegative over the set $X = \mathbb{R}^n$, then we have $\beta \in \operatorname{conv} \mathcal{A}$. That is, the exponent vector of the term with negative coefficients lies in the convex hull of the exponent vectors of terms with positive coefficients, see Figure 3. This property can get lost for other sets X. For example, for $X = \mathbb{R}_+$ and $\mathcal{T} = \{(0), (1)\}$, the signomial $f : \mathbb{R} \to \mathbb{R}$, $f(x) = \exp(x) - \exp(0)$ is contained in $C_X(\mathcal{T})$. The left picture in Figure 4 shows the graph of f and the right picture the support points.

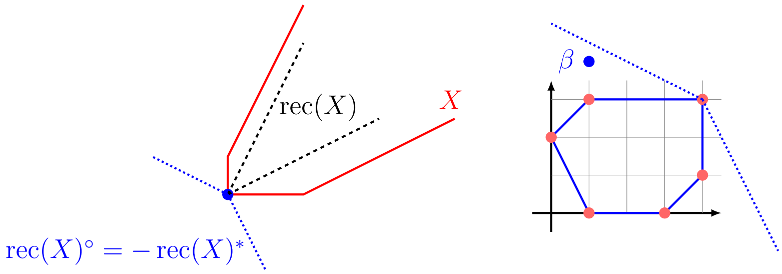

FIGURE 5. The left picture visualizes the inequality $\sigma_X(-\mathcal{T}\nu) < \infty$ in terms of the condition that $-\mathcal{T}\nu$ is contained in the relative interior of $\operatorname{rec}(X)^\circ$. The right picture shows the cone in which the term with negative coefficient has to lie. For example, the blue dot would be a possible support point for the negative coordinate β.

To characterize this change in geometry and combinatorics, recall that the *recession cone* $\operatorname{rec}(X)$ of a convex set X is defined as

$$\operatorname{rec}(X) := \{y \,:\, \exists x \in X \text{ such that } x + \lambda y \in X \text{ for all } \lambda \geq 0\}.$$

For a given cone S, let $S^\circ := \{y \in \mathbb{R}^n \,:\, \langle x, y\rangle \leq 0 \text{ for all } x \in S\}$ denote the *polar cone* of S. Observe that the polar cone and the dual cone are related through $S^\circ = -S^*$. For $\mathcal{T} := \mathcal{A} \cup \{\beta\}$ and a given convex set X, the inequality $\sigma_X(-\mathcal{T}\nu) < \infty$ implies the weaker condition $\beta \in \operatorname{conv} \mathcal{A} + \operatorname{rec}(X)^\circ \;=\; \operatorname{conv} \mathcal{A} - \operatorname{rec}(X)^*$, see Figure 5.

Membership to the conditional SAGE cone can now be formulated as a relative entropy program as well. For disjoint $\emptyset \neq \mathcal{A} \subset \mathbb{R}^n$ and $\mathcal{B} \subset \mathbb{R}^n$, write

$$C_X(\mathcal{A}, \mathcal{B}) \;:=\; \sum_{\beta \in \mathcal{B}} C_{X,\mathrm{AGE}}(\mathcal{A} \cup \mathcal{B} \setminus \{\beta\}, \beta).$$

It holds $C_X(\mathcal{A}, \mathcal{B}) \;=\; \left\{f = \sum_{\alpha \in \mathcal{A}} c_\alpha e^{\langle \alpha, x\rangle} + \sum_{\beta \in \mathcal{B}} c_\beta e^{\langle \beta, x\rangle} \in C_X(\mathcal{A} \cup \mathcal{B}) \,:\, c_\alpha \geq 0 \text{ for } \alpha \in \mathcal{A}\right\}$.

Theorem 11. [25] *A signomial*

$$f = \sum_{\alpha \in \mathcal{A}} c_\alpha e^{\langle \alpha, x\rangle} + \sum_{\beta \in \mathcal{B}} c_\beta e^{\langle \beta, x\rangle}$$

with $c_\alpha \geq 0$ for $\alpha \in \mathcal{A}$ and $c_\beta < 0$ for $\beta \in \mathcal{B}$ is contained in $C_X(\mathcal{A}, \mathcal{B})$ if and only if for every $\beta \in \mathcal{B}$ there exist $c^{(\beta)} \in \mathbb{R}_+^{\mathcal{A}}$ and $\nu^{(\beta)} \in \mathbb{R}_+^{\mathcal{A}}$ such that

$$\begin{aligned} \sigma_X(-(\mathcal{A} \cup \beta)\nu) + D(\nu^{(\beta)}, e \cdot c^{(\beta)}) \;&\leq\; c_\beta \quad \textit{for } \beta \in \mathcal{B}, \\ \sum_{\beta \in \mathcal{B}} c_\alpha^{(\beta)} \;&\leq\; c_\alpha \quad \textit{for } \alpha \in \mathcal{A}. \end{aligned}$$

Since the formulations of the conditional SAGE approach use the support function, efficient numerical computation requires sets X for which the support function can be computed efficiently. A natural class where this is possible is provided by polyhedral sets X given by linear inequalities.

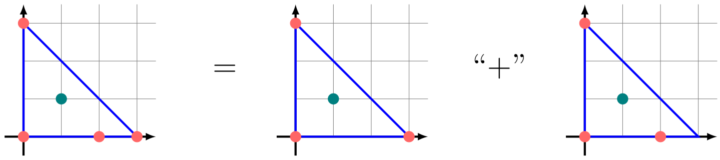

FIGURE 6. The support of f and its summands in Example 12.

4. The circuit view for unconstrained AM/GM optimization

Revealing the structure of the SAGE cone and of the conditional SAGE cone relies on the decomposition of signomials into simpler signomials. In this section, we present the central ideas for the case of unconstrained AM/GM optimization. The decomposition property manifests itself on the level of the dual vector ν in the entropy condition for nonnegativity. The linear (in-)equalities in the entropy condition offer a polyhedral view and an access through generators known as simplicial circuits. We begin with an example.

Example 12. The nonnegative function $f = 7e^0 + 3e^{2x} + e^{3x} + 3e^{3y} - 9e^{x+y}$ decomposes as

$$f \;=\; \frac{1}{2}\left(2e^0 + 2e^{3x} + 2e^{3y} - 6e^{x+y}\right) + \frac{1}{2}\left(12e^0 + 6e^{2x} + 4e^{3y} - 12e^{x+y}\right)$$

into two non-proportional nonnegative AGE signomials. The support sets of f and the summands are depicted in Figure 6.

Denote by $\mathbf{1}$ the all-ones vector and by supp the support of a vector, i.e., the index set of its non-zero components.

Definition 13. A nonzero vector $\nu^* \in \{\nu \in \mathbb{R}^{\mathcal{T}} : \langle \mathbf{1}, \nu \rangle = 0\}$ with $\mathcal{T}\nu^* = 0$ is called a *simplicial circuit* if it is minimally supported and has exactly one negative component, named $(\nu^*)^-$. Here, minimally supported means that there does not exist $\nu' \in \mathbb{R}^{\mathcal{T}} \setminus \{0\}$ with $\operatorname{supp} \nu' \subsetneq \operatorname{supp} \nu^*$, $\langle \mathbf{1}, \nu' \rangle = 0$ and $\mathcal{T}\nu' = 0$. A simplicial circuit is called *normalized* if $\nu^*_\beta = -1$, where $\beta := (\nu^*)^-$.

From a combinatorial viewpoint, we can consider the support sets of the positive coefficients of circuits. In matroid theory, these support sets constitute the set of circuits in the affine-linear matroid on the ground set $\mathcal{T}$. When working with normalized simplicial circuits, we often use the symbol λ rather than ν. Denote by $\Lambda(\mathcal{T})$ the set of normalized simplicial circuits. For every simplicial circuit $\lambda \in \Lambda(\mathcal{T})$ with $\lambda_\beta = -1$, Proposition 5 describes a natural set of signomials, whose nonnegativity is witnessed by λ. Formally, let $C(\mathcal{T}, \lambda)$ be the λ-*witnessed AGE cone*

$$\left\{ \sum_{\alpha \in \mathcal{T}} c_\alpha \exp(\langle \alpha, x \rangle) \;:\; \prod_{\alpha \in \mathcal{A} \text{ with } \lambda_\alpha > 0} \left(\frac{c_\alpha}{\lambda_\alpha} \right)^{\lambda_\alpha} \geq -c_\beta, \; c_\alpha \geq 0 \text{ for } \alpha \in \mathcal{T} \setminus \{\beta\} \right\},$$

where $\beta := \lambda^-$. The SAGE cone admits the following decomposition as a Minkowski sum of λ-witnessed AGE cones and of exponential monomials (i.e., signomials whose exponent vector α has a single nonnegative entry).

Theorem 14. *For $\Lambda(\mathcal{T}) \neq \emptyset$, the SAGE cone decomposes as*

$$C(\mathcal{T}) \;=\; \sum_{\lambda \in \Lambda(\mathcal{T})} C(\mathcal{T}, \lambda) + \sum_{\alpha \in \mathcal{T}} \mathbb{R}_+ \cdot \exp(\langle \alpha, x \rangle).$$

As a consequence, every signomial in the SAGE cone $C(\mathcal{T})$ can be written as a non-negative sum of circuit-witnessed signomials and of exponential monomials.

This was shown by Wang [36] in the polynomial setting and by Murray, Chandrasekaran, Wierman [24]. It reveals the two views on the SAGE cone. The first view comes from the definition of the SAGE cone in terms of AGE signomials. Equivalently, we can regard the SAGE cone as being non-negatively generated by the circuit-witnessed signomials and by exponential monomials.

Remark 15. In the polynomial setting, a variety of works used the latter viewpoint in their definition of a cone for AM/GM optimization. Prominently, Iliman and de Wolff [14] have employed the notion of the cone of SONC polynomials ("*Sums of Non-negative Circuit polynomial*"). When considering polynomials over $\mathbb{R}^n_{>0}$, the exponential monomials from the signomial setting become ordinary monomials. In the adaption to polynomials over $\mathbb{R}^n$, which also has to take into account the signs in other orthants, the monomials become monomial squares. The dual SONC cone was studied in [11].

In order to give the idea for the decomposition in Theorem 14, we approach the situation from the linear condition within the relative entropy condition of Theorem 7.

Lemma 16 (Decomposition Lemma)**.** *Let*

$$f = \sum_{\alpha \in \mathcal{A}} c_\alpha \exp(\langle \alpha, x \rangle) + c_\beta \exp(\langle \beta, x \rangle)$$

be a signomial in $C_{\mathrm{AGE}}(\mathcal{A}, \beta)$ and ν satisfy the relative entropy condition for f. If ν can be written as a convex combination $\nu = \sum_{i=1}^k \theta_i \nu^{(i)}$ of non-proportional $\nu^{(i)} \in N_\beta$, then f has a decomposition into non-proportional signomials in $C_{\mathrm{AGE}}(\mathcal{A}, \beta)$.

Before the proof, we illustrate the statement.

Example 17. We adapt the earlier Example 12, taking $f = 2e^0 + 3e^{2x} + e^{3x} + 3e^{3y} - 9e^{x+y}$. Here, the dual vector $\nu = (2, 3, 1, 3, -9)$ certifies the relative entropy condition. This specific situation is rather simple, because the coefficients sum to zero, which leads to a root at the origin. Writing $\nu = \frac{1}{2}(2, 6, 0, 4, -12) + \frac{1}{2}(2, 0, 2, 2, -6)$ gives the decomposition

$$f \;=\; \frac{1}{2}\left(2e^0 + 6e^{2x} + 4e^{3y} - 12e^{x+y}\right) + \frac{1}{2}\left(2e^0 + 2e^{3x} + 2e^{3y} - 6e^{x+y}\right).$$

The first summand $\nu' := \frac{1}{2}(2, 6, 0, 4, -12)$ of ν is a simplicial circuit, so that Lemma 16 does not apply on ν'. Indeed, the signomial $\frac{1}{2}(2e^0 + 6e^{2x} + 4e^{3y} - 12e^{x+y})$ cannot be decomposed into two non-proportional signomials, such that in these two signomials only the exponential monomial $12e^{x+y}$ has a negative coefficient and only the exponential monomials e^0, $6e^{2x}$ and $4e^{3y}$ can have a nonzero, positive coefficient. The same holds true for the second summand of ν and the second summand of f.

Proof of Lemma 16. Denote by $\nu^+ := \{\alpha \in \mathcal{A} \ : \ \nu_\alpha > 0\}$ the positive support of ν and set $\mathcal{T} = \mathcal{A} \cup \{\beta\}$. For the given $f = \sum_{\alpha \in \mathcal{T}} c_\alpha \exp(\langle \alpha, x \rangle)$, construct vectors $c^{(i)}$ for $1 \le i \le k$ by

$$c_\alpha^{(i)} \ = \ \begin{cases} (c_\alpha/\nu_\alpha)\nu_\alpha^{(i)} & \text{if } \alpha \in \nu^+ \\ 0 & \text{otherwise} \end{cases} \qquad \text{for all } \alpha \in \mathcal{T} \setminus \{\beta\}$$

and by setting $c_\beta^{(i)} = D(\nu_{\setminus \beta}^{(i)}, e c_{\setminus \beta}^{(i)})$. The coefficient vectors $c^{(i)}$ define nonnegative AGE signomials. It remains to show that $\sum_{i=1}^k \theta_i c^{(i)} \le c$.

For indices $\alpha \in \nu^+$, we have equality by construction, and for indices $\alpha \in \operatorname{supp} c \setminus \operatorname{supp} \nu$, we have $\sum_{i=1}^k \theta_i c_\alpha^{(i)} = 0 \le c_\alpha$. Now consider the index β. By construction, $\nu_\alpha^{(i)}/c_\alpha^{(i)} = \nu_\alpha / c_\alpha$ for $\alpha \in \mathcal{T} \setminus \{\beta\}$, which gives

$$\sum_{i=1}^k \theta_i D(\nu_{\setminus \beta}^{(i)}, e c_{\setminus \beta}^{(i)}) \ = \ \sum_{i=1}^k \theta_i \sum_{\alpha \in \mathcal{A}} \nu_\alpha^{(i)} \ln \frac{\nu_\alpha^{(i)}}{e \cdot c_\alpha^{(i)}} \ = \ D(\nu_{\setminus \beta}, e c_{\setminus \beta}).$$

Hence,

$$\sum_{i=1}^k \theta_i c_\beta^{(i)} \ = \ \sum_{i=1}^k \theta_i D(\nu_{\setminus \beta}^{(i)}, e c_{\setminus \beta}^{(i)}) \ = \ D(\nu_{\setminus \beta}, e c_{\setminus \beta}) \le c_\beta.$$

□

Example 18. We consider circuits in the one-dimensional space $\mathbb{R}$. Let $\mathcal{T} = \{\alpha_1, \ldots, \alpha_m\} \subset \mathbb{R}$. Then the simplicial circuits are recognized as the union of the edge generators of the polyhedral cones N_β for $\beta \in \mathcal{T}$. The set of normalized circuits can then be determined as

$$\lambda \ = \ \frac{\alpha_k - \alpha_j}{\alpha_k - \alpha_i} e_i - e_j + \frac{\alpha_j - \alpha_i}{\alpha_k - \alpha_i} e_k \quad \text{for } i < j < k,$$

where e_i denotes the i-th unit vector in $\mathbb{R}^m$. Applying Theorem 14 gives a Minkowski decomposition of the univariate SAGE cone with ground set $\mathcal{T}$.

Using the circuit concept, membership in the SAGE cone can be also certified by a second-order cone program (see Averkov [2] in the polynomial setting, Magron, Wang [18] for a computational view and Naumann and the current author [27] for the extension of second-order representability to the primal and the dual $\mathcal{S}$-cone) or a power cone (Papp [33]).

A principal question is whether further decompositions are possible by employing different negative terms. This will be treated in Section 6.

5. Sublinear circuits

In the previous sections, we have discussed the circuit concept for unconstrained AM/GM optimization and we have presented the framework of constrained AM/GM optimization. Strikingly, a generalized circuit concept can also be established for the case of constrained AM/GM optimization. These generalized circuits are called sublinear circuits and the setup is visualized in Figure 7.

Let $X \subset \mathbb{R}^n$ be a convex set and $\mathcal{T} \subset \mathbb{R}^n$ be the finite ground support, where we assume that the functions $x \mapsto \exp(\langle \alpha, x \rangle)$, $\alpha \in \mathcal{T}$, are linearly independent on X. We consider the conditional AGE cone $C_{X,\mathrm{AGE}}(\mathcal{A}, \beta)$ and the conditional SAGE cone $C_X(\mathcal{T})$ as defined

Unconstrained AM/GM optimization	$\longrightarrow$	Conditional AM/GM optimization
$\downarrow$		$\downarrow$
Circuits	$\longrightarrow$	Sublinear circuits

FIGURE 7. The role of the sublinear circuits.

in (3.1) and (3.2). The primary goal of a circuit concept is to decompose signomials into simpler signomials and thus to decompose a cone under consideration. This will be an essential ingredient for studying the geometry of the conditional SAGE cone for a given set X and we will achieve a Minkowski decomposition. Further refinements will then yield irredundant decompositions, both for the unconstrained and the constrained case, and characterizations of the extreme rays.

In [26], the following concept of sublinear circuits has been developed to resolve these questions. For $\beta \in \mathcal{T}$, recall the notion $N_\beta = \{\nu \in \mathbb{R}^{\mathcal{T}} : \nu_{\setminus\beta} \geq \mathbf{0}, \sum_{\alpha\in\mathcal{T}} \nu_\alpha = 0\}$. We consider the following generalization of the circuit notion from Definition 13. Recall that the support function σ_X is sublinear, that is, it satisfies $\sigma_X(\mu_1 z_2 + \mu_2 z_2) \leq \mu_1 \sigma_X(z_1) + \mu_2 \sigma_X(z_2)$ for $\mu_1, \mu_2 > 0$ and $z_1, z_2 \in \mathbb{R}^n$.

Definition 19. A non-zero vector $\nu^* \in N_\beta$ is called a *sublinear circuit of* $\mathcal{T}$ *with respect to* X (for short, X*-circuit*) if

(1) $\sigma_X(-\mathcal{T}\nu^*) < \infty$,
(2) whenever a mapping $\nu \mapsto \sigma_X(-\mathcal{T}\nu)$ is linear on a two-dimensional cone in N_β, then ν^* is not in the relative interior of that cone.

Denote by $\Lambda_X(\mathcal{T})$ the set of all normalized X-circuits of $\mathcal{T}$, i.e., circuits with entry -1 in the negative coordinate.

In more geometric terms, Definition 19 can be equivalently expressed as follows. For given $\beta \in \mathcal{T}$, consider the convex cone

$$\begin{aligned} P &:= \operatorname{pos}\{(\nu, \sigma_X(-\mathcal{T}\nu)) \;:\; \nu \in N_\beta,\ \sigma_X(-\mathcal{T}\nu) < \infty\} \\ &= \{(\nu, \sigma_X(-\mathcal{T}\nu)) \;:\; \nu \in N_\beta,\ \sigma_X(-\mathcal{T}\nu) < \infty\} \cup \{\mathbf{0}\}, \end{aligned}$$

where pos denotes the positive hull. This cone can be shown to be closed and pointed. Then a vector $\nu^* \in N_\beta$ is an X-circuit if and only if $(\nu^*, \sigma_X(-\mathcal{T}\nu^*))$ spans an extreme ray of P. For the case of polyhedral X, this characterization of X-circuits is straightforward to see and for non-polyhedral X, the convex-geometric details can be found in [26, Theorem 3.6].

As explained in the following, the structure of the sublinear circuits generalizes the structure of the affine-linear matroid which appears in the circuits of the unconstrained case.

Example 20. 1. Let $X = [-1,1]^2$ and $\mathcal{T} = \{(0,0)^T, (1,0)^T, (0,1)^T\}$. Since X is compact, the condition $\sigma_X(-\mathcal{T}\nu) < \infty$ is always satisfied. The set of normalized circuits is $\Lambda_X(\mathcal{T}) = \{(-1,1,0)^T, (-1,0,1)^T, (1,-1,0)^T, (0,-1,1)^T, (1,0,-1)^T, (0,1,-1)^T\}$. Namely, for $\beta = (0,0)^T$, the closed convex cone

$$\begin{aligned} P &= \operatorname{pos}\{(\nu, \sigma_X(-\mathcal{T}\nu)) \ : \ \nu_{(0,0)} < 0, \ \nu_{(1,0)} \geq 0, \ \nu_{(0,1)} \geq 0\} \\ &= \operatorname{pos}\{(\nu, \sigma_X((-\nu_{(1,0)}, -\nu_{(0,1)})^T)) \ : \ \nu_{(0,0)} < 0, \ \nu_{(1,0)} \geq 0, \ \nu_{(0,1)} \geq 0\} \\ &= \operatorname{pos}\{(\nu, (-1,-1) \cdot (\nu_{(1,0)}, \nu_{(0,1)})^T) \ : \ \nu_{(0,0)} < 0, \ \nu_{(1,0)} \geq 0, \ \nu_{(0,1)} \geq 0\} \\ &= \operatorname{pos}\{(-1,1,0,-1)^T, (-1,0,1,-1)^T\} \end{aligned}$$

is generated by the vectors $(\nu, \sigma_X(-\mathcal{T}\nu))$, where ν runs over $(-1,1,0)^T$ and $(-1,0,1)^T$. Similarly, the other elements of $\Lambda_X(\mathcal{T})$ result from the cases $\beta = (1,0)^T$ and $\beta = (0,1)^T$.

2. While we are mostly interested in polyhedral sets X, it is instructive to consider also non-polyhedral situations. Let $X = \mathbb{B}^2 = \{x \in \mathbb{R}^2 \ : \ x_1^2 + x_2^2 \leq 1\}$ be the unit disk in $\mathbb{R}^2$ and let $\mathcal{T} = \{(0,0)^T, (1,0)^T, (0,1)^T\}$. Then the set of circuits is the infinite set $N_1 \cup N_2 \cup N_3$, where

$$N_i \ = \ \{\nu \in \mathbb{R}^3 \setminus \{\mathbf{0}\} \ : \ \nu_1 + \nu_2 + \nu_3 = 0, \ \nu_j \geq 0 \text{ for } j \neq i\}. \tag{5.1}$$

3. Let $X = \mathbb{B}^2 \cap \{x \in \mathbb{R}^2 \ : \ x_2 \geq 0\}$ be the upper unit half disk in $\mathbb{R}^2$ and again $\mathcal{T} = \{(0,0)^T, (1,0)^T, (0,1)^T\}$. Then the set of circuits is $N_1 \cup N_2$, where N_i is defined as in (5.1). Note that $(-1, 1/2, 1/2)$ is not a circuit, because the decomposition into the sum $(-1/2, 1/2, 0)^T + (-1/2, 0, 1/2)^T$ shows a violation of the second condition in Definition 19.

In order to illustrate the suitability of Definition 19, recall from Section 3 that the first condition captures the candidate positions of the support of the negative term. For the special case $X = \mathbb{R}^n$, the condition $\sigma_X(-\mathcal{T}\nu^*) < \infty$ means $\sup_{x \in \mathbb{R}^n} (-\mathcal{T}\nu^*)^T x < \infty$, which is equivalent to $\mathcal{T}\nu^* = \mathbf{0}$. For $\nu^*_\beta = -1$, the components $(\nu^*)_{\alpha \in \mathcal{T} \setminus \{\beta\}}$ are the coefficients of a convex combination of β with respect to $\mathcal{T} \setminus \{\beta\}$. Further, for $X = \mathbb{R}^n$, the second condition in Definition 19 tells us that ν^* is a simplicial circuit of the affine matroid with ground set $\mathcal{T} \subset \mathbb{R}^n$.

In order to explain the relevance of the second condition in Definition 19, let $\mathcal{T} = \mathcal{A} \cup \{\beta\}$ and $f = \sum_{\alpha \in \mathcal{T}} c_\alpha \exp(\langle \alpha, x \rangle) \in C_{X,\mathrm{AGE}}(\mathcal{A}, \beta)$. The Decomposition Lemma 16 can be generalized to capture also the conditional SAGE situation.

Lemma 21 (Decomposition Lemma for conditional SAGE)**.** *Let ν satisfy the relative entropy condition for a signomial f. If ν can be written as a convex combination $\nu = \sum_{i=1}^k \theta_i \nu^{(i)}$ of non-proportional $\nu^{(i)} \in N_\beta$ and $\tilde{\nu} \mapsto \sigma_X(-\mathcal{T}\tilde{\nu})$ is linear on $\operatorname{conv}\{\nu^{(i)}\}_{i=1}^k$, then f can be decomposed as a sum of two non-proportional signomials in $C_{X,\mathrm{AGE}}(\mathcal{A}, \beta)$.*

Proof of Lemma 21. We generalize the proof of Lemma 16. For the given $f = \sum_{\alpha \in \mathcal{T}} c_\alpha \exp(\langle \alpha, x \rangle)$, define vectors $c^{(i)}$ by

$$c_\alpha^{(i)} \ = \ \begin{cases} (c_\alpha / \nu_\alpha) \nu_\alpha^{(i)} & \text{if } \alpha \in \nu^+ \\ 0 & \text{otherwise} \end{cases} \qquad \text{for all } \alpha \in \mathcal{T} \setminus \{\beta\},$$

and $c_\beta^{(i)} = \sigma_X(-\mathcal{T}\nu^{(i)}) + D(\nu^{(i)}_{\setminus \beta}, e c^{(i)}_{\setminus \beta})$. The coefficient vectors $c^{(i)}$ define nonnegative X-AGE signomials. It remains to be shown that $\sum_{i=1}^k \theta_i c^{(i)} \leq c$. For $\alpha \in \nu^+$ and $\alpha \in \operatorname{supp} c \setminus \operatorname{supp} \nu$, this is immediately done as in the proof of Theorem 16. For the index β, we have

$$\sigma_X(-\mathcal{T}\nu) \;=\; \sigma_X\left(-\mathcal{T}(\textstyle\sum_{i=1}^k \theta_i \nu^{(i)})\right) \;=\; \displaystyle\sum_{i=1}^k \theta_i \sigma_X(-\mathcal{T}\nu^{(i)}).$$

Now $\nu_\alpha^{(i)}/c_\alpha^{(i)} = \nu_\alpha/c_\alpha$, gives $\sum_{i=1}^k \theta_i D(\nu_{\backslash\beta}^{(i)}, ec_{\backslash\beta}^{(i)}) = D(\nu_{\backslash\beta}, ec_{\backslash\beta})$. Hence,

$$\begin{aligned}\sum_{i=1}^k \theta_i c_\beta^{(i)} &= \sum_{i=1}^k \theta_i \left(\sigma_X(-\mathcal{T}\nu^{(i)}) + D(\nu_{\backslash\beta}^{(i)}, ec_{\backslash\beta}^{(i)})\right) \\ &= \sigma_X(-\mathcal{T}\nu) + D(\nu_{\backslash\beta}, ec_{\backslash\beta}) \;\leq\; c_\beta.\end{aligned}$$

□

Among the basic properties of sublinear circuits are:

Proposition 22. *1. If $\nu \in N_\beta$ is an X-circuit, then ν^+ ($= \operatorname{supp} \nu \setminus \{\beta\}$) is affinely independent.*
2. For polyhedral X, the number of X-circuits is finite.

The following example illustrates that the combinatorial structure of the sublinear circuits depends on the constraint set X.

Example 23. (Dependency of sublinear circuits on X.) Let $X^{(1)} = \mathbb{R}$, $X^{(2)} = \mathbb{R}_+$ and $X^{(3)} = [-1, 1]$, and let $\mathcal{T} = \{0, 1, 2\}$. The corresponding sets $\Lambda^{(1)}$, $\Lambda^{(2)}$ and $\Lambda^{(3)}$ of normalized sublinear circuits are

$$\begin{aligned}\Lambda^{(1)} &= (1/2, -1, 1/2)^T, \\ \Lambda^{(2)} &= \Lambda^{(1)} \cup \{(0, -1, 1)^T, (-1, 0, 1)^T, (-1, 1, 0)^T\}, \\ \Lambda^{(3)} &= \Lambda^{(2)} \cup \{(0, 1, -1)^T, (1, 0, -1)^T, (1, -1, 0)^T\}.\end{aligned}$$

Now we consider the decomposition of the conditional SAGE cone. In generalization of Section 4, every normalized vector $\lambda \in N_\beta$ with $\sigma_X(-\mathcal{T}\lambda) < \infty$ naturally induces a set of nonnegative signomials and that set of signomials can be described through an explicit condition.

Lemma 24. *Let $\lambda \in N_\beta$ with $\lambda_\beta = -1$ and $\sigma_X(-\mathcal{T}\lambda) < \infty$. Further, let $f = \sum_{\alpha \in \mathcal{T}} c_\alpha \exp(\langle \alpha, x\rangle)$ with $c_\alpha \geq 0$ for $\alpha \in \mathcal{T} \setminus \{\beta\}$ and*

$$\prod_{\alpha \in \lambda^+} \left(\frac{c_\alpha}{\lambda_\alpha}\right)^{\lambda_\alpha} \geq -c_\beta \exp\left(\sigma_X(-\mathcal{T}\lambda)\right), \tag{5.2}$$

where $\beta := \lambda^-$ and $\lambda^+ = \{\alpha \in \mathcal{A} \;:\; \lambda_\alpha > 0\}$ denotes the positive support*. Then f is nonnegative.*

Proof. We show that (5.2) is satisfied if and only if there exists some $\nu \in N_\beta \setminus \{\mathbf{0}\}$ such that the relative entropy condition in Corollary 9 is satisfied. Using $\nu = s\lambda$ with some $s \geq 0$, we have $s = |\nu_\beta|$. Hence,

$$\begin{aligned}D(\nu_{\backslash\beta}, ec_{\backslash\beta}) + \sigma_X(-\mathcal{T}\nu) &= \sum_{\alpha \in \lambda^+} \nu_\alpha \ln \frac{\nu_\alpha}{ec_\alpha} + |\nu_\beta| \sigma_X(-\mathcal{T}\lambda) \\ &= \sum_{\alpha \in \lambda^+} \left(\nu_\alpha \ln \frac{\nu_\alpha}{e\tilde{c}_\alpha}\right),\end{aligned} \tag{5.3}$$

where we have used $|\nu_\beta| = \sum_{\alpha\in\mathcal{A}\,:\,\lambda_\alpha>0} \nu_\alpha$ as well as the scaled coefficients $\tilde{c}_\alpha := c_\alpha \exp(-\sigma_X(-\mathcal{T}\lambda))$.

Using the unconstrained version in Theorem 10, we know that there exists some $\nu = s\lambda$ such that (5.3) is less than or equal to c_β if and only if

$$\prod_{\alpha\in\lambda^+} \left(\frac{\tilde{c}_\alpha}{\lambda_\alpha}\right)^{\lambda_\alpha} \geq -c_\beta.$$

Since the left-hand side evaluates to

$$\prod_{\alpha\in\lambda^+} \left(\frac{c_\alpha}{\lambda_\alpha}\right)^{\lambda_\alpha} \prod_{\alpha\in\lambda^+} \exp(-\sigma_X(-\mathcal{T}\lambda))^{\lambda_\alpha} = \prod_{\alpha\in\lambda^+} \left(\frac{c_\alpha}{\lambda_\alpha}\right)^{\lambda_\alpha} \exp(-\sigma_X(-\mathcal{T}\lambda)),$$

the claim follows. □

Let the λ-*witnessed AGE cone* $C_X(\mathcal{T},\lambda)$ be defined as the set of signomials $\sum_{\alpha\in\mathcal{T}} c_\alpha \exp(\langle\alpha,x\rangle)$ with $c_\alpha \geq 0$ for $\alpha \in \mathcal{T}\setminus\{\beta\}$ and which satisfy (5.2). By Lemma 24, the signomials in $C_X(\mathcal{T},\lambda)$ are nonnegative X-AGE signomials. For polyhedral X, the conditional SAGE cone can be decomposed as a Minkowski sum with respect to the λ-witnessed cones of the sublinear circuits.

Theorem 25. [26] *For polyhedral X, the conditional SAGE cone decomposes as*

$$C_X(\mathcal{T}) \;=\; \sum_{\lambda\in\Lambda_X(\mathcal{T})} C_X(\mathcal{T},\lambda) + \sum_{\alpha\in\mathcal{T}} \mathbb{R}_+ \cdot \exp(\langle\alpha,x\rangle).$$

Using the concept of sublinear circuits, the results on second-order representability and on power cone representability of the SAGE cone mentioned at the end of Section 4 can be generalized to the constrained case. If X is a polyhedron, then $C_X(\mathcal{T})$ is power cone representable. If in addition $\mathcal{T}^T X$ is *rational*, then $C_X(\mathcal{T})$ is second-order representable, see [26].

6. Irredundant decompositions

We consider irredundant decompositions both for the SAGE cone and for the conditional SAGE cone. We begin with the observation that, as a consequence of the definitions, simplicial circuits and sublinear circuits cannot be further decomposed on their supports. This raises the question whether a decomposition is possible on a larger support. Somewhat surprisingly, the answer is yes.

Example 26. The signomial $f(x,y) = e^0 + e^{3x} + e^{3y} - 3e^{x+y}$ is globally nonnegative. On the ground set $\mathcal{T} = \{(0,0)^T, (3,0)^T, (0,3)^T, (1,1)^T\}$, the vector $\nu = (1,1,1,-3)^T$ is a simplicial circuit and thus f cannot be decomposed any further into non-proportional nonnegative signomials in the SAGE cone $C(\mathcal{T})$. If the ground set $\mathcal{T}$ also contains the point $(0,1)^T$, then the exponential e^y $(= e^{0\cdot x+1\cdot y})$ is available and f can be decomposed into two non-proportional, nonnegative circuit signomials,

$$f \;=\; \left(e^{3x} + \frac{1}{2}e^{3y} + \frac{3}{2}e^{y} - 3e^{x+y}\right) + \left(e^0 + \frac{1}{2}e^{3y} - \frac{3}{2}e^{y}\right).$$

reduced for $\mathcal{T} = \left\{ \binom{0}{0}, \binom{3}{0}, \binom{0}{3}, \binom{1}{1} \right\}$

not reduced for $\mathcal{T} = \left\{ \binom{0}{0}, \binom{3}{0}, \binom{0}{3}, \binom{1}{1}, \binom{0}{1} \right\}$

FIGURE 8. Reduced circuits.

To see the nonnegativity of the summands, we can verify the condition in Proposition 5 through

$$\left(\frac{1}{1/3}\right)^{\frac{1}{3}} \cdot \left(\frac{1/2}{1/6}\right)^{\frac{1}{6}} \cdot \left(\frac{3/2}{1/2}\right)^{\frac{1}{2}} = 3^{\frac{1}{3}} \cdot 3^{\frac{1}{6}} \cdot 3^{\frac{1}{2}} = 3,$$
$$\left(\frac{1}{2/3}\right)^{\frac{2}{3}} \cdot \left(\frac{1/2}{1/3}\right)^{\frac{1}{3}} = \left(\frac{3}{2}\right)^{\frac{2}{3}} \cdot \left(\frac{3}{2}\right)^{\frac{1}{3}} = \frac{3}{2}.$$

In [16], the concept of reduced simplicial circuits (for short, reduced circuits) has been introduced.

Definition 27. A simplicial circuit ν is *reduced* if the support points contain no additional element of $\mathcal{T}$ in their convex hull. For the set of normalized reduced circuits on the ground set $\mathcal{T}$, we use the notation $\Lambda^\star(\mathcal{T})$.

See Figure 8 for an illustration. The concept of reduced circuits provides an irredundant decomposition:

Theorem 28. [16] *For a finite support set T, we have*

$$C(\mathcal{T}) = \sum_{\lambda \in \Lambda^\star(\mathcal{T})} C(\mathcal{T}, \lambda) + \sum_{\alpha \in \mathcal{T}} \mathbb{R}_+ \cdot \exp(\langle \alpha, x \rangle)$$

and there is no proper subset $\Lambda \subsetneq \Lambda^\star(\mathcal{T})$ such that $C(\mathcal{T}) = \sum_{\lambda \in \Lambda} C(\mathcal{T}, \lambda) + \sum_{\alpha \in \mathcal{T}} \mathbb{R}_+ \cdot \exp(\langle \alpha, x \rangle)$.

Example 29. Let $\mathcal{T} = \{(0,0)^T, (4,0)^T, (2,4)^T, (1,1)^T, (3,1)^T\}$. The convex hull of $\mathcal{T}$ is a triangle. The points $(1,1)^T$ and $(3,1)^T$ are contained in the interior of the triangle. There are four simplicial circuits and all of them are two-dimensional. Two of them have the three vertices of $\text{conv}(\mathcal{T})$ as outer vertices. Since the two simplicial circuits whose outer vertices are from $\text{conv}(\mathcal{T})$ are not reduced, only the other two simplicial circuits are reduced.

Example 30. For $\mathcal{T} = \{(0,0)^T, (3,0)^T, (0,3)^T, (1,1)^T, (0,1)^T\}$, we compute a decomposition of the non-reduced circuit $\lambda = \left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}, -1, 0\right)^T$. As initial step, we determine another simplicial circuit whose positive support is contained in the (positive or negative) support of the given simplicial circuit. For example, we can choose $\lambda' = (\frac{2}{3}, 0, \frac{1}{3}, 0, -1)^T$. Then we determine the maximal τ such that $\lambda - \tau\lambda'$ is a simplicial circuit and the maximal τ' such that $\lambda' - \tau'\lambda$ is a simplicial circuit. The given circuit can be decomposed into these two newly computed circuits.

In our situation, we obtain $\tau = \frac{1}{2}$ and thus $\nu := \lambda - \tau\lambda' = (0, \frac{1}{3}, \frac{1}{6}, -1, \frac{1}{2})^T$. Further, $\tau' = 0$ and thus $\nu' := \lambda' - \tau\lambda = (\frac{2}{3}, 0, \frac{1}{3}, 0, -1)^T$. Using ν and ν', we observe the decomposition $\lambda = \nu + \frac{1}{2}\nu'$.

Reduced circuits provide an essential tool for studying the SAGE cone. In particular, they allow to characterize the extreme rays of the SAGE cone.

Theorem 31. [16] *The set $\mathcal{E}(\mathcal{T})$ of extreme rays of the SAGE cone $C(\mathcal{T})$ with support $\mathcal{T}$ is*

$$\mathcal{E}(\mathcal{T}) \; = \; \bigcup_{\substack{\lambda \in \Lambda^\star(\mathcal{T}) \\ \beta := \lambda^-}} \left\{ \sum_{\alpha \in \lambda^+} c_\alpha e^{\langle \alpha, x \rangle} - \prod_{\alpha \in \lambda^+} \left(\frac{c_\alpha}{\lambda_\alpha} \right)^{\lambda_\alpha} e^{\langle \beta, x \rangle} \; : \; c_\alpha > 0 \; \forall \alpha \in \lambda^+ \right\}$$
$$\cup \bigcup_{\beta \in \mathcal{T}} \left\{ c e^{\langle \beta, x \rangle} \; : \; c \in \mathbb{R}_+ \right\}.$$

Algebraic aspects of the boundary of the SAGE cone, such as the degree of the algebraic boundary, have been studied by Forsgård and de Wolff [12]. The reducibility concept and the irredundant decomposition in Theorem 28 can be generalized to the conditional SAGE cone. We start with an example.

Example 32. Let $X = [0, \infty)$ and $\mathcal{T} = \{0, 1, 2\}$.

1) For $f = -e^0 + e^{2x}$, choosing $\nu = (-1, 0, 1)^T$ works as a dual certificate to certify the nonnegativity through the entropy condition. Writing $\nu = \frac{1}{2}(-2, 2, 0)^T + \frac{1}{2}(0, -2, 2)^T$ gives the decomposition

$$f \; = \; \frac{1}{2}(-2e^0 + 2e^x) + \frac{1}{2}(-2e^x + 2e^{2x}).$$

2) For $f = -e^x + e^{2x}$, choosing $\nu = (0, -1, 1)^T$ works to certify the nonnegativity. Writing $\nu = \frac{1}{2}(2, -4, 2)^T + \frac{1}{2}(-2, 2, 0)^T$ gives the decomposition

$$f \; = \; \frac{1}{2}(2e^0 - 4e^x + 2e^{2x}) + \frac{1}{2}(-2e^0 + 2e^x).$$

Formally, reduced sublinear circuits are defined as follows.

Definition 33. The reduced X-circuits of $\mathcal{T}$ are the X-circuits ν^* such that $(\nu^*, \sigma_X(-\mathcal{T}\nu^*))$ generates an extreme ray of

$$\operatorname{pos}\left(\left\{ \left(\nu, \sigma_X(-\mathcal{T}\nu) \right) \; : \; \lambda \in \Lambda_X(\mathcal{T}) \right\} \cup \{(\mathbf{0}, 1)\} \right).$$

The set of normalized reduced X-circuits is denoted by $\Lambda_X^\star(\mathcal{T})$.

The reduced sublinear circuits are sufficient to generate the full conditional SAGE cone in the following sense. For a polyhedral set X, this can be stated more simply as an irreducible Minkowski decomposition.

Theorem 34. [26] *For a finite support set $\mathcal{T}$, we have*

$$C_X(\mathcal{T}) \; = \; \operatorname{cl}\left(\operatorname{conv}\left\{ C_X(\mathcal{T}, \lambda) \; : \; \lambda \in \Lambda_X^\star(\mathcal{T}) \right\}\right) + \sum_{\alpha \in \mathcal{T}} \mathbb{R}_+ \cdot \exp(\langle \alpha, x \rangle).$$

For polyhedral X, the conditional SAGE cone decomposes as the finite Minkowski sum

$$C_X(\mathcal{T}) \;=\; \sum_{\lambda \in \Lambda_X^\star(\mathcal{T})} C_X(\mathcal{T}, \lambda) + \sum_{\alpha \in \mathcal{T}} \mathbb{R}_+ \cdot \exp(\langle \alpha, x \rangle)$$

and there does not not exist a proper subset $\Lambda \subsetneq \Lambda_X^\star(\mathcal{T})$ such that $C_X(\mathcal{T}) = \sum_{\lambda \in \Lambda} C_X(\mathcal{T}, \lambda) + \sum_{\alpha \in \mathcal{T}} \mathbb{R}_+ \cdot \exp(\langle \alpha, x \rangle)$.

7. Further developments

Let us mention some further research directions on the SAGE cone and the conditional SAGE cone. Symmetry reduction for AM/GM-based optimization has been studied and also computationally evaluated by Moustrou, Naumann, Riener et al. [21]. Recently, extensions of the conditional SAGE approach towards hierarchies and Positivstellensätze [35] and to additional non-convex constraints [10] have been given. The latter work also compares the computations in the software system `Sageopt` [23] to semidefinite hierarchies based on sum of squares and moments.

The combinatorial structure of the sublinear circuits is also rather open. For some results concerning polyhedral sets X see [28]. That work gives some necessary and sufficient conditions for an element ν to be a (reduced) X-circuit and discusses distinguished classes, such as $X = \mathbb{R}_+^n$ and $X = [-1, 1]^n$. Moreover, further understanding of the conditional SAGE cone, such as its algebraic boundary, through the sublinear circuits remains to be done. From the practical point of view, combining the conditional SAGE cone into hybrid techniques with other existing optimization techniques for nonnegativity certificates, such as sums of squares, appears to be a relevant task.

Acknowledgement. This expository article took its origin in the presentation material developed for a minicourse held at the Second POEMA (Polynomial Optimization, Efficiency Through Moments and Algebra) Learning Week, September 2021, Toulouse/hybrid. Thanks to Constantin Ickstadt, Nadine Defoßa as well as the anonymous reviewers for helpful comments.

References

[1] A. A. Ahmadi and A. Majumdar. DSOS and SDSOS optimization: more tractable alternatives to sum of squares and semidefinite optimization. *SIAM J. Appl. Algebra Geom.*, 3(2):193–230, 2019.

[2] G. Averkov. Optimal size of linear matrix inequalities in semidefinite approaches to polynomial optimization. *SIAM J. Appl. Algebra Geom.*, 3(1):128–151, 2019.

[3] S. Boyd, S.-J. Kim, L. Vandenberghe, and A. Hassibi. A tutorial on geometric programming. *Optimization and Engineering*, 8:67–127, 2007.

[4] V. Chandrasekaran and P. Shah. Relative entropy relaxations for signomial optimization. *SIAM J. Optim.*, 26(2):1147–1173, 2016.

[5] V. Chandrasekaran and P. Shah. Relative entropy optimization and its applications. *Math. Program., Ser. A*, 161(1-2):1–32, 2017.

[6] R. Chares. *Cones and Interior-Point Algorithms for Structured Convex Optimization Involving Powers and Exponentials*. PhD thesis, Université Catholique de Louvain, 2009.

[7] S. Diamond and S. Boyd. CVXPY: A Python-embedded modeling language for convex optimization. *J. Machine Learning Research*, 17(83):1–5, 2016.

[8] A. Domahidi, E. Chu, and S. Boyd. ECOS: An SOCP solver for embedded systems. In *Proc. European Control Conference 2013, Zürich*, pages 3071–3076. IEEE, 2013.
[9] M. Dressler, S. Iliman, and T. de Wolff. A Positivstellensatz for sums of nonnegative circuit polynomials. *SIAM J. Appl. Algebra Geom.*, 1(1):536–555, 2017.
[10] M. Dressler and R. Murray. Algebraic perspectives on signomial optimization. *SIAM J. Appl. Algebra Geom.*, 6(4):650–684, 2022.
[11] M. Dressler, H. Naumann, and T. Theobald. The dual cone of sums of non-negative circuit polynomials. *Adv. Geom.*, 21(2):227–236, 2021.
[12] J. Forsgård and T. de Wolff. The algebraic boundary of the SONC cone. *SIAM J. Appl. Algebra Geom.*, 6:468–502, 2022.
[13] D. Handelman. Representing polynomials by positive linear functions on compact convex polyhedra. *Pacific J. Math.*, 132(1):35–62, 1988.
[14] S. Iliman and T. de Wolff. Amoebas, nonnegative polynomials and sums of squares supported on circuits. *Res. Math. Sci.*, 3(paper no. 9), 2016.
[15] O. Karaca, G. Darivianakis, P. Beuchat, A. Georghiou, and J. Lygeros. The REPOP toolbox: Tackling polynomial optimization using relative entropy relaxations. In *20th IFAC World Congress, IFAC PapersOnLine*, volume 50(1), pages 11652–11657. Elsevier, 2017.
[16] L. Katthän, H. Naumann, and T. Theobald. A unified framework of SAGE and SONC polynomials and its duality theory. *Math. Comput.*, 90:1297–1322, 2021.
[17] J.B. Lasserre. *Moments, Positive Polynomials and their Applications*. Imperial College Press, London, 2010.
[18] V. Magron and J. Wang. SONC optimization and exact nonnegativity certificates via second-order cone programming. *J. Symb. Comput.*, 115:346–370, 2023.
[19] V. Magron and J. Wang. *Sparse Polynomial Optimization: Theory and Practice*. World Scientific, 2023.
[20] MOSEK. MOSEK Modeling Cookbook 3.3.0. Online, https://docs.mosek.com/modeling-cookbook/index.html, 2022.
[21] P. Moustrou, H. Naumann, C. Riener, T. Theobald, and H. Verdure. Symmetry reduction in AM/GM-based optimization. *SIAM J. Optim.*, 32:765–785, 2022.
[22] S. Müller, E. Feliu, G. Regensburger, C. Conradi, A. Shiu, and A. Dickenstein. Sign conditions for injectivity of generalized polynomial maps with applications to chemical reaction networks and real algebraic geometry. *Found. Comp. Math.*, 16(1):69–97, 2015.
[23] R. Murray. Sageopt 0.5.3, 2020. DOI:10.5281/ZENODO.4017991.
[24] R. Murray, V. Chandrasekaran, and A. Wierman. Newton polytopes and relative entropy optimization. *Found. Comput. Math.*, 21:1703–1737, 2021.
[25] R. Murray, V. Chandrasekaran, and A. Wierman. Signomial and polynomial optimization via relative entropy and partial dualization. *Math. Program. Comput.*, 13:257–295, 2021.
[26] R. Murray, H. Naumann, and T. Theobald. Sublinear circuits and the constrained signomial nonnegativity problem. *Math. Program.*, 198:471–505, 2023.
[27] H. Naumann and T. Theobald. The $\mathcal{S}$-cone and a primal-dual view on second-order representability. *Beiträge Algebra Geom. (Special issue on the 50th anniversary of the journal)*, 62:229–249, 2021.
[28] H. Naumann and T. Theobald. Sublinear circuits for polyhedral sets. *Vietnam J. Math. (Special issue on the honor of Bernd Sturmfels)*, 50:447–468, 2022.
[29] Y. Nesterov. Constructing self-concordant barriers for convex cones. CORE discussion paper no. 2006/30, 2006.
[30] Y. Nesterov. *Lectures on Convex Optimization*. Second edition. Springer, 2018.
[31] C. Nowzari, V. M. Preciado, and G. J. Pappas. Optimal resource allocation for control of networked epidemic models. *IEEE Transactions on Control of Network Systems*, 4(2):159–169, 2017.
[32] C. Pantea, H. Koeppl, and G. Craciun. Global injectivity and multiple equilibria in uni- and bi-molecular reaction networks. *Discrete Contin. Dyn. Syst., Ser. B*, 17(6):2153–2170, 2012.

[33] D. Papp. Duality of sum of nonnegative circuit polynomials and optimal SONC bounds. *J. Symb. Comput.*, 114:246–266, 2023.
[34] B. Reznick. Forms derived from the arithmetic-geometric inequality. *Math. Annalen*, 283(3):431–464, 1989.
[35] A. H. Wang, P. Jaini, Y. Yu, and P. Poupart. A Positivstellensatz for conditional SAGE signomials. Preprint, arXiv:2003.03731, 2020.
[36] J. Wang. Nonnegative polynomials and circuit polynomials. *SIAM J. Appl. Algebra Geom.*, 6(2):111–133, 2022.
[37] J. Wang, V. Magron, J.-B. Lasserre, and N. H. A. Mai. CS-TSSOS: correlative and term sparsity for large-scale polynomial optimization. *ACM Trans. Math. Softw.*, 48(4):1–26, 2022.
[38] M. York, W. Hoburg, and M. Drela. Turbofan engine sizing and tradeoff analysis via signomial programming. *Journal of Aircraft*, 55(3):988–1003, 2018.

Goethe-Universität, FB 12 – Institut für Mathematik, Postfach 11 19 32, D–60054 Frankfurt am Main, Germany

Email address: `theobald@math.uni-frankfurt.de`