

Tractable approximate robust geometric programming

Kan-Lin Hsiung · Seung-Jean Kim · Stephen Boyd

Received: 24 March 2006 / Accepted: 9 August 2007 / Published online: 4 October 2007
© Springer Science+Business Media, LLC 2007

Abstract The optimal solution of a geometric program (GP) can be sensitive to variations in the problem data. Robust geometric programming can systematically alleviate the sensitivity problem by explicitly incorporating a model of data uncertainty in a GP and optimizing for the worst-case scenario under this model. However, it is not known whether a general robust GP can be reformulated as a tractable optimization problem that interior-point or other algorithms can efficiently solve. In this paper we propose an approximation method that seeks a compromise between solution accuracy and computational efficiency.

The method is based on approximating the robust GP as a robust linear program (LP), by replacing each nonlinear constraint function with a piecewise-linear (PWL) convex approximation. With a polyhedral or ellipsoidal description of the uncertain data, the resulting robust LP can be formulated as a standard convex optimization problem that interior-point methods can solve. The drawback of this basic method is that the number of terms in the PWL approximations required to obtain an acceptable approximation error can be very large. To overcome the “curse of dimensionality” that arises in directly approximating the nonlinear constraint functions in the original robust GP, we form a conservative approximation of the original robust GP, which contains only bivariate constraint functions. We show how to find globally optimal PWL approximations of these bivariate constraint functions.

K.-L. Hsiung (✉) · S.-J. Kim · S. Boyd
Information Systems Laboratory, Electrical Engineering Department, Stanford University, Stanford,
CA 94305-9510, USA
e-mail: klhsiung@stanford.edu

S.-J. Kim
e-mail: sjkim@stanford.edu

S. Boyd
e-mail: boyd@stanford.edu

Keywords Geometric programming · Linear programming · Piecewise-linear function · Robust geometric programming · Robust linear programming · Robust optimization

1 Introduction

1.1 Geometric programming

The convex function $\mathbf{lse} : \mathbf{R}^k \rightarrow \mathbf{R}$, defined as

$$\mathbf{lse}(z_1, \dots, z_k) = \log(e^{z_1} + \dots + e^{z_k}), \quad (1)$$

is called the (k -term) *log-sum-exp* function. (We use the same notation, no matter what k is; the context will always unambiguously determine the number of exponential terms.) When $k = 1$, the log-sum-exp function reduces to the identity.

A *geometric program* (in convex form) has the form

$$\begin{aligned} & \text{minimize } c^T y \\ & \text{subject to } \mathbf{lse}(A_i y + b_i) \leq 0, \quad i = 1, \dots, m, \\ & \quad \quad \quad Gy + h = 0, \end{aligned} \quad (2)$$

where the optimization variable is $y \in \mathbf{R}^n$ and the problem data are $A_i \in \mathbf{R}^{K_i \times n}$, $b_i \in \mathbf{R}^{K_i}$, $c \in \mathbf{R}^n$, $G \in \mathbf{R}^{l \times n}$, and $h \in \mathbf{R}^l$. We call the inequality constraints in the GP (2) *log-sum-exp (inequality) constraints*. In many applications, GPs arise in *posynomial form*, and are then transformed by a standard change of coordinates and constraint functions to the convex form (2); see Appendix 1. This transformation does not in any way change the problem data, which are the same for the posynomial form and convex form problems.

Geometric programming has been used in various fields since the late 1960s; early applications of geometric programming can be found in the books Avriel (1980), Duffin et al. (1967), Zener (1971) and the survey papers Ecker (1980), Peterson (1976), Boyd et al. (2007). More recent applications can be found in various fields including circuit design (Boyd et al. 2005; Chen et al. 2000; Dawson et al. 2001; Daems et al. 2003; Hershenson 2002; Hershenson et al. 2001; Mohan et al. 2000; Sapatnekar 1996; Singh et al. 2005; Sapatnekar et al. 1993; Young et al. 2001), chemical process control (Wall et al. 1986), environment quality control (Greenberg 1995), resource allocation in communication systems (Dutta and Rama 1992), information theory (Chiang and Boyd 2004; Karlof and Chang 1997), power control of wireless communication networks (Kandukuri and Boyd 2002; O'Neill et al. 2006), queue proportional scheduling in fading broadcast channels (Seong et al. 2006), and statistics (Mazumdar and Jefferson 1983).

Algorithms for solving geometric programs appeared in the late 1960s, and research on this topic continued until the early 1990s; see, e.g., Avriel et al. (1975), Rajpogal and Bricker (1990). A huge improvement in computational efficiency was achieved in 1994, when Nesterov and Nemirovsky developed provably efficient

interior-point methods for many nonlinear convex optimization problems, including GPs (Nesterov and Nemirovsky 1994). A bit later, Kortanek, Xu, and Ye developed a primal-dual interior-point method for geometric programming, with efficiency approaching that of interior-point linear programming solvers (Kortanek et al. 1997).

1.2 Robust geometric programming

In *robust geometric programming* (RGP), we include an explicit model of uncertainty or variation in the data that defines the GP. We assume that the problem data (A_i, b_i) depend affinely on a vector of uncertain parameters u , that belongs to a set $\mathcal{U} \subseteq \mathbf{R}^L$:

$$(\tilde{A}_i(u), \tilde{b}_i(u)) = \left(A_i^0 + \sum_{j=1}^L u_j A_i^j, b_i^0 + \sum_{j=1}^L u_j b_i^j \right), \quad u \in \mathcal{U} \subseteq \mathbf{R}^L. \quad (3)$$

The data variation is described by $A_i^j \in \mathbf{R}^{K_i \times n}$, $b_i^j \in \mathbf{R}^{K_i}$, and the uncertainty set \mathcal{U} . We assume that all of these are known.

We consider two types of uncertainty sets. One is *polyhedral uncertainty*, in which \mathcal{U} is a polyhedron, *i.e.*, the intersection of a finite number of halfspaces:

$$\mathcal{U} = \{u \in \mathbf{R}^L \mid Du \preceq d\}, \quad (4)$$

where $d \in \mathbf{R}^K$, $D \in \mathbf{R}^{K \times L}$, and the symbol \preceq denotes the componentwise inequality between two vectors: $w \preceq v$ means $w_i \leq v_i$ for all i . The other is *ellipsoidal uncertainty*, in which \mathcal{U} is an ellipsoid:

$$\mathcal{U} = \{\bar{u} + P\rho \mid \|\rho\|_2 \leq 1, \rho \in \mathbf{R}^L\}, \quad (5)$$

where $\bar{u} \in \mathbf{R}^L$ and $P \in \mathbf{R}^{L \times L}$. Here, the matrix P describes the variation in u and can be singular, in order to model the situation when the variation in u is restricted to a subspace. Note that due to the affine structure in (3), the ellipsoid uncertainty set \mathcal{U} can be transformed to a unit ball (*i.e.*, P can be assumed to be an identity matrix) without loss of generality.

A (worst-case) robust GP (RGP) has the form

$$\begin{aligned} & \text{minimize } c^T y \\ & \text{subject to } \sup_{u \in \mathcal{U}} \mathbf{lse}(\tilde{A}_i(u)y + \tilde{b}_i(u)) \leq 0, \quad i = 1, \dots, m, \\ & Gy + h = 0. \end{aligned} \quad (6)$$

The inequality constraints in the RGP (6) are called *robust log-sum-exp (inequality) constraints*.

The RGP (6) is a special type of robust convex optimization problem; see, *e.g.*, Ben-Tal and Nemirovski (1998) for more on robust convex optimization. Unlike the various types of robust convex optimization problems that have been studied in the literature (*e.g.*, Ben-Tal and Nemirovski 1999; Ben-Tal et al. 2002; Ghaoui and Lebret 1997, 1998; Goldfarb and Iyengar 2003; Boni et al. 2007), the

computational tractability of the RGP (6) is not clear; it is not yet known whether one can reformulate a general RGP as a tractable optimization problem that interior-point or other algorithms can efficiently solve.

1.3 Brief overview and outline

We first observe that a log-sum-exp function can be approximated arbitrarily well by a piecewise-linear (PWL) convex function. Using these approximations, the RGP can be approximated arbitrarily well as a robust LP, with polyhedral or ellipsoidal data uncertainty. Since robust LPs, with polyhedral or ellipsoidal uncertainty, can be tractably solved (see Appendix 2), this gives us an approximation method for the RGP. In fact, this general approach can be used for any robust convex optimization problem with polyhedral or ellipsoidal uncertainty. Piecewise-linear approximation has been used in prior work on approximation methods for nonlinear convex optimization problems, since it allows us to approximately solve a nonlinear convex problem by solving a linear program; see, *e.g.*, Ben-Tal and Nemirovski (2001), Feijoo and Meyer (1988), Glineur (2000), Thakur (1978).

The problem with the basic PWL approach is that the number of terms needed in a PWL approximation of the log-sum-exp function (1), to obtain a given level of accuracy, grows rapidly with the dimension k . Thus, the size of the resulting robust LP is prohibitively large, unless all K_i are small. To overcome this “curse of dimensionality”, we propose the following approach. We first replace the RGP with a new RGP, in which each log-sum-exp function has only one or two terms. This transformation to a two-term GP is exact for a nonrobust GP, and conservative for a RGP. We then use the PWL approximation method on the reduced RGP.

In Sect. 2, we show how PWL approximation of the constraint functions in the RGP (6) leads to a robust LP. We also describe how to approximate a general RGP with a more tractable RGP which contains only bivariate constraint functions.

In Sect. 3, we develop a constructive algorithm to solve the best PWL convex lower and upper approximation problems for the bivariate log-sum-exp function. Some numerical examples are presented in Sect. 4. Our conclusions are given in Sect. 5. Supplementary material is collected in the appendices.

2 Solving robust GPs via PWL approximation

2.1 Robust LP approximation

Suppose we have PWL lower and upper bounds on the log-sum-exp function in the i th constraint of the RGP (6),

$$\max_{j=1,\dots,I_i} \{ \underline{f}_{ij}^T y + \underline{g}_{ij} \} \leq \mathbf{lse}(y) \leq \max_{j=1,\dots,I_i} \{ \overline{f}_{ij}^T y + \overline{g}_{ij} \}, \quad \forall y \in \mathbf{R}^{K_i},$$

where $\underline{f}_{ij}, \bar{f}_{ij} \in \mathbf{R}^{K_i}$ and $\underline{g}_{ij}, \bar{g}_{ij} \in \mathbf{R}$. Replacing the log-sum-exp functions in the RGP (6) with the PWL bounds above, we obtain the two problems

$$\begin{aligned} & \text{minimize } c^T y \\ & \text{subject to } \sup_{u \in \mathcal{U}} \max_{j=1, \dots, J_i} \{ \bar{f}_{ij}^T \tilde{A}_i(u) y + \bar{f}_{ij}^T \tilde{b}_i(u) + \bar{g}_{ij} \} \leq 0, \quad i = 1, \dots, m, \quad (7) \\ & \quad Gy + h = 0, \end{aligned}$$

and

$$\begin{aligned} & \text{minimize } c^T y \\ & \text{subject to } \sup_{u \in \mathcal{U}} \max_{j=1, \dots, I_i} \{ \underline{f}_{ij}^T \tilde{A}_i(u) y + \underline{f}_{ij}^T \tilde{b}_i(u) + \underline{g}_{ij} \} \leq 0, \quad i = 1, \dots, m, \quad (8) \\ & \quad Gy + h = 0. \end{aligned}$$

These problems can be reformulated as the robust LPs

$$\begin{aligned} & \text{minimize } c^T y \\ & \text{subject to } \sup_{u \in \mathcal{U}} \{ \bar{f}_{ij}^T \tilde{A}_i(u) y + \bar{f}_{ij}^T \tilde{b}_i(u) + \bar{g}_{ij} \} \leq 0, \quad i = 1, \dots, m, \quad j = 1, \dots, J_i, \\ & \quad Gy + h = 0, \quad (9) \end{aligned}$$

and

$$\begin{aligned} & \text{minimize } c^T y \\ & \text{subject to } \sup_{u \in \mathcal{U}} \{ \underline{f}_{ij}^T \tilde{A}_i(u) y + \underline{f}_{ij}^T \tilde{b}_i(u) + \underline{g}_{ij} \} \leq 0, \quad i = 1, \dots, m, \quad j = 1, \dots, I_i, \\ & \quad Gy + h = 0. \quad (10) \end{aligned}$$

With a polyhedral uncertainty set, these can be cast as (larger) LPs, and for ellipsoidal uncertainty sets, they can be cast as SOCPs; see Appendix 2.

Note that an optimal solution, say \bar{y} , of the robust LP (9) is also a feasible solution to the RGP (6). In other words, the robust LP (9) gives a conservative approximation of the RGP (6). The robust LP (10) has the opposite property: its feasible set covers the feasible set of the RGP (6). Therefore, the optimal value of the robust LP (10), say, $c^T \bar{y}$, gives a lower bound on the optimal value of the original RGP (6), and in particular, allows us to bound the error in the feasible, suboptimal point \bar{y} , for the RGP. In other words, we have

$$0 \leq c^T (\bar{y} - y^*) \leq c^T (\bar{y} - \underline{y}), \quad (11)$$

where y^* is an optimal solution of the RGP. Finally, it is not difficult to see that as the PWL convex approximations of the log-sum-exp functions are made finer, the optimal values of the robust LPs (9) and (10) get closer to that of the RGP (6).

2.2 Tractable robust GP approximation

The RGP (6) can be reformulated as another RGP

$$\begin{aligned}
 & \text{minimize } \bar{c}^T \eta \\
 & \text{subject to } \sup_{u \in \mathcal{U}} \mathbf{lse}((\bar{a}_{i1} + \bar{B}_{i1}u)^T \eta, \dots, (\bar{a}_{iK_i} + \bar{B}_{iK_i}u)^T \eta) \leq 0, \quad i = 1, \dots, m, \\
 & \quad \bar{G}\eta + \bar{h} = 0
 \end{aligned} \tag{12}$$

with the optimization variables $\eta = (y, t) \in \mathbf{R}^n \times \mathbf{R}$. Here the problem data

$$\begin{aligned}
 \bar{c} \in \mathbf{R}^{n+1}, \quad \bar{G} \in \mathbf{R}^{(l+1) \times (n+1)}, \quad \bar{h} \in \mathbf{R}^{l+1}, \\
 \bar{a}_{i_s} \in \mathbf{R}^{n+1}, \quad \bar{B}_{i_s} \in \mathbf{R}^{(n+1) \times L}
 \end{aligned}$$

can be readily obtained from the problem data of the RGP (6); see Appendix 3 for the details. The RGP (6) and (12) are equivalent: $\bar{y} \in \mathbf{R}^n$ is feasible to (6) if and only if $(\bar{y}, \bar{t}) \in \mathbf{R}^{n+1}$ is feasible to (12) for some $\bar{t} \in \mathbf{R}$. In the following we form a conservative approximation of the RGP (12), in which all the nonlinear constraint functions are bivariate.

Consider a k -term robust log-sum-exp constraint in the following generic form:

$$\sup_{u \in \mathcal{U}} \mathbf{lse}((a_1 + B_1u)^T \eta, \dots, (a_k + B_ku)^T \eta) \leq 0, \tag{13}$$

where $a_i \in \mathbf{R}^{n+1}$, $B_i \in \mathbf{R}^{(n+1) \times L}$. An approximate reduction procedure, described in Appendix 4, shows that $\eta \in \mathbf{R}^{n+1}$ satisfies (13) if there exists $z = (z_1, \dots, z_{k-2}) \in \mathbf{R}^{k-2}$ such that (η, z) satisfies the following system of $k - 1$ two-term robust log-sum-exp constraints:

$$\begin{aligned}
 & \sup_{u \in \mathcal{U}} \mathbf{lse}((a_1 + B_1u)^T \eta, z_1) \leq 0, \\
 & \sup_{u \in \mathcal{U}} \mathbf{lse}((a_{s+1} + B_{s+1}u)^T \eta - z_s, z_{s+1} - z_s) \leq 0, \quad s = 1, \dots, k - 3, \\
 & \sup_{u \in \mathcal{U}} \mathbf{lse}((a_{k-1} + B_{k-1}u)^T \eta - z_{k-2}, (a_k + B_ku)^T \eta - z_{k-2}) \leq 0,
 \end{aligned} \tag{14}$$

in which all the constraint functions are bivariate. We will call (14) a “two-term (conservative) approximation” of the k -term robust log-sum-exp constraint (13).

The idea of tractable RGP approximation is simple: we replace every robust log-sum-exp constraint (with more than two terms) by its two-term conservative approximation to obtain a “two-term RGP”, which gives a conservative approximation of the original RGP. Although with more variables and constraints, the two-term RGP is much more tractable, in the sense that we can approximate the bivariate log-sum-exp function well with a small number of hyperplanes, as described in Sect. 3. Then the two-term RGP can be further solved via robust LP approximation, as shown in Sect. 2.1.

Now we give an exact expression of the two-term RGP approximation. First note that a one-term robust log-sum-exp constraint is simply a robust linear inequality.

Since no PWL approximation for a one-term constraint is necessary, we can simply keep all the one-term constraints of a RGP in its two-term RGP approximation (and the consequent robust LP approximation). Therefore for simplicity, in the following we assume all the robust log-sum-exp constraints in RGP (12) have at least two terms, *i.e.*, $K_i \geq 2, i = 1, \dots, m$. The two-term RGP has the form

$$\begin{aligned} &\text{minimize } \hat{c}^T x \\ &\text{subject to } \sup_{u \in \mathcal{U}} \text{lse}((\hat{a}_i^1 + \hat{B}_i^1 u)^T x, (\hat{a}_i^2 + \hat{B}_i^2 u)^T x) \leq 0, \quad i = 1, \dots, K_c, \\ &\quad \hat{G}x + \hat{h} = 0, \end{aligned} \tag{15}$$

where the optimization variables are $x = (y, t, z) \in \mathbf{R}^n \times \mathbf{R} \times \mathbf{R}^{K_v}$, and the problem data are

$$\begin{aligned} \hat{a}_i^j &\in \mathbf{R}^{n+K_v+1}, \quad \hat{B}_i^j \in \mathbf{R}^{(n+K_v+1) \times L}, \quad i = 1, \dots, K_c, \quad j = 1, 2, \\ \hat{c} &= (\bar{c}, 0) \in \mathbf{R}^{n+K_v+1}, \quad \hat{G} = [\bar{G} \ 0] \in \mathbf{R}^{(l+1) \times (n+K_v+1)}, \quad \hat{h} = \bar{h} \in \mathbf{R}^{l+1}. \end{aligned}$$

Here $K_v = \sum_{i=1}^m (K_i - 2)$ is the number of additional variables and $K_c = \sum_{i=1}^m (K_i - 1)$ is the number of two-term log-sum-exp constraints.

With general uncertainty structures, the RGP (15) is a conservative approximation of the original RGP (6). In other words, if $\hat{x} = (\hat{y}, \hat{t}, \hat{z}) \in \mathbf{R}^n \times \mathbf{R} \times \mathbf{R}^{K_v}$ is feasible to (15), \hat{y} is feasible to (6). Hence the optimal value of the two-term RGP (15), if feasible, is an upper bound on that of the RGP (6).

3 PWL approximation of two-term log-sum-exp function

There has been growing interest in approximation and interpolation with convexity constraints (Beatson 1981, 1982; Gao et al. 1995; Hu 1991; McAllister and Roullier 1978). However, relatively little attention has been paid to the best PWL convex approximation problem for multivariate, or even bivariate, convex functions. (A heuristic method, based on the K -means clustering algorithm, is developed in Magnani and Boyd (2006).) In this section, the problem of finding the best PWL convex approximation of the two-term (*i.e.*, bivariate) log-sum-exp function is solved and a constructive algorithm is provided.

3.1 Definitions

Let $\text{int } X$ denote the interior of $X \subseteq \mathbf{R}^m$. A function $h : \mathbf{R}^m \rightarrow \mathbf{R}$ is called (*r-term*) *piecewise-linear* if there exists a partition of \mathbf{R}^m as

$$\mathbf{R}^m = X_1 \cup X_2 \cup \dots \cup X_r,$$

where $\text{int } X_i \neq \emptyset$ and $\text{int } X_i \cap \text{int } X_j = \emptyset$ for $i \neq j$, and a family of affine functions $a_1^T x + b_1, \dots, a_r^T x + b_r$ such that $h(x) = a_i^T x + b_i$ for $x \in X_i$. If an *r-term* PWL function h is convex, it can be expressed as the maximum of r affine functions:

$h(x) = \max\{a_1^T x + b_1, \dots, a_r^T x + b_r\}$. (See, e.g., Boyd and Vandenberghe (2004).) Let \mathcal{P}_r^m denote the set of r -term PWL convex functions from \mathbf{R}^m into \mathbf{R} . Note that $h \in \mathcal{P}_r^1$ if and only if there exist $x_i, i = 1, \dots, r - 1$ and $a_i, b_i, i = 1, \dots, r$ with $x_1 < \dots < x_{r-1}$ and $a_1 < \dots < a_r$ such that h can be expressed as

$$h(x) = \begin{cases} a_1 x + b_1, & x \in (-\infty, x_1], \\ a_i x + b_i, & x \in [x_{i-1}, x_i], \quad i = 2, \dots, r - 1, \\ a_r x + b_r, & x \in [x_{r-1}, \infty). \end{cases}$$

The points x_1, \dots, x_{r-1} are called the *break points* of h , and the affine functions $a_i x + b_i, i = 1, \dots, r$ are called the *segments*.

Let f be a continuous function from \mathbf{R}^m into \mathbf{R} . A function $h : \mathbf{R}^m \rightarrow \mathbf{R}$ is called an *r -term PWL convex lower (respectively, upper) approximation* to f if $h \in \mathcal{P}_r^m$ and $h(x) \leq f(x)$ (respectively, $h(x) \geq f(x)$) for all $x \in \mathbf{R}^m$. An r -term PWL convex lower (respectively, upper) approximation $\underline{f}_r \in \mathcal{P}_r^m$ (respectively, $\overline{f}_r \in \mathcal{P}_r^m$) to f is called a *best r -term PWL convex lower (respectively, upper) approximation* if it has the minimum approximation error in the uniform norm among all r -term PWL convex lower (respectively, upper) approximations to f , which is denoted by $\underline{\epsilon}_f(r)$ (respectively, $\overline{\epsilon}_f(r)$):

$$\begin{aligned} \underline{\epsilon}_f(r) &= \sup_{x \in \mathbf{R}^m} (f(x) - \underline{f}_r(x)) \\ &= \inf_{h \in \mathcal{P}_r^m} \left\{ \sup_{x \in \mathbf{R}^m} (f(x) - h(x)) \mid h(x) \leq f(x), \forall x \in \mathbf{R}^m \right\}, \\ \overline{\epsilon}_f(r) &= \sup_{x \in \mathbf{R}^m} (\overline{f}_r(x) - f(x)) \\ &= \inf_{h \in \mathcal{P}_r^m} \left\{ \sup_{x \in \mathbf{R}^m} (h(x) - f(x)) \mid h(x) \geq f(x), \forall x \in \mathbf{R}^m \right\}. \end{aligned}$$

3.2 Best PWL approximation of two-term log-sum-exp function

3.2.1 Equivalent univariate best approximation problem

Finding the best r -term PWL convex approximation to the two-term log-sum-exp function is a “bivariate” best approximation problem over \mathcal{P}_r^2 . In the following we show that this bivariate best approximation problem can be simplified as an equivalent “univariate” best approximation problem over \mathcal{P}_r^1 .

We define the function $\phi : \mathbf{R} \rightarrow \mathbf{R}$ as

$$\phi(x) = \log(1 + e^x). \tag{16}$$

Note that ϕ satisfies

$$\lim_{x \rightarrow -\infty} \phi(x) = \lim_{x \rightarrow \infty} (\phi(x) - x) = 0. \tag{17}$$

Thus,

$$\underline{\epsilon}_\phi(1) = \inf_{(a,b) \in \mathbf{R}^2} \sup_{x \in \mathbf{R}} (\phi(x) - ax - b) = \infty, \tag{18}$$

$$\epsilon_{\phi}(2) = \sup_{x \in \mathbf{R}} (\phi(x) - \max\{0, x\}) = \log 2. \tag{19}$$

Now, note that the two-term log-sum-exp function can be expressed as

$$\mathbf{lse}(y_1, y_2) = y_1 + \phi(y_2 - y_1) = y_2 + \phi(y_1 - y_2), \quad \forall (y_1, y_2) \in \mathbf{R}^2. \tag{20}$$

Therefore it follows from (18–20) that the two-term log-sum-exp function cannot be approximated by a single affine function with a finite approximation error over \mathbf{R}^2 , but has the unique best two-term PWL convex lower approximation $\underline{h}_2 : \mathbf{R}^2 \rightarrow \mathbf{R}$ and upper approximation $\bar{h}_2 : \mathbf{R}^2 \rightarrow \mathbf{R}$ defined as $\underline{h}_2(y_1, y_2) = \max\{y_1, y_2\}$ and $\bar{h}_2(y_1, y_2) = \max\{y_1 + \log 2, y_2 + \log 2\}$ respectively.

From now on, we restrict our discussion to the case $r \geq 3$. The following proposition establishes the uniqueness and some useful properties of the best r -term PWL convex lower approximation $\underline{\phi}_r$ to ϕ for $r \geq 3$.

Proposition 1 *For $r \geq 3$, there exist x_1, \dots, x_{r-1} and $(\underline{a}_i^*, \underline{b}_i^*) \in \mathbf{R}^2, i = 1, \dots, r - 2$ with*

$$x_1 < \dots < x_{r-1}, \quad 0 < \underline{a}_1^* < \underline{a}_2^* < \dots < \underline{a}_{r-2}^* < 1, \tag{21}$$

$$\underline{a}_i^* + \underline{a}_{r-i-1}^* = 1, \quad \underline{b}_i^* = \underline{b}_{r-i-1}^*, \quad i = 1, \dots, r - 2, \tag{22}$$

such that the function ϕ has the unique best r -term PWL convex lower approximation $\underline{\phi}_r$ defined as

$$\underline{\phi}_r(x) = \begin{cases} 0, & x \in (-\infty, x_1], \\ \underline{a}_i^*x + \underline{b}_i^*, & x \in [x_i, x_{i+1}], \quad i = 1, \dots, r - 2, \\ x, & x \in [x_{r-1}, \infty). \end{cases} \tag{23}$$

Moreover, there exist $\tilde{x}_1, \dots, \tilde{x}_{r-2} \in \mathbf{R}$ which satisfy

$$x_1 < \tilde{x}_1 < x_2 < \tilde{x}_2 < \dots < x_{r-2} < \tilde{x}_{r-2} < x_{r-1}$$

such that the segments $\underline{a}_1^*x + \underline{b}_1^*, \dots, \underline{a}_{r-2}^*x + \underline{b}_{r-2}^*$ are tangent to ϕ at the points $\tilde{x}_1, \dots, \tilde{x}_{r-2}$ respectively. Finally, the maximum approximation error occurs only at the break points of $\underline{\phi}_r$:

$$\begin{aligned} \phi(x) - \underline{\phi}_r(x) &< \epsilon_{\phi}(r), \quad x \notin \{x_1, \dots, x_{r-1}\}, \\ \phi(x_i) - \underline{\phi}_r(x_i) &= \epsilon_{\phi}(r), \quad i = 1, \dots, r - 1. \end{aligned}$$

The proof of Proposition 1 is straightforward but lengthy, due to many cases and sub-cases that have to be probed. The reader interested in the complete proof is referred to Hsiung et al. (2006).

As a consequence of Proposition 1 and (20), we have the following corollary.

Corollary 1 For $r \geq 3$, the unique best r -term PWL convex lower approximation $\underline{h}_r: \mathbf{R}^2 \rightarrow \mathbf{R}$ of the two-term log-sum-exp function is

$$\underline{h}_r(y_1, y_2) = \max\{y_1, \underline{a}_{r-2}^*y_1 + \underline{a}_1^*y_2 + \underline{b}_1^*, \underline{a}_{r-3}^*y_1 + \underline{a}_2^*y_2 + \underline{b}_2^*, \dots, \underline{a}_1^*y_1 + \underline{a}_{r-2}^*y_2 + \underline{b}_{r-2}^*, y_2\} \tag{24}$$

and the unique best r -term PWL convex upper approximation $\bar{h}_r: \mathbf{R}^2 \rightarrow \mathbf{R}$ is

$$\bar{h}_r(y_1, y_2) = \underline{h}_r(y_1, y_2) + \underline{\epsilon}_\phi(r), \tag{25}$$

where $a_i^*, b_i^*, i = 1, \dots, r - 2$ are the coefficients of the segments of $\underline{\phi}_r$ defined in (23).

The proof is given in Appendix 5.

This corollary shows that both the best r -term PWL convex upper and lower approximations to the two-term log-sum-exp function can be readily obtained, provided that $\underline{\phi}_r$ is given. Hence we can restrict our attention on solving the best PWL convex lower approximation problem for the univariate function ϕ .

3.2.2 Constructive algorithm

Proposition 1 implies that a function $h \in \mathcal{P}_r^1$ ($r \geq 3$) with $r - 1$ break points $x_1 < \dots < x_{r-1}$ solves the best PWL convex lower approximation problem for ϕ with approximation error $\epsilon \in (0, \log 2)$ (i.e., $h \equiv \underline{\phi}_r$ and $\epsilon = \underline{\epsilon}_\phi(r)$) if and only if

$$h(x) \leq \phi(x), \quad \forall x \in \mathbf{R}, \tag{26}$$

$$\lim_{x \rightarrow -\infty} h(x) - \phi(x) = 0, \quad \lim_{x \rightarrow \infty} h(x) - \phi(x) = 0, \tag{27}$$

$$h(x_i) - \phi(x_i) = \epsilon, \quad i = 2, \dots, r - 2, \tag{28}$$

$$x_1 = \log(e^\epsilon - 1), \tag{29}$$

$$x_{r-1} = -\log(e^\epsilon - 1), \tag{30}$$

and there exist $\tilde{x}_1, \dots, \tilde{x}_{r-2} \in \mathbf{R}$ such that

$$h(\tilde{x}_i) - \phi(\tilde{x}_i) = 0, \quad i = 1, \dots, r - 2, \tag{31}$$

$$x_1 < \tilde{x}_1 < x_2 < \tilde{x}_2 < \dots < x_{r-2} < \tilde{x}_{r-2} < x_{r-1}. \tag{32}$$

Using these properties of the best r -term best PWL convex lower approximation, for any given $\epsilon \in (0, \log 2)$ and $r \geq 3$, the following algorithm can verify if $\epsilon = \underline{\epsilon}_\phi(r)$ holds.

given $\epsilon \in (0, \log 2)$, $r \geq 3$
 define $\underline{x}_\epsilon = \log(e^\epsilon - 1)$ and $\bar{x}_\epsilon = -\log(e^\epsilon - 1)$
 $k := 1, x_1^\epsilon := \underline{x}_\epsilon$
repeat

1. find the line $y = a_k^\epsilon x + b_k^\epsilon$ passing through the point $(x_k^\epsilon, \phi(x_k^\epsilon) - \epsilon)$ and tangent to the curve $y = \phi(x)$ at a point $(\tilde{x}_k^\epsilon, \phi(\tilde{x}_k^\epsilon))$ with $\tilde{x}_k^\epsilon > x_k^\epsilon$

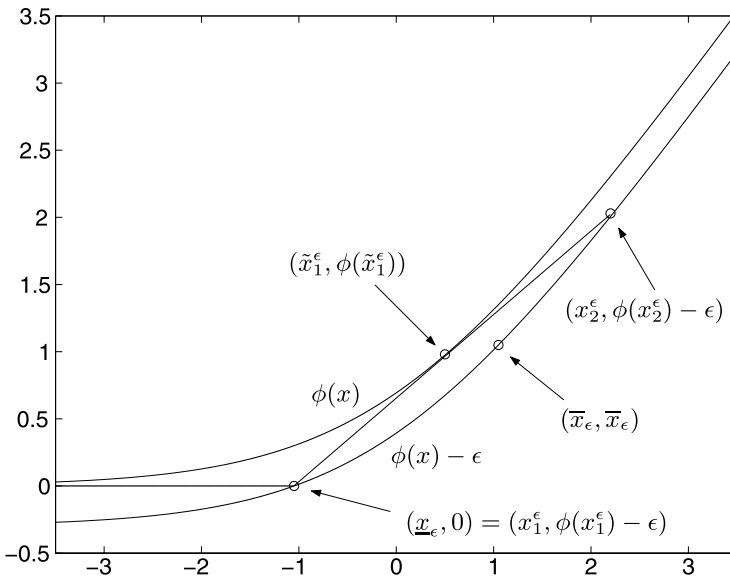


Fig. 1 An illustration of the algorithm which checks if $\epsilon = \epsilon_{\phi}(r)$ holds for given $\epsilon \in (0, \log 2)$ and $r \geq 3$. In this example we let $\epsilon = 0.3$ and $r = 3$. Since $x_2^\epsilon > \bar{x}_\epsilon$, we can conclude that $\epsilon_{\phi}(3) < 0.3$

2. find $x_{k+1}^\epsilon > \tilde{x}_k^\epsilon$ such that $a_k^\epsilon x_{k+1}^\epsilon + b_k^\epsilon = \phi(x_{k+1}^\epsilon) - \epsilon$
3. $k := k + 1$

until $k \geq r - 1$

This algorithm is illustrated in Fig. 1.

Now, define an r -term PWL convex function $h^\epsilon : \mathbf{R} \rightarrow \mathbf{R}$ as

$$h^\epsilon(x) = \max\{0, a_1^\epsilon x + b_1^\epsilon, \dots, a_{r-2}^\epsilon x + b_{r-2}^\epsilon, x\}.$$

Note that $x_1^\epsilon < \dots < x_{r-1}^\epsilon$ and h^ϵ satisfy (26–29), and $\tilde{x}_1^\epsilon < \dots < \tilde{x}_{r-2}^\epsilon$ satisfies (31–32). Thus $h^\epsilon \equiv \underline{\phi}_r$ if and only if (30) holds, which further implies

$$\epsilon = \epsilon_{\phi}(r) \iff x_{r-1}^\epsilon = \bar{x}_\epsilon. \tag{33}$$

Moreover, (30) implies

$$\epsilon < \epsilon_{\phi}(r) \iff x_{r-1}^\epsilon < \bar{x}_\epsilon, \tag{34}$$

$$\epsilon > \epsilon_{\phi}(r) \iff x_{r-1}^\epsilon > \bar{x}_\epsilon. \tag{35}$$

Observing (33–35), we can see that the following simple bisection algorithm finds $\epsilon_{\phi}(r)$ and $\underline{\phi}_r$ for any given $r \geq 3$.

given $r \geq 3$ and $\delta > 0$
 $\underline{\epsilon} := 0$ and $\bar{\epsilon} := \log 2$
repeat

1. $\epsilon := (\underline{\epsilon} + \bar{\epsilon})/2$
2. find the points $\underline{x}_\epsilon, \bar{x}_\epsilon$, the segments $a_k^\epsilon x + b_k^\epsilon, k = 1, \dots, r - 1$, and the break points $x_k^\epsilon, k = 1, \dots, r - 1$ by the algorithm described above
3. if $x_{r-1}^\epsilon > \bar{x}_\epsilon, \bar{\epsilon} := \epsilon$; otherwise, $\underline{\epsilon} := \epsilon$

until $|x_{r-1}^\epsilon - \bar{x}_\epsilon| \leq \delta$

let $\epsilon^\delta = \epsilon$ and define an r -term PWL convex function $\underline{\phi}_r^\delta : \mathbf{R} \rightarrow \mathbf{R}$ as

$$\underline{\phi}_r^\delta(x) = \max\{0, a_1^\epsilon x + b_1^\epsilon, \dots, a_{r-2}^\epsilon x + b_{r-2}^\epsilon, x\}$$

Here, it is easy to see that

$$\limsup_{\delta \rightarrow 0} \sup_{x \in \mathbf{R}} |\underline{\phi}_r^\delta(x) - \underline{\phi}_r(x)| = 0, \quad \lim_{\delta \rightarrow 0} \epsilon^\delta = \underline{\epsilon}_\phi(r),$$

i.e., $\delta > 0$ controls the tolerance.

3.2.3 Some approximation results

Table 1 shows the best r -term PWL convex lower approximation to the two-term log-sum-exp function for $r = 2, \dots, 5$ and the corresponding approximation error $\underline{\epsilon}_\phi(r)$. As will be shown in Sect. 4, the approximation method described in Sect. 2.2 with the five-term PWL convex lower approximation provides a quite accurate approximate solution for the RGP (6). In practical applications we are usually interested in r in the range $5 \leq r \leq 10$, but we can estimate the error decay rate for large r . Figure 2 shows the optimal error $\underline{\epsilon}_\phi(r)$ for $2 \leq r \leq 1000$. We observe that the curve is almost linear in log-log scale, and using a least-squares fit to the data points $(\log r, \log \underline{\epsilon}_\phi(r)), r = 2, \dots, 1000$, we obtain

$$\log \underline{\epsilon}_\phi(r) \approx -2.0215 \log r + 0.3457.$$

In normal scale,

$$\underline{\epsilon}_\phi(r) \approx \frac{1.4130}{r^{2.0215}} \leq \frac{\sqrt{2}}{r^2}.$$

4 Numerical examples

In the following we use some simple RGP numerical examples to demonstrate the robust LP approximation method described in Sect. 2.1. Practical engineering applications, such as power control in lognormal shadowing channels (Hsiung et al. 2005) and robust analog/RF circuit design (Yang et al. 2005), have been reported to reveal the effectiveness of the tractable robust GP approximation method proposed in Sect. 2.2.

Table 1 Some best PWL convex lower approximations to the two-term log-sum-exp function

r	Approximation Error $\underline{\epsilon}_\phi(r)$	Best r -Term PWL Convex Lower Approximation $\underline{\phi}_r$
2	0.693	$\max\{ y_1, y_2 \}$
3	0.223	$\max\{ y_1, 0.500y_1 + 0.500y_2 + 0.693, y_2 \}$
4	0.109	$\max\{ y_1, 0.271y_1 + 0.729y_2 + 0.584, 0.729y_1 + 0.271y_2 + 0.584, y_2 \}$
5	0.065	$\max\{ y_1, 0.167y_1 + 0.833y_2 + 0.450, 0.500y_1 + 0.500y_2 + 0.693, 0.833y_1 + 0.167y_2 + 0.450, y_2 \}$

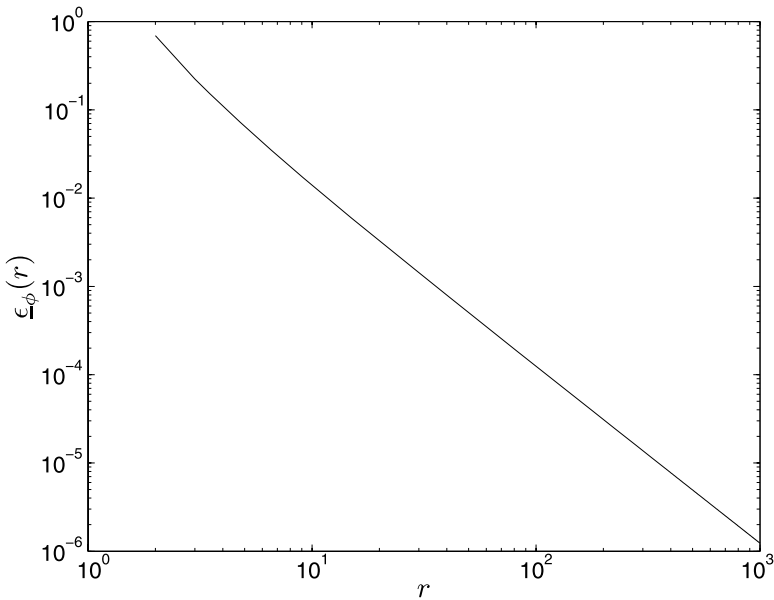


Fig. 2 Approximation error $\underline{\epsilon}_\phi(r)$ vs. the degree of PWL approximation r in log-log scale: $r = 2, \dots, 1000$

4.1 Two random families

We consider the following RGP, with 500 optimization variables, 500 two-term log-sum-exp inequality constraints, and no equality constraints:

$$\begin{aligned}
 R_L: \quad & \text{minimize } c^T y \\
 & \text{subject to } \sup_{u \in \mathcal{U}} \text{lse}((a_i^1 + B_i^1 u)^T y, (a_i^2 + B_i^2 u)^T y) \leq 0, \quad i = 1, \dots, 500.
 \end{aligned}
 \tag{36}$$

The optimization variable is $y \in \mathbf{R}^{500}$, $u \in \mathbf{R}^L$ represents the uncertain problem data, B_i^1 and B_i^2 are sparse matrices in $\mathbf{R}^{500 \times L}$, and

$$c = \mathbf{1} \in \mathbf{R}^{500}, \quad a_i^1 = a_i^2 = -\mathbf{1} \in \mathbf{R}^{500}.$$

Here, $\mathbf{1}$ is the vector with all entries one. The uncertainty set $\mathcal{U} \subseteq \mathbf{R}^L$ is given by the box in \mathbf{R}^L :

$$\mathcal{U} = \{u \in \mathbf{R}^L \mid \|u\|_\infty \leq 1\}, \tag{37}$$

where $\|u\|_\infty$ denotes the ℓ_∞ -norm of u .

We generated 20 feasible instances, R_5^1, \dots, R_5^{20} , of the RGP (36) with $L = 5$, by randomly generating the sparse matrices $B_i^1, B_i^2 \in \mathbf{R}^{500 \times 5}, i = 1, \dots, 500$ with sparsity density 0.1 and nonzero entries independently uniformly distributed on the interval $[-1, 1]$. The family $\{R_5^1, \dots, R_5^{20}\}$ is denoted by \mathcal{F}_5 . With $L = 20$, we also generated a family \mathcal{F}_{20} of 20 feasible instances, $R_{20}^1, \dots, R_{20}^{20}$, in a similar way.

4.2 Approximation results

Before presenting the approximation results for the two random families \mathcal{F}_5 and \mathcal{F}_{20} , we describe the error measure associated with the approximation method described in this paper.

Suppose the r -term PWL approximation of the two-term log-sum-exp function is used to obtain approximate solutions of the RGP (36). We call r the *degree of PWL approximation*, and call the solution \bar{y}_r of the robust LP (7) corresponding to the RGP (36) the *r-term upper approximate solution* and the solution \underline{y}_r of the robust LP (8) the *r-term lower approximate solution*. Let \bar{y}_r and y^* be an r -term upper approximate solution and an exact optimal solution of the RGP (36) respectively. Then, $e^{c^T y^*}$ is the optimal value of the corresponding RGP in *posynomial* form. To express the difference between $e^{c^T y^*}$ and $e^{c^T \bar{y}_r}$, we use the fractional difference in percentage α , given by

$$\alpha = 100 \left(\frac{e^{c^T \bar{y}_r}}{e^{c^T y^*}} - 1 \right) = 100(e^{c^T (\bar{y}_r - y^*)} - 1).$$

We call the value α the *r-term PWL approximation error (in percentage)* of the RGP (36).

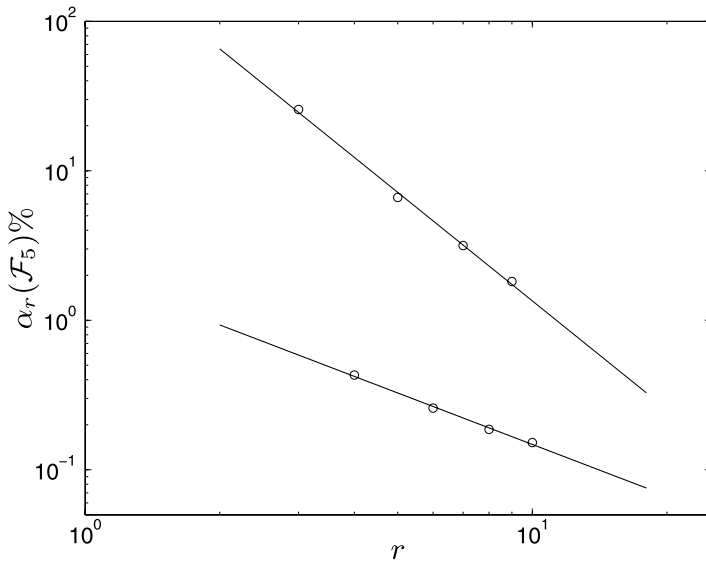


Fig. 3 Approximation results for the random family \mathcal{F}_5 : the degree of PWL approximation r vs. the mean $\alpha_r(\mathcal{F}_5)$ of the r -term PWL approximation errors in log-log scale. The *upper solid line* is obtained from linear least-squares fitting of the data points $(\log r, \log \alpha_r(\mathcal{F}_5))$, $r = 3, 5, 7, 9$ (shown as circles), while the *lower one* is obtained from linear least-squares fitting of the data points $(\log r, \log \alpha_r(\mathcal{F}_5))$, $r = 4, 6, 8, 10$

We first describe the approximation results for \mathcal{F}_5 . For each $r = 3, \dots, 10$, we found the r -term upper approximate solutions $\bar{y}_r(1), \dots, \bar{y}_r(20)$ of the randomly generated instances R_5^1, \dots, R_5^{20} . We also found the exact optimal solutions $y^*(1), \dots, y^*(20)$ of the instances, by solving the equivalent GPs with 16,000 inequality constraints obtained by replicating the inequality constraints for all vertices of the uncertainty box \mathcal{U} in (37).

Figure 3 shows the degree of PWL approximation r vs. the mean $\alpha_r(\mathcal{F}_5)$ of the r -term PWL approximation errors $100(e^{c^T(\bar{y}_r(i)-y^*(i))} - 1)$, $i = 1, \dots, 20$, where

$$\alpha_r(\mathcal{F}_5) = \frac{1}{20} \sum_{i=1}^{20} 100(e^{c^T(\bar{y}_r(i)-y^*(i))} - 1).$$

This figure shows that, in the region of interest, $\alpha_r(\mathcal{F}_5)$ decreases faster than quadratically with increasing r , since $\alpha_r(\mathcal{F}_5)$, $r = 3, 5, 7, 9$ decrease faster than quadratically. The variance of the r -term PWL approximation errors $100(e^{c^T(\bar{y}_r(i)-y^*(i))} - 1)$, $i = 1, \dots, 20$ was found to be less than 10^{-6} , regardless of r . The four-term PWL convex upper approximation therefore provides an approximate solution with less than 1% approximation error *quite consistently* for each of the randomly generated instances R_5^1, \dots, R_5^{20} .

Note that $\alpha_r(\mathcal{F}_5)$ does not decrease monotonically with increasing r . This is mainly because it does not necessarily hold that

$$r_1 \geq r_2 \implies \bar{h}_{r_2}(y_1, y_2) \geq \bar{h}_{r_1}(y_1, y_2) \geq \mathbf{lse}(y_1, y_2), \quad \forall (y_1, y_2) \in \mathbf{R}^2,$$

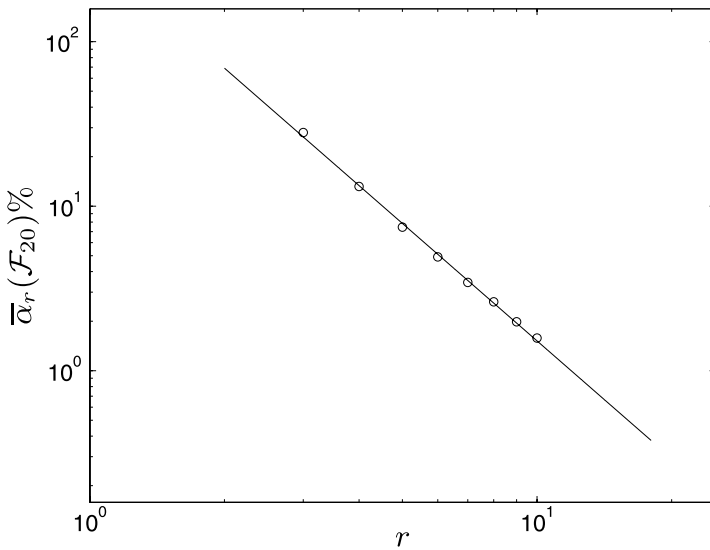


Fig. 4 Approximation results for the random family \mathcal{F}_{20} : the degree of PWL approximation r vs. the upper bound $\bar{\alpha}_r(\mathcal{F}_{20})$ on the mean $\alpha_r(\mathcal{F}_{20})$ of the r -term PWL approximation errors in log-log scale. The *solid line* is obtained from linear least-squares fitting of the data points $(\log r, \log \bar{\alpha}_r(\mathcal{F}_{20}))$, $r = 3, 4, \dots, 10$, shown as *circles*

although

$$r_1 \geq r_2 \implies \sup_{(y_1, y_2) \in \mathbf{R}^2} (\bar{h}_{r_1}(y_1, y_2) - \mathbf{lse}(y_1, y_2)) < \sup_{(y_1, y_2) \in \mathbf{R}^2} (\bar{h}_{r_2}(y_1, y_2) - \mathbf{lse}(y_1, y_2)),$$

where \bar{h}_r denotes the best r -term PWL convex upper approximation to the two-term log-sum-exp function.

We next describe the approximation results for \mathcal{F}_{20} . For each $r = 3, \dots, 10$, we found the r -term upper approximate solutions $\bar{y}_r(1), \dots, \bar{y}_r(20)$ of the randomly generated instances $R_{20}^1, \dots, R_{20}^{20}$. Replicating the inequality constraints for all the vertices was not possible for the random family \mathcal{F}_{20} , since the corresponding uncertainty box \mathcal{U} has approximately 10^6 vertices. Thus, it is too expensive to find the optimal solutions $y^*(1), \dots, y^*(20)$ of the instances $R_{20}^1, \dots, R_{20}^{20}$. Instead, we found the r -term lower approximate solutions $\underline{y}_r(1), \dots, \underline{y}_r(20)$ of the instances $R_{20}^1, \dots, R_{20}^{20}$ for each $r = 3, \dots, 10$.

Note from (11) that

$$0 \leq e^{c^T(\bar{y}_r(i) - y^*(i))} - 1 \leq e^{c^T(\bar{y}_r(i) - \underline{y}_r(i))} - 1, \quad i = 1, \dots, 20.$$

The mean $\bar{\alpha}_r(\mathcal{F}_{20})$ of the r -term approximation errors $e^{c^T(\bar{y}_r(i) - \underline{y}_r(i))} - 1, i = 1, \dots, 20$ is therefore an upper bound on the mean $\alpha_r(\mathcal{F}_{20})$ of the r -term approxi-

mation errors $e^{c^T(\bar{y}_r(i)-y^*(i))} - 1, i = 1, \dots, 20$:

$$\begin{aligned} \bar{\alpha}_r(\mathcal{F}_{20}) &= \frac{1}{20} \sum_{i=1}^{20} 100(e^{c^T(\bar{y}_r(i)-y_r(i))} - 1) \geq \alpha_r(\mathcal{F}_{20}) \\ &= \frac{1}{20} \sum_{i=1}^{20} 100(e^{c^T(\bar{y}_r(i)-y^*(i))} - 1). \end{aligned}$$

Figure 4 shows the degree of PWL approximation r vs. $\bar{\alpha}_r(\mathcal{F}_{20})$. This figure shows that, in the region of interest, $\bar{\alpha}_r(\mathcal{F}_{20})$ decreases faster than quadratically with increasing r . The variance of the upper bounds $100(e^{c^T(\bar{y}_r(i)-y_r(i))} - 1), i = 1, \dots, 20$ was found to be less than 10^{-4} , regardless of r . The seven-term PWL convex upper approximation therefore provides an approximate solution with less than 5% approximation error consistently for each of the instances $R_{20}^1, \dots, R_{20}^{20}$.

5 Conclusions

We have described an approximation method for a RGP with polyhedral or ellipsoidal uncertainty. The approximation method is based on conservatively approximating the original RGP (6) with a more tractable robust two-term GP in which every nonlinear function in the constraints is bivariate. The idea can be extended to a (small) k -term RGP approximation in which every nonlinear function in the constraints has at most k exponential terms. The extension relies on accurate PWL approximations of k -term log-sum-exp functions. We are currently working on the extension using the heuristic for PWL approximation of convex functions developed in Magnani and Boyd (2006).

Acknowledgements The authors are grateful to anonymous reviewers for helpful comments. This material is based upon work supported by the National Science Foundation under grants #0423905 and (through October 2005) #0140700, by the Air Force Office of Scientific Research under grant #F49620-01-1-0365, by MARCO Focus center for Circuit & System Solutions contract #2003-CT-888, by MIT DARPA contract #N00014-05-1-0700, and by the Post-doctoral Fellowship Program of Korea Science and Engineering Foundation (KOSEF). The authors thank Alessandro Magnani for helpful comments and suggestions.

Appendix 1: Convex formulation of GP

Let \mathbf{R}_{++}^n denote the set of real n -vectors whose components are positive. Let x_1, \dots, x_n be n real positive variables. A function $f : \mathbf{R}_{++}^n \rightarrow \mathbf{R}$, defined as

$$f(x) = d \prod_{j=1}^n x_j^{a_j}, \tag{38}$$

where $d > 0$ and $a_j \in \mathbf{R}$, is called a *monomial*. A sum of monomials, i.e., a function of the form

$$f(x) = \sum_{k=1}^K d_k \prod_{j=1}^n x_j^{a_{jk}}, \tag{39}$$

where $d_k > 0$ and $a_{jk} \in \mathbf{R}$, is called a *posynomial* (with K terms).

An optimization problem of the form

$$\begin{aligned} &\text{minimize } f_0(x) \\ &\text{subject to } f_i(x) \leq 1, \quad i = 1, \dots, m, \\ &\quad \quad \quad h_i(x) = 1, \quad i = 1, \dots, l, \end{aligned} \tag{40}$$

where f_0, \dots, f_m are posynomials and h_1, \dots, h_l are monomials, is called a *geometric program in posynomial form*. Here, the constraints $x_i > 0, i = 1, \dots, n$ are implicit. The corresponding robust convex optimization problem is called a *RGP in posynomial form*.

We assume without loss of generality that the objective function f_0 is a monomial whose coefficient is one:

$$f_0(x) = \prod_{j=1}^n x_j^{c_j}.$$

If f_0 is not a monomial, we can equivalently reformulate the GP (40) as the following GP whose objective function is a monomial:

$$\begin{aligned} &\text{minimize } t \\ &\text{subject to } f_0(x)t^{-1} \leq 1, \\ &\quad \quad \quad f_i(x) \leq 1, \quad i = 1, \dots, m, \\ &\quad \quad \quad h_i(x) = 1, \quad i = 1, \dots, l, \end{aligned}$$

where $(x, t) \in \mathbf{R}_{++}^n \times \mathbf{R}_{++}$ are the optimization variables.

GPs in posynomial form are not (in general) convex optimization problems, but they can be reformulated as convex problems by a change of variables and a transformation of the objective and constraint functions. To show this, we define new variables $y_i = \log x_i$, and take the logarithm of the posynomial f of x given in (39) to get

$$\tilde{f}(y) = \log(f(e^{y_1}, \dots, e^{y_n})) = \log\left(\sum_{i=1}^K e^{a_i^T y + b_i}\right) = \mathbf{lse}(a_1^T y + b_1, \dots, a_K^T y + b_K),$$

where $a_k = (a_{1k}, \dots, a_{nk}) \in \mathbf{R}^n$ and $b_k = \log d_k$, i.e., a posynomial becomes a sum of exponentials of affine functions after the change of variables. (Note that if the posynomial f is a monomial, then the transformed function \tilde{f} is an affine function.) This converts the original GP (40) into a GP:

$$\begin{aligned} &\text{minimize } c^T y \\ &\text{subject to } \mathbf{lse}(a_{i1}^T y + b_{i1}, \dots, a_{iK}^T y + b_{iK}) \leq 0, \quad i = 1, \dots, m, \\ &\quad \quad \quad g_i^T y + h_i = 0, \quad i = 1, \dots, l, \end{aligned} \tag{41}$$

where $a_{ij} \in \mathbf{R}^n, i = 1, \dots, m, j = 1, \dots, K_i$ contain the exponents of the posynomial inequality constraints, $c \in \mathbf{R}^n$ contains the exponents of the monomial objective function of the original GP, and $g_i \in \mathbf{R}^n, i = 1, \dots, l$ contain the exponents of the monomial equality constraints of the original GP.

Appendix 2: Robust linear programming

Consider the robust LP

$$\begin{aligned} &\text{minimize } c^T x \\ &\text{subject to } \sup_{u \in \mathcal{U}} (\bar{a}_i + B_i u)^T x + b_i \leq 0, \quad i = 1, \dots, m, \end{aligned} \tag{42}$$

where the optimization variable is $x \in \mathbf{R}^n, u \in \mathbf{R}^L$ represents the uncertain problem data, the set $\mathcal{U} \subseteq \mathbf{R}^L$ describes the uncertainty in u , and $c \in \mathbf{R}^n, \bar{a}_i \in \mathbf{R}^n, B_i \in \mathbf{R}^{n \times L}, b \in \mathbf{R}^m$. When the uncertainty set \mathcal{U} is given by a bounded polyhedron or an ellipsoid, the robust LP (42) can be cast as a standard convex optimization problem, as shown below.

7.1 Polyhedral uncertainty

Let the uncertainty set \mathcal{U} be a polyhedron:

$$\mathcal{U} = \{u \in \mathbf{R}^L \mid Du \preceq d\},$$

where $D \in \mathbf{R}^{K \times L}$ and $d \in \mathbf{R}^K$. We assume that \mathcal{U} is non-empty and bounded.

Using the duality theorem for linear programming, we can equivalently reformulate the robust LP (42) as the following LP:

$$\begin{aligned} &\text{minimize } c^T x \\ &\text{subject to } D^T z_i = B_i^T x, \quad i = 1, \dots, m, \\ &\quad \bar{a}_i^T x + d^T z_i + b_i \leq 0, \quad i = 1, \dots, m, \\ &\quad z_i \geq 0, \quad i = 1, \dots, m, \end{aligned} \tag{43}$$

where the optimization variables are $(x, z_1, \dots, z_m) \in \mathbf{R}^n \times \mathbf{R}^K \times \dots \times \mathbf{R}^K$.

7.2 Ellipsoidal uncertainty

Without loss of generality, we assume that the uncertainty set \mathcal{U} is a unit ball:

$$\mathcal{U} = \{u \in \mathbf{R}^L \mid \|u\|_2 \leq 1\}.$$

Then, the robust LP (42) can be cast as the second-order cone program

$$\begin{aligned} &\text{minimize } c^T x \\ &\text{subject to } \bar{a}_i^T x + \|B_i^T x\|_2 + b_i \leq 0, \quad i = 1, \dots, m. \end{aligned}$$

See, e.g., Lobo et al. (1998) for details.

Appendix 3: Reformulation of the robust GP

We start with reformulating the RGP (6) as the equivalent RGP

$$\begin{aligned} &\text{minimize } \bar{c}^T \begin{bmatrix} y \\ t \end{bmatrix} \\ &\text{subject to } \sup_{u \in \mathcal{U}} \mathbf{lse} \left(\left(\tilde{A}_i^0 + \sum_{j=1}^L u_j \tilde{A}_i^j \right) \begin{bmatrix} y \\ t \end{bmatrix} \right) \leq 0, \quad i = 1, \dots, m, \quad (44) \\ &\quad \tilde{G} \begin{bmatrix} y \\ t \end{bmatrix} + \bar{h} = 0, \end{aligned}$$

where $(y, t) \in \mathbf{R}^n \times \mathbf{R}$ are the optimization variables, and the problem data are

$$\bar{c} = (c, 0) \in \mathbf{R}^{n+1}, \quad \tilde{G} = \begin{bmatrix} G & 0 \\ 0 & 1 \end{bmatrix} \in \mathbf{R}^{(l+1) \times (n+1)}, \quad \bar{h} = \begin{bmatrix} h \\ -1 \end{bmatrix} \in \mathbf{R}^{l+1},$$

$$\tilde{A}_i^j = \begin{bmatrix} A_i^j & b_i^j \end{bmatrix} \in \mathbf{R}^{K_i \times (n+1)}, \quad i = 1, \dots, m, \quad j = 0, 1, \dots, L.$$

Denote the s th row of \tilde{A}_i^j as \tilde{a}_{is}^{jT} , $s = 1, \dots, K_i$, *i.e.*,

$$\tilde{A}_i^j = \begin{bmatrix} \tilde{a}_{i1}^{jT} \\ \vdots \\ \tilde{a}_{iK_i}^{jT} \end{bmatrix} \in \mathbf{R}^{K_i \times (n+1)}, \quad i = 1, \dots, m, \quad j = 0, 1, \dots, L.$$

Then the RGP (44) can be readily rewritten as the equivalent RGP (12) with the optimization variables $\eta = (y, t) \in \mathbf{R}^n \times \mathbf{R}$ and the problem data

$$\begin{aligned} \bar{a}_{is} &= \tilde{a}_{is}^0 \in \mathbf{R}^{n+1}, \\ \bar{B}_{is} &= [\tilde{a}_{is}^1 \dots \tilde{a}_{is}^L] \in \mathbf{R}^{(n+1) \times L}, \quad s = 1, \dots, K_i, \quad i = 1, \dots, m. \end{aligned}$$

Appendix 4: Details of the two-term robust GP approximation

Consider a k -term log-sum-exp constraint:

$$\sup_{u \in \mathcal{U}} \mathbf{lse}((a_1 + B_1 u)^T \eta, \dots, (a_k + B_k u)^T \eta) \leq 0,$$

where $a_i \in \mathbf{R}^{n+1}$, $B_i \in \mathbf{R}^{(n+1) \times L}$. It is easy to see that

$$\begin{aligned} &\sup_{u \in \mathcal{U}} \mathbf{lse}((a_1 + B_1 u)^T \eta, \dots, (a_k + B_k u)^T \eta) \\ &= \sup_{u \in \mathcal{U}} \mathbf{lse}((a_1 + B_1 u)^T \eta, \mathbf{lse}((a_2 + B_2 u)^T \eta, \dots, (a_k + B_k u)^T \eta)) \\ &\leq \sup_{u \in \mathcal{U}} \mathbf{lse} \left((a_1 + B_1 u)^T \eta, \sup_{u \in \mathcal{U}} \mathbf{lse}((a_2 + B_2 u)^T \eta, \dots, (a_k + B_k u)^T \eta) \right). \end{aligned}$$

Therefore a sufficient condition for the k -term robust log-sum-exp constraint (13) is that there exists $z_1 \in \mathbf{R}$ such that

$$\sup_{u \in \mathcal{U}} \mathbf{lse}((a_1 + B_1u)^T \eta, z_1) \leq 0, \quad \sup_{u \in \mathcal{U}} \mathbf{lse}((a_2 + B_2u)^T \eta, \dots, (a_k + B_ku)^T \eta) \leq z_1. \tag{45}$$

Similarly, since

$$\begin{aligned} & \sup_{u \in \mathcal{U}} \mathbf{lse}((a_2 + B_2u)^T \eta, \dots, (a_k + B_ku)^T \eta) \\ & \leq \sup_{u \in \mathcal{U}} \mathbf{lse} \left((a_2 + B_2u)^T \eta, \sup_{u \in \mathcal{U}} \mathbf{lse}((a_3 + B_3u)^T \eta, \dots, (a_k + B_ku)^T \eta) \right), \end{aligned}$$

a sufficient condition for (45) is that there exist $z_1, z_2 \in \mathbf{R}$ such that

$$\begin{aligned} & \sup_{u \in \mathcal{U}} \mathbf{lse}((a_1 + B_1u)^T \eta, z_1) \leq 0, \\ & \sup_{u \in \mathcal{U}} \mathbf{lse}((a_2 + B_2u)^T \eta, z_2) \leq z_1, \\ & \sup_{u \in \mathcal{U}} \mathbf{lse}((a_3 + B_3u)^T \eta, \dots, (a_k + B_ku)^T \eta) \leq z_2. \end{aligned}$$

Now it is easy to see that η satisfies (13) if there exists $z = (z_1, \dots, z_{k-2}) \in \mathbf{R}^{k-2}$ such that (η, z) satisfies the system of $k - 1$ two-term robust log-sum-exp constraints:

$$\begin{aligned} & \sup_{u \in \mathcal{U}} \mathbf{lse}((a_1 + B_1u)^T \eta, z_1) \leq 0, \\ & \sup_{u \in \mathcal{U}} \mathbf{lse}((a_{s+1} + B_{s+1}u)^T \eta, z_{s+1}) \leq z_s, \quad s = 1, \dots, k - 3, \\ & \sup_{u \in \mathcal{U}} \mathbf{lse}((a_{k-1} + B_{k-1}u)^T \eta, (a_k + B_ku)^T \eta) \leq z_{k-2}, \end{aligned}$$

which is obviously equivalent to (14).

Appendix 5: Proof of Corollary 1

The best PWL convex lower approximation problem for the two-term log-sum-exp function can be formulated as

$$\begin{aligned} & \text{minimize} \quad \sup_{(y_1, y_2) \in \mathbf{R}^2} \left(\mathbf{lse}(y_1, y_2) - \max_{i=1, \dots, r} \{f_{i1}y_1 + f_{i2}y_2 + g_i\} \right) \\ & \text{subject to} \quad \mathbf{lse}(y_1, y_2) \geq \max_{i=1, \dots, r} \{f_{i1}y_1 + f_{i2}y_2 + g_i\}, \quad \forall (y_1, y_2) \in \mathbf{R}^2, \end{aligned} \tag{46}$$

where $f_{i1}, f_{i2}, g_i \in \mathbf{R}$, $i = 1, \dots, r$ are the optimization variables. Here, note from (20) that

$$\begin{aligned} \mathbf{lse}(y_1, y_2) &= \max_{i=1, \dots, r} \{f_{i1}y_1 + f_{i2}y_2 + g_i\} \\ &= y_1 + \phi(y_2 - y_1) - \max_{i=1, \dots, r} \{f_{i1}y_1 + f_{i2}y_2 + g_i\} \\ &= \phi(y_2 - y_1) - \max_{i=1, \dots, r} \{(f_{i1} + f_{i2} - 1)y_1 + f_{i2}(y_2 - y_1) + g_i\}. \end{aligned}$$

Obviously, if

$$\sup_{(y_1, y_2) \in \mathbf{R}^2} \left(\mathbf{lse}(y_1, y_2) - \max_{i=1, \dots, r} \{f_{i1}y_1 + f_{i2}y_2 + g_i\} \right) < \infty,$$

then $f_{i1} + f_{i2} = 1, i = 1, \dots, r$. Hence (46) is equivalent to

$$\begin{aligned} &\text{minimize } \sup_{(y_1, y_2) \in \mathbf{R}^2} \left(y_1 + \phi(y_2 - y_1) - \max_{i=1, \dots, r} \{f_{i1}y_1 + f_{i2}y_2 + g_i\} \right) \\ &\text{subject to } y_1 + \phi(y_2 - y_1) \geq \max_{i=1, \dots, r} \{f_{i1}y_1 + f_{i2}y_2 + g_i\}, \quad \forall (y_1, y_2) \in \mathbf{R}^2, \quad (47) \\ &f_{i1} + f_{i2} = 1, \quad i = 1, \dots, r. \end{aligned}$$

This optimization problem is further equivalent to

$$\begin{aligned} &\text{minimize } \sup_{x \in \mathbf{R}} (\phi(x) - \max_{i=1, \dots, r} \{c_i x + d_i\}) \\ &\text{subject to } \phi(x) \geq \max_{i=1, \dots, r} \{c_i x + d_i\}, \quad \forall x \in \mathbf{R} \end{aligned} \quad (48)$$

in which $c_i, d_i \in \mathbf{R}, i = 1, \dots, r$ are the optimization variables. If $c_i^*, d_i^* \in \mathbf{R}, i = 1, \dots, r$ solve (48), then $f_{i1}^* = 1 - c_i^*, f_{i2}^* = c_i^*, g_i^* = d_i^*, i = 1, \dots, r$ solve (47). Conversely, if $f_{i1}^*, f_{i2}^*, g_i^* \in \mathbf{R}, i = 1, \dots, r$ solve (47), then $c_i^* = 1 - f_{i1}^* = f_{i2}^*, d_i^* = g_i^*, i = 1, \dots, r$ solve (48). Moreover, (47) and (48) have the same optimal value. Hence it is obvious from Proposition 1 that the two-term log-sum-exp function has the unique best r -term PWL convex lower approximation \underline{h}_r , given by (24).

We next show that the best r -term PWL convex upper approximation \bar{h}_r to the two-term log-sum-exp function can be obtained from (25). To see this, we cast the optimization problem (46) as

$$\begin{aligned} &\text{minimize } \epsilon \\ &\text{subject to } \max_{i=1, \dots, r} \{f_{i1}y_1 + f_{i2}y_2 + g_i\} \leq \mathbf{lse}(y_1, y_2), \quad \forall (y_1, y_2) \in \mathbf{R}^2, \quad (49) \\ &\mathbf{lse}(y_1, y_2) \leq \max_{i=1, \dots, r} \{f_{i1}y_1 + f_{i2}y_2 + g_i\} + \epsilon, \quad \forall (y_1, y_2) \in \mathbf{R}^2, \end{aligned}$$

which is obviously equivalent to

$$\begin{aligned} &\text{minimize } \epsilon \\ &\text{subject to } \max_{i=1, \dots, r} \{\tilde{f}_{i1}y_1 + \tilde{f}_{i2}y_2 + \tilde{g}_i\} - \epsilon \leq \mathbf{lse}(y_1, y_2), \quad \forall (y_1, y_2) \in \mathbf{R}^2, \quad (50) \\ &\mathbf{lse}(y_1, y_2) \leq \max_{i=1, \dots, r} \{\tilde{f}_{i1}y_1 + \tilde{f}_{i2}y_2 + \tilde{g}_i\}, \quad \forall (y_1, y_2) \in \mathbf{R}^2. \end{aligned}$$

If $\underline{\epsilon}, \underline{f}_{i1}, \underline{f}_{i2}, \underline{g}_i, i = 1, \dots, r$ solve (49) and $\bar{\epsilon}, \bar{f}_{i1}, \bar{f}_{i2}, \bar{g}_i, i = 1, \dots, r$ solve (50) respectively, then $\underline{\epsilon} = \bar{\epsilon} = \epsilon_\phi(r), \underline{f}_{i1} = \bar{f}_{i1}, \underline{f}_{i2} = \bar{f}_{i2}, \underline{g}_i = \bar{g}_i + \underline{\epsilon}, i = 1, \dots, r$. Here, note that the best PWL convex upper approximation problem for the two-term log-sum-exp function can be formulated exactly as (50). Finally, it is easy to see from the uniqueness of the best r -term PWL convex lower approximation to ϕ that the two-term log-sum-exp function has the unique best r -term PWL convex upper approximation \bar{h}_r , given by (25).

References

- Avriel M (ed) (1980) Advances in geometric programming. In: Mathematical concept and methods in science and engineering, vol 21. Plenum, New York
- Avriel M, Dembo R, Passy U (1975) Solution of generalized geometric programs. *Int J Numer Methods Eng* 9:149–168
- Beatson R (1981) Convex approximation by splines. *SIAM J Math Anal* 12:549–559
- Beatson R (1982) Monotone and convex approximation by splines: error estimates and a curve fitting algorithm. *SIAM J Numer Anal* 19:1278–1285
- Ben-Tal A, Nemirovski A (1998) Robust convex optimization. *Math Oper Res* 23(4):769–805
- Ben-Tal A, Nemirovski A (1999) Robust solutions of uncertain linear programs. *Oper Res Lett* 25:1–13
- Ben-Tal A, Nemirovski A (2001) On polyhedral approximations of the second-order cone. *Math Oper Res* 26(2):193–205
- Ben-Tal A, Nemirovski A, Roos C (2002) Robust solutions of uncertain quadratic and conic-quadratic problems. *SIAM J Optim* 13(2):535–560
- Boni O, Ben-Tal A, Nemirovski A (2007) Robust solutions to conic quadratic problems and their applications. *Optim Eng*, to appear
- Boyd S, Kim S-J, Patil D, Horowitz M (2005) Digital circuit sizing via geometric programming. *Oper Res* 53(6):899–932
- Boyd S, Kim S-J, Vandenberghe L, Hassibi A (2007) A tutorial on geometric programming. *Optim Eng* 8(1):67–127
- Boyd S, Vandenberghe L (2004) Convex optimization. Cambridge University Press, Cambridge
- Chen W, Hsieh C-T, Pedram M (2000) Simultaneous gate sizing and placement. *IEEE Trans Comput-Aided Des Integr Circuits Syst* 19(2):206–214
- Chiang M, Boyd S (2004) Geometric programming duals of channel capacity and rate distortion. *IEEE Trans Inf Theory* 50(2):245–258
- Daems W, Gielen G, Sansen W (2003) Simulation-based generation of posynomial performance models for the sizing of analog integrated circuits. *IEEE Trans Comput-Aided Des Integr Circuits Syst* 22(5):517–534
- Dawson J, Boyd S, Hershenson M, Lee T (2001) Optimal allocation of local feedback in multistage amplifiers via geometric programming. *IEEE Trans Circuits Syst I* 48(1):1–11
- Duffin R, Peterson E, Zener C (1967) Geometric programming—theory and application. Wiley, New York
- Dutta A, Rama D (1992) An optimization model of communications satellite planning. *IEEE Trans Commun* 40(9):1463–1473
- Ecker J (1980) Geometric programming: methods, computations and applications. *SIAM Rev* 22(3):338–362
- Feijoo B, Meyer P (1988) Piecewise-linear approximation methods for nonseparable convex optimization. *Manag Sci* 34(3):411–419
- Gao B, Newman D, Popov V (1995) Convex approximation by rational functions. *SIAM J Math Anal* 26(2):488–499
- Ghaoui LE, Lebret H (1997) Robust solutions to least-squares problems with uncertain data. *SIAM J Matrix Anal Appl* 18(4):1035–1064
- Ghaoui LE, Lebret H (1998) Robust solutions to uncertain semidefinite programs. *SIAM J Optim* 9(1):33–52
- Glione F (2000) Polyhedral approximation of the second-order cone: computational experiments. IMAGE Technical Report 0001, Faculté Polytechnique de Mons, Mons, Belgium

- Goldfarb D, Iyengar G (2003) Robust convex quadratically constrained programming. *Math Program* 97(3):495–515
- Greenberg H (1995) Mathematical programming models for environmental quality control. *Oper Res* 43(4):578–622
- Hershenson M (2002) Design of pipeline analog-to-digital converters via geometric programming. In: *Proceedings of the IEEE/ACM international conference on computer aided design*, San Jose, CA, pp 317–324
- Hershenson M, Boyd S, Lee TH (2001) Optimal design of a CMOS op-amp via geometric programming. *IEEE Trans Comput-Aided Des Integr Circuits Syst* 20(1):1–21
- Hsiung K-L, Kim S-J, Boyd S (2005) Power control in lognormal fading wireless channels with uptime probability specifications via robust geometric programming. In: *Proceedings 2005 IEEE American control conference*, Portland, OR, vol 6, pp 3955–3959
- Hsiung K-L, Kim S-J, Boyd S (2006) Tractable approximate robust geometric programming. Technical Report of Department of Electrical Engineering, Stanford University. Available at <http://www.stanford.edu/~boyd/rgp.html>
- Hu H (1991) Convexity preserving approximation by free knot splines. *SIAM J Math Anal* 22:1183–1191
- Kandukuri S, Boyd S (2002) Optimal power control in interference-limited fading wireless channels with outage-probability specifications. *IEEE Trans Wirel Commun* 1(1):46–55
- Karlov J, Chang Y (1997) Optimal permutation codes for the Gaussian channel. *IEEE Trans Inf Theory* 43(1):356–358
- Kortanek K, Xu X, Ye Y (1997) An infeasible interior-point algorithm for solving primal and dual geometric programs. *Math Program* 76(1):155–181
- Lobo M, Vandenberghe L, Boyd S, Lebret H (1998) Applications of second-order cone programming. *Linear Algebra Appl* 284(1–3):193–228
- Magnani A, Boyd S (2006) Convex piecewise-linear fitting. *Optim Eng*, submitted. Available at http://www.stanford.edu/~boyd/cvx_pwl_fit.html
- Mazumdar M, Jefferson T (1983) Maximum likelihood estimates for multinomial probabilities via geometric programming. *Biometrika* 70(1):257–261
- McAllister D, Roullier J (1978) Interpolation by convex quadratic splines. *Math Comput* 32:1154–1162
- Mohan S, Hershenson M, Boyd S, Lee T (2000) Bandwidth extension in CMOS with optimized on-chip inductors. *IEEE J Solid-State Circuits* 35(3):346–355
- Nesterov Y, Nemirovsky A (1994) Interior-point polynomial methods in convex programming. *Studies in applied mathematics*, vol 13. SIAM, Philadelphia
- O’Neill D, Julian D, Boyd S (2006) Seeking Foschini’s genie: optimal rates and powers in wireless networks. *IEEE Trans Veh Technol* (accepted for publication). Available at http://www.stanford.edu/~boyd/foschini_genie.html
- Peterson E (1976) Geometric programming. *SIAM Rev* 18(1):1–51
- Rajpogal J, Bricker D (1990) Posynomial geometric programming as a special case of semi-infinite linear programming. *J Optim Theory Appl* 66:444–475
- Sapatnekar S (1996) Wire sizing as a convex optimization problem: exploring the area-delay tradeoff. *IEEE Trans Comput-Aided Des Integr Circuits Syst* 15:1001–1011
- Sapatnekar S, Rao V, Vaidya P, Kang S (1993) An exact solution to the transistor sizing problem for CMOS circuits using convex optimization. *IEEE Trans Comput-Aided Des Integr Circuits Syst* 12(11):1621–1634
- Seong K, Narasimhan R, Cioffi JM (2006) Queue proportional scheduling via geometric programming in fading broadcast channels. *IEEE J Sel Areas Commun* 24(8):1593–1602
- Singh J, Nookala V, Luo Z-Q, Sapatnekar SS (2005) Robust gate sizing by geometric programming. In: *Proceedings 42nd ACM/IEEE design automation conference*, Anaheim, CA, pp 315–320
- Thakur L (1978) Error analysis for convex separable programs: the piecewise linear approximation and the bounds on the optimal objective value. *SIAM J Appl Math* 34(4):704–714
- Wall T, Greening D, Woolsey R (1986) Solving complex chemical equilibria using a geometric-programming based technique. *Oper Res* 34(3):345–355
- Yang X, Hsiung K-L, Li X, Nausieda I, Boyd S, Pileggi L (2005) OPERA: optimization with ellipsoidal uncertainty for robust analog IC design. In: *Proceedings 42nd IEEE/ACM design automation conference*, Anaheim, CA, pp 632–637
- Young F, Chu C, Luk W, Wong Y (2001) Handling soft modules in general nonslicing floorplan using Lagrangian relaxation. *IEEE Trans Comput-Aided Des Integr Circuits Syst* 20(5):687–629
- Zener C (1971) *Engineering design by geometric programming*. Wiley, New York