RELATIVE ENTROPY RELAXATIONS FOR SIGNOMIAL OPTIMIZATION[*]

VENKAT CHANDRASEKARAN[†] AND PARIKSHIT SHAH[‡]

Abstract. Signomial programs (SPs) are optimization problems specified in terms of signomials, which are weighted sums of exponentials composed with linear functionals of a decision variable. SPs are nonconvex optimization problems in general, and families of NP-hard problems can be reduced to SPs. In this paper we describe a hierarchy of convex relaxations to obtain successively tighter lower bounds of the optimal value of SPs. This sequence of lower bounds is computed by solving increasingly larger-sized relative entropy optimization problems, which are convex programs specified in terms of linear and relative entropy functions. Our approach relies crucially on the observation that the relative entropy function, by virtue of its joint convexity with respect to both arguments, provides a convex parametrization of certain sets of globally nonnegative signomials with efficiently computable nonnegativity certificates via the arithmetic-geometric-mean inequality. By appealing to representation theorems from real algebraic geometry, we show that our sequences of lower bounds converge to the global optima for broad classes of SPs. Finally, we also demonstrate the effectiveness of our methods via numerical experiments.

Key words. arithmetic-geometric-mean inequality, convex optimization, geometric programming, global optimization, real algebraic geometry

AMS subject classifications. 90C25, 90C26

DOI. 10.1137/140988978

1. Introduction. A *signomial* is a weighted sum of exponentials composed with linear functionals of a variable $\mathbf{x} \in \mathbb{R}^n$:

$$(1.1) \qquad f(\mathbf{x}) = \sum_{j=1}^{\ell} \mathbf{c}_j \exp\left\{ \boldsymbol{\alpha}^{(j)'} \mathbf{x} \right\}.$$

Here $\mathbf{c}_j \in \mathbb{R}$ and $\boldsymbol{\alpha}^{(j)} \in \mathbb{R}^n$ are fixed parameters.[1] A signomial program (SP) is an optimization problem in which a signomial is minimized subject to constraints on signomials, all in a decision variable $\mathbf{x}$ [10, 11]. SPs are nonconvex in general, and they include NP-hard problems as special cases [5]. Consequently, we do not expect to obtain globally optimal solutions of general SPs in a computationally tractable manner. This paper describes a hierarchy of convex relaxations based on relative entropy optimization for obtaining lower bounds of the optimal value of an SP.

Geometric programs (GPs) constitute a prominent subclass of SPs in which a *posynomial*—a signomial with positive coefficients, i.e., the $\mathbf{c}_j$'s are all positive—is minimized subject to upper-bound constraints on posynomials [3, 11]. GPs are convex

[†]Departments of Computing and Mathematical Sciences and of Electrical Engineering, California Institute of Technology, Pasadena, CA 91125 (venkatc@caltech.edu).

[‡]Yahoo! Research and Wisconsin Institutes for Discovery at the University of Wisconsin, Madison, WI 53715 (pshah@discovery.wisc.edu).

[1]In the literature [11], signomials are typically parametrized somewhat differently as weighted sums of generalized "monomials." A monomial consists of a product of variables, each raised to an arbitrary real power, and the variables only take on positive values. It is straightforward to transform such functions to sums of exponentials of the type (1.1) discussed in this paper; see [3, 11].

optimization problems that can be solved efficiently, and they have been successfully employed in a wide range of applications such as power control in communication systems [5], circuit design [2], approximations to the matrix permanent [16], and the computation of capacities of point-to-point communication channels [6]. However, the additional flexibility provided by SPs via constraints on arbitrary signomials rather than just posynomials is useful in a range of problems to which GPs are not directly applicable. Examples include resource allocation in networks [5], control problems involving chemical processes (see the extensive reference list in [17]), spatial frame design [33], and certain nonlinear flow problems in graphs [21].

1.1. Our contributions. Central to the development in this paper is a view of global minimization that is grounded in duality: globally minimizing a signomial is computationally equivalent to the problem of certifying global nonnegativity of a signomial. Although certifying global nonnegativity of a general signomial is a computationally intractable problem, one can appeal to GP duality [11] to show that certifying nonnegativity of a signomial with all but one coefficient being positive can be accomplished in a computationally tractable manner. These observations suggest a natural sufficient condition for certifying nonnegativity of general signomials via certificates that can be computed efficiently. Specifically, if a signomial $f(\mathbf{x})$ can be decomposed as $f(\mathbf{x}) = \sum_{i=1}^{k} f_i(\mathbf{x})$, where each $f_i(\mathbf{x})$ is a globally nonnegative signomial with at most one negative coefficient, then $f(\mathbf{x})$ is clearly nonnegative. We refer to a decomposition of this form as a sum-of-arithmetic-geometric-mean exponentials (SAGE) decomposition and to the associated functions $f(\mathbf{x})$ that are decomposable in this manner as SAGE functions. The reason for this terminology is that certifying global nonnegativity of each of the functions $f_i(\mathbf{x})$ is accomplished by verifying an appropriate arithmetic-geometric-mean (AM/GM) inequality as described in section 2.1.

The key insight underlying our approach in section 2 is that computing a SAGE decomposition of $f(\mathbf{x})$ can be cast as a feasibility problem involving a system of linear inequalities as well as inequalities specifying upper bounds on *relative entropy* functions. Recall that the relative entropy function is defined as follows for $\boldsymbol{\nu}, \boldsymbol{\lambda}$ in the nonnegative orthant $\mathbb{R}_+^{\ell}$:

$$(1.2) \qquad D(\boldsymbol{\nu}, \boldsymbol{\lambda}) = \sum_{j=1}^{\ell} \boldsymbol{\nu}_j \log\left(\frac{\boldsymbol{\nu}_j}{\boldsymbol{\lambda}_j}\right).$$

This function is jointly convex in both arguments [4, p. 90], and the associated feasibility problem of finding a SAGE decomposition can be solved efficiently via convex optimization. As discussed in section 3.1, this method can be employed to obtain lower bounds on the original signomial $f(\mathbf{x})$ by solving a tractable convex program involving linear as well as relative entropy functions.

In section 3 we also describe a principled framework to obtain a family of increasingly tighter lower bounds for general (constrained and unconstrained) SPs. This family of bounds is obtained by solving hierarchies of successively larger convex programs based on relative entropy optimization; these hierarchies are derived by considering a sequence of tighter nonnegativity certificates for a signomial over a set defined by signomial constraints. A prominent example of hierarchies of tractable convex programs being employed for intractable problems is in the setting of polynomial optimization problems, for which SDP relaxations have been developed based on sum-of-squares techniques [15, 19, 20]. However, those methods are not directly relevant to SPs for several reasons, and we highlight these distinctions in section 4.2.

The hierarchy of convex relaxations that we describe in section 3 has several notable features. First, GPs are solved exactly by the first level in this hierarchy; thus, our hierarchy has the desirable property that "easy problem instances remain easy." Second, the family of lower bounds is invariant under a natural transformation of the problem data. Specifically, the optimal value of an SP remains unchanged under the application of a nonsingular linear transformation simultaneously to all the parameters $\boldsymbol{\alpha}^{(j)}$ that appear in the exponents of the signomials in an SP (both in the objective and in the constraints). The hierarchy of relative entropy relaxations described in section 3 leads to bounds that are invariant under such transformations. Third, it is desirable that any procedure for obtaining lower bounds of the optimal value of an SP be robust to small perturbations of the exponents $\boldsymbol{\alpha}^{(j)}$ in an SP. As discussed in section 4.2, this is also a feature of our proposed approach. Fourth, for broad families of SPs our approach leads to a convergent sequence of lower bounds, i.e., our hierarchy approximates the optimal value of the SP arbitrarily well from below (see Theorems 4.1 and 4.2). Such a property of a hierarchy is usually referred to as *completeness*. At various stages in the paper we discuss numerical experiments that illustrate the effectiveness of our methods.

Related work. Several researchers have developed strategies for optimizing SPs based on variants of branch-and-bound methods [17] as well as heuristic techniques based on successive approximations that can be solved via linear or geometric programming [3, 5, 28, 29]. The framework presented in this paper is qualitatively different as it is based on solving convex optimization problems involving linear and relative entropy functions to obtain guaranteed bounds on the optimal value of SPs. Our approach is founded on the insight that the joint convexity of the relative entropy function leads to an effective convex parametrization of certain families of globally nonnegative signomials.

1.2. Paper outline. Section 2 describes the SAGE decomposition for certifying nonnegativity of a signomial based on relative entropy optimization and its connection to the AM/GM inequality. This approach underlies the subsequent development in section 3, in which a hierarchy of relative entropy relaxations is proposed for general SPs. In section 4, we provide theoretical support for our hierarchies of relative entropy relaxations by proving that they provide a sequence of lower bounds that converges to the global minimum for a broad class of SPs. Throughout the paper, we describe numerical experiments in which relative entropy optimization techniques are employed to obtain bounds on some stylized SPs. These results were obtained using a basic interior-point solver written in MATLAB by following the discussions in [4]. We conclude with a discussion of potential future directions in section 5. We prove Theorems 4.1 and 4.2 in the appendix.

Notation. In ℓ-dimensional space we denote the nonnegative orthant by $\mathbb{R}_+^\ell$, the positive orthant by $\mathbb{R}_{++}^\ell$, rational vectors by $\mathbb{Q}^\ell$, and nonnegative rational vectors by $\mathbb{Q}_+^\ell$. We denote the nonnegative integers by $\mathbb{Z}_+$ and the positive integers by $\mathbb{Z}_{++}$. Given a vector $\mathbf{w} \in \mathbb{R}^\ell$, we denote the ith coordinate by $\mathbf{w}_i$ and denote the vector formed by removing the ith coordinate from $\mathbf{w}$ by $\mathbf{w}_{\backslash i} \in \mathbb{R}^{\ell-1}$. By convention, we assume that $0 \log \frac{0}{\lambda} = 0$ for any $\lambda \in \mathbb{R}_+$ and that $\nu \log \frac{\nu}{0}$ equals 0 if $\nu = 0$ and ∞ if $\nu > 0$. Given a collection of vectors $\{\boldsymbol{\alpha}^{(j)}\}_{j=1}^\ell \subset \mathbb{R}^n$, we define the set $E_p(\{\boldsymbol{\alpha}^{(j)}\}_{j=1}^\ell) \subset \mathbb{R}^n$ for $p \in \mathbb{Z}_+$ as follows:

$$E_p(\{\boldsymbol{\alpha}^{(j)}\}_{j=1}^\ell) = \left\{ \sum_{j=1}^\ell \lambda_j \boldsymbol{\alpha}^{(j)} \;\middle|\; \sum_{j=1}^\ell \lambda_j \leq p, \; \lambda_j \in \mathbb{Z}_+ \right\}.$$

2. SAGE decomposition and relative entropy optimization. In this section we describe a sufficient condition to certify global nonnegativity of a signomial f by decomposing it into a sum of terms $f = \sum_i f_i$, where each f_i has an efficiently computed nonnegativity certificate based on relative entropy optimization. This method serves as the basic building block in the next section on hierarchies of relative entropy relaxations for general SPs.

2.1. Signomials with one negative term and the AM/GM inequality. The main observation underlying our methods is that certifying the nonnegativity of a signomial with at most one negative coefficient can be accomplished efficiently.

DEFINITION 2.1. *A globally nonnegative signomial with at most one negative coefficient is called an* AM/GM-*exponential.*

The reason for this terminology is that certifying the nonnegativity of an AM/GM-exponential corresponds to verifying an AM/GM inequality. We develop this connection more precisely in what follows. Consider a signomial $g(\mathbf{x}) = \beta \exp\{\boldsymbol{\alpha}'\mathbf{x}\} + \sum_{j=1}^{\ell} \mathbf{c}_j \exp\{[\mathbf{Q}'\mathbf{x}]_j\}$ with at most one negative coefficient—here, $\mathbf{c} \in \mathbb{R}_+^{\ell}$, $\beta \in \mathbb{R}$, $\mathbf{Q} \in \mathbb{R}^{n \times \ell}$, and $\boldsymbol{\alpha} \in \mathbb{R}^n$. The signomial $g(\mathbf{x})$ being an AM/GM-exponential is equivalent to the existence of a $\boldsymbol{\nu} \in \mathbb{R}^{\ell}$ satisfying the following conditions:

$$(2.1) \qquad D(\boldsymbol{\nu}, e\mathbf{c}) - \beta \leq 0, \ \boldsymbol{\nu} \in \mathbb{R}_+^{\ell}, \ \mathbf{Q}\boldsymbol{\nu} = (\mathbf{1}'\boldsymbol{\nu})\boldsymbol{\alpha}.$$

This equivalence follows as a consequence of strong duality applied to a suitable convex program.

LEMMA 2.2. *Let* $g(\mathbf{x}) = \beta \exp\{\boldsymbol{\alpha}'\mathbf{x}\} + \sum_{j=1}^{\ell} \mathbf{c}_j \exp\{[\mathbf{Q}'\mathbf{x}]_j\}$, *where* $\mathbf{c} \in \mathbb{R}_+^{\ell}$, $\beta \in \mathbb{R}$, $\mathbf{Q} \in \mathbb{R}^{n \times \ell}$, *and* $\boldsymbol{\alpha} \in \mathbb{R}^n$. *Then* $g(\mathbf{x}) \geq 0$ *for all* $\mathbf{x} \in \mathbb{R}^n$ *if and only if there exists* $\boldsymbol{\nu} \in \mathbb{R}^{\ell}$ *satisfying the conditions* (2.1).

Proof. The signomial $g(\mathbf{x})$ being globally nonnegative is equivalent to the signomial $g(\mathbf{x}) \exp\{-\boldsymbol{\alpha}'\mathbf{x}\}$ being globally nonnegative, which in turn is equivalent to $\sum_{j=1}^{\ell} \mathbf{c}_j \exp\{[\mathbf{Q}'\mathbf{x}]_j - \boldsymbol{\alpha}'\mathbf{x}\} \geq -\beta$ for all $\mathbf{x} \in \mathbb{R}^n$. To certify a lower bound on the convex function $\sum_{j=1}^{\ell} \mathbf{c}_j \exp\{[\mathbf{Q}'\mathbf{x}]_j - \boldsymbol{\alpha}'\mathbf{x}\}$, consider the following convex program:

$$\inf_{\mathbf{x} \in \mathbb{R}^n, \ \mathbf{t} \in \mathbb{R}^{\ell}} \mathbf{c}'\mathbf{t} \quad \text{s.t.} \quad \exp\{[\mathbf{Q}'\mathbf{x}]_j - \boldsymbol{\alpha}'\mathbf{x}\} \leq \mathbf{t}_j, \ j = 1, \dots, \ell.$$

This primal problem satisfies Slater's conditions [26], and its Lagrangian dual is

$$\sup_{\boldsymbol{\nu} \in \mathbb{R}^{\ell}} -D(\boldsymbol{\nu}, e\mathbf{c}) \quad \text{s.t.} \quad \boldsymbol{\nu} \in \mathbb{R}_+^{\ell}, \ \mathbf{Q}\boldsymbol{\nu} = (\mathbf{1}'\boldsymbol{\nu})\boldsymbol{\alpha}.$$

The result follows as strong duality holds and the dual optimum is attained. $\square$

A $\boldsymbol{\nu} \in \mathbb{R}_+^\ell \backslash \mathbf{0}$ satisfying the conditions (2.1) may also be interpreted as a certificate of the nonnegativity of the AM/GM-exponential $g(\mathbf{x})$ via the AM/GM inequality.[2] Specifically, we have that

$$\sum_{j=1}^{\ell} \mathbf{c}_j \exp\{[\mathbf{Q}'\mathbf{x}]_j\} \overset{\text{(i)}}{\geq} \prod_{j=1}^{\ell} \left(\frac{\mathbf{c}_j \exp\{[\mathbf{Q}'\mathbf{x}]_j\}}{\boldsymbol{\nu}_j/(\mathbf{1}'\boldsymbol{\nu})} \right)^{\frac{\boldsymbol{\nu}_j}{\mathbf{1}'\boldsymbol{\nu}}} \overset{\text{(ii)}}{=} \prod_{j=1}^{\ell} \left(\frac{\mathbf{c}_j}{\boldsymbol{\nu}_j/(\mathbf{1}'\boldsymbol{\nu})} \right)^{\frac{\boldsymbol{\nu}_j}{\mathbf{1}'\boldsymbol{\nu}}} \exp\{\boldsymbol{\alpha}'\mathbf{x}\}.$$

Here inequality (i) is a consequence of the AM/GM inequality applied with weights $\frac{\boldsymbol{\nu}}{\mathbf{1}'\boldsymbol{\nu}}$, and equation (ii) follows from (2.1). Next, we observe that:

$$\prod_{j=1}^{\ell} \left(\frac{\mathbf{c}_j}{\boldsymbol{\nu}_j/(\mathbf{1}'\boldsymbol{\nu})} \right)^{\frac{\boldsymbol{\nu}_j}{\mathbf{1}'\boldsymbol{\nu}}} \overset{\text{(iii)}}{=} \exp\left\{ -D(\tfrac{\boldsymbol{\nu}}{\mathbf{1}'\boldsymbol{\nu}}, \mathbf{c}) \right\} \overset{\text{(iv)}}{\geq} -\xi[D(\tfrac{\boldsymbol{\nu}}{\mathbf{1}'\boldsymbol{\nu}}, \mathbf{c}) + \log(\xi) - 1] \; \forall \xi \in \mathbb{R}_+$$

$$\overset{\text{(v)}}{=} -\xi D(\tfrac{\boldsymbol{\nu}}{\mathbf{1}'\boldsymbol{\nu}}, \tfrac{e}{\xi}\mathbf{c}), \; \forall \xi \in \mathbb{R}_+ \overset{\text{(vi)}}{=} -D(\tfrac{\xi\boldsymbol{\nu}}{\mathbf{1}'\boldsymbol{\nu}}, e\mathbf{c}) \; \forall \xi \in \mathbb{R}_+$$

$$\overset{\text{(vii)}}{\geq} -D(\boldsymbol{\nu}, e\mathbf{c}) \overset{\text{(viii)}}{\geq} -\beta.$$

Here (iii) follows from the definition of the relative entropy function (1.2), inequality (iv) follows from the observation that the exponential function is the convex conjugate of the negative entropy function—i.e., $\exp\{-\rho\} = \sup_{\xi \in \mathbb{R}_+} -\xi[\rho + \log(\xi) - 1]$—(v), (vi) follow from the properties of the relative entropy function (1.2) (in particular, by noting that this function is positively homogenous), inequality (vii) follows by setting $\xi = \mathbf{1}'\boldsymbol{\nu}$, and finally inequality (viii) follows from (2.1). Hence the existence of $\boldsymbol{\nu} \in \mathbb{R}^\ell$ satisfying (2.1) certifies the nonnegativity of the signomial $g(\mathbf{x})$. The preceding discussion parallels some of the early expositions on GP duality based on the AM/GM inequality [11], although our formulation (2.1) is parametrized differently compared to the following formulation in the GP literature [5, 11]:

$$(2.2) \qquad D(\boldsymbol{\nu}, \mathbf{c}) + \log(-\beta) \leq 0, \; \boldsymbol{\nu} \in \mathbb{R}_+^\ell, \mathbf{1}'\boldsymbol{\nu} = 1, \; \mathbf{Q}\boldsymbol{\nu} = \boldsymbol{\alpha}.$$

In comparison to (2.1), the parametrization (2.2) is not jointly convex in $(\boldsymbol{\nu}, \mathbf{c}, \beta)$. As discussed after Lemma 2.2, the joint convexity of the inequality $D(\boldsymbol{\nu}, e\mathbf{c}) - \beta \leq 0$ in (2.1) with respect to $(\boldsymbol{\nu}, \mathbf{c}, \beta)$ plays a crucial role in our subsequent development.

Example 2.1. Consider the function $g(\mathbf{x}) = \exp\{\mathbf{x}_1\} + \exp\{\mathbf{x}_2\} - \exp\{(\frac{1}{2} + \epsilon)\mathbf{x}_1 + (\frac{1}{2} - \epsilon)\mathbf{x}_2\}$ for a fixed $\epsilon \in [-\frac{1}{2}, \frac{1}{2}]$. It is easily seen that $\boldsymbol{\nu} = (\frac{1}{2} + \epsilon, \frac{1}{2} - \epsilon)'$ satisfies the conditions (2.1):

$$D((\tfrac{1}{2} + \epsilon, \tfrac{1}{2} - \epsilon)', e\mathbf{1}) - (-1) = D((\tfrac{1}{2} + \epsilon, \tfrac{1}{2} - \epsilon)', \mathbf{1}) \leq 0,$$
$$I_{2\times2} \, \boldsymbol{\nu} = (\tfrac{1}{2} + \epsilon, \tfrac{1}{2} - \epsilon)'.$$

Here $I_{2\times2}$ is the 2×2 identity matrix and the first inequality follows from the fact that $D((\frac{1}{2} + \epsilon, \frac{1}{2} - \epsilon)', \mathbf{1})$ is the negative entropy of a probability vector, which is always nonpositive [8, p. 15]. Consequently, $\boldsymbol{\nu}$ certifies the nonnegativity of $g(\mathbf{x})$.

In describing these results, we have implicitly viewed $\mathbf{c} \in \mathbb{R}_+^\ell, \beta \in \mathbb{R}$ as *fixed* coefficients. However, the conditions (2.1) are *convex* with respect to $\mathbf{c}, \beta$ by virtue of the joint convexity of the relative entropy function. Consequently, the *set* of AM/GM-exponentials with respect to a collection of exponents can be efficiently parametrized

---

[2]The conditions (2.1) being satisfied with $\boldsymbol{\nu} = \mathbf{0}$ corresponds to the situation that $\beta \geq 0$.

as a convex set given by linear and relative entropy inequalities:

$$\mathcal{C}_{\mathrm{AGE}}\big(\boldsymbol{\alpha}^{(1)},\ldots,\boldsymbol{\alpha}^{(\ell)};\boldsymbol{\alpha}^{(0)}\big)$$

$$= \left\{ (\mathbf{c},\beta) \in \mathbb{R}_+^\ell \times \mathbb{R} \mid g(\mathbf{x}) = \beta\exp\{\boldsymbol{\alpha}^{(0)'}\mathbf{x}\} + \sum_{j=1}^{\ell} \mathbf{c}_j \exp\{\boldsymbol{\alpha}^{(j)'}\mathbf{x}\}, \right.$$

$$\left. g(\mathbf{x}) \text{ an AM/GM-exponential} \right\}$$

$$(2.3) \quad = \left\{ (\mathbf{c},\beta) \in \mathbb{R}_+^\ell \times \mathbb{R} \mid \exists \boldsymbol{\nu} \in \mathbb{R}_+^\ell \text{ s.t. } D(\boldsymbol{\nu},e\mathbf{c}) \le \beta, \sum_{j=1}^{\ell} \boldsymbol{\alpha}^{(j)}\boldsymbol{\nu}_j = (\mathbf{1}'\boldsymbol{\nu})\boldsymbol{\alpha}^{(0)} \right\}.$$

Based on this representation, one can see that the set $\mathcal{C}_{\mathrm{AGE}}(\boldsymbol{\alpha}^{(1)},\ldots,\boldsymbol{\alpha}^{(\ell)};\boldsymbol{\alpha}^{(0)})$ is a convex cone. This parametrization plays a crucial role in the next subsection and beyond as we describe computationally efficient methods based on relative entropy optimization (*jointly* over both arguments) to obtain bounds for SPs.

2.2. SAGE decomposition of a general signomial. If a signomial consists of more than one negative term, a natural sufficient condition for certifying nonnegativity is to express the signomial as a sum of AM/GM-exponentials as follows.

DEFINITION 2.3. *Given a finite collection of vectors $\mathcal{M} \subset \mathbb{R}^n$, the set of* sum-of-AM/GM-exponentials *with respect to $\mathcal{M}$ is defined as*

$$\mathrm{SAGE}(\mathcal{M}) = \left\{ f \mid f = \sum_{i=1}^m f_i, \ f_i \text{ an AM/GM-exponential with exponents in } \mathcal{M} \right\}.$$

A signomial $f \in \mathrm{SAGE}(\mathcal{M})$ is called a SAGE *function* with respect to $\mathcal{M}$, and a *decomposition of f as a sum of AM/GM-exponentials is called a* SAGE *decomposition.*

As each term in a SAGE decomposition is globally nonnegative, it is easily seen that SAGE functions are globally nonnegative signomials.

Example 2.2. Consider a signomial $f(\mathbf{x}) = \sum_i p_i(\mathbf{x})g_i(\mathbf{x})$, where each $p_i(\mathbf{x})$ is a posynomial and each $g_i(\mathbf{x})$ is an AM/GM-exponential. If each $p_i(\mathbf{x}), g_i(\mathbf{x})$ is defined with respect to exponents $\{\boldsymbol{\alpha}^{(j)}\}_{j=1}^\ell$, one can check that f is a SAGE function with respect to exponents $E_2(\{\boldsymbol{\alpha}^{(j)}\}_{j=1}^\ell)$.

Example 2.3. As another example, consider the parametrized family of signomials $f_{a,b}(\mathbf{x}_1,\mathbf{x}_2) = \exp\{\mathbf{x}_1\}+\exp\{\mathbf{x}_2\}-a\exp\{\delta\mathbf{x}_1+(1-\delta)\mathbf{x}_2\}-b\exp\{(1-\delta)\mathbf{x}_1+\delta\mathbf{x}_2\}$, where we set $\delta = \pi/4$. Figure 1 shows a plot of the values of $(a,b) \in \mathbb{R}_+^2$ for which $f_{a,b}(\mathbf{x}_1,\mathbf{x}_2)$ is a SAGE function with respect to the exponents $\{(1,0)',(0,1)',(\delta,1-\delta)',(1-\delta,\delta)'\}$.

Example 2.4. Reznick considered (homogenous) polynomials with at most one negative coefficient that can be certified as being globally nonnegative via the AM/GM-inequality and denoted such polynomials as agiforms [24]. Examples include the Motzkin form $p(\mathbf{y}_1,\mathbf{y}_2,\mathbf{y}_3) = \mathbf{y}_1^2\mathbf{y}_2^4 + \mathbf{y}_1^4\mathbf{y}_2^2 + \mathbf{y}_3^6 - 3\mathbf{y}_1^2\mathbf{y}_2^2\mathbf{y}_3^2$ as well as forms due to Choi and Lam [7] such as $q(\mathbf{y}_1,\mathbf{y}_2,\mathbf{y}_3) = \mathbf{y}_1^4\mathbf{y}_2^2 + \mathbf{y}_2^4\mathbf{y}_3^2 + \mathbf{y}_3^4\mathbf{y}_1^2 - 3\mathbf{y}_1^2\mathbf{y}_2^2\mathbf{y}_3^2$ and
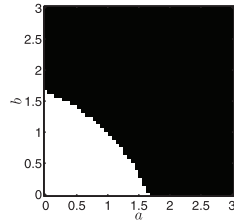
FIG. 1. *The locations in white denote those values of $(a, b) \in \mathbb{R}_+^2$ for which the signomial $f_{a,b}(\mathbf{x}_1, \mathbf{x}_2) = \exp\{\mathbf{x}_1\} + \exp\{\mathbf{x}_2\} - a \exp\{\delta \mathbf{x}_1 + (1-\delta)\mathbf{x}_2\} - b \exp\{(1-\delta)\mathbf{x}_1 + \delta \mathbf{x}_2\}$ with $\delta = \pi/4$ is a SAGE function. The intercept is around 1.682.*

$r(\mathbf{y}_1, \mathbf{y}_2, \mathbf{y}_3, \mathbf{y}_4) = \mathbf{y}_1^4 + \mathbf{y}_2^2 \mathbf{y}_3^2 + \mathbf{y}_2^2 \mathbf{y}_4^2 + \mathbf{y}_3^2 \mathbf{y}_4^2 - 4\mathbf{y}_1 \mathbf{y}_2 \mathbf{y}_3 \mathbf{y}_4$. None of these nonnegative polynomials are sum-of-squares polynomials. However, after applying the transformation $\exp\{\mathbf{x}_i\} \leftarrow \mathbf{y}_i$, the resulting signomials are SAGE (in fact, they are just AM/GM-exponentials). This provides a proof of the nonnegativity of the polynomials p, q, r over the orthant (as the transformation $\exp\{\mathbf{x}_i\} \leftarrow \mathbf{y}_i$ is only valid for $\mathbf{y}_i > 0$); however, due to their special structure, the polynomials p, q, r being nonnegative over the orthant is equivalent to p, q, r being globally nonnegative. Thus, the notion of agiforms as introduced by Reznick may essentially be viewed as polynomial reformulations of AM/GM-exponentials with positive integer exponents.

Although SAGE functions are globally nonnegative, not all nonnegative signomials are decomposable as SAGE functions. The following simple example illustrates the point that SAGE decomposability is only a sufficient condition for nonnegativity.

Example 2.5. Consider the signomial $f(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3) = (\exp\{\mathbf{x}_1\} - \exp\{\mathbf{x}_2\} - \exp\{\mathbf{x}_3\})^2$. This signomial is clearly nonnegative but one can check that it is not a SAGE function.

For $\mathcal{M} = \{\boldsymbol{\alpha}^{(j)}\}_{j=1}^{\ell}$, we have that

$$\text{SAGE}\left(\{\boldsymbol{\alpha}^{(j)}\}_{j=1}^{\ell}\right) = \left\{ f \mid f(\mathbf{x}) = \sum_{i=1}^{\ell} f_i(\mathbf{x}), \text{ each } f_i(\mathbf{x}) \geq 0 \ \forall \mathbf{x} \in \mathbb{R}^n \right.$$

(2.4)

$$\left. f_i(\mathbf{x}) = \sum_{j=1}^{\ell} \tilde{\mathbf{c}}_j^{(i)} \exp\left\{\boldsymbol{\alpha}^{(j)'} \mathbf{x}\right\} \text{ with } \tilde{\mathbf{c}}_{\backslash i}^{(i)} \in \mathbb{R}_+^{\ell-1} \right\}.$$

In order to obtain an efficient description of $\text{SAGE}(\{\boldsymbol{\alpha}^{(j)}\}_{j=1}^{\ell})$ we consider the following set of coefficients (analogous to $\mathcal{C}_{\text{AGE}}(\boldsymbol{\alpha}^{(1)}, \dots, \boldsymbol{\alpha}^{(\ell)}; \boldsymbol{\alpha}^{(0)})$ in (2.3)):

(2.5)

$$\mathcal{C}_{\text{SAGE}}\left(\boldsymbol{\alpha}^{(1)}, \dots, \boldsymbol{\alpha}^{(\ell)}\right) = \left\{ \mathbf{c} \in \mathbb{R}^{\ell} \ \middle| \ \sum_{j=1}^{\ell} \mathbf{c}_j \exp\{\boldsymbol{\alpha}^{(j)'} \mathbf{x}\} \in \text{SAGE}\left(\{\boldsymbol{\alpha}^{(j)}\}_{j=1}^{\ell}\right) \right\}.$$

The next proposition summarizes the relevant properties of SAGE functions by giving an explicit representation of the set $\mathcal{C}_{\text{SAGE}}(\boldsymbol{\alpha}^{(1)}, \dots, \boldsymbol{\alpha}^{(\ell)})$.

PROPOSITION 2.4. *Fix a collection of exponents* $\{\boldsymbol{\alpha}^{(j)}\}_{j=1}^{\ell} \subset \mathbb{R}^n$. *Then we have*

$$\mathcal{C}_{\mathrm{SAGE}}\left(\boldsymbol{\alpha}^{(1)}, \ldots, \boldsymbol{\alpha}^{(\ell)}\right)$$

$$= \left\{ \mathbf{c} \in \mathbb{R}^{\ell} \mid \exists \mathbf{c}^{(j)} \in \mathbb{R}^{\ell}, \ j = 1, \ldots, \ell \text{ s.t. } \mathbf{c} = \sum_{j=1}^{\ell} \mathbf{c}^{(j)}, \right.$$

$$\left. \begin{pmatrix} \mathbf{c}_{\backslash j}^{(j)} \\ \mathbf{c}_j^{(j)} \end{pmatrix} \in \mathcal{C}_{\mathrm{AGE}}\left(\boldsymbol{\alpha}^{(1)}, \ldots, \boldsymbol{\alpha}^{(j-1)}, \boldsymbol{\alpha}^{(j+1)}, \ldots, \boldsymbol{\alpha}^{(\ell)}; \boldsymbol{\alpha}^{(j)}\right) \right\}$$

$$= \left\{ \mathbf{c} \in \mathbb{R}^{\ell} \mid \exists \mathbf{c}^{(j)} \in \mathbb{R}^{\ell}, \boldsymbol{\nu}^{(j)} \in \mathbb{R}^{\ell}, \ j = 1, \ldots, \ell \text{ s.t. } \mathbf{c} = \sum_{j=1}^{\ell} \mathbf{c}^{(j)} \right.$$

$$\sum_{i=1}^{\ell} \boldsymbol{\alpha}^{(i)} \boldsymbol{\nu}_i^{(j)} = \mathbf{0}, \ \boldsymbol{\nu}_j^{(j)} = -\mathbf{1}' \boldsymbol{\nu}_{\backslash j}^{(j)}, \ j = 1, \ldots, \ell$$

(2.6) $$\left. D\left(\boldsymbol{\nu}_{\backslash j}^{(j)}, e \mathbf{c}_{\backslash j}^{(j)}\right) - \mathbf{c}_j^{(j)} \le 0, \ \boldsymbol{\nu}_{\backslash j}^{(j)} \in \mathbb{R}_+^{\ell-1}, \ j = 1, \ldots, \ell \right\}.$$

The set $\mathcal{C}_{\mathrm{SAGE}}(\boldsymbol{\alpha}^{(1)}, \ldots, \boldsymbol{\alpha}^{(\ell)})$ *is a convex cone, and its dual is given by*

(2.7) $$\mathcal{C}_{\mathrm{SAGE}}^{\star}\left(\boldsymbol{\alpha}^{(1)}, \ldots, \boldsymbol{\alpha}^{(\ell)}\right) = \left\{ \mathbf{v} \in \mathbb{R}_+^{\ell} \mid \exists \boldsymbol{\tau}^{(j)} \in \mathbb{R}^n, \ j = 1, \ldots, \ell \text{ s.t.} \right.$$

$$\left. \mathbf{v}_i \log\left(\frac{\mathbf{v}_i}{\mathbf{v}_j}\right) \le \left(\boldsymbol{\alpha}^{(i)} - \boldsymbol{\alpha}^{(j)}\right)' \boldsymbol{\tau}^{(i)} \ \forall i, j \right\}.$$

Proof. For a coefficient vector $\mathbf{c} \in \mathbb{R}^{\ell}$ to belong to $\mathcal{C}_{\mathrm{SAGE}}(\boldsymbol{\alpha}^{(1)}, \ldots, \boldsymbol{\alpha}^{(\ell)})$, we have from (2.5) that $f(\mathbf{x}) = \sum_{j=1}^{\ell} \mathbf{c}_j \exp\{\boldsymbol{\alpha}^{(j)'}\mathbf{x}\}$ must be a SAGE function with respect to the exponents $\{\boldsymbol{\alpha}^{(j)}\}_{j=1}^{\ell}$. The description of $\mathcal{C}_{\mathrm{SAGE}}(\boldsymbol{\alpha}^{(1)}, \ldots, \boldsymbol{\alpha}^{(\ell)})$ in (2.6) then follows directly from Lemma 2.2 and from (2.4).

It is also clear that $\mathcal{C}_{\mathrm{SAGE}}(\boldsymbol{\alpha}^{(1)}, \ldots, \boldsymbol{\alpha}^{(\ell)})$ is a convex cone, because it can be viewed as the projection of a convex cone from (2.6). Specifically, the inequality $D(\boldsymbol{\nu}_{\backslash j}^{(j)}, e\mathbf{c}_{\backslash j}^{(j)}) - \mathbf{c}_j^{(j)} \le 0$ defines a convex cone because the relative entropy function is positively homogenous and convex jointly in both arguments, and all the other constraints define convex cones in the variables $\mathbf{c}, \mathbf{c}^{(j)}, \boldsymbol{\nu}^{(j)}$.

Finally, the description of the dual cone $\mathcal{C}_{\mathrm{SAGE}}^{\star}(\boldsymbol{\alpha}^{(1)}, \ldots, \boldsymbol{\alpha}^{(\ell)})$ follows from a straightforward calculation based on the observation that $\mathcal{C}_{\mathrm{SAGE}}(\boldsymbol{\alpha}^{(1)}, \ldots, \boldsymbol{\alpha}^{(\ell)})$ is a sum of convex cones $\mathcal{C}_{\mathrm{AGE}}(\boldsymbol{\alpha}^{(1)}, \ldots, \boldsymbol{\alpha}^{(j-1)}, \boldsymbol{\alpha}^{(j+1)}, \ldots, \boldsymbol{\alpha}^{(\ell)}; \boldsymbol{\alpha}^{(j)})$ corresponding to AM/GM exponentials; therefore, the dual $\mathcal{C}_{\mathrm{SAGE}}^{\star}(\boldsymbol{\alpha}^{(1)}, \ldots, \boldsymbol{\alpha}^{(\ell)})$ is the intersection of the duals $\mathcal{C}_{\mathrm{AGE}}^{\star}(\boldsymbol{\alpha}^{(1)}, \ldots, \boldsymbol{\alpha}^{(j-1)}, \boldsymbol{\alpha}^{(j+1)}, \ldots, \boldsymbol{\alpha}^{(\ell)}; \boldsymbol{\alpha}^{(j)})$. $\square$

Thus, the set of SAGE functions with respect to a set of exponents can be efficiently characterized based on this proposition. In section 3, we employ this result to develop computationally tractable methods for obtaining lower bounds on signomials. The characterization of the cone $\mathcal{C}_{\mathrm{SAGE}}(\boldsymbol{\alpha}^{(1)}, \ldots, \boldsymbol{\alpha}^{(\ell)})$ in Proposition 2.4 also clarifies an appealing invariance property of SAGE functions as follows.

COROLLARY 2.5. *Fix a collection of exponents* $\{\boldsymbol{\alpha}^{(j)}\}_{j=1}^{\ell} \subset \mathbb{R}^n$. *For any nonsingular matrix* $M \in \mathbb{R}^{n \times n}$, *we have that*

$$\mathcal{C}_{\mathrm{SAGE}}\left(\boldsymbol{\alpha}^{(1)}, \ldots, \boldsymbol{\alpha}^{(\ell)}\right) = \mathcal{C}_{\mathrm{SAGE}}\left(M\boldsymbol{\alpha}^{(1)}, \ldots, M\boldsymbol{\alpha}^{(\ell)}\right).$$

Proof. This result follows by noting that the only appearance of the exponents $\{\boldsymbol{\alpha}^{(j)}\}_{j=1}^{\ell}$ in the description of $\mathcal{C}_{\text{SAGE}}(\boldsymbol{\alpha}^{(1)}, \ldots, \boldsymbol{\alpha}^{(\ell)})$ in (2.6) is in the equation $\sum_{i=1}^{\ell} \boldsymbol{\alpha}^{(i)} \boldsymbol{\nu}_i^{(j)} = \mathbf{0}$. A $\boldsymbol{\nu}^{(j)} \in \mathbb{R}^{\ell}$ satisfies this constraint if and only if the same $\boldsymbol{\nu}^{(j)}$ also satisfies $\sum_{i=1}^{\ell} M \boldsymbol{\alpha}^{(i)} \boldsymbol{\nu}_i^{(j)} = \mathbf{0}$. This concludes the proof. $\qquad \square$

Notice that the nonnegativity of a signomial is preserved under the application of a nonsingular linear transformation simultaneously to all the exponents; more generally, the optimal value of an SP is unchanged under a nonsingular transformation applied simultaneously to all the exponents in the objective and all the constraints. The above invariance property of $\mathcal{C}_{\text{SAGE}}(\boldsymbol{\alpha}^{(1)}, \ldots, \boldsymbol{\alpha}^{(\ell)})$ ensures that our lower bounds for SPs in section 3 based on SAGE decompositions are also similarly invariant under a simultaneous linear transformation applied to all the exponents in the SP.

2.3. Remarks on nonnegative signomials. Consider a signomial $f(\mathbf{x}) = \sum_{j=1}^{\ell} \mathbf{c}_j \exp\{\boldsymbol{\alpha}^{(j)'}\mathbf{x}\}$ with respect to exponents $\{\boldsymbol{\alpha}^{(j)}\}_{j=1}^{\ell}$, and let the coefficients $\mathbf{c}_j$ be nonzero for each $j = 1, \ldots \ell$. If $f(\mathbf{x}) \geq 0$ for all $\mathbf{x} \in \mathbb{R}^n$, then we can make some observations about the coefficients $\mathbf{c}_j$'s based on the structure of the set of exponents. Specifically, suppose $\boldsymbol{\alpha}^{(1)}$ is an extreme point of the convex hull of the exponents $\text{conv}(\{\boldsymbol{\alpha}^{(j)}\}_{j=1}^{\ell})$; we will denote such an exponent as an *extremal exponent*. Then we can conclude from the separation theorem in convex analysis [26] that there exists $\mathbf{u} \in \mathbb{R}^n$ such that $\mathbf{u}'\boldsymbol{\alpha}^{(1)} = 1$ and $\mathbf{u}'\boldsymbol{\alpha}^{(j)} < 1$ for $j = 2, \ldots, \ell$. Consequently, $\exp\{\gamma\boldsymbol{\alpha}^{(1)'}\mathbf{u}\}$ dominates $\exp\{\gamma\boldsymbol{\alpha}^{(j)'}\mathbf{u}\}$ for $j = 2, \ldots, \ell$ as $\gamma \to \infty$. Based on the nonnegativity of $f(\mathbf{x})$, we conclude that $\mathbf{c}_1 > 0$. Thus, a coefficient $\mathbf{c}_j$ must be positive if the corresponding exponent $\boldsymbol{\alpha}^{(j)}$ is an extremal exponent.

These observations imply that the cone of coefficients of AM/GM-exponentials $\mathcal{C}_{\text{AGE}}(\boldsymbol{\alpha}^{(1)}, \ldots, \boldsymbol{\alpha}^{(\ell)}; \boldsymbol{\alpha}^{(0)}) = \mathbb{R}_+^{\ell+1}$ if $\boldsymbol{\alpha}^{(0)} \notin \text{conv}(\{\boldsymbol{\alpha}^{(j)}\}_{j=1}^{\ell})$; this follows from the point that AM/GM-exponentials *exactly* correspond to nonnegative signomials with at most one negative coefficient. Consequently, one can exclude such posynomials in the description of SAGE functions in (2.4), which also leads to simpler representations in Proposition 2.4. These modifications may provide computational benefits if one wishes to certify nonnegativity of a signomial with the property that many of the exponents are extremal exponents.

3. Hierarchies of relative entropy relaxations for SPs. In this section we describe hierarchies of convex relaxations to obtain lower bounds for SPs. Our development parallels the literature on convex relaxations for polynomial optimization [15, 19, 20], but with an important distinction—our approach for SPs is based on relative entropy optimization derived via SAGE decompositions, while relaxations for polynomial optimization are based on SDPs derived via sum-of-squares decompositions. We give theoretical support for our hierarchies in section 4.1 by showing that they provide convergent sequences of lower bounds to the optimal value for broad classes of SPs. Finally, we present a dual perspective of our methods, which leads to a technique for recognizing when our lower bounds are tight and for extracting optimal solutions.

3.1. A hierarchy for unconstrained SPs. Globally minimizing a signomial $f(\mathbf{x})$ is equivalent to maximizing a lower bound of $f(\mathbf{x})$:

$$(3.1) \qquad f_\star = \inf_{\mathbf{x} \in \mathbb{R}^n} f(\mathbf{x}) = \sup_{\gamma \in \mathbb{R}} \gamma \text{ s.t. } f(\mathbf{x}) - \gamma \geq 0 \; \forall \mathbf{x} \in \mathbb{R}^n.$$

Replacing the global nonnegativity condition on $f(\mathbf{x}) - \gamma$ by a sufficient condition for nonnegativity gives lower bounds on the optimal value $f_\star$. SAGE functions provide

a natural sufficient condition for global nonnegativity leading to a computationally tractable approach for obtaining lower bounds, and we describe this relaxation next.

To state things concretely, fix a set of exponents $\{\boldsymbol{\alpha}^{(j)}\}_{j=1}^{\ell} \subset \mathbb{R}^n$ with $\boldsymbol{\alpha}^{(1)} = \mathbf{0}$ and consider a signomial $f(\mathbf{x}) = \sum_{j=1}^{\ell} \mathbf{c}_j \exp\{\boldsymbol{\alpha}^{(j)'}\mathbf{x}\}$ defined with respect to these exponents. (If f contains no constant term, then set the corresponding coefficient $\mathbf{c}_1 = 0$.) In order to produce a lower bound on the global minimum $f_\star$ of f, we consider the following convex relaxation based on SAGE decompositions:

$$(3.2) \qquad f_{\text{SAGE}} = \sup_{\gamma \in \mathbb{R}} \gamma \quad \text{s.t.} \ f(\mathbf{x}) - \gamma \in \text{SAGE}\left(\{\boldsymbol{\alpha}^{(j)}\}_{j=1}^{\ell}\right).$$

Based on the preceding discussion, it is clear that $f_{\text{SAGE}} \leq f_\star$. Furthermore, computing f_{SAGE} can be reformulated as follows via Proposition 2.4:

$$(3.3) \qquad f_{\text{SAGE}} = \sup_{\gamma \in \mathbb{R}} \gamma \quad \text{s.t.} \ (\mathbf{c}_1 - \gamma, \mathbf{c}_2, \ldots, \mathbf{c}_\ell) \in \mathcal{C}_{\text{SAGE}}(\mathbf{0}, \boldsymbol{\alpha}^{(2)}, \ldots, \boldsymbol{\alpha}^{(\ell)}),$$

where $\mathcal{C}_{\text{SAGE}}(\mathbf{0}, \boldsymbol{\alpha}^{(2)}, \ldots, \boldsymbol{\alpha}^{(\ell)})$ can be described as in (2.6). Thus, the bound f_{SAGE} is the optimal value of a tractable convex program with linear and joint relative entropy constraints.

If $f(\mathbf{x})$ is a posynomial, then $f_{\text{SAGE}} = f_\star$; this follows from the fact that $f(\mathbf{x}) - \gamma$ is a SAGE function if and only if $f(\mathbf{x}) - \gamma \geq 0$ for all $\mathbf{x} \in \mathbb{R}^n$ because $f(\mathbf{x}) - \gamma$ has all but one coefficient being positive. Consequently, the SAGE relaxation is exact for unconstrained GPs, which is reassuring as tractable problem instances (GPs are convex optimization problems that can be solved efficiently) can be computed effectively in our framework. More generally, we illustrate the utility of SAGE relaxations for minimizing signomials with multiple negative coefficients in the following examples.

Example 3.1. We consider signomials in three variables with seven terms of the form $f(\mathbf{x}) = \sum_{j=1}^{7} \mathbf{c}_j \exp\{\boldsymbol{\alpha}^{(j)'}\mathbf{x}\}$ with $\mathbf{x} \in \mathbb{R}^3$ and the following parameters fixed: $\mathbf{c}_1 = 0$, $\mathbf{c}_2 = \mathbf{c}_3 = \mathbf{c}_4 = 10$, $\boldsymbol{\alpha}^{(1)} = (0,0,0)'$, $\boldsymbol{\alpha}^{(2)} = (10.2, 0, 0)'$, $\boldsymbol{\alpha}^{(3)} = (0, 9.8, 0)'$, and $\boldsymbol{\alpha}^{(4)} = (0, 0, 8.2)'$. The exponents $\boldsymbol{\alpha}^{(5)}, \boldsymbol{\alpha}^{(6)}, \boldsymbol{\alpha}^{(7)} \in \mathbb{R}^3$ are chosen to be random vectors with entries distributed uniformly in $[0, 3]$, and the coefficients $\mathbf{c}_5, \mathbf{c}_6, \mathbf{c}_7$ are chosen to be random Gaussians with mean 0 and standard deviation 10. We employ this construction—relatively large exponents $\boldsymbol{\alpha}^{(2)}, \boldsymbol{\alpha}^{(3)}, \boldsymbol{\alpha}^{(4)}$ in comparison to $\boldsymbol{\alpha}^{(5)}, \boldsymbol{\alpha}^{(6)}, \boldsymbol{\alpha}^{(7)}$, and the corresponding positive coefficients $\mathbf{c}_2 = \mathbf{c}_3 = \mathbf{c}_4 = 10$—to obtain signomials that are bounded below. An example of a signomial generated in this manner is

$$(3.4) \qquad \begin{aligned} f(\mathbf{x}) = {} & 10\exp\{10.2\mathbf{x}_1\} + 10\exp\{9.8\mathbf{x}_2\} + 10\exp\{8.2\mathbf{x}_3\} \\ & - 14.6794\exp\{1.5089\mathbf{x}_1 + 1.0981\mathbf{x}_2 + 1.3419\mathbf{x}_3\} \\ & - 7.8601\exp\{1.0857\mathbf{x}_1 + 1.9069\mathbf{x}_2 + 1.6192\mathbf{x}_3\} \\ & + 8.7838\exp\{1.0459\mathbf{x}_1 + 0.0492\mathbf{x}_2 + 1.6245\mathbf{x}_3\}. \end{aligned}$$

The SAGE relaxation (3.3) applied to $f(\mathbf{x})$ gives the lower bound $f_{\text{SAGE}} \approx -0.9747$. By applying a technique presented in section 3.4, we obtain the point $\mathbf{x}^\star = (-0.3020, -0.2586, -0.4010)'$ with $f(\mathbf{x}^\star) \approx -0.9747$. Consequently, the lower bound f_{SAGE} is tight in this case. As another instance of the construction described here, consider

the randomly generated signomial

$$
\begin{aligned}
f(\mathbf{x}) = {} & 10 \exp\{10.2\mathbf{x}_1\} + 10 \exp\{9.8\mathbf{x}_2\} + 10 \exp\{8.2\mathbf{x}_3\} \\
& + 7.5907 \exp\{1.9864\mathbf{x}_1 + 0.2010\mathbf{x}_2 + 1.0855\mathbf{x}_3\} \\
& - 10.9888 \exp\{2.8242\mathbf{x}_1 + 1.9355\mathbf{x}_2 + 2.0503\mathbf{x}_3\} \\
& - 13.9164 \exp\{0.1828\mathbf{x}_1 + 2.7772\mathbf{x}_2 + 1.9001\mathbf{x}_3\}.
\end{aligned}
$$

(3.5)

In this case the SAGE relaxation (3.3) gives a lower bound of -1.426, and the technique from section 3.4 gives an upper bound of -0.739. This suggests that the SAGE lower bound may not be tight, and we see in what follows that this is indeed the case (Example 3.2). More generally, we generated 80 random signomials according to the above description, and the SAGE relaxation was tight in 63% of the cases, while there was a gap in the remaining cases.

These examples demonstrate that there are several cases in which the lower bound f_{SAGE} equals the global minimum $f_\star$. There are also situations in which the lower bound f_{SAGE} does not equal $f_\star$, which is not unexpected as global minimization of signomials is in general a computationally intractable problem. Motivated by such cases, we next describe a methodology for obtaining a sequence of increasingly tighter lower bounds based on improved sufficient conditions for the global nonnegativity of the signomial $f(\mathbf{x}) - \gamma$. These improved lower bounds require the solution of successively larger-sized relative entropy optimization problems.

Specifically, for a set of exponents $\{\boldsymbol{\alpha}^{(j)}\}_{j=1}^\ell \subset \mathbb{R}^n$ with $\boldsymbol{\alpha}^{(1)} = \mathbf{0}$ and a signomial $f(\mathbf{x}) = \sum_{j=1}^\ell \mathbf{c}_j \exp\{\boldsymbol{\alpha}^{(j)'}\mathbf{x}\}$ defined with respect to these exponents, we replace the SAGE condition in (3.2) by a stronger sufficient condition for the global nonnegativity of $f(\mathbf{x}) - \gamma$, leading to the following lower bound for each $p \in \mathbb{Z}_+$:
(3.6)

$$
f_{\mathrm{SAGE}}^{(p)} = \sup_{\gamma \in \mathbb{R}} \ \gamma \quad \text{s.t.} \quad \left(\sum_{j=1}^\ell \exp\{\boldsymbol{\alpha}^{(j)'}\mathbf{x}\} \right)^p [f(\mathbf{x}) - \gamma] \in \mathrm{SAGE}(E_{p+1}(\{\boldsymbol{\alpha}^{(j)}\}_{j=1}^\ell)).
$$

The multiplier term $(\sum_{j=1}^\ell \exp\{\boldsymbol{\alpha}^{(j)'}\mathbf{x}\})^p$ is globally nonnegative for all $\mathbf{x} \in \mathbb{R}^n$, and consequently the constraint above is a sufficient condition for the global nonnegativity of $f(\mathbf{x}) - \gamma$. This in turn implies that $f_{\mathrm{SAGE}}^{(p)} \le f_\star$ for all $p \in \mathbb{Z}_+$. One can check that the bound f_{SAGE} of (3.3) is equal to $f_{\mathrm{SAGE}}^{(0)}$ in (3.6). Computing $f_{\mathrm{SAGE}}^{(p)}$ for each p can be recast as a relative entropy optimization problem by appealing to Proposition 2.4 and to the fact that the coefficients of the signomial $(\sum_{j=1}^\ell \exp\{\boldsymbol{\alpha}^{(j)'}\mathbf{x}\})^p [f(\mathbf{x}) - \gamma]$ are linear functionals of γ. For example, for the case $p = 1$, the signomial $(\sum_{\tilde{j}=1}^\ell \exp\{\boldsymbol{\alpha}^{(\tilde{j})'}\mathbf{x}\})[f(\mathbf{x}) - \gamma] \in \mathrm{SAGE}(E_2(\{\boldsymbol{\alpha}^{(j)}\}_{j=1}^\ell))$ is equivalent to $\sum_{j,\tilde{j}=1}^\ell \mathbf{c}_{j,\tilde{j}} \exp\{[\boldsymbol{\alpha}^{(j)} + \boldsymbol{\alpha}^{(\tilde{j})}]'\mathbf{x}\} \in \mathrm{SAGE}(E_2(\{\boldsymbol{\alpha}^{(j)}\}_{j=1}^\ell))$ with $\mathbf{c}_{j,\tilde{j}} = \mathbf{c}_1 - \gamma$ whenever $\tilde{j} = 1$ and $\mathbf{c}_{j,\tilde{j}} = \mathbf{c}_j$ otherwise; the latter SAGE condition can be reformulated using linear and relative entropy constraints on $\mathbf{c}_{j,\tilde{j}}$, and in turn on γ, via Proposition 2.4. The next result states that $f_{\mathrm{SAGE}}^{(p)}$ is a nondecreasing function of p; this improvement carries a computational cost as the size of the associated relative entropy program grows with p.

LEMMA 3.1. *Fix a collection of exponents $\{\boldsymbol{\alpha}^{(j)}\}_{j=1}^\ell \subset \mathbb{R}^n$ with $\boldsymbol{\alpha}^{(1)} = \mathbf{0}$, and consider a signomial $f(\mathbf{x})$ defined with respect to these exponents. We have that $f_{\mathrm{SAGE}}^{(p)} \le f_{\mathrm{SAGE}}^{(p+1)}$ for all $p \in \mathbb{Z}_+$.*

Proof. For a fixed value of γ, let $g_p(\mathbf{x}) = (\sum_{j=1}^{\ell} \exp\{\boldsymbol{\alpha}^{(j)'}\mathbf{x}\})^p [f(\mathbf{x}) - \gamma]$ and let $g_{p+1}(\mathbf{x}) = (\sum_{j=1}^{\ell} \exp\{\boldsymbol{\alpha}^{(j)'}\mathbf{x}\})^{p+1}[f(\mathbf{x}) - \gamma]$. We observe that

$$g_{p+1}(\mathbf{x}) = \sum_{j=1}^{\ell} \exp\{\boldsymbol{\alpha}^{(j)'}\mathbf{x}\} g_p(\mathbf{x})$$

and that $g_p(\mathbf{x}) \in \mathrm{SAGE}(E_{p+1}(\{\boldsymbol{\alpha}^{(j)}\}_{j=1}^{\ell}))$ implies that $\exp\{\boldsymbol{\alpha}^{(j)'}\mathbf{x}\} g_p(\mathbf{x})$ belongs to $\mathrm{SAGE}(E_{p+2}(\{\boldsymbol{\alpha}^{(j)}\}_{j=1}^{\ell}))$. Consequently, $g_p(\mathbf{x}) \in \mathrm{SAGE}(E_{p+1}(\{\boldsymbol{\alpha}^{(j)}\}_{j=1}^{\ell}))$ implies that $g_{p+1}(\mathbf{x}) \in \mathrm{SAGE}(E_{p+2}(\{\boldsymbol{\alpha}^{(j)}\}_{j=1}^{\ell}))$ as $\mathrm{SAGE}(E_{p+2}(\{\boldsymbol{\alpha}^{(j)}\}_{j=1}^{\ell}))$ is closed under addition. Since this implication is true for each fixed γ, we have the desired result. $\square$

In section 4.1, we prove that the sequence $\{f_{\mathrm{SAGE}}^{(p)}\}$ converges to $f_\star$ as $p \to \infty$ for suitable families of signomials. Consequently, the methodology described in this section leads to a convergent sequence of lower bounds, computed by solving increasingly larger relative entropy optimization problems.

Example 3.2. Consider the signomial $f(\mathbf{x})$ from (3.5) in Example 3.1. The lower bound $f_{\mathrm{SAGE}}^{(0)} = -1.426$, and this is not tight. The next level of the hierarchy presented above leads to the improved bound $f_{\mathrm{SAGE}}^{(1)} = -1.395$.

3.2. From unconstrained SPs to constrained SPs: Algebraic certificates of compactness. Our approach to developing relaxations for constrained SPs is qualitatively different from that in the unconstrained case. Specifically, the hierarchy of relaxations that we develop for constrained SPs in section 3.3 does not directly specialize to the hierarchy developed for unconstrained SPs in section 3.1. The reason for this discrepancy is that our development in the constrained setting is based on the assumption that the constraint set is a compact set. Concretely, fix a collection of exponents $\{\boldsymbol{\alpha}^{(j)}\}_{j=1}^{\ell} \subset \mathbb{R}^n$, and let $f(\mathbf{x})$ and $C = \{g_i(\mathbf{x})\}_{i=1}^{m}$ be signomials with respect to these exponents. Consider the constrained SP

$$(3.7) \qquad f_\star = \inf_{\mathbf{x} \in \mathbb{R}^n} \ f(\mathbf{x}) \quad \text{s.t. } g_i(\mathbf{x}) \geq 0, \ i = 1, \dots, m.$$

The hierarchy we propose in section 3.3 is shown to be complete in section 4.1 (see Theorem 4.2) under the premise that the constraint set $\mathcal{K}_C = \{\mathbf{x} \in \mathbb{R}^n \mid g(\mathbf{x}) \geq 0$ for all $g \in C\}$ is compact. Crucially, the proof of completeness relies on the assumption that the compactness of $\mathcal{K}_C$ is explicitly enforced via inequalities of the form $U \geq \exp\{\boldsymbol{\alpha}^{(j)'}\mathbf{x}\} \geq L$, $j = 1, \dots, \ell$ (for appropriate $U, L \in \mathbb{R}_{++}$) in the list of constraints $C = \{g_i(\mathbf{x})\}_{i=1}^{m}$. Such explicit enforcements of compactness enable us to appeal to representation theorems from real algebraic geometry that require appropriate Archimedeanity assumptions (see the survey [18]), a point that is also significant in the development of hierarchies of SDP relaxations based on sum-of-squares decompositions for polynomial optimization problems [15, 19, 20, 23, 27].

As such a natural "unconstrained" counterpart of the constrained case is a setting in which we wish to minimize a signomial subject to a constraint set of the form $\{\mathbf{x} \in \mathbb{R}^n \mid U \geq \exp\{\boldsymbol{\alpha}^{(j)'}\mathbf{x}\} \geq L, \ j = 1, \dots, \ell\}$ for some fixed $U, L \in \mathbb{R}_{++}$. In particular, if we have some additional knowledge that a global minimizer of a signomial belongs to a set of this form for some known U, L, then the methods developed in section 3.3 provide an alternative hierarchy of convergent lower bounds to those in section 3.1.

3.3. A hierarchy for constrained SPs. Here we describe a hierarchy of relative entropy relaxations based on SAGE decompositions for constrained SPs. In analogy with the unconstrained case, we can recast (3.7) as

$$(3.8) \qquad f_\star = \sup_{\gamma \in \mathbb{R}} \gamma \quad \text{s.t. } f(\mathbf{x}) - \gamma \geq 0 \ \forall \mathbf{x} \in \mathcal{K}_C.$$

As the original constrained SP is computationally intractable to solve in general, certifying nonnegativity of a signomial over a constraint set defined by signomials is also intractable. In order to obtain lower bounds on $f_\star$ that are tractable to compute, our approach is again to proceed by employing efficiently computable sufficient conditions for certifying nonnegativity of a signomial $f(\mathbf{x}) - \gamma$ over a constraint set $\mathcal{K}_C$ defined by signomial inequalities.

A basic approach based on weak duality for bounding (3.8) is to replace the constraint by the following sufficient condition for nonnegativity:

$$f_{WD} = \sup_{\gamma \in \mathbb{R}, \boldsymbol{\mu} \in \mathbb{R}^m_+} \gamma \quad \text{s.t. } f(\mathbf{x}) - \gamma - \sum_{i=1}^{m} \boldsymbol{\mu}_i g_i(\mathbf{x}) \geq 0 \ \forall \mathbf{x} \in \mathbb{R}^n.$$

As the nonnegativity condition here is a sufficient condition for the constraint in (3.8), we note that $f_{WD} \leq f_\star$; this is the standard argument underlying weak duality. However, there are two difficulties with this approach. The first is that the constraint still involves the computationally intractable problem of checking nonnegativity of an arbitrary signomial. The second difficulty, independent of the first one, is that the underlying SP is nonconvex, and therefore we do not expect strong duality to hold; as such, even if one could compute f_{WD} it is in general the case that $f_{WD} < f_\star$. To remedy the first difficulty, we replace the nonnegativity condition by one based on SAGE decomposability. To address the second challenge we add valid inequalities that are consequences of the original set of constraints; although this idea of adding appropriate redundant constraints is an old one, it has been employed to particularly powerful effect in polynomial optimization problems [15, 19, 20].

Formally, consider the set of signomials defined as products of the original set of constraint functions $C = \{g_i(\mathbf{x})\}_{i=1}^m$ in (3.8):

$$(3.9) \qquad R_q(C) = \left\{ \prod_{k=1}^{q} h_k \mid h_k \in \{1\} \cup C \right\}.$$

Here 1 represents the signomial that is identically equal to one for all $\mathbf{x} \in \mathbb{R}^n$. Based on this definition, we consider the following relaxation of (3.8):
(3.10)

$$f_{\text{SAGE}}^{(p,q)} = \sup_{\substack{\gamma \in \mathbb{R}, \\ s_h(\mathbf{x}) \in \text{SAGE}(E_p(\{\boldsymbol{\alpha}^{(j)}\}_{j=1}^\ell))}} \gamma$$

$$\text{s.t.} \qquad f(\mathbf{x}) - \gamma - \sum_{h(\mathbf{x}) \in R_q(C)} s_h(\mathbf{x}) h(\mathbf{x})$$

$$\in \text{SAGE}(E_{p+q}(\{\boldsymbol{\alpha}^{(j)}\}_{j=1}^\ell)).$$

For a fixed q, the number of terms in the set $R_q(C)$ is at most $(\ell+1)^q$, and consequently the sum consists of a finite number of terms. Moreover, computing $f_{\text{SAGE}}^{(p,q)}$ can be recast as a relative entropy optimization problem via Proposition 2.4 as the coefficients of

the signomial $f(\mathbf{x}) - \gamma - \sum_{h \in R_q(C)} s_h(\mathbf{x})h(\mathbf{x})$ are linear functionals of γ and the coefficients of the SAGE functions $s_h(\mathbf{x})$. Note that for $h \in R_q(C)$ and $s_h(\mathbf{x}) \in$ SAGE(E_p), the signomial $s_h(\mathbf{x})h(\mathbf{x})$ is nonnegative on the set $\mathcal{K}_C$, thus corresponding to a valid constraint. Hence, for a fixed value of γ and fixed SAGE functions $s_h(\mathbf{x})$, the condition $f(\mathbf{x}) - \gamma - \sum_{h \in R_q(C)} s_h(\mathbf{x})h(\mathbf{x}) \in$ SAGE(E_{p+q}) is a sufficient condition for the nonnegativity of $f(\mathbf{x}) - \gamma$ on the constraint set $\mathcal{K}_C$. Consequently we have that $f_{\text{SAGE}}^{(p,q)} \leq f_\star$ for all $p, q \in \mathbb{Z}_+$.

The next result records the fact that $f_{\text{SAGE}}^{(p,q)}$ provides increasingly tighter bounds as p, q grow larger. This improvement comes at the expense of solving larger relative entropy optimization problems for bigger values of p, q.

LEMMA 3.2. *Fix a set of exponents* $\{\boldsymbol{\alpha}^{(j)}\}_{j=1}^{\ell} \subset \mathbb{R}^n$ *with* $\boldsymbol{\alpha}^{(1)} = \mathbf{0}$, *and let* $f(\mathbf{x})$ *and* $C = \{g_i(\mathbf{x})\}_{i=1}^{m}$ *be signomials with respect to these exponents. Then we have that* $f_{\text{SAGE}}^{(p,q)} \leq f_{\text{SAGE}}^{(p',q')}$ *for* $p \leq p', q \leq q'$ *with* $p, p', q, q' \in \mathbb{Z}_+$.

Proof. This result follows from the observations that SAGE$(E_p) \subset$ SAGE$(E_{p'})$, $R_q(C) \subset R_{q'}(C)$, and SAGE$(E_{p+q}) \subset$ SAGE$(E_{p'+q'})$ if $p \leq p', q \leq q'$. $\qquad\square$

In section 4.1 we show that for SPs with compact constraint sets, $f_{\text{SAGE}}^{(p,q)} \to f_\star$ as $p, q \to \infty$. These results demonstrate the value in the redundant constraints employed in the bounds $f_{\text{SAGE}}^{(p,q)}$. We also note that GPs are solved exactly at the first level of the hierarchy, i.e., the lower bound $f_{\text{SAGE}}^{(0,1)}$ equals the global minimum for GPs.

Example 3.3. As our first example, consider a constrained SP in which the objective to be minimized is the signomial (3.4) from Example 3.1, and the constraint set is convex as follows:

$$\{\mathbf{x} \in \mathbb{R}^3 \mid 8 \exp\{10.2\mathbf{x}_1\} + 8 \exp\{9.8\mathbf{x}_2\} + 8 \exp\{8.2\mathbf{x}_3\}$$
$$+ 6.4 \exp\{1.0857\mathbf{x}_1 + 1.9069\mathbf{x}_2 + 1.6192\mathbf{x}_3\} \leq 1\}.$$

Notice that the exponents in $f(\mathbf{x})$ in (3.4) and ones in the constraint here are common. Although this constraint set is a convex set, this problem is not a GP because the objective function is a signomial consisting of two negative coefficients. Applying the first level of the SAGE hierarchy described in this subsection gives the lower bound $f_{\text{SAGE}}^{(0,1)} = -0.6147$. This bound is tight, and it is achieved at the feasible point $\mathbf{x}^\star = (-0.4312, -0.3823, -0.6504)'$.

Example 3.4. As our next example, consider a constrained SP in which the objective to be minimized is again the signomial (3.4) from Example 3.1, and the constraint is given by following signomial:

(3.11)
$$\begin{aligned} g(\mathbf{x}) = &- 8 \exp\{10.2\mathbf{x}_1\} - 8 \exp\{9.8\mathbf{x}_2\} - 8 \exp\{8.2\mathbf{x}_3\} \\ &+ 0.7410 \exp\{1.5089\mathbf{x}_1 + 1.0981\mathbf{x}_2 + 1.3419\mathbf{x}_3\} \\ &- 0.4492 \exp\{1.0857\mathbf{x}_1 + 1.9069\mathbf{x}_2 + 1.6192\mathbf{x}_3\} \\ &+ 1.4240 \exp\{1.0459\mathbf{x}_1 + 0.0492\mathbf{x}_2 + 1.6245\mathbf{x}_3\}. \end{aligned}$$

This constraint was chosen in a similar manner to the signomials in Example 3.1. Keeping the exponents fixed to the same values as those in (3.4), the first three coefficients are chosen to be equal to -8 and the last three are chosen to be random Gaussians with mean 0 and standard deviation 1. Thus, we wish to minimize $f(\mathbf{x})$ from (3.4) subject to the constraint that $g(\mathbf{x}) \geq 0$. The first level of the hierarchy for constrained SPs gives us the lower bound $f_{\text{SAGE}}^{(0,1)} = -0.7372$. This bound is also tight, and it is achieved at $\mathbf{x}^\star = (0.0073, 0.0065, 0.0130)'$.

3.4. Dual perspectives and extracting optimal solutions. The hierarchies of relaxations described in the previous subsections have an appealing dual perspective. Focusing on the unconstrained case for simplicity, the problem of minimizing a signomial $f(\mathbf{x}) = \sum_{j=1}^{\ell} \mathbf{c}_j \exp\{\boldsymbol{\alpha}^{(j)'}\mathbf{x}\}$ with $\boldsymbol{\alpha}^{(1)} = \mathbf{0}$ can be recast as

$$f_\star = \inf_{\mathbf{v} \in \mathbb{R}^\ell} \mathbf{c}'\mathbf{v} \quad \text{s.t. } \mathbf{v} \in \{(1, \exp\{\boldsymbol{\alpha}^{(2)'}\mathbf{x}\}, \ldots, \exp\{\boldsymbol{\alpha}^{(\ell)'}\mathbf{x}\}) \mid \mathbf{x} \in \mathbb{R}^n\}.$$

As the objective is linear, this problem can equivalently be expressed as a convex program by convexifying the constraint set to $\mathrm{conv}\{(1, \exp\{\boldsymbol{\alpha}^{(2)'}\mathbf{x}\}, \ldots, \exp\{\boldsymbol{\alpha}^{(\ell)'}\mathbf{x}\}) \mid \mathbf{x} \in \mathbb{R}^n\}$. However, checking membership in $\mathrm{conv}\{(1, \exp\{\boldsymbol{\alpha}^{(2)'}\mathbf{x}\}, \ldots, \exp\{\boldsymbol{\alpha}^{(\ell)'}\mathbf{x}\}) \mid \mathbf{x} \in \mathbb{R}^n\}$ is in general an intractable problem.

The methods described in section 3.1 can be viewed as providing a sequence of tractable *outer convex approximations* to the set $\{(1, \exp\{\boldsymbol{\alpha}^{(2)'}\mathbf{x}\}, \ldots, \exp\{\boldsymbol{\alpha}^{(\ell)'}\mathbf{x}\}) \mid \mathbf{x} \in \mathbb{R}^n\}$. Specifically, the dual corresponding to the SAGE lower bound $f_{\mathrm{SAGE}}^{(p)}$ (3.6) is

$$(3.12) \qquad f_{\mathrm{SAGE}}^{(p)} = \inf_{\mathbf{v} \in \mathbb{R}^\ell} \mathbf{c}'\mathbf{v} \quad \text{s.t. } \mathbf{v} \in \mathcal{W}_p(\mathbf{0}, \boldsymbol{\alpha}^{(2)}, \ldots, \boldsymbol{\alpha}^{(\ell)}),$$

where each $\mathcal{W}_p(\mathbf{0}, \boldsymbol{\alpha}^{(2)}, \ldots, \boldsymbol{\alpha}^{(\ell)}) \subset \mathbb{R}^\ell$ is a convex set that contains $\{(1, \exp\{\boldsymbol{\alpha}^{(2)'}\mathbf{x}\}, \ldots, \exp\{\boldsymbol{\alpha}^{(\ell)'}\mathbf{x}\}) \mid \mathbf{x} \in \mathbb{R}^n\}$. The set $\mathcal{W}_0(\mathbf{0}, \boldsymbol{\alpha}^{(2)}, \ldots, \boldsymbol{\alpha}^{(\ell)})$ corresponding to $f_{\mathrm{SAGE}}^{(0)}$ (3.6) is

$$(3.13) \ \ \mathcal{W}_0(\mathbf{0}, \boldsymbol{\alpha}^{(2)}, \ldots, \boldsymbol{\alpha}^{(\ell)}) = \left\{\mathbf{v} \in \mathbb{R}_+^\ell \mid \mathbf{v}_1 = 1 \text{ and } \mathbf{v} \in \mathcal{C}_{\mathrm{SAGE}}^\star(\mathbf{0}, \boldsymbol{\alpha}^{(2)}, \ldots, \boldsymbol{\alpha}^{(\ell)})\right\},$$

where $\mathcal{C}_{\mathrm{SAGE}}^\star(\mathbf{0}, \boldsymbol{\alpha}^{(2)}, \ldots, \boldsymbol{\alpha}^{(\ell)})$ is characterized in (2.7). The next result shows that $\mathcal{W}_0(\mathbf{0}, \boldsymbol{\alpha}^{(2)}, \ldots, \boldsymbol{\alpha}^{(\ell)})$ is an outer convex approximation of the set $\{(1, \exp\{\boldsymbol{\alpha}^{(2)'}\mathbf{x}\}, \ldots, \exp\{\boldsymbol{\alpha}^{(\ell)'}\mathbf{x}\}) \mid \mathbf{x} \in \mathbb{R}^n\}$.

LEMMA 3.3. *Fix a set of exponents $\{\boldsymbol{\alpha}^{(j)}\}_{j=1}^\ell \subset \mathbb{R}^n$ with $\boldsymbol{\alpha}^{(1)} = \mathbf{0}$, and consider the convex set $\mathcal{W}_0(\mathbf{0}, \boldsymbol{\alpha}^{(2)}, \ldots, \boldsymbol{\alpha}^{(\ell)})$ described in (3.13). Then we have that* $\{(1, \exp\{\boldsymbol{\alpha}^{(2)'}\mathbf{x}\}, \ldots, \exp\{\boldsymbol{\alpha}^{(\ell)'}\mathbf{x}\}) \mid \mathbf{x} \in \mathbb{R}^n\} \subseteq \mathcal{W}_0(\mathbf{0}, \boldsymbol{\alpha}^{(2)}, \ldots, \boldsymbol{\alpha}^{(\ell)})$.

Proof. For each $\mathbf{x} \in \mathbb{R}^n$, we note that $\mathbf{v} = (1, \exp\{\boldsymbol{\alpha}^{(2)'}\mathbf{x}\}, \ldots, \exp\{\boldsymbol{\alpha}^{(\ell)'}\mathbf{x}\})' \in \mathcal{W}_0(\mathbf{0}, \boldsymbol{\alpha}^{(2)}, \ldots, \boldsymbol{\alpha}^{(\ell)})$ by setting $\boldsymbol{\tau}^{(i)} = \exp\{\boldsymbol{\alpha}^{(i)'}\mathbf{x}\}\mathbf{x}$ in the characterization (2.7) of $\mathcal{C}_{\mathrm{SAGE}}^\star(\mathbf{0}, \boldsymbol{\alpha}^{(2)}, \ldots, \boldsymbol{\alpha}^{(\ell)})$. $\square$

This lemma gives an alternative justification via duality that $f_{\mathrm{SAGE}}^{(0)} \leq f_\star$. The higher levels of the hierarchy in section 3.1 corresponding to the improved bounds $f_{\mathrm{SAGE}}^{(p)}$ provide successively tighter convex approximations of the set $\{(1, \exp\{\boldsymbol{\alpha}^{(2)'}\mathbf{x}\}, \ldots, \exp\{\boldsymbol{\alpha}^{(\ell)'}\mathbf{x}\}) \mid \mathbf{x} \in \mathbb{R}^n\}$. For example, one can check that the set $\mathcal{W}_1(\mathbf{0}, \boldsymbol{\alpha}^{(2)}, \ldots, \boldsymbol{\alpha}^{(\ell)})$ is given as follows:

$$\mathcal{W}_1(\mathbf{0}, \boldsymbol{\alpha}^{(2)}, \ldots, \boldsymbol{\alpha}^{(\ell)}) = \left\{ \mathbf{v} \in \mathbb{R}_+^\ell \mid \mathbf{v}_1 = 1, \ \exists \boldsymbol{\lambda}^{(i)} \in \mathbb{R}^\ell, i = 1, \ldots, \ell, \ \text{s.t.} \sum_{i=1}^\ell \boldsymbol{\lambda}^{(i)} = \mathbf{v}, \right.$$

$$\left. [\boldsymbol{\lambda}^{(1)'}, \ldots, \boldsymbol{\lambda}^{(\ell)'}]' \in \mathcal{C}_{\mathrm{SAGE}}^\star\left(\boldsymbol{\zeta}^{(1,1)}, \ldots, \boldsymbol{\zeta}^{(\ell,\ell)}\right) \right\},$$

where $\boldsymbol{\zeta}^{(i,j)} = \boldsymbol{\alpha}^{(i)} + \boldsymbol{\alpha}^{(j)}$ for $i, j = 1, \ldots, \ell$.

The dual perspective presented here also provides a method to check whether the lower bound $f_{\mathrm{SAGE}}^{(p)}$ is tight, i.e., $f_{\mathrm{SAGE}}^{(p)} = f_\star$. Focussing on the first level of the hierarchy with $p = 0$, suppose an optimal solution $\hat{\mathbf{v}} \in \mathbb{R}_+^\ell$ of the convex program (3.12) is such that there exists a corresponding point $\hat{\mathbf{x}} \in \mathbb{R}^n$ so that $\hat{\mathbf{v}}_j = \exp\{\boldsymbol{\alpha}^{(j)'}\hat{\mathbf{x}}\}$ for each $j = 1, \ldots, \ell$. Then it follows that $f_{\mathrm{SAGE}}^{(0)} = f_\star$ because the optimal solution $\hat{\mathbf{v}}$ belongs to the set $\{(1, \exp\{\boldsymbol{\alpha}^{(2)'}\mathbf{x}\}, \ldots, \exp\{\boldsymbol{\alpha}^{(\ell)'}\mathbf{x}\}) \mid \mathbf{x} \in \mathbb{R}^n\}$, and the point $\hat{\mathbf{x}} \in \mathbb{R}^n$ is the optimal solution of the original SP (3.1). If $\hat{\mathbf{v}} \notin \{(1, \exp\{\boldsymbol{\alpha}^{(2)'}\mathbf{x}\}, \ldots, \exp\{\boldsymbol{\alpha}^{(\ell)'}\mathbf{x}\}) \mid \mathbf{x} \in \mathbb{R}^n\}$, then a natural heuristic to attempt to obtain a good approximation of a minimizer of the original SP is to project $\log(\hat{\mathbf{v}})$ onto the image of the $\ell \times n$ matrix formed by setting the rows to be the $\boldsymbol{\alpha}^{(j)}$'s. Denoting this projection by $\boldsymbol{\delta} \in \mathbb{R}^\ell$, we obtain an upper bound of the optimal value of the SP (3.1) by computing $\sum_{j=1}^\ell \mathbf{c}_j \exp\{\boldsymbol{\delta}_j\}$. This is the technique employed in Example 3.1 to produce upper bounds.

The dual perspective in the constrained case is conceptually similar. Fix a set of exponents $\{\boldsymbol{\alpha}^{(j)}\}_{j=1}^\ell \subset \mathbb{R}^n$ with $\boldsymbol{\alpha}^{(1)} = \mathbf{0}$, let $f(\mathbf{x}) = \sum_{j=1}^\ell \mathbf{c}_j^{(f)} \exp\{\boldsymbol{\alpha}^{(j)'}\mathbf{x}\}$ be the objective, and let $C = \{g_i(\mathbf{x})\}_{i=1}^m$ specify the set of signomials defining the constraints with $g_i(\mathbf{x}) = \sum_{j=1}^\ell \mathbf{c}_j^{(g_i)} \exp\{\boldsymbol{\alpha}^{(j)'}\mathbf{x}\}$. One can recast the constrained SP (3.7) as

$$
\begin{aligned}
f_\star = \inf_{\mathbf{v} \in \mathbb{R}_+^\ell} \quad & \mathbf{c}^{(f)'}\mathbf{v} \\
\text{s.t.} \quad & \mathbf{v} \in \{(1, \exp\{\boldsymbol{\alpha}^{(2)'}\mathbf{x}\}, \ldots, \exp\{\boldsymbol{\alpha}^{(\ell)'}\mathbf{x}\}) \mid \mathbf{x} \in \mathbb{R}^n\}, \\
& \mathbf{c}^{(g_i)'}\mathbf{v} \geq 0, \ i = 1, \ldots, m.
\end{aligned}
$$

As the objective function is linear, the constraint set can be replaced by its convex hull. By considering the duals of the convex programs (3.10) corresponding to $f_{\mathrm{SAGE}}^{(p,q)}$, one can again interpret SAGE relaxations as providing a sequence of computationally tractable convex outer approximations of the above constraint set:

$$
(3.14) \qquad f_{\mathrm{SAGE}}^{(p,q)} = \inf_{\mathbf{v} \in \mathbb{R}^\ell} \ \mathbf{c}^{(f)'}\mathbf{v} \quad \text{s.t.} \ \mathbf{v} \in \mathcal{W}_{p,q}(\mathbf{0}, \boldsymbol{\alpha}^{(2)}, \ldots, \boldsymbol{\alpha}^{(\ell)}; \mathbf{c}^{(g_1)}, \ldots, \mathbf{c}^{(g_m)}).
$$

That is, $\mathcal{W}_{p,q}(\mathbf{0}, \boldsymbol{\alpha}^{(2)}, \ldots, \boldsymbol{\alpha}^{(\ell)}; \mathbf{c}^{(g_1)}, \ldots, \mathbf{c}^{(g_m)})$ can be viewed as a convex outer approximation of the set $\{(1, \exp\{\boldsymbol{\alpha}^{(2)'}\mathbf{x}\}, \ldots, \exp\{\boldsymbol{\alpha}^{(\ell)'}\mathbf{x}\}) \mid \mathbf{x} \in \mathbb{R}^n\} \cap \{\mathbf{v} \mid \mathbf{c}^{(g_i)'}\mathbf{v} \geq 0, \ i = 1, \ldots, m\}$. For example, the first level of the hierarchy corresponding to $f^{(0,1)}$ is given by

$$
\begin{aligned}
\mathcal{W}_{0,1}(\mathbf{0}, \boldsymbol{\alpha}^{(2)}, \ldots, \boldsymbol{\alpha}^{(\ell)}; \mathbf{c}^{(g_1)}, \ldots, \mathbf{c}^{(g_m)}) = \Big\{ & \mathbf{v} \in \mathbb{R}_+^\ell \mid \mathbf{v}_1 = 1, \\
& \mathbf{v}'\mathbf{c}^{(g_i)} \geq 0, \ i = 1, \ldots, m, \\
& \mathbf{v} \in \mathcal{C}_{\mathrm{SAGE}}^\star(\mathbf{0}, \boldsymbol{\alpha}^{(2)}, \ldots, \boldsymbol{\alpha}^{(\ell)}) \Big\}.
\end{aligned}
$$

Analogous to Lemma 3.3, one can show that $\mathcal{W}_{0,1}(\mathbf{0}, \boldsymbol{\alpha}^{(2)}, \ldots, \boldsymbol{\alpha}^{(\ell)}; \mathbf{c}^{(g_1)}, \ldots, \mathbf{c}^{(g_m)})$ is indeed a convex outer approximation of the set $\{(1, \exp\{\boldsymbol{\alpha}^{(2)'}\mathbf{x}\}, \ldots, \exp\{\boldsymbol{\alpha}^{(\ell)'}\mathbf{x}\}) \mid \mathbf{x} \in \mathbb{R}^n\} \cap \{\mathbf{v} \mid \mathbf{c}^{(g_i)'}\mathbf{v} \geq 0, \ i = 1, \ldots, m\}$. As with the unconstrained setting, one can check whether $f_{\mathrm{SAGE}}^{(p,q)} = f_\star$ by verifying that the optimal solution of the dual problem (3.14) lies in the set $\{(1, \exp\{\boldsymbol{\alpha}^{(2)'}\mathbf{x}\}, , \ldots, \exp\{\boldsymbol{\alpha}^{(\ell)'}\mathbf{x}\}) \mid \mathbf{x} \in \mathbb{R}^n\} \cap \{\mathbf{v} \mid \mathbf{c}^{(g_i)'}\mathbf{v} \geq 0, \ i = 1, \ldots, m\}$.

4. Completeness of hierarchies. In this section, we show that the hierarchies of lower bounds presented in the previous section converge to the optimal value for suitable classes of SPs. We also contrast the methods developed in this paper with those based on SDPs for polynomial optimization problems [15, 19, 20], highlighting the point that relative entropy relaxations based on SAGE decompositions are better suited to SPs than SDP relaxations based on sum-of-squares techniques.

4.1. Completeness of relative entropy hierarchies for SPs with rational exponents. Our completeness results are stated for SPs in which the signomials consist of rational exponent vectors. Specifically, the high-level idea behind our proofs is as follows: We transform the SPs to appropriate polynomial optimization problems over the nonnegative orthant by suitably clearing denominators in the exponents of the SP, then apply Positivstellensatz results from real algebraic geometry to obtain representations of positive polynomials [14, 18, 32] in terms of SAGE functions (which can be viewed as nonnegative polynomials over the orthant modulo an ideal), and finally transform the results back to derive a consequence in terms of signomials. We emphasize that the transformation of an SP to a polynomial optimization problem by clearing denominators in the exponents is purely an element of our proof technique due to the appeal to representation theorems from real algebraic geometry. Our algorithmic methodology described in section 3 does not require such clearing of denominators, and we highlight the significance of this point in section 4.2.

Unconstrained SPs. Our transformation of SPs with rational exponents leads to polynomial optimization problems over the positive orthant intersected with an algebraic variety. However, our appeal to positivstellensatz results from real algebraic geometry gives representations of polynomials over the *nonnegative* orthant intersected with an algebraic variety. Further, these representation theorems also entail certain Archimedeanity conditions as discussed in Appendix A.1. In order to overcome both these issues, we require an additional condition on the exponents as elaborated in the following result.

THEOREM 4.1. *Let* $\{\boldsymbol{\alpha}^{(j)}\}_{j=1}^{\ell} \subset \mathbb{Q}^n$ *be a collection of rational exponents with the following properties: (a) the first n exponents $\{\boldsymbol{\alpha}^{(j)}\}_{j=1}^{n}$ are linearly independent, (b) $\boldsymbol{\alpha}^{(n+1)} = \mathbf{0}$, and (c) the remaining exponents $\{\boldsymbol{\alpha}^{(j)}\}_{j=n+2}^{\ell}$ lie in the convex hull $\mathrm{conv}(\{\boldsymbol{\alpha}^{(i)}\}_{i=1}^{n+1})$, and $\boldsymbol{\alpha}^{(j)} \notin \mathrm{conv}(\{\boldsymbol{\alpha}^{(i)}\}_{i=1}^{n})$ for $j = n+2, \ldots, \ell$. Suppose $f(\mathbf{x})$ is a signomial with respect to these exponents such that $\inf_{\mathbf{x} \in \mathbb{R}^n} f(\mathbf{x}) > 0$ and that $\mathbf{c}_j > 0$ for each $j \in \{1, \ldots, n\}$. Then there exists $p \in \mathbb{Z}_+$ such that $(\sum_{j=1}^{\ell} \exp\{\boldsymbol{\alpha}^{(j)'}\mathbf{x}\})^p f(\mathbf{x}) \in \mathrm{SAGE}(E_{p+1}(\{\boldsymbol{\alpha}^{(j)}\}_{j=1}^{\ell}))$.*

Note. The condition that $\mathbf{c}_j > 0$ for each $j = 1, \ldots, n$ is not restrictive because nonzero coefficients corresponding to extremal exponents must be positive for any globally nonnegative signomial; indeed, if any of these $\mathbf{c}_j$'s for $j = 1, \ldots, n$ is negative, then the signomial $f(\mathbf{x})$ is unbounded below (see section 2.3). The main purpose of this condition is to ensure that $\mathbf{c}_j \neq 0$ for each $j = 1, \ldots, n$.

This theorem is similar in the spirit of Polya–Reznick-type results on rational sum-of-squares representations of positive polynomials with uniform denominators [22, 25]. Without loss of generality one can take the first n exponent vectors $\{\boldsymbol{\alpha}^{(j)}\}_{j=1}^{\ell}$ to be linearly independent to satisfy condition (a); generalizing our discussion to address the case in which there are fewer than n independent vectors in $\{\boldsymbol{\alpha}^{(j)}\}_{j=1}^{\ell}$ is straightforward. Further, one can also add the zero vector to the collection of exponents $\{\boldsymbol{\alpha}^{(j)}\}_{j=1}^{\ell}$ to satisfy condition (b). Condition (c) is the key assumption in this theorem as it enables us to overcome the two difficulties described above preceding

the statement of Theorem 4.1 (see Appendix A.4). Observe also that the structure of the exponents in the signomials considered in Example 3.1 satisfy the conditions of Theorem 4.1.

To see that Theorem 4.1 yields completeness for unconstrained SPs, suppose $f(\mathbf{x})$ is a signomial with respect to a set of exponents $\{\boldsymbol{\alpha}^{(j)}\}_{j=1}^{\ell}$ satisfying the assumptions of Theorem 4.1, and let $f_{\star}$ denote the global minimum of $f(\mathbf{x})$. For each $\epsilon < f_{\star}$, the signomial $f(\mathbf{x}) - \epsilon$ is strictly positive for all $\mathbf{x} \in \mathbb{R}^n$. Theorem 4.1 asserts that there exists a sufficiently large $p_{\epsilon} \in \mathbb{Z}_+$ such that $(\sum_{j=1}^{\ell} \exp\{\boldsymbol{\alpha}^{(j)'}\mathbf{x}\})^{p_{\epsilon}}[f(\mathbf{x}) - \epsilon] \in$ SAGE$(E_{p_{\epsilon}+1}(\{\boldsymbol{\alpha}^{(j)}\}_{j=1}^{\ell}))$. Combined with the monotonicity result of Lemma 3.1, one can conclude that $f_{\mathrm{SAGE}}^{(p)} \to f_{\star}$ as $p \to \infty$.

Constrained SPs. Our next result states that the hierarchy of lower bounds $f_{\mathrm{SAGE}}^{(p,q)}$ for constrained SPs also converges to the global optimum if the constraint set is compact. Consider an SP specified in terms of an objective signomial $f(\mathbf{x})$ and constraint signomials $C = \{g_i(\mathbf{x})\}_{i=1}^m$, all defined with respect to a collection of exponents $\{\boldsymbol{\alpha}^{(j)}\}_{j=1}^{\ell} \subset \mathbb{R}^n$. As discussed in section 3.2, the compactness of the constraint set $\mathcal{K}_C = \{\mathbf{x} \mid g_i(\mathbf{x}) \geq 0, \ i = 1, \ldots, m\}$ must be explicitly certified via the presence of redundant inequalities of the form $U \geq \exp\{\boldsymbol{\alpha}^{(j)'}\mathbf{x}\} \geq L, \ j = 1, \ldots, \ell$ for suitable $U, L \in \mathbb{R}_{++}$ in the list of constraints C. If $\mathcal{K}_C$ is a compact set, then such redundant constraints can always be added for appropriately large U and small L. (The compactness of the constraint set ensures that some of the difficulties in the unconstrained case—the Archimedeanity requirement and the distinction between positive versus nonnegative orthant in appealing to representation theorems—do not present obstacles in the constrained setting; specifically, we do not require additional conditions on the exponents beyond the assumption that they are rational vectors).

THEOREM 4.2. *Fix a set of* rational *exponents* $\{\boldsymbol{\alpha}^{(j)}\}_{j=1}^{\ell} \subset \mathbb{Q}^n$, *and let* $f(\mathbf{x})$ *and* $C = \{g_i(\mathbf{x})\}_{i=1}^m$ *be signomials with respect to these exponents. Let* $R_q(C)$ *be as defined in* (3.9), *and let the constraint set be* $\mathcal{K}_C = \{\mathbf{x} \mid g_i(\mathbf{x}) \geq 0, \ i = 1, \ldots, m\}$. *Suppose inequalities of the form* $U \geq \exp\{\boldsymbol{\alpha}^{(j)'}\mathbf{x}\} \geq L, \ j = 1, \ldots, \ell$ *for* $U, L \in \mathbb{R}_{++}$ *are formally specified in the list of constraints* C *(explicitly serving as witnesses of the compactness of* $\mathcal{K}_C$), *and suppose* $f(\mathbf{x})$ *is strictly positive for all* $\mathbf{x} \in \mathcal{K}_C$. *Then there exist* $p, q \in \mathbb{Z}_+$ *and SAGE functions* $s_h(\mathbf{x}) \in$ SAGE$(E_p(\{\boldsymbol{\alpha}^{(j)}\}_{j=1}^{\ell}))$ *indexed by* $h \in R_q(C)$ *such that* $f(\mathbf{x}) - \sum_{h(\mathbf{x}) \in R_q(C)} s_h(\mathbf{x})h(\mathbf{x}) \in$ SAGE$(E_{p+q}(\{\boldsymbol{\alpha}^{(j)}\}_{j=1}^{\ell}))$.

By proceeding with a line of reasoning as in the unconstrained case, one can check that the sequence of lower bounds $f_{\mathrm{SAGE}}^{(p,q)}$ converges to the global optimum of a constrained SP with rational exponents and a compact constraint set. Note that Theorem 4.2 also implies that the methods described in section 3.3 can be employed to find a minimizer of a signomial over a constraint set of the form $\{\mathbf{x} \in \mathbb{R}^n \mid U \geq \exp\{\boldsymbol{\alpha}^{(j)'}\mathbf{x}\} \geq L, \ j = 1, \ldots, \ell\}$ for $U, L \in \mathbb{R}_{++}$ (or for unconstrained SPs when bounds on the location of a global minimizer are known).

Remarks. The proof of Theorem 4.1 uses the fact that the set of posynomials forms a preprime (i.e., closed under addition and multiplication) and that posynomials are SAGE functions. The proof of Theorem 4.2 appeals to the property that the set of SAGE functions is closed under multiplication by posynomials (formally, that the set of SAGE functions forms a module over the preprime of posynomials). Although these results provide theoretical support for the methods proposed in section 3, our proof techniques do not reflect the full strength of relative entropy relaxations. For example, one of the appealing features of relative entropy relaxations for SPs—not revealed by

the proofs of Theorems 4.1 and 4.2—is that they are robust to small perturbations of the exponents of the underlying SP. We discuss these points in section 4.2; see Proposition 4.3 and Example 4.1.

4.2. Contrast with polynomial optimization. A natural point of comparison with the framework presented in this paper is the literature on polynomial optimization, which involves the minimization of a multivariate polynomial subject to constraints specified by multivariate polynomials. As with SPs, polynomial optimization problems are also nonconvex in general, and they include families of NP-hard problems. Parrilo [19, 20] and Lasserre [15] describe computationally feasible methods to obtain lower bounds for polynomial optimization problems. These techniques rely on nonnegativity certificates for polynomials based on sum-of-squares decompositions [14, 18, 32] and the observation by Shor [30] that checking if a polynomial is a sum-of-squares can be recast as an SDP feasibility problem.

SPs with rational exponents can be transformed to polynomial optimization problems over the nonnegative orthant by clearing denominators in the exponents. This transformation generally leads to polynomials of very large degrees, thus making sum-of-squares techniques ill-suited for general SPs. Example 2.1 gives a concrete illustration of this point. The signomial $f(\mathbf{x}) = \exp\{\mathbf{x}_1\} + \exp\{\mathbf{x}_2\} - \exp\{(\frac{1}{2}+\epsilon)\mathbf{x}_1 + (\frac{1}{2}-\epsilon)\mathbf{x}_2\}$ for a fixed $\epsilon \in [-\frac{1}{2}, \frac{1}{2}]$ can be efficiently certified as being globally nonnegative by solving a suitable relative entropy feasibility problem. Now suppose that $\epsilon = \frac{p}{q}$ is a rational number in $[-\frac{1}{2}, \frac{1}{2}]$. One can then transform the question of nonnegativity of $f(\mathbf{x})$ to a problem of polynomial nonnegativity. In particular by setting $\mathbf{z}_1 = \exp\{\mathbf{x}_1\}, \mathbf{z}_2 = \exp\{\mathbf{x}_2\}, \mathbf{z}_3 = \exp\{(\frac{1}{2} + \frac{p}{q})\mathbf{x}_1 + (\frac{1}{2} - \frac{p}{q})\mathbf{x}_2\}$, we would like to certify the nonnegativity of the polynomial $\tilde{f}(\mathbf{z}_1, \mathbf{z}_2, \mathbf{z}_3) = \mathbf{z}_1 + \mathbf{z}_2 - \mathbf{z}_3$ over the nonnegative orthant $\mathbb{R}^3_+$ modulo the ideal generated by the polynomial $\mathbf{z}_3^{2q} - \mathbf{z}_1^{q+2p}\mathbf{z}_2^{q-2p}$. If q is large, the corresponding certificates based on sum-of-squares methods can be of very large degree (see [15, 19, 20] for more details about constructing these certificates), which in turn require solution of large SDPs. On the other hand, there exists a short certificate for the nonnegativity of $f(\mathbf{x})$ via the AM/GM inequality, which is the basic insight we exploit to develop tractable convex relaxations for SPs.

More broadly, relative entropy relaxations for SPs also have the virtue that the bounds they provide are generally robust to small perturbations of the exponents. Specifically, given a set of exponents $\{\boldsymbol{\alpha}^{(j)}\}_{j=1}^{\ell} \subset \mathbb{R}^n$, we say that a SAGE cone $\mathcal{C}_{\mathrm{SAGE}}(\boldsymbol{\alpha}^{(1)}, \ldots, \boldsymbol{\alpha}^{(\ell)})$ is robust to small perturbations of the exponents $\{\boldsymbol{\alpha}^{(j)}\}_{j=1}^{\ell}$ when the following condition holds: if a coefficient vector $\mathbf{c} \in \mathbb{R}^\ell$ belongs to the interior of the cone $\mathcal{C}_{\mathrm{SAGE}}(\boldsymbol{\alpha}^{(1)}, \ldots, \boldsymbol{\alpha}^{(\ell)})$ for a collection of exponents $\{\boldsymbol{\alpha}^{(j)}\}_{j=1}^{\ell}$, then $\mathbf{c}$ also belongs to the cone $\mathcal{C}_{\mathrm{SAGE}}(\tilde{\boldsymbol{\alpha}}^{(1)}, \ldots, \tilde{\boldsymbol{\alpha}}^{(\ell)})$ for all sufficiently small perturbations $\tilde{\boldsymbol{\alpha}}^{(j)}$ of each $\boldsymbol{\alpha}^{(j)}$. As the SAGE cone $\mathcal{C}_{\mathrm{SAGE}}$ is a direct sum of AM/GM exponential cones $\mathcal{C}_{\mathrm{AGE}}$, the robustness of each of the component AM/GM exponential cones $\mathcal{C}_{\mathrm{AGE}}$ implies the robustness of $\mathcal{C}_{\mathrm{SAGE}}$. The following proposition shows that cones of AM/GM exponentials are robust to small perturbations of the exponents for a broad class of exponents. The basic observation underlying this result is that the exponents $\{\boldsymbol{\alpha}^{(j)}\}_{j=1}^{\ell}$ appear only as linear constraints in the characterization of the cones $\mathcal{C}_{\mathrm{SAGE}}$ and $\mathcal{C}_{\mathrm{AGE}}$ based on (2.3) and Proposition 2.4.

PROPOSITION 4.3. *Fix a set of exponents $\{\boldsymbol{\alpha}^{(j)}\}_{j=0}^{\ell} \subset \mathbb{R}^n$, and let A denote the $n \times (\ell+1)$ matrix with the $\boldsymbol{\alpha}^{(j)}$'s as the columns. Let $(\mathbf{c}, \beta) \in \mathbb{R}^\ell_+ \times \mathbb{R}$ be any point in the interior of the cone $\mathcal{C}_{\mathrm{AGE}}(\boldsymbol{\alpha}^{(1)}, \ldots, \boldsymbol{\alpha}^{(\ell)}; \boldsymbol{\alpha}^{(0)})$. Suppose $\boldsymbol{\alpha}^{(0)}$ is not contained in the convex hull of any subset of $k \leq n$ of the exponents $\{\boldsymbol{\alpha}^{(j)}\}_{j=1}^{\ell}$. Then there exists*

an open set $\mathcal{S} \subset \mathbb{R}^{n \times (\ell+1)}$ containing A with the following property: For any $\tilde{A} \in \mathcal{S}$ with columns $\{\tilde{\boldsymbol{\alpha}}^{(j)}\}_{j=0}^{\ell}$, we have that $(\mathbf{c}, \beta)$ belongs to $\mathcal{C}_{\mathrm{AGE}}(\tilde{\boldsymbol{\alpha}}^{(1)}, \ldots, \tilde{\boldsymbol{\alpha}}^{(\ell)}; \tilde{\boldsymbol{\alpha}}^{(0)})$.

Proof sketch. We provide a brief outline of the proof. As $(\mathbf{c}, \beta)$ belongs to the interior of the cone $\mathcal{C}_{\mathrm{AGE}}(\boldsymbol{\alpha}^{(1)}, \ldots, \boldsymbol{\alpha}^{(\ell)}; \boldsymbol{\alpha}^{(0)})$, the entries of $\mathbf{c}$ must be strictly positive. Further, we have from Lemma 2.2 that there exists a $\boldsymbol{\nu} \in \mathbb{R}_+^{\ell}$ such that $D(\boldsymbol{\nu}, e\mathbf{c}) < \beta$ and that $\sum_{j=1}^{\ell} \boldsymbol{\alpha}^{(j)} \nu_j = (\mathbf{1}'\boldsymbol{\nu})\boldsymbol{\alpha}^{(0)}$. Due to the assumption that $\boldsymbol{\alpha}^{(0)}$ is not contained in the convex hull of any subset of $k \leq n$ of the exponents $\{\boldsymbol{\alpha}^{(j)}\}_{j=1}^{\ell}$, we have that $\boldsymbol{\nu}$ must contain at least $n+1$ nonzeros. Therefore, for all sufficiently small perturbations $\{\tilde{\boldsymbol{\alpha}}^{(j)}\}_{j=0}^{\ell}$, there exists a corresponding $\tilde{\boldsymbol{\nu}}$ close to $\boldsymbol{\nu}$ such that $\sum_{j=1}^{\ell}(\tilde{\boldsymbol{\alpha}}^{(j)} - \tilde{\boldsymbol{\alpha}}^{(0)})\tilde{\nu}_j = \mathbf{0}$. (If some of the coordinates of the original $\boldsymbol{\nu}$ are zero, then there exists a nearby $\tilde{\boldsymbol{\nu}}$ with zeros in the same coordinates satisfying $\sum_{j=1}^{\ell}(\tilde{\boldsymbol{\alpha}}^{(j)} - \tilde{\boldsymbol{\alpha}}^{(0)})\tilde{\nu}_j = \mathbf{0}$ for all sufficiently small perturbations $\{\tilde{\boldsymbol{\alpha}}^{(j)}\}_{j=0}^{\ell}$.) The relative entropy function $D(\cdot, e\mathbf{c})$ is continuous with respect to the first argument if the second argument $\mathbf{c}$ is fixed and contains strictly positive entries. Consequently, we have that $D(\tilde{\boldsymbol{\nu}}, e\mathbf{c}) < \beta$. $\square$

This result demonstrates that the bounds $f_{\mathrm{SAGE}}^{(0)}$ (3.6) and $f_{\mathrm{SAGE}}^{(0,1)}$ (3.10) based on SAGE decompositions are robust to small perturbations of the exponents in the underlying SPs; one can also carry out a similar analysis to show that the bounds $f_{\mathrm{SAGE}}^{(p)}$ (3.6) and $f_{\mathrm{SAGE}}^{(p,q)}$ (3.10) corresponding to higher levels of the hierarchies are robust to small perturbations of the exponents in the underlying SPs.

Example 4.1. Consider the following signomial obtained by perturbing the exponents of the signomial (3.4) in Example 3.1 (but leaving the coefficients unchanged):

$$
\begin{aligned}
f(\mathbf{x}) = {} & 10 \exp\{10.2070\mathbf{x}_1 + 0.0082\mathbf{x}_2 - 0.0039\mathbf{x}_3\} \\
& + 10 \exp\{-0.0081\mathbf{x}_1 + 9.8024\mathbf{x}_2 - 0.0097\mathbf{x}_3\} \\
& + 10 \exp\{0.0070\mathbf{x}_1 - 0.0156\mathbf{x}_2 + 8.1923\mathbf{x}_3\} \\
& - 14.6794 \exp\{1.5296\mathbf{x}_1 + 1.0927\mathbf{x}_2 + 1.3441\mathbf{x}_3\} \\
& - 7.8601 \exp\{1.0750\mathbf{x}_1 + 1.9108\mathbf{x}_2 + 1.6339\mathbf{x}_3\} \\
& + 8.7838 \exp\{1.0513\mathbf{x}_1 + 0.0571\mathbf{x}_2 + 1.6188\mathbf{x}_3\}.
\end{aligned}
$$

The SAGE lower bound is tight for the perturbed signomial specified here, with the optimal value being equal to -0.9458, and the optimal solution being $\mathbf{x}^\star = (-0.3016, -0.2605, -0.4013)'$. Recall that the SAGE lower bound was also tight for the signomial (3.4) of Example 3.1, with the optimal value being equal to -0.9747 and the optimal solution being $\mathbf{x}^\star = (-0.3020, -0.2586, -0.4010)'$. Hence, this example provides numerical evidence for the robustness of our relative entropy relaxation methods with respect to small perturbations of the exponents.

Approaching this SP as a polynomial optimization problem clearly illustrates the shortcomings of that viewpoint, as small changes in the exponents can lead to very different polynomials after clearing denominators in the exponents. In turn, the quality of the bounds and amount of computation required to obtain them via SDP relaxations based on sum-of-squares techniques can vary dramatically for small changes in the exponents. However, relative entropy relaxations for SPs based on SAGE decompositions are well-behaved under such small perturbations of the exponents.

5. Discussion. A number of directions for further investigation arise from this work, and we mention some of these here.

Implications for polynomial optimization. The methods developed in this paper may offer a useful addition to SDP relaxation frameworks based on sum-of-squares decompositions for polynomial optimization problems over the nonnegative orthant (equivalently, polynomial optimization problems involving even polynomials). For example, the Motzkin polynomial is certified as being globally nonnegative via the AM/GM inequality, although this polynomial cannot directly be expressed as a sum-of-squares (see Example 2.4). More generally, the methods in this paper may offer effective certificates of global nonnegativity for certain "sparse" polynomials of large degree, which are sometimes referred to as "fewnomials" in the literature [13]; this point is also exploited in [12], where the authors employ GPs to compute lower bounds on certain multivariate polynomials more efficiently than SDPs based on sum-of-squares techniques. In fact one can combine the SDP and the relative entropy relaxation frameworks, based on sum-of-squares and SAGE decompositions, respectively, to obtain alternative hierarchies for polynomial optimization problems. Such a combined hierarchy would provide bounds that are at least as good as those produced by SDP relaxations, although this improvement comes at an increased computational cost; it is of interest to investigate and quantify these comparisons.

Efficient transcendental representations of semialgebraic sets. Consider the set $S_d = \{(a,b) \in \mathbb{R}^2 \mid x^{2d} + ax^2 + b \geq 0 \text{ for all } x \in \mathbb{R}\}$ for $d \in \mathbb{Z}_{++}$. This set is convex and semialgebraic for each $d \in \mathbb{Z}_{++}$, and it is SDP representable (i.e., expressible as the projection of a slice of the cone of $w_d \times w_d$ positive semidefinite matrices, where $w_d \in \mathbb{Z}_{++}$ may depend on d) based on the observation that a nonnegative univariate polynomial can be expressed as a sum-of-squares. However, the algebraic degree of the boundary of S_d grows with d; thus, the size w_d of the smallest SDP description of S_d must grow with d. On the other hand, S_d can be represented efficiently via relative entropy inequalities as $S_d = \{(a,b) \in \mathbb{R} \times \mathbb{R}_+ \mid \exists \boldsymbol{\nu} \in \mathbb{R}_+^2 \text{ s.t. } D(\boldsymbol{\nu}, e(1,b)') \leq a, \ (d-1)\boldsymbol{\nu}_1 = \boldsymbol{\nu}_2\}$. Note that the size of this description does not grow with d. Consequently, there are semialgebraic sets (indeed, SDP representable sets) that are more efficiently specified via transcendental representations. Underlying this discussion is the insight that the AM/GM inequality offers a different proof system than sum-of-squares methods for certifying nonnegativity. Specifically, the length of a sum-of-squares proof of the AM/GM inequality for a *fixed* set of rational weights can be very large depending on the denominators of the rational weights. On the other hand, relative entropy methods offer an effective approach for *computing* a set of weights that certifies nonnegativity via the AM/GM inequality. Investigating the respective power of these different proof systems may yield new insights into the possibility of efficient transcendental representations of convex semialgebraic sets.

Effective bounds on quality of approximation. A third direction is to provide useful bounds on the size of a relaxation required in order to achieve a specified quality of approximation to the global minimum. A first step toward addressing this question is to achieve a deeper understanding of the effect of perturbations in the exponents of an SP on the eventual bounds computed at a certain level of our hierarchy. As discussed in section 4.1, the techniques employed in our proofs of Theorems 4.1 and 4.2 do not provide guidance on the effects of such perturbations, and consequently a different set of ideas may be required to address this question. A second point that would arise in an investigation of the quality of approximation provided by a SAGE relaxation is the relative volume of the cone of SAGE signomials with respect to the cone of nonnegative signomials (for a fixed set of exponents). In the context of sum-of-

squares polynomials versus nonnegative polynomials, Blekherman has characterized such relative volumes in an asymptotic sense [1], and it would be of interest to carry out a similar study in the context of signomials.

Numerical scalability. Finally, the deployment of relative entropy relaxations for SPs arising in practice requires the efficient solution of large instances of relative entropy optimization problems. Relative entropy optimization problems can be viewed as a special case of conic optimization problems with respect to the so-called power cone. Motivated by the fact that this cone is nonsymmetric, recent efforts have developed effective interior-point methods that scale well to moderate instances [31]. In order to scale to massive instances, it is desirable to develop first-order algorithms as well as specialized methods that exploit structure in the relative entropy relaxations presented in this paper.

Appendix A. We begin by briefly stating the relevant representation theorems from the real algebraic geometry literature—the one we use is due to Krivine [14]; see [23, 27, 32] for other prominent examples. Then we describe a transformation of SPs with rational exponent vectors to polynomial optimization problems. Finally, we put these steps together to give proofs of Theorems 4.1 and 4.2.

A.1. Representation theorems for Archimedean modules over preprimes. We primarily draw from the wonderful exposition in [18]. Let $\mathbb{R}[\mathbf{Y}_1, \dots, \mathbf{Y}_\ell]$ denote the polynomial ring in indeterminates $\mathbf{Y}$ with real coefficients, and let $\mathbb{Q}$ denote the rational numbers. The first object that plays a critical role in our development is a preprime in a polynomial ring.

DEFINITION A.1. *A preprime $P \subset \mathbb{R}[\mathbf{Y}_1, \dots, \mathbf{Y}_\ell]$ is a subset containing $\mathbb{Q}$ that is closed under addition and multiplication. Further, a preprime P is* Archimedean *if for each $r \in \mathbb{R}[\mathbf{Y}_1, \dots, \mathbf{Y}_\ell]$ there exists $m \in \mathbb{Z}$ such that $r + m \in P$.*

For a preprime $P \subset \mathbb{R}[\mathbf{Y}_1, \dots, \mathbf{Y}_\ell]$, the *ring of bounded elements of P with respect to $\mathbb{R}[\mathbf{Y}_1, \dots, \mathbf{Y}_\ell]$* is defined as

$$H_P = \{r \in \mathbb{R}[\mathbf{Y}_1, \dots, \mathbf{Y}_\ell] \mid \exists m \in \mathbb{Z} \text{ s.t. } m \pm r \in P\}.$$

From [18, Prop. 5.1.3], the preprime P is Archimedean if and only if $H_P = \mathbb{R}[\mathbf{Y}_1, \dots, \mathbf{Y}_\ell]$. The next algebraic structure that is central to our discussion is a module over a preprime.

DEFINITION A.2. *Let $P \subset \mathbb{R}[\mathbf{Y}_1, \dots, \mathbf{Y}_\ell]$ be a preprime. Then $M \subset \mathbb{R}[\mathbf{Y}_1, \dots, \mathbf{Y}_\ell]$ is a P-module if it is closed under addition, is closed under multiplication by an element of P, and contains 1. As with preprimes, a P-module M is* Archimedean *if for each $r \in \mathbb{R}[\mathbf{Y}_1, \dots, \mathbf{Y}_\ell]$ there exists $m \in \mathbb{Z}$ such that $r + m \in M$.*

Note that $1 \in M$ for a P-module M implies that $P \subset M$. Consequently, if a preprime P is Archimedean, then any P-module M is also Archimedean. Finally, P itself is a P-module. The following result due, in various forms, to several authors is the main foundation for our proofs of Theorems 4.1 and 4.2.

THEOREM A.3 (see [14, 18]). *Let $P \subset \mathbb{R}[\mathbf{Y}_1, \dots, \mathbf{Y}_\ell]$ be an Archimedean preprime and let M be a P-module. Let $\mathcal{K}_M$ denote the set of points in $\mathbb{R}^\ell$ on which every element of M is nonnegative:*

$$\mathcal{K}_M = \{\mathbf{y} \in \mathbb{R}^\ell \mid r(\mathbf{y}) \geq 0 \ \forall r \in M\}.$$

If $s \in \mathbb{R}[\mathbf{Y}_1, \ldots, \mathbf{Y}_\ell]$ is strictly positive on the set of points $\mathcal{K}_M$, then $s \in M$:

$$s(\mathbf{y}) > 0 \ \forall \mathbf{y} \in \mathcal{K}_M \Rightarrow s \in M.$$

The converse direction, i.e., $s \in M$ implies $s(\mathbf{y}) \geq 0$ for all $\mathbf{y} \in \mathcal{K}_M$, follows directly from the definition of $\mathcal{K}_M$. Results of the form of Theorem A.3 are known as representation theorems because the P-module M is usually constructed explicitly, and the positivity of a polynomial s on the nonnegativity set $\mathcal{K}_M$ associated to M implies that s has an explicit representation.

A.2. Implicitization and parametrization of SPs with rational exponents. In this section, we present an elementary result that allows us to transform certain families of SPs with rational exponents to polynomial optimization problems over the nonnegative orthant. This result is employed in our proofs of Theorems 4.1 and 4.2. Specifically, consider the problem of verifying positivity of $\inf_{\mathbf{x} \in \mathcal{S}} f(\mathbf{x})$ for $\mathcal{S} = \{\mathbf{x} \in \mathbb{R}^n \mid g_i(\mathbf{x}) \geq 0, \ i = 1, \ldots, m\}$, where $f(\mathbf{x}), \{g_i(\mathbf{x})\}_{i=1}^m$ are signomials with respect to exponents $\{\boldsymbol{\alpha}^{(j)}\}_{j=1}^\ell \subset \mathbb{Q}^n$. Letting $f(\mathbf{x}) = \sum_{j=1}^\ell \mathbf{c}_j^{(f)'} \exp\{\boldsymbol{\alpha}^{(j)'}\mathbf{x}\}$ and letting each $g_i(\mathbf{x}) = \sum_{j=1}^\ell \mathbf{c}_j^{(g_i)'} \exp\{\boldsymbol{\alpha}^{(j)'}\mathbf{x}\}$, the positivity of $\inf_{\mathbf{x} \in \mathcal{S}} f(\mathbf{x})$ with $\mathcal{S} = \{\mathbf{x} \in \mathbb{R}^n \mid g_i(\mathbf{x}) \geq 0, \ i = 1, \ldots, m\}$ is equivalent to $\inf_{\mathbf{y} \in \mathcal{T}_{++}} \mathbf{c}^{(f)'}\mathbf{y} > 0$ for

$$
(A.1) \qquad
\begin{aligned}
\mathcal{T}_{++} &= \{\mathbf{y} \in \mathbb{R}^\ell \mid \mathbf{c}^{(g_i)'}\mathbf{y} \geq 0, \ i = 1, \ldots, m\} \\
&\quad \cap \{\mathbf{y} \in \mathbb{R}_{++}^\ell \mid \exists \mathbf{x} \in \mathbb{R}^n \text{ s.t. } \mathbf{y}_j = \exp\{\boldsymbol{\alpha}^{(j)'}\mathbf{x}\}, \ j = 1, \ldots, \ell\}.
\end{aligned}
$$

A.2.1. Implicitization. For rational exponents $\{\boldsymbol{\alpha}^{(j)}\}_{j=1}^\ell \in \mathbb{Q}^n$, the set $\{\mathbf{y} \in \mathbb{R}_{++}^\ell \mid \exists \mathbf{x} \in \mathbb{R}^n \text{ s.t. } \mathbf{y}_j = \exp\{\boldsymbol{\alpha}^{(j)'}\mathbf{x}\}, \ j = 1, \ldots, \ell\}$ can be implicitized via polynomial equations as follows. Suppose without loss of generality that the first n of the exponent vectors $\{\boldsymbol{\alpha}^{(j)}\}_{j=1}^\ell$ are linearly independent. Consequently, the remaining exponents $\{\boldsymbol{\alpha}^{(j)}\}_{j=n+1}^\ell$ can be expressed as linear combinations with *rational* coefficients of the first n exponents. As a result, each $\mathbf{y}_j$ for $j = n+1, \ldots, \ell$ can be expressed as $\mathbf{y}_j = \prod_{k=1}^n \mathbf{y}_k^{\gamma_k^{(j)}}$ for rational vectors $\boldsymbol{\gamma}^{(j)} \in \mathbb{Q}^n$, $j = n+1, \ldots, \ell$. By moving terms involving negative powers from the right to the left side of each of these equations and then suitably clearing denominators in the exponents (valid manipulations over the positive orthant $\mathbb{R}_{++}^\ell$), the set $\mathcal{T}_{++}$ from (A.1) can be characterized as

$$
(A.2) \qquad
\begin{aligned}
\mathcal{T}_{++} &= \{\mathbf{y} \in \mathbb{R}^\ell \mid \mathbf{c}^{(g_i)'}\mathbf{y} \geq 0, \ i = 1, \ldots, m\} \\
&\quad \cap \{\mathbf{y} \in \mathbb{R}_{++}^\ell \mid v_j(\mathbf{y}) - w_j(\mathbf{y}) = 0, \ j = n+1, \ldots, \ell\}
\end{aligned}
$$

for monomials $v_j(\mathbf{y}), w_j(\mathbf{y}), j = n+1, \ldots, \ell$. The restriction to the positive orthant $\mathbb{R}_{++}^\ell$ in the second set here is crucial as the manipulations we describe are only valid over the positive orthant.

In our appeal to Theorem A.3, we consider the positivity of $\inf_{\mathbf{y} \in \mathcal{T}_+} \mathbf{c}^{(f)'}\mathbf{y}$ for

$$
(A.3) \qquad
\begin{aligned}
\mathcal{T}_+ &= \{\mathbf{y} \in \mathbb{R}^\ell \mid \mathbf{c}^{(g_i)'}\mathbf{y} \geq 0, \ i = 1, \ldots, m\} \\
&\quad \cap \{\mathbf{y} \in \mathbb{R}_+^\ell \mid v_j(\mathbf{y}) - w_j(\mathbf{y}) = 0, \ j = n+1, \ldots, \ell\}.
\end{aligned}
$$

However, the implicitization operation is, in general, only valid over the positive orthant $\mathbb{R}_{++}^\ell$. Consequently, we require that the closure of the set $\mathcal{T}_{++}$ is equal to $\mathcal{T}_+$, as this would imply that $\inf_{\mathbf{y} \in \mathcal{T}_+} \mathbf{c}^{(f)'}\mathbf{y} = \inf_{\mathbf{y} \in \mathcal{T}_{++}} \mathbf{c}^{(f)'}\mathbf{y}$. The next two results

give conditions under which $\mathcal{T}_+ = \mathrm{cl}(\mathcal{T}_{++})$; these correspond to the assumptions in Theorems 4.1 and 4.2.

LEMMA A.4. *Fix a set of exponents $\{\boldsymbol{\alpha}^{(j)}\}_{j=1}^{\ell} \in \mathbb{Q}^n$ with $\{\boldsymbol{\alpha}^{(j)}\}_{j=1}^{n}$ being linearly independent, and consider an SP with transformed constraint sets $\mathcal{T}_{++}$ and $\mathcal{T}_+$ as defined in (A.2), (A.3) for some set of coefficients $\mathbf{c}^{(g_i)} \in \mathbb{R}^{\ell}, i = 1, \ldots, m$. If $\{\mathbf{y} \in \mathbb{R}^{\ell} \mid \mathbf{c}^{(g_i)'}\mathbf{y} \geq 0, \ i = 1, \ldots, m\} \subset \mathbb{R}_{++}^{\ell}$, then $\mathrm{cl}(\mathcal{T}_{++}) = \mathcal{T}_+$.*

Proof. One can check that $\mathcal{T}_{++} = \mathcal{T}_+$. The result then follows from the observation that $\mathcal{T}_+$ is a closed set. $\qquad\square$

LEMMA A.5. *Fix a set of exponents $\{\boldsymbol{\alpha}^{(j)}\}_{j=1}^{\ell} \in \mathbb{Q}^n$ with $\{\boldsymbol{\alpha}^{(j)}\}_{j=1}^{n}$ being linearly independent. If the remaining vectors $\{\boldsymbol{\alpha}^{(j)}\}_{j=n+1}^{\ell}$ lie in the cone generated by $\{\boldsymbol{\alpha}^{(j)}\}_{j=1}^{n}$ and if $\mathcal{T}_{++} = \{\mathbf{y} \in \mathbb{R}_{++}^{\ell} \mid v_j(\mathbf{y}) - w_j(\mathbf{y}) = 0, \ j = n+1, \ldots, \ell\}$, then $\mathrm{cl}(\mathcal{T}_{++}) = \mathcal{T}_+$, where these sets are defined in (A.2), (A.3).*

Proof. As $\{\boldsymbol{\alpha}^{(j)}\}_{j=n+1}^{\ell}$ lie in the cone generated by $\{\boldsymbol{\alpha}^{(j)}\}_{j=1}^{n}$, the constraint set $\mathcal{T}_{++} = \{\mathbf{y} \in \mathbb{R}_{++}^{\ell} \mid v_j(\mathbf{y}) - w_j(\mathbf{y}) = 0, \ j = n+1, \ldots, \ell\}$ is specified via monomials $v_j(\mathbf{y}), w_j(\mathbf{y}), j = n+1, \ldots, \ell$ with the property that each $v_j(\mathbf{y})$ is a monomial consisting *only* of the variable $\mathbf{y}_j$ (i.e., just $\mathbf{y}_j$ raised to a positive integer) and each $w_j(\mathbf{y})$ is a monomial consisting of (a subset of) the variables $\mathbf{y}_1, \ldots, \mathbf{y}_n$. This follows directly from the assumption that $\{\boldsymbol{\alpha}^{(j)}\}_{j=n+1}^{\ell}$ lie in the cone generated by $\{\boldsymbol{\alpha}^{(j)}\}_{j=1}^{n}$, and from the construction of the monomials $v_j(\mathbf{y}), w_j(\mathbf{y})$ preceding (A.2). Consequently, one can check that for each point in $\mathcal{T}_+$ that lies on a face of the nonnegative orthant $\mathbb{R}_+^{\ell}$, the point can be approached as a limit of a sequence in $\mathcal{T}_{++}$. Hence, $\mathrm{cl}(\mathcal{T}_{++}) = \mathcal{T}_+$. $\qquad\square$

After the implicitization step and the appeal to Theorem A.3 to obtain representations of positive polynomials, we finally derive a conclusion in terms of signomials as follows. We apply the reverse transformation $\mathbf{y}_j \leftarrow \exp\{\boldsymbol{\alpha}^{(j)'}\mathbf{x}\}$ for each $j = 1, \ldots, \ell$, which is a valid parametrization of the system of polynomial equations $v_j(\mathbf{y}) - w_j(\mathbf{y}) = 0, j = n+1, \ldots, \ell$ for $\mathbf{y} \in \mathbb{R}_{++}^{\ell}$. The validity of this step relies on the fact that $\{\boldsymbol{\alpha}^{(j)}\}_{j=1}^{n}$ are linearly independent.

A.2.2. SAGE functions and redundant constraints as polynomials. The above discussion implies that a nonnegative signomial $\sum_{j=1}^{\ell} \mathbf{c}_j \exp\{\boldsymbol{\alpha}^{(j)'}\mathbf{x}\}$ defined with respect to rational exponents $\{\boldsymbol{\alpha}^{(j)}\}_{j=1}^{\ell} \subset \mathbb{Q}^n$ can also be viewed as a linear function $\mathbf{c}'\mathbf{y}$ that is nonnegative over $\mathrm{cl}(\mathcal{T}_{++})$, where $\mathcal{T}_{++} = \{\mathbf{y} \in \mathbb{R}_{++}^{\ell} \mid v_j(\mathbf{y}) = w_j(\mathbf{y}), j = n+1, \ldots, \ell\}$ is constructed as in the description preceding (A.2). A SAGE function is simply a nonnegative linear function over $\mathrm{cl}(\mathcal{T}_{++})$ with an efficiently computable nonnegativity certificate. In a similar vein, nonnegative signomials (SAGE functions) defined with respect to exponents in $E_p(\{\boldsymbol{\alpha}^{(j)}\}_{j=1}^{\ell})$ can be viewed as degree-p polynomials that are nonnegative (efficiently certified as being nonnegative) over $\mathrm{cl}(\mathcal{T}_{++})$ and vice versa. As a result of this correspondence, we use overloaded terminology to refer to SAGE functions both in the domain of signomials (distinguished by the use of the variable $\mathbf{x}$) and in the domain of polynomials (using the variable $\mathbf{y}$). In an analogous manner, we also view elements of the set of redundant constraints $R_q(C)$ as degree-q polynomials in the variables $\mathbf{y}_1, \ldots, \mathbf{y}_{\ell}$ that are constrained to be nonnegative over $\mathrm{cl}(\mathcal{T}_{++})$.

A.3. Proof of Theorem 4.2. Let $f(\mathbf{x}) = \sum_{j=1}^{\ell} \mathbf{c}_j^{(f)} \exp\{\boldsymbol{\alpha}^{(j)'}\mathbf{x}\}$ and let each $g_i(\mathbf{x}) = \sum_{j=1}^{\ell} \mathbf{c}_j^{(g_i)} \exp\{\boldsymbol{\alpha}^{(j)'}\mathbf{x}\}$ for $i = 1, \ldots, m$. Set $\mathbf{y}_j = \exp\{\boldsymbol{\alpha}^{(j)'}\mathbf{x}\}$ for $j =$

$1, \ldots, \ell$. The assumption that constraints of the form $U \geq \exp\{\boldsymbol{\alpha}^{(j)'}\mathbf{x}\} \geq L$ for each $j = 1, \ldots, \ell$ with $U, L \in \mathbb{R}_{++}$ are explicitly specified in C implies that $\{\mathbf{y} \in \mathbb{R}^\ell \mid \mathbf{c}^{(g_i)'}\mathbf{y} \geq 0, \ i = 1, \ldots, m\} \subset \mathbb{R}^\ell_{++}$. Hence, we appeal to Lemma A.4 to conclude that $\inf_{\mathbf{y} \in \mathcal{T}_{++}} \mathbf{c}^{(f)'}\mathbf{y} = \inf_{\mathbf{x} \in \mathcal{K}_C} f(\mathbf{x}) > 0$ is equivalent to $\inf_{\mathbf{y} \in \mathcal{T}_+} \mathbf{c}^{(f)'}\mathbf{y} > 0$, where the set $\mathcal{T}_+ = \{\mathbf{y} \in \mathbb{R}^\ell \mid \mathbf{c}^{(g_i)'}\mathbf{y} \geq 0, \ i = 1, \ldots, m\} \cap \{\mathbf{y} \in \mathbb{R}^\ell_+ \mid v_j(\mathbf{y}) - w_j(\mathbf{y}) = 0, \ j = n+1, \ldots, \ell\}$ for suitable monomials v_j, w_j (A.3).

Let $\tilde{P} \subset \mathbb{R}[\mathbf{Y}_1, \ldots, \mathbf{Y}_\ell]$ be the preprime generated by $\mathbb{R}_+, \mathbf{y}_1, \ldots, \mathbf{y}_\ell$, and $\mathbf{c}^{(g_1)'}\mathbf{y}$, $\ldots, \mathbf{c}^{(g_m)'}\mathbf{y}$. We note that $\tilde{P}$ is Archimedean as it contains $\mathbf{y}_j$ and $U - \mathbf{y}_j$ for each $j = 1, \ldots, \ell$, and hence the generators of the preprime are in the ring of bounded elements. Consider the preprime $P = \tilde{P} + I$, where I is the ideal generated by the polynomials $v_j(\mathbf{y}) - w_j(\mathbf{y})$ for $j = n+1, \ldots, \ell$. As $\tilde{P}$ is Archimedean, so is P.

Next consider the set $\tilde{M}$ formed by taking finite sums of elements from the set $\{s(\mathbf{y})h(\mathbf{y}) \mid s(\mathbf{y}) \in \mathrm{SAGE}(E_p(\{\boldsymbol{\alpha}^{(j)}\}_{j=1}^\ell)), \ h(\mathbf{y}) \in R_q(C), \ p, q \in \mathbb{Z}_+\}$; see section A.2.2 for discussion on this overloaded notation. The set of SAGE functions is closed under addition and also under multiplication by posynomials; thus, one can check that $\tilde{M}$ is a $\tilde{P}$-module. Consequently, $M = \tilde{M} + I$ is a P-module. Further, M is Archimedean as P is Archimedean. Finally, we observe that the nonnegativity set $\mathcal{K}_M = \{\mathbf{y} \in \mathbb{R}^\ell \mid r(\mathbf{y}) \geq 0 \text{ for all } r \in M\}$ is equal to $\mathcal{T}_+$ as the constraints specified by the polynomials in M are implied by those that specify the constraints in $\mathcal{T}_+$. Hence, we apply Theorem A.3 to conclude that $\mathbf{c}^{(f)'}\mathbf{y} \in M$. By substituting back $\mathbf{y}_j \leftarrow \exp\{\boldsymbol{\alpha}^{(j)'}\mathbf{x}\}$, the term corresponding to the ideal I vanishes and we obtain the desired result. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

A.4. Proof of Theorem 4.1. We begin with the following result on a constrained analogue of Polya's theorem [18, 22] regarding representations of forms that are positive on the simplex; a generalization of this statement is proved in [9], but the specialization stated here suffices for our purposes.

PROPOSITION A.6 (see [9]). *Let $u_k(\mathbf{y})$ for $k = 0, \ldots, s$ be a collection of homogeneous polynomials in $\mathbb{R}[\mathbf{Y}_1, \ldots, \mathbf{Y}_\ell]$. Let $\mathcal{V} = \{\mathbf{y} \in \mathbb{R}^\ell_+ \mid u_k(\mathbf{y}) = 0, k = 1, \ldots, s\}$. If $u_0(\mathbf{y}) > 0$ for all $\mathbf{y} \in \mathcal{V}\backslash\{\mathbf{0}\}$, then there exists $p \in \mathbb{Z}_+$ such that $(\sum_{j=1}^\ell \mathbf{y}_i)^p u_0(\mathbf{y}) = \tau(\mathbf{y}) + \theta(\mathbf{y})$ where $\tau(\mathbf{y})$ is a polynomial consisting of all nonnegative coefficients and $\theta(\mathbf{y})$ lies in the ideal generated by the polynomials $\{u_k(\mathbf{y})\}_{k=1}^s$.*

We use this result after suitably transforming our unconstrained SP to a polynomial optimization problem over the nonnegative orthant.

Proof of Theorem 4.1. The main idea in this proof is to transform the signomial $f(\mathbf{x}) = \sum_{j=1}^\ell \mathbf{c}_j \exp\{\boldsymbol{\alpha}^{(j)'}\mathbf{x}\}$ appropriately so that we can apply Proposition A.6.

The unconstrained SP $\inf_{\mathbf{x} \in \mathbb{R}^n} f(\mathbf{x})$ can be transformed to $\inf_{\mathbf{y} \in \mathcal{T}_+} \mathbf{c}'\mathbf{y}$ based on Lemma A.5, where $\mathcal{T}_+ = \{\mathbf{y} \in \mathbb{R}^\ell_+ \mid v_j(\mathbf{y}) = w(\mathbf{y}), j = n+1, \ldots, \ell\}$ and the construction of the monomials $v_j(\mathbf{y}), w_j(\mathbf{y})$ is described in the discussion preceding (A.2). We note that the monomials $v_j(\mathbf{y}), w_j(\mathbf{y})$ for $j = n+1, \ldots, \ell$ satisfy several properties:

(i) We have that $v_{n+1}(\mathbf{y}) = \mathbf{y}_{n+1}$ and $w_{n+1}(\mathbf{y}) = 1$ because $\boldsymbol{\alpha}^{(n+1)} = \mathbf{0}$.

(ii) For each $j = n+2, \ldots, \ell$, the monomial $v_j(\mathbf{y})$ is *only* a function of $\mathbf{y}_j$ and the monomial $w_j(\mathbf{y})$ is *only* a function of $\mathbf{y}_1, \ldots, \mathbf{y}_n$. This follows because each $\boldsymbol{\alpha}^{(j)}$ is a convex combination with rational weights of the exponents $\{\boldsymbol{\alpha}^{(i)}\}_{i=1}^{n+1}$ for $j = n+2, \ldots, \ell$.

(iii) Finally, $\deg(v_j) > \deg(w_j)$ for each $j = n+2, \ldots, \ell$, where $\deg(\cdot)$ denotes the degree of a polynomial. This is a consequence of the assumption that $\boldsymbol{\alpha}^{(j)} \notin \mathrm{conv}(\{\boldsymbol{\alpha}^{(i)}\}_{i=1}^n)$ for each $j = n+2, \ldots, \ell$.

Now define modified monomials:

$$\text{(A.4)} \qquad \tilde{w}_j(\mathbf{y}) = w_j(\mathbf{y})\mathbf{y}_{n+1}^{\deg(v_j)-\deg(w_j)}, \ j = n+2,\ldots,\ell,$$

and consider the set $\tilde{\mathcal{T}}_+ = \{\mathbf{y} \in \mathbb{R}_+^{\ell} \mid v_j(\mathbf{y}) = \tilde{w}_j(\mathbf{y}), j = n+2,\ldots,\ell\}$. The difference between $\tilde{\mathcal{T}}_+$ and $\mathcal{T}_+$ is that each $w_j(\mathbf{y})$ is replaced by $\tilde{w}_j(\mathbf{y})$, and there is no constraint corresponding to $j = n+1$ (i.e., the constraint $\mathbf{y}_{n+1} = 1$ from property (i) above is removed). Further, note that the polynomials $v_j(\mathbf{y}) - \tilde{w}_j(\mathbf{y})$ for $j = n+2,\ldots,\ell$ are homogenous polynomials based on property (iii) above.

As the polynomials $v_j(\mathbf{y}) - \tilde{w}_j(\mathbf{y})$ for $j = n+2,\ldots,\ell$ are homogenous, we intend to apply Proposition A.6 to the polynomial optimization problem $\inf_{\mathbf{y}\in\tilde{\mathcal{T}}_+} \mathbf{c}'\mathbf{y}$. In order to do so, we need to verify that $\mathbf{c}'\mathbf{y} > 0$ for all $\mathbf{y} \in \tilde{\mathcal{T}}_+\backslash\{\mathbf{0}\}$.

First, suppose that $\mathbf{y} \in \tilde{\mathcal{T}}_+\backslash\{\mathbf{0}\}$ with $\mathbf{y}_{n+1} > 0$. As the polynomials $v_j(\mathbf{y}) - \tilde{w}_j(\mathbf{y})$ for $j = n+2,\ldots,\ell$ and the linear function $\mathbf{c}'\mathbf{y}$ are homogenous, we may scale $\mathbf{y}$ so that $\mathbf{y}_{n+1} = 1$ since this not affect our conclusions about the positivity of $\mathbf{c}'\mathbf{y}$. Setting $\mathbf{y}_{n+1} = 1$ corresponds to verifying the positivity of $\mathbf{c}'\mathbf{y}$ over $\mathcal{T}_+$. As $\inf_{\mathbf{y}\in\mathcal{T}_+} \mathbf{c}'\mathbf{y} = \inf_{\mathbf{x}\in\mathbb{R}^n} f(\mathbf{x}) > 0$, we have that $\mathbf{c}'\mathbf{y} > 0$ for all $\mathbf{y} \in \tilde{\mathcal{T}}_+\backslash\{\mathbf{0}\}$ whenever $\mathbf{y}_{n+1} \neq 0$.

Next, suppose that $\mathbf{y} \in \tilde{\mathcal{T}}_+\backslash\{\mathbf{0}\}$ with $\mathbf{y}_{n+1} = 0$. The definition of $\tilde{\mathcal{T}}_+$ directly implies that each $\mathbf{y}_j = 0$ for $j = n+2,\ldots,\ell$, because $v_j(\mathbf{y})$ for $j = n+2,\ldots,\ell$ is only a monomial consisting of $\mathbf{y}_j$ raised to a positive integer (from property (ii) above) and $\tilde{w}_j(\mathbf{y})$ consists of $\mathbf{y}_{n+1}$ raised to a positive integer (from property (ii) above and from (A.4)). Consequently, we need to verify the positivity of $\mathbf{c}'\mathbf{y}$ with $\mathbf{y} \in \mathbb{R}_+^{\ell}\backslash\{\mathbf{0}\}$ and with $\mathbf{y}_j = 0$ for $j = n+1,\ldots,\ell$. Since $\mathbf{c}_j > 0$ for $j = 1,\ldots,n$ by assumption, it follows that $\mathbf{c}'\mathbf{y} > 0$ for all $\mathbf{y} \in \tilde{\mathcal{T}}_+\backslash\{\mathbf{0}\}$ with $\mathbf{y}_{n+1} = 0$.

Combining these observations, we appeal to Proposition A.6 to conclude that $(\sum_{j=1}^{\ell} \mathbf{y}_i)^p \mathbf{c}'\mathbf{y} = \tau(\mathbf{y}) + \theta(\mathbf{y})$, where $\tau(\mathbf{y})$ is a polynomial consisting of all nonnegative coefficients and $\theta(\mathbf{y})$ lies in the ideal generated by the polynomials $\{v_j(\mathbf{y}) - \tilde{w}_j(\mathbf{y})\}_{j=n+2}^{\ell}$. By substituting $\mathbf{y}_{n+1} \leftarrow 1$ and $\mathbf{y}_j \leftarrow \exp\{\boldsymbol{\alpha}^{(j)'}\mathbf{y}\}$ for $j = n+2,\ldots,\ell$, and by noting that posynomials are SAGE functions, we have the desired result. $\quad\square$

REFERENCES

[1] G. Blekherman, *There are significantly more nonnegative polynomials than sums of squares*, Israel J. Math., 153 (2006), pp. 355–380.
[2] S. Boyd, S. J. Kim, D. Patil, and M. Horowitz, *Digital circuit optimization via geometric programming*, Oper. Res., 53 (2005), pp. 899–932.
[3] S. Boyd, S. J. Kim, L. Vandenberghe, and A. Hassibi, *A tutorial on geometric programming*, Optim. Eng., 8 (2007), pp. 67–127.
[4] S. P. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge University Press, Cambridge, 2004.
[5] M. Chiang, *Geometric programming for communication systems*, Found. Trends Commun. Inform. Theory, 2 (2005), pp. 1–154.
[6] M. Chiang and S. Boyd, *Geometric programming duals of channel capacity and rate distortion*, IEEE Trans. Inform. Theory, 50 (2004), pp. 245–258.
[7] M. D. Choi and T. Y. Lam, *Extremal positive semidefinite forms*, Math. Ann., 231 (1977), pp. 1–18.
[8] T. Cover and J. Thomas, *Elements of Information Theory*, Wiley, New York, 2006.
[9] P. J. C. Dickinson and J. Povh, *On an extension of Polya's positivstellensatz*, J. Global Optim., (2014).
[10] R. J. Duffin and E. L. Peterson, *Geometric programming with signomials*, J. Optim. Theory Appl., 11 (1973), pp. 3–35.

[11] R. J. Duffin, E. L. Peterson, and C. M. Zener, *Geometric Programming: Theory and Application*, Wiley, New York, 1967.

[12] M. Ghasemi and M. Marshall, *Lower bounds for polynomials using geometric programming*, SIAM J. Optim., 22 (2012), pp. 460–473.

[13] A. G. Khovanskii, *Fewnomials*, AMS, Providence, RI, 1991.

[14] J. L. Krivine, *Anneaux préordonnés*, J. Anal. Math., 12 (1964), pp. 307–326.

[15] J. B. Lasserre, *Global optimization with polynomials and the problem of moments*, SIAM J. Optim., 11 (2001), pp. 796–817.

[16] N. Linial, A. Samorodnitsky, and A. Wigderson, *A deterministic strongly polynomial algorithm for matrix scaling and approximate permanents*, Combinatorica, 20 (2000).

[17] C. D. Maranas and C. A. Floudas, *Global optimization in generalized geometric programming*, Comput. Chem. Eng., 21 (1997), pp. 351–369.

[18] M. Marshall, *Positive Polynomials and Sums of Squares*, AMS, Providence, RI, 2008.

[19] P. A. Parrilo, *Structured Semidefinite Programs and Semialgebraic Geometry Methods in Robustness and Optimization*, Ph.D. thesis, California Institute of Technology, Pasadena, CA, 2000.

[20] P. A. Parrilo, *Semidefinite programming relaxations for semialgebraic problems*, Math. Program., 96 (2003), pp. 293–320.

[21] E. L. Peterson, *Geometric programming*, SIAM Rev., 18 (1976), pp. 1–51.

[22] G. Polya, *Uber positive darstellung von polynomen*, Vierterhartschrift Naturforschenden Gessellschaft Zurich, 73 (1928), pp. 141–145.

[23] M. Putinar, *Positive polynomials on compact semi-algebraic sets*, Indiana Univ. Math. J., 42 (1993), pp. 969–984.

[24] B. Reznick, *Forms derived from the arithmetic-geometric inequality*, Math. Ann., 283 (1989), pp. 431–464.

[25] B. Reznick, *Uniform Denominators in Hilbert's seventeenth problem*, Math. Z., 220 (1995), pp. 75–97.

[26] R. T. Rockafellar, *Convex Analysis*, Princeton University Press, Princeton, NJ, 1970.

[27] K. Schmudgen, *The K-moment problem for compact semi-algebraic sets*, Ann. of Math., 289 (1991), pp. 203–206.

[28] P. Shen, *Linearization method of global optimization for generalized geometric programming*, Appl. Math. Comput., 162 (2005), pp. 353–370.

[29] H. D. Sherali, *Global optimization of nonconvex polynomial programming problems having rational exponents*, J. Global Optim., 12 (1998), pp. 267–283.

[30] N. Z. Shor, *Class of global minimum bounds of polynomial functions*, Cybernetics, 23 (1987), pp. 731–734.

[31] A. Skajaa and Y. Ye, *A homogeneous interior-point algorithm for nonsymmetric convex conic optimization*, Math. Program., 150 (2014), pp. 391–422.

[32] G. Stengle, *A nullstellensatz and a positivstellensatz in semialgebraic geometry*, Ann. of Math., 207 (1974), pp. 87–97.

[33] K. Yun and C. Xi, *Second-order method of generalized geometric programming for spatial frame optimization*, Comput. Methods Appl. Mech. Engrg., 141 (1997), pp. 117–123.