

# Relative entropy optimization and its applications

Venkat Chandrasekaran<sup>1</sup> · Parikshit Shah<sup>2</sup>

Received: 14 January 2015 / Accepted: 22 February 2016

© Springer-Verlag Berlin Heidelberg and Mathematical Optimization Society 2016

**Abstract** In this expository article, we study optimization problems specified via linear and relative entropy inequalities. Such relative entropy programs (REPs) are convex optimization problems as the relative entropy function is jointly convex with respect to both its arguments. Prominent families of convex programs such as geometric programs (GPs), second-order cone programs, and entropy maximization problems are special cases of REPs, although REPs are more general than these classes of problems. We provide solutions based on REPs to a range of problems such as permanent maximization, robust optimization formulations of GPs, and hitting-time estimation in dynamical systems. We survey previous approaches to some of these problems and the limitations of those methods, and we highlight the more powerful generalizations afforded by REPs. We conclude with a discussion of quantum analogs of the relative entropy function, including a review of the similarities and distinctions with respect to the classical case. We also describe a stylized application of quantum relative entropy optimization that exploits the joint convexity of the quantum relative entropy function.

**Keywords** Matrix permanent · Robust optimization · Dynamical systems · Quantum information · Quantum channel capacity · Shannon entropy · Von-Neumann entropy · Araki–Umegaki relative entropy · Golden–Thompson inequality · Optimization over non-commuting variables

---

✉ Venkat Chandrasekaran  
venkatc@caltech.edu

Parikshit Shah  
pshah@discovery.wisc.edu

<sup>1</sup> Departments of Computing and Mathematical Sciences and of Electrical Engineering, California Institute of Technology, Pasadena, CA 91125, USA

<sup>2</sup> Wisconsin Institutes for Discovery, University of Wisconsin, Madison, WI 53715, USA

**Mathematics Subject Classification** 81P45 · 90C25 · 94A15 · 94A17

## 1 Introduction

The relative entropy  $d(\mathbf{v}, \boldsymbol{\lambda}) = \sum_{i=1}^n v_i \log(v_i/\lambda_i)$  between two nonnegative vectors  $\mathbf{v}, \boldsymbol{\lambda} \in \mathbb{R}_+^n$  plays a prominent role in information theory and in statistics in the characterization of the performance of a variety of inferential procedures as well as in proofs of a number of fundamental inequalities [16]. In this expository article, we focus on the computational aspects of this function by considering *relative entropy programs* in which the objective and constraints are specified in terms of linear and relative entropy inequalities.

### 1.1 Relative entropy programs

Formally, relative entropy programs (REPs) are conic optimization problems in which a linear functional of a decision variable is minimized subject to linear constraints as well as a conic constraint specified by a *relative entropy cone*. The relative entropy cone is defined for triples  $(\mathbf{v}, \boldsymbol{\lambda}, \boldsymbol{\delta}) \in \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n$  via Cartesian products of the following elementary cone  $\mathcal{RE}_1 \subset \mathbb{R}^3$ :

$$\mathcal{RE}_1 = \left\{ (\nu, \lambda, \delta) \in \mathbb{R}_+ \times \mathbb{R}_+ \times \mathbb{R} \mid \nu \log\left(\frac{\nu}{\lambda}\right) \leq \delta \right\}. \quad (1)$$

As the function  $\nu \log(\nu/\lambda)$  is the perspective of the negative logarithm, which is a convex function, this cone is convex. More generally, the relative entropy cone  $\mathcal{RE}_n \subset \mathbb{R}^{3n}$  is defined as follows:

$$\mathcal{RE}_n = \mathcal{RE}_1^{\otimes n} = \left\{ (\mathbf{v}, \boldsymbol{\lambda}, \boldsymbol{\delta}) \in \mathbb{R}_+^n \times \mathbb{R}_+^n \times \mathbb{R}^n \mid v_i \log\left(\frac{v_i}{\lambda_i}\right) \leq \delta_i, \forall i \right\}. \quad (2)$$

Sublevel sets of the relative entropy function  $d(\mathbf{v}, \boldsymbol{\lambda})$  between two nonnegative vectors  $\mathbf{v}, \boldsymbol{\lambda} \in \mathbb{R}_+^n$  can be represented using the relative entropy cone  $\mathcal{RE}_n$  and a linear equality as follows:

$$d(\mathbf{v}, \boldsymbol{\lambda}) \leq t \Leftrightarrow \exists \boldsymbol{\delta} \in \mathbb{R}^n \text{ s.t. } (\mathbf{v}, \boldsymbol{\lambda}, \boldsymbol{\delta}) \in \mathcal{RE}_n, \mathbf{1}'\boldsymbol{\delta} = t.$$

REPs can be solved efficiently to a desired accuracy via interior-point methods due to the existence of computationally tractable barrier functions for the convex function  $\nu \log(\nu/\lambda)$  for  $\nu, \lambda \geq 0$  [44].

The relative entropy cone  $\mathcal{RE}_1$  is a reparametrization of the *exponential cone*  $\{(\nu, \lambda, \delta) \in \mathbb{R}_+ \times \mathbb{R}_+ \times \mathbb{R} \mid \exp(\frac{\delta}{\nu}) \leq \frac{\lambda}{\nu}\}$ , which has been studied previously [24]. One can easily see this based on the following relation for any  $(\nu, \lambda, \delta) \in \mathbb{R}_+ \times \mathbb{R}_+ \times \mathbb{R}$ :

$$\exp\left(\frac{\delta}{\nu}\right) \leq \frac{\lambda}{\nu} \Leftrightarrow (\nu, \lambda, -\delta) \in \mathcal{RE}_1.$$

However, we stick with our description (1) of  $\mathcal{RE}_1$  because it leads to a transparent generalization in the sequel to quantum relative entropy optimization problems (see Sect. 1.3 for a brief introduction and Sect. 4 for more details).

REPs offer a common generalization of a number of prominent families of convex optimization problems such as geometric programming (GP) [11,20], second-order cone programming (SOCP) [7,42], and entropy maximization.

*GPs as a special case of REPs* A GP in convex form is an optimization problem in which the objective and the constraints are specified by positive sums of exponentials:

$$\begin{aligned} \inf_{\mathbf{x} \in \mathbb{R}^n} \quad & \sum_{i=1}^k \mathbf{c}_i^{(0)} \exp(\mathbf{a}^{(i)'} \mathbf{x}) \\ \text{s.t.} \quad & \sum_{i=1}^k \mathbf{c}_i^{(j)} \exp(\mathbf{a}^{(i)'} \mathbf{x}) \leq 1, \quad j = 1, \dots, m. \end{aligned} \quad (3)$$

Here the exponents  $\mathbf{a}^{(i)} \in \mathbb{R}^n, i = 1, \dots, k$  and the coefficients  $\mathbf{c}^{(j)} \in \mathbb{R}_+^k, j = 0, \dots, m$  are fixed parameters. Applications of GPs include the computation of information-theoretic quantities [15], digital circuit gate sizing [10], chemical process control [55], matrix scaling and approximate permanent computation [41], entropy maximization problems in statistical learning [18], and power control in communication systems [14]. The GP above can be reformulated as follows based on the observation that each coefficient vector  $\mathbf{c}^{(j)}$  consists of nonnegative entries:

$$\begin{aligned} \inf_{\substack{\mathbf{x} \in \mathbb{R}^n \\ \mathbf{y}, \mathbf{z} \in \mathbb{R}^k}} \quad & \mathbf{c}^{(0)'} \mathbf{z} \\ \text{s.t.} \quad & \mathbf{c}^{(j)'} \mathbf{z} \leq 1, \quad j = 1, \dots, m \\ & \mathbf{y}_i \leq \log(\mathbf{z}_i), \quad i = 1, \dots, k \\ & \mathbf{a}^{(i)'} \mathbf{x} = \mathbf{y}_i, \quad i = 1, \dots, k. \end{aligned} \quad (4)$$

The set described by the constraints  $\mathbf{y}_i \leq \log(\mathbf{z}_i), i = 1, \dots, k$  can be specified using the relative entropy cone and affine constraints as  $(\mathbf{1}, \mathbf{z}, -\mathbf{y}) \in \mathcal{RE}_n$ , and consequently GPs are a subclass of REPs. We also refer the reader to [24] in which GPs are viewed as a special case of optimization problems involving the exponential cone.

*SOCPs as a special case of REPs* An SOCP is a conic optimization problem with respect to the following second-order, or Lorentz, cone [7] in  $\mathbb{R}^{k+1}$ :

$$L_k = \left\{ (\mathbf{x}, y) \in \mathbb{R}^k \times \mathbb{R} \mid \sqrt{\mathbf{x}_1^2 + \dots + \mathbf{x}_k^2} \leq y \right\}. \quad (5)$$

Applications of SOCPs include filter design in signal processing [46], portfolio optimization [42], truss system design [4], and robust least-squares problems in statistical inference [22]. It is well-known that any second-order cone  $L_k$  for  $k \geq 2$  can be represented via suitable Cartesian products of  $L_2$  [6]. In the “Appendix”, we show that the second-order cone  $L_2 \in \mathbb{R}^3$  can be specified using linear and relative entropy inequalities:

$$L_2 = \left\{ (\mathbf{x}, y) \in \mathbb{R}^2 \times \mathbb{R} \mid y - \mathbf{x}_1 \geq 0, y + \mathbf{x}_1 \geq 0 \text{ and } \exists v \in \mathbb{R}_+ \text{ s.t.} \right. \\ \left. v \log \left( \frac{v}{y + \mathbf{x}_1} \right) + v \log \left( \frac{v}{y - \mathbf{x}_1} \right) - 2v \leq 2\mathbf{x}_2 \right. \\ \left. v \log \left( \frac{v}{y + \mathbf{x}_1} \right) + v \log \left( \frac{v}{y - \mathbf{x}_1} \right) - 2v \leq -2\mathbf{x}_2 \right\}. \quad (6)$$

Consequently, SOCPs are also a special case of REPs.

Recall that linear programs (LPs) are a special case both of GPs and of SOCPs; as a result, we have that REPs also contain LPs as a special case. The relation between semidefinite programs (SDPs) and REPs is less clear. It is still an open question whether REPs contain SDPs as a special case. In the other direction, SDPs do not contain REPs as a special case; this follows from the observation that the boundary of the constraint set of an REP is not algebraic in general, whereas constraint sets of SDPs have algebraic boundaries.

As an illustration of the consequence of the observation that REPs contain GPs and SOCPs as special cases, we note that a variety of regularized logistic regression problems in machine learning can be recast as REPs. Logistic regression methods are widely used for identifying classifiers  $f : \mathbb{R}^k \rightarrow \{-1, +1\}$  of data points in  $\mathbb{R}^k$  based on a collection of labelled observations  $\{(\mathbf{x}^{(i)}, y^{(i)})\}_{i=1}^n \subset \mathbb{R}^k \times \{-1, +1\}$  [17, 32]. To identify good classification functions of the form  $f(\mathbf{x}) = \text{sign}(\mathbf{w}'\mathbf{x} + b)$  for some parameters  $\mathbf{w} \in \mathbb{R}^k, b \in \mathbb{R}$ , regularized logistic regression techniques entail the solution of convex relaxation problems of the following form:

$$\inf_{\mathbf{w} \in \mathbb{R}^k, b \in \mathbb{R}} \sum_{i=1}^n \log \left( 1 + \exp \left\{ -y^{(i)} \left[ \mathbf{w}'\mathbf{x}^{(i)} + b \right] \right\} \right) + \mu r(\mathbf{w})$$

for  $\mu > 0$  and for suitable regularization functions  $r : \mathbb{R}^k \rightarrow \mathbb{R}$ . Examples of convex regularization functions that are widely used in practice include  $r(\mathbf{w}) = \|\mathbf{w}\|_1$  and  $r(\mathbf{w}) = \|\mathbf{w}\|_2^2$ . Sublevel sets of the logistic loss function  $\log(1 + \exp\{a\})$  can be represented within the GP framework. In particular, for any  $(a, t) \in \mathbb{R}^2$ , we have that:

$$\log(1 + \exp\{a\}) \leq t \Leftrightarrow \exp\{-t\} + \exp\{a - t\} \leq 1.$$

On the other hand, sublevel sets of the regularizers  $\|\mathbf{w}\|_1$  and  $\|\mathbf{w}\|_2^2$  can be described via LPs and SOCPs. Consequently, regularized logistic regression problems with the  $\ell_1$ -norm or the squared- $\ell_2$ -norm regularizer are examples of REPs.

## 1.2 Applications of relative entropy optimization

Although REPs contain GPs and SOCPs as special cases, the relative entropy framework is more general than either of these classes of convex programs. Indeed, REPs are useful for solving a range of problems to which these other classes of convex

programs are not directly applicable. We illustrate the utility of REPs in the following diverse application domains:

1. *Permanent maximization* Given a collection of matrices, find the one with the largest permanent. Computing the permanent of a matrix is a well-studied problem that is believed to be computationally intractable. Therefore, we seek approximate solutions to the problem of permanent maximization.
2. *Robust GPs* The solution of a GP (3) is sensitive to the input parameters of the problem. Compute GPs within a robust optimization framework so that the solutions offer robustness to variability in the problem parameters.
3. *Hitting-times in dynamical systems* Given a linear dynamical system consisting of multiple modes and a region of feasible starting points, compute the smallest time required to hit a specified target set from an arbitrary feasible starting point.

We give a formal mathematical description of each of these problems in Sects. 2 and 3, and we describe computationally tractable solutions based on REPs. Variants of these questions as well as more restrictive formulations than the ones we consider have been investigated in previous work, with some techniques proposed for solving these problems based on GPs. We survey this prior literature, highlighting the limitations of previous approaches and emphasizing the more powerful generalizations afforded by the relative entropy formalism.

In Sect. 2, we describe an approximation algorithm for the permanent maximization problem based on an REP relaxation. We bound the quality of the approximation provided by our method via Van Der Waerden's inequality. We contrast our discussion with previous work in which GPs have been employed to approximate the permanent of a fixed matrix [41]. In Sect. 3 we describe an REP-based method for robust GPs in which the coefficients  $\mathbf{c}^{(j)}$ 's and the exponents  $\mathbf{a}^{(i)}$ 's in a GP (3) are not known precisely, but instead lie in some known uncertainty set. Our technique enables *exact* and *tractable* solutions of robust GPs for a significantly broader class of uncertainty sets than those considered in prior work [5, 36]. We illustrate the power of these REP-based approaches for robust GPs by employing them to compute hitting-times in dynamical systems (these problems can be recast as robust GPs). In recent work Han et al. [31] have also employed our reformulations of robust GPs (based on an earlier preprint of our work [12]) to optimally allocate resources to control the worst-case spread of epidemics in a network; here the exact network is unknown and it belongs to an uncertainty set.

As yet another application of REPs, we note that REP-based relaxations are useful for obtaining bounds on the optimal value in signomial programming problems [20]. Signomial programs are a generalization of GPs in which the entries of the coefficient vectors  $\mathbf{c}^{(j)}$ 's in (3) in both the objective and the constraints can take on positive and negative values. Sums of exponentials in which the coefficients are both positive and negative are no longer convex functions. As such, signomial programs are intractable to solve in general (unlike GPs), and NP-hard problems can be reduced to special cases of signomial programs. In a separate paper [13], we describe a family of tractable REP-based convex relaxations for computing lower bounds in general signomial programs. In the present manuscript we do not discuss this application further, and we refer the interested reader to [13] for more details.

### 1.3 From relative entropy to quantum relative entropy

Building on our discussion of relative entropy optimization problems and their applications, we consider optimization problems specified in terms of the quantum relative entropy function in Sect. 4. The quantum relative entropy function is a matrix generalization of the function  $\nu \log(\nu/\lambda)$ , and the domain of each of its arguments is the cone of positive semidefinite matrices:

$$D(\mathbf{N}, \mathbf{A}) = \text{Tr}[\mathbf{N} \log(\mathbf{N}) - \mathbf{N} \log(\mathbf{A})]. \quad (7)$$

Here  $\log$  refers to the matrix logarithm. As this function is convex and positively homogenous, its epigraph defines a natural matrix analog of the relative entropy cone  $\mathcal{RE}_1$  from (1):

$$\mathcal{QRE}_d = \left\{ (\mathbf{N}, \mathbf{A}, \delta) \in \mathbb{S}_+^d \times \mathbb{S}_+^d \times \mathbb{R} \mid D(\mathbf{N}, \mathbf{A}) \leq \delta \right\}. \quad (8)$$

Here  $\mathbb{S}_+^d$  denotes the cone of  $d \times d$  positive semidefinite matrices in the space  $\mathbb{S}^d \cong \mathbb{R}^{(d^2+d)/2}$  of  $d \times d$  symmetric matrices. Hence the convex cone  $\mathcal{QRE}_d \subset \mathbb{S}^d \times \mathbb{S}^d \times \mathbb{R} \cong \mathbb{R}^{d^2+d+1}$  and the case  $d = 1$  corresponds to  $\mathcal{RE}_1$ . Quantum relative entropy programs (QREPs) are conic optimization problems specified with respect to the quantum relative cone  $\mathcal{QRE}_d$ .

The focus of Sect. 4 is on applications of QREPs. Indeed, a broader objective of this expository article is to initiate an investigation of the utility of the quantum relative entropy function from a mathematical programming perspective. We begin by surveying previously-studied applications of Von-Neumann entropy optimization, which is a class of convex programs obtained by restricting the second argument of the quantum relative entropy function (7) to be the identity matrix. We also describe a Von-Neumann entropy optimization approach for obtaining bounds on certain capacities associated with quantum channels; as a demonstration, in Sect. 4.3 we provide a comparison between a “classical-to-quantum” capacity of a quantum channel and the capacity of a purely classical channel induced by the quantum channel (i.e., one is restricted to send and receive only classical information). Finally, we describe a stylized application of a QREP that exploits the full convexity of the quantum relative entropy function with respect to both its arguments. This application illustrates some of the key distinctions, especially in the context of convex duality, between the classical and quantum cases.

### 1.4 Remarks and paper outline

As a general remark, we note that there are several types of relative entropy functions (and associated entropies) that have been studied in the literature in both the classical and quantum settings. In this article, we restrict our attention in the classical case to the relative entropy function  $\nu \log(\nu/\lambda)$  corresponding to Shannon entropy [16]. In the quantum setting, the function  $D(\mathbf{N}, \mathbf{A})$  from (7) is called the Araki–Umegaki relative entropy [45], and it gives the Von-Neumann entropy when suitably restricted.

This paper is structured as follows. In Sect. 2, we describe our REP relaxation for the permanent maximization problem, and in Sect. 3 we discuss REP-based approaches for robust GPs and the application of these techniques to the computation of hitting-times in dynamical systems. In Sect. 4 we describe QREPs and their applications, highlighting the similarities and distinctions with respect to the classical case. We conclude with a brief discussion and open questions in Sect. 5.

## 1.5 Notational convention

The nonnegative orthant in  $\mathbb{R}^n$  is denoted by  $\mathbb{R}_+^n$ . The space of  $n \times n$  symmetric matrices is denoted by  $\mathbb{S}^n$  (in particular,  $\mathbb{S}^n \cong \mathbb{R}^{\binom{n+1}{2}}$ ) and the cone of  $n \times n$  positive semidefinite matrices is denoted by  $\mathbb{S}_+^n$ . For vectors  $\mathbf{y}, \mathbf{z} \in \mathbb{R}^n$  we denote elementwise ordering by  $\mathbf{y} \leq \mathbf{z}$  to specify that  $y_i \leq z_i$  for  $i = 1, \dots, n$ . We denote positive semidefinite ordering by  $\leq$  so that  $\mathbf{M} \leq \mathbf{N}$  implies  $\mathbf{N} - \mathbf{M} \in \mathbb{S}_+^n$  for symmetric matrices  $\mathbf{M}, \mathbf{N} \in \mathbb{S}^n$ . Finally, for any real-valued function  $f$  with domain  $\mathcal{D} \subseteq \mathbb{R}^n$ , the *Fenchel* or *convex conjugate*  $f^*$  is given by [48]:

$$f^*(\mathbf{v}) = \sup\{\mathbf{x}'\mathbf{v} - f(\mathbf{x}) \mid \mathbf{x} \in \mathcal{D}\} \quad (9)$$

for  $\mathbf{v} \in \mathbb{R}^n$ .

## 2 An approximation algorithm for permanent maximization via relative entropy optimization

The permanent of a matrix  $\mathbf{M} \in \mathbb{R}^{n \times n}$  is defined as:

$$\text{perm}(\mathbf{M}) \triangleq \sum_{\sigma \in \mathfrak{S}_n} \prod_{i=1}^n \mathbf{M}_{i, \sigma(i)}, \quad (10)$$

where  $\mathfrak{S}_n$  refers to the set of all permutations of elements from the set  $\{1, \dots, n\}$ . The matrix permanent arises in combinatorics as the sum over weighted perfect matchings in a bipartite graph [43], in geometry as the mixed volume of hyperrectangles [8], and in multivariate statistics in the computation of order statistics [1]. In this section we consider the problem of maximizing the permanent over a family of matrices. This problem has received some attention for families of positive semidefinite matrices with specified eigenvalues [19, 28], but all of those works have tended to seek analytical solutions for very special instances of this problem. Here we consider permanent maximization over general convex subsets of the set of nonnegative matrices. (As a parallel, recall that SDPs are useful for maximizing the determinant over an affine section of the cone of symmetric positive semidefinite matrices [54].) Permanent maximization is relevant for designing a bipartite network in which the average weighted matching is to be maximized subject to additional topology constraints on the network. This problem also arises in geometry in finding a configuration of a set of hyperrectangles so that their mixed volume is maximized.

## 2.1 Approximating the permanent of a nonnegative matrix

To begin with, we note that even computing the permanent of a fixed matrix is a #P-hard problem and it is therefore considered to be intractable in general [53]; accordingly a large body of research has investigated approximate computation of the permanent [2, 30, 38, 41]. The class of elementwise nonnegative matrices has particularly received attention as these arise most commonly in application domains (e.g., bipartite graphs with nonnegative edge weights). For such matrices, several *deterministic* polynomial-time approximation algorithms provide exponential-factor (in the size  $n$  of the matrix) approximations of the permanent, e.g. [41]. The approach in [41] describes a technique based on solving a GP to approximate the permanent of a nonnegative matrix. The approximation guarantees in [41] are based on Van Der Waerden's inequality, which states that the matrix  $\mathbf{M} = \frac{1}{n} \mathbf{1}\mathbf{1}'$  has the smallest permanent among all  $n \times n$  doubly stochastic matrices (nonnegative matrices with all row-sums and column-sums equal to one). This inequality was proved originally in the 1980s [21, 23], and Gurvits recently gave a very simple and elegant proof of this result [29]. In what follows, we describe the approach of [41], but using the terminology developed by Gurvits in [29].

The permanent of a matrix  $\mathbf{M} \in \mathbb{R}_+^{n \times n}$  can be defined in terms of a particular coefficient of the following homogenous polynomial in  $\mathbf{y} \in \mathbb{R}_+^n$ :

$$p_{\mathbf{M}}(\mathbf{y}_1, \dots, \mathbf{y}_n) = \prod_{i=1}^n \left( \sum_{j=1}^n \mathbf{M}_{i,j} \mathbf{y}_j \right). \quad (11)$$

Specifically, the permanent of  $\mathbf{M}$  is the coefficient corresponding to the  $\mathbf{y}_1 \cdots \mathbf{y}_n$  monomial term of  $p_{\mathbf{M}}(\mathbf{y}_1, \dots, \mathbf{y}_n)$ :

$$\text{perm}(\mathbf{M}) = \frac{\partial^n p_{\mathbf{M}}(\mathbf{y}_1, \dots, \mathbf{y}_n)}{\partial \mathbf{y}_1 \cdots \partial \mathbf{y}_n}.$$

In his proof of Van Der Waerden's inequality, Gurvits defines the *capacity* of a homogeneous polynomial  $p(\mathbf{y}_1, \dots, \mathbf{y}_n)$  of degree  $n$  over  $\mathbf{y} \in \mathbb{R}_+^n$  as follows:

$$\text{cap}(p) \triangleq \inf_{\mathbf{y} \in \mathbb{R}_+^n} \frac{p(\mathbf{y})}{\mathbf{y}_1 \cdots \mathbf{y}_n} = \inf_{\mathbf{y} \in \mathbb{R}_+^n, \mathbf{y}_1 \cdots \mathbf{y}_n = 1} p(\mathbf{y}). \quad (12)$$

Gurvits then proves the following result:

**Theorem 1** [29] *For any matrix  $\mathbf{M} \in \mathbb{R}_+^{n \times n}$ , we have that:*

$$\frac{n!}{n^n} \text{cap}(p_{\mathbf{M}}) \leq \text{perm}(\mathbf{M}) \leq \text{cap}(p_{\mathbf{M}}).$$

Here the polynomial  $p_{\mathbf{M}}$  and its capacity  $\text{cap}(p_{\mathbf{M}})$  are as defined in (11) and (12). Further if each column of  $\mathbf{M}$  has at most  $k$  nonzeros, then the factor in the lower bound can be improved from  $\frac{n!}{n^n}$  to  $\left(\frac{k-1}{k}\right)^{(k-1)n}$ .



Gurvits in fact proves a more general statement involving so-called stable polynomials, but the above restricted version will suffice for our purposes. The upper-bound in this statement is straightforward to prove; it is the lower bound that is the key technical novelty. Thus, if one could compute the capacity of the polynomial  $p_{\mathbf{M}}$  associated to a nonnegative matrix  $\mathbf{M}$ , then one can obtain an exponential-factor approximation of the permanent of  $\mathbf{M}$  as  $\frac{n!}{n^n} \approx \exp(-n)$ . In order to compute the capacity of  $p_{\mathbf{M}}$  via GP, we apply the transformation  $\mathbf{x}_i = \log(\mathbf{y}_i)$ ,  $i = 1, \dots, n$  in (12) and solve the following program:<sup>1</sup>

$$\begin{aligned} \log(\text{cap}(p_{\mathbf{M}})) &= \inf_{\mathbf{x} \in \mathbb{R}^n} \sum_{i=1}^n \log \left( \sum_{j=1}^n \mathbf{M}_{i,j} \exp(\mathbf{x}_j) \right) \text{ s.t. } \mathbf{1}'\mathbf{x} = 0 \\ &= \inf_{\substack{\mathbf{x} \in \mathbb{R}^n \\ \boldsymbol{\beta} \in \mathbb{R}^n}} \sum_{i,j=1}^n \mathbf{M}_{i,j} \exp(\mathbf{x}_j - \beta_i - 1) + \mathbf{1}'\boldsymbol{\beta} \text{ s.t. } \mathbf{1}'\mathbf{x} = 0. \end{aligned} \quad (13)$$

Here we obtain the second formulation by appealing to the following easily-verified variational characterization of the logarithm, which holds for  $\gamma > 0$ :

$$\log(\gamma) = \inf_{\beta \in \mathbb{R}} e^{-\beta-1} \gamma + \beta. \quad (14)$$

## 2.2 An approximation algorithm for permanent maximization

Focusing on a set  $\mathcal{M} \subset \mathbb{R}_+^{n \times n}$  of entrywise nonnegative matrices, our proposal to approximately maximize the permanent is to find  $\mathbf{M} \in \mathcal{M}$  with maximum capacity  $\text{cap}(p_{\mathbf{M}})$ , which leads to the following consequence of Theorem 1:

**Proposition 1** *Let  $\mathcal{M} \in \mathbb{R}_+^{n \times n}$  be a set of nonnegative matrices, and consider the following two quantities:*

$$\begin{aligned} \hat{\mathbf{M}}_{\text{perm}} &= \arg \sup_{\mathbf{M} \in \mathcal{M}} \text{perm}(\mathbf{M}) \\ \hat{\mathbf{M}}_{\text{cap}} &= \arg \sup_{\mathbf{M} \in \mathcal{M}} \text{cap}(p_{\mathbf{M}}) \end{aligned}$$

*Then we have that*

$$\frac{n!}{n^n} \text{perm}(\hat{\mathbf{M}}_{\text{perm}}) \leq \frac{n!}{n^n} \text{cap}(p_{\hat{\mathbf{M}}_{\text{cap}}}) \leq \text{perm}(\hat{\mathbf{M}}_{\text{perm}}).$$

*The factor in the lower bound can be improved from  $\frac{n!}{n^n}$  to  $\left(\frac{k-1}{k}\right)^{(k-1)n}$  if every matrix in  $\mathcal{M}$  has at most  $k$  nonzeros in each column.*

<sup>1</sup> Such a logarithmic transformation of the variables is also employed in converting a GP specified in terms of non-convex posynomial functions to a GP in convex form; see [11, 20] for more details.

*Proof* The proof follows from a direct application of Theorem 1:

$$\begin{aligned} \frac{n!}{n^n} \text{perm}(\hat{\mathbf{M}}_{\text{perm}}) &\leq \frac{n!}{n^n} \text{cap}(p_{\hat{\mathbf{M}}_{\text{perm}}}) \\ &\leq \frac{n!}{n^n} \text{cap}(p_{\hat{\mathbf{M}}_{\text{cap}}}) \\ &\leq \text{perm}(\hat{\mathbf{M}}_{\text{cap}}) \\ &\leq \text{perm}(\hat{\mathbf{M}}_{\text{perm}}). \end{aligned}$$

The first and third inequalities are a result of Theorem 1, and the second and fourth inequalities follow from the definitions of  $\hat{\mathbf{M}}_{\text{cap}}$  and of  $\hat{\mathbf{M}}_{\text{perm}}$  respectively. The improvement in the lower bound also follows from Theorem 1.  $\square$

In summary, maximizing the capacity with respect to the family  $\mathcal{M}$  gives a matrix in  $\mathcal{M}$  that approximately maximizes the permanent over all matrices in  $\mathcal{M}$ . As the computation of the capacity  $\text{cap}(p_{\mathbf{M}})$  of a *fixed* matrix  $\mathbf{M} \in \mathcal{M}$  involves the solution of a GP, the maximization of  $\text{cap}(p_{\mathbf{M}})$  over the set  $\mathcal{M}$  involves the *maximization of the optimal value* of the GP (13) over  $\mathbf{M} \in \mathcal{M}$ :

$$\sup_{\mathbf{M} \in \mathcal{M}} \log(\text{cap}(p_{\mathbf{M}})) = \sup_{\mathbf{M} \in \mathcal{M}} \left[ \inf_{\substack{\mathbf{x} \in \mathbb{R}^n, \mathbf{1}'\mathbf{x}=0 \\ \boldsymbol{\beta} \in \mathbb{R}^n}} \sum_{i,j=1}^n \mathbf{M}_{i,j} \exp(\mathbf{x}_j - \beta_i - 1) + \mathbf{1}'\boldsymbol{\beta} \right] \quad (15)$$

Rewriting the inner convex program as an REP (as in Sect. 1.1 in the introduction), we have that:

$$\begin{aligned} \log(\text{cap}(p_{\mathbf{M}})) &= \inf_{\substack{\mathbf{x}, \boldsymbol{\beta} \in \mathbb{R}^n \\ \mathbf{Y} \in \mathbb{R}_+^{n \times n}}} \quad \text{Tr}(\mathbf{M}\mathbf{Y}) + \mathbf{1}'\boldsymbol{\beta} \\ \text{s.t.} \quad &\mathbf{1}'\mathbf{x} = 0 \\ &\mathbf{x}_j - \beta_i - 1 \leq \log(\mathbf{Y}_{i,j}), \quad i, j = 1, \dots, n. \end{aligned}$$

The dual of this REP reformulation of the GP (13) is the following optimization problem based on a straightforward calculation:

$$\begin{aligned} \log(\text{cap}(p_{\mathbf{M}})) &= \sup_{\Delta \in \mathbb{R}_+^{n \times n}} \quad - \sum_{i,j=1}^n \Delta_{i,j} \log\left(\frac{\Delta_{i,j}}{\mathbf{M}_{i,j}}\right) \\ \text{s.t.} \quad &\sum_{j=1}^n \Delta_{i,j} = 1, \quad i = 1, \dots, n \\ &\sum_{i=1}^n \Delta_{i,1} = \dots = \sum_{i=1}^n \Delta_{i,n}. \end{aligned} \quad (16)$$

The constraint on the last line requires that the column-sums be equal to each other. As the relative entropy function  $d(\Delta, \mathbf{M}) = \sum_{i,j=1}^n \Delta_{i,j} \log \left( \frac{\Delta_{i,j}}{\mathbf{M}_{i,j}} \right)$  is convex with respect to  $\Delta$ , this dual optimization problem is a convex program and it can be expressed as an REP.

However, the function  $d(\Delta, \mathbf{M})$  is *jointly* convex with respect to  $(\Delta, \mathbf{M})$ . Consequently, one can optimize the objective function in the dual problem (16) above jointly with respect to  $(\Delta, \mathbf{M})$ . This observation leads to the following reformulation of (15):

$$\begin{aligned} \sup_{\mathbf{M} \in \mathcal{M}} \log(\text{cap}(p_{\mathbf{M}})) &= \sup_{\mathbf{M}, \Delta \in \mathbb{R}_+^{n \times n}} - \sum_{i,j=1}^n \Delta_{i,j} \log \left( \frac{\Delta_{i,j}}{\mathbf{M}_{i,j}} \right) \\ \text{s.t.} \quad &\mathbf{M} \in \mathcal{M} \\ &\sum_{j=1}^n \Delta_{i,j} = 1, \quad i = 1, \dots, n \\ &\sum_{i=1}^n \Delta_{i,1} = \dots = \sum_{i=1}^n \Delta_{i,n}. \end{aligned}$$

If the set  $\mathcal{M} \subset \mathbb{R}_+^{n \times n}$  is convex, this problem is a convex program. Further, if  $\mathcal{M}$  can be represented tractably, then this convex program can be solved efficiently. We record these observations in the following proposition:

**Proposition 2** *Suppose that  $\mathcal{M} \subset \mathbb{R}_+^{n \times n}$  has a conic representation:*

$$\mathcal{M} = \{ \mathbf{M} \in \mathbb{R}^{n \times n} \mid \mathbf{M} \in \mathbb{R}_+^{n \times n}, \exists \mathbf{z} \in \mathbb{R}^m \text{ s.t. } \mathbf{g} + \mathcal{L}(\mathbf{M}, \mathbf{z}) \in \mathcal{K} \}$$

for  $\mathcal{L} : \mathbb{R}^{n \times n} \times \mathbb{R}^m \rightarrow \mathbb{R}^\ell$  a linear operator,  $\mathbf{g} \in \mathbb{R}^\ell$ , and  $\mathcal{K} \subset \mathbb{R}^\ell$  a convex cone. Then the problem of maximizing capacity with respect to the set  $\mathcal{M}$  can be reformulated as follows:

$$\begin{aligned} \sup_{\mathbf{M} \in \mathcal{M}} \log(\text{cap}(p_{\mathbf{M}})) &= \sup_{\mathbf{M}, \Delta \in \mathbb{R}_+^{n \times n}, \mathbf{z} \in \mathbb{R}^m} - \sum_{i,j=1}^n \Delta_{i,j} \log \left( \frac{\Delta_{i,j}}{\mathbf{M}_{i,j}} \right) \\ \text{s.t.} \quad &\mathbf{g} + \mathcal{L}(\mathbf{M}, \mathbf{z}) \in \mathcal{K} \\ &\sum_{j=1}^n \Delta_{i,j} = 1, \quad i = 1, \dots, n \\ &\sum_{i=1}^n \Delta_{i,1} = \dots = \sum_{i=1}^n \Delta_{i,n}. \end{aligned}$$

Suppose the set  $\mathcal{M}$  has a tractable LP, GP, or SOCP representation so that the convex cone  $\mathcal{K}$  in the above proposition is the relative entropy cone. Then the convex program in Proposition 2 is an REP that can be solved efficiently. If the set  $\mathcal{M}$  has a tractable SDP representation, then  $\mathcal{K}$  is the cone of positive semidefinite matrices; in such cases, the program in Proposition 2 is no longer an REP, but it can still be solved efficiently via interior-point methods by combining logarithmic barrier functions for the positive semidefinite cone and for the relative entropy cone [44].

### 3 Robust GP and hitting-time estimation in dynamical systems

As our next application, we describe the utility of REPs in addressing the problem of computing solutions to GPs that are robust to uncertainty in the input parameters

of the GP. Robust GPs arise in power control problems in communication systems [14] as well as in robust digital circuit gate sizing [10]. Specifically, a robust GP is an optimization problem in which a positive sum of exponentials is minimized subject to affine constraints and constraints of the following form in a decision variable  $\mathbf{x} \in \mathbb{R}^n$ :

$$\sup_{[\mathbf{c}, \mathbf{a}^{(1)}, \dots, \mathbf{a}^{(k)}] \in \mathcal{U}} \sum_{i=1}^k \mathbf{c}_i \exp(\mathbf{a}^{(i)'} \mathbf{x}) \leq 1. \quad (17)$$

Here the set  $\mathcal{U} \subset \mathbb{R}_+^k \times \mathbb{R}^{nk}$  specifies the possible uncertainty in the coefficients  $\mathbf{c}$  and in the exponents  $\mathbf{a}^{(1)}, \dots, \mathbf{a}^{(k)}$ . In principle, constraints of the type (17) specify convex sets in a decision variable  $\mathbf{x}$  because they can be viewed as the intersection of a (possibly infinite) collection of convex constraints. However, this observation does not lead directly to efficient methods for the numerical solution of robust GPs, as constraints of the form (17) are not tractable to describe in general. For example, if  $\mathcal{U}$  is a finite set then the constraint (17) reduces to a finite collection of constraints on positive sums of exponentials; however, if  $\mathcal{U}$  consists of infinitely many elements, then the expression  $\sup_{[\mathbf{c}, \mathbf{a}^{(1)}, \dots, \mathbf{a}^{(k)}] \in \mathcal{U}} \sum_{i=1}^k \mathbf{c}_i \exp(\mathbf{a}^{(i)'} \mathbf{x})$  may not be efficiently specified in closed form. Thus, the objective in robust GPs is to obtain tractable reformulations of constraints of the form (17) via a small, finite number of inequalities with a tractable description.

Such optimization problems in which one seeks solutions that are robust to parameter uncertainty have been extensively investigated in the field of robust optimization [3, 5], and exact, tractable reformulations of robust convex programs are available in a number of settings, e.g., for robust LPs. However, progress has been limited in the context of robust GPs. We discuss prior work on this problem in Sect. 3.1, highlighting some of the shortcomings, and we describe the more powerful generalizations afforded by REP-based reformulations in Sect. 3.2. In Sect. 3.3, we illustrate the utility of these reformulations in estimating hitting-times in dynamical systems.

### 3.1 GP-based reformulations of robust GPs

In their seminal work on robust convex optimization [5], Ben-Tal and Nemirovski obtained an exact, tractable reformulation of robust GPs in which a very restricted form of coefficient uncertainty is allowed in a positive sum-of-exponentials function—specifically, they assume that the uncertainty set  $\mathcal{U}$  is decomposed as  $\mathcal{U} = \mathcal{C} \times \{\bar{\mathbf{a}}^{(1)}\} \times \dots \times \{\bar{\mathbf{a}}^{(k)}\}$ , where each  $\bar{\mathbf{a}}^{(i)} \in \mathbb{R}^n$ ,  $i = 1, \dots, k$ , and  $\mathcal{C} \subset \mathbb{R}_+^k$  is a convex ellipsoid in  $\mathbb{R}_+^k$  specified by a quadratic form defined by an elementwise nonnegative matrix. Thus, the exponents are assumed to be known exactly (i.e., there is *no uncertainty* in these exponents) and the uncertainty in the coefficients is specified by a very particular type of ellipsoidal set. The reformulation given in [5] for such robust GPs is itself a GP with additional variables, which can be solved efficiently.

In subsequent work, Hsiung et al. [36] considered sum-of-exponentials functions with the coefficients absorbed into the exponent as follows:

$$\sup_{[\mathbf{d}, \mathbf{a}^{(1)}, \dots, \mathbf{a}^{(k)}] \in \mathcal{D}} \sum_{i=1}^k \exp(\mathbf{a}^{(i)'} \mathbf{x} + \mathbf{d}_i) \leq 1$$

For such constraints with  $\mathcal{D}$  being either a polyhedral set or an ellipsoidal set, Hsiung et al. [36] obtain tractable but inexact reformulations via piecewise linear approximations, with the reformulations again being GPs.

A reason for the limitations in these previous works—very restrictive forms of uncertainty sets in [5] and inexact reformulations in [36]—is that they considered GP-based reformulations of the inequality (17). In the next section, we discuss exact and tractable REP-based reformulations of the inequality (17) for a general class of uncertainty sets.

### 3.2 Relative entropy reformulations of robust GPs

We describe exact and efficient REP-based reformulations of robust GPs for settings in which the uncertainty set  $\mathcal{U}$  is decoupled as follows:

$$\mathcal{U} = \mathcal{C} \times \mathcal{E}^{(1)} \times \dots \times \mathcal{E}^{(k)}, \quad (18)$$

where  $\mathcal{C} \subset \mathbb{R}_+^k$  specifies uncertainty in the coefficients and each  $\mathcal{E}^{(i)} \subset \mathbb{R}^n, i = 1, \dots, k$  specifies uncertainty in the  $i$ 'th exponent. Although this decomposition imposes a restriction on the types of uncertainty sets  $\mathcal{U}$  we consider, our methodology allows for flexible specifications of the sets  $\mathcal{C}$  and  $\mathcal{E}^{(1)}, \dots, \mathcal{E}^{(k)}$  as each of these can essentially be arbitrary convex sets with a tractable description.

For such forms of uncertainty, the following proposition provides an exact and efficient REP reformulation of (17) based on an appeal to convex duality:

**Proposition 3** *Suppose  $\mathcal{C} \subset \mathbb{R}^k$  is a convex set of the form:*

$$\mathcal{C} = \{\mathbf{c} \in \mathbb{R}^k \mid \exists \mathbf{z} \in \mathbb{R}^{m_{\mathcal{C}}} \text{ s.t. } \mathbf{c} \geq \mathbf{0}, \mathbf{g} + \mathfrak{F}_{\mathcal{C}}(\mathbf{c}) + \mathfrak{H}_{\mathcal{C}}(\mathbf{z}) \in \mathcal{K}_{\mathcal{C}}\}$$

*for a convex cone  $\mathcal{K}_{\mathcal{C}} \subset \mathbb{R}^{\ell_{\mathcal{C}}}$ , a linear map  $\mathfrak{F}_{\mathcal{C}} : \mathbb{R}^k \rightarrow \mathbb{R}^{\ell_{\mathcal{C}}}$ , a linear map  $\mathfrak{H}_{\mathcal{C}} : \mathbb{R}^{m_{\mathcal{C}}} \rightarrow \mathbb{R}^{\ell_{\mathcal{C}}}$ , and  $\mathbf{g} \in \mathbb{R}^{\ell_{\mathcal{C}}}$ , and suppose each  $\mathcal{E}^{(i)} \subset \mathbb{R}^n$  for  $i = 1, \dots, k$  is a set of the form:*

$$\mathcal{E}^{(i)} = \left\{ \mathbf{q} \in \mathbb{R}^n \mid \exists \mathbf{z} \in \mathbb{R}^{m_i} \text{ s.t. } \mathbf{g}^{(i)} + \mathfrak{F}^{(i)}(\mathbf{q}) + \mathfrak{H}^{(i)}(\mathbf{z}) \in \mathcal{K}^{(i)} \right\}$$

*for convex cones  $\mathcal{K}^{(i)} \subset \mathbb{R}^{\ell_i}$ , linear maps  $\mathfrak{F}^{(i)} : \mathbb{R}^n \rightarrow \mathbb{R}^{\ell_i}$ , linear maps  $\mathfrak{H}^{(i)} : \mathbb{R}^{m_i} \rightarrow \mathbb{R}^{\ell_i}$ , and  $\mathbf{g}^{(i)} \in \mathbb{R}^{\ell_i}$ . Further, assume that there exists a point  $(\bar{\mathbf{c}}, \bar{\mathbf{z}}) \in \mathbb{R}^k \times \mathbb{R}^{m_{\mathcal{C}}}$  satisfying the conic and nonnegativity constraints in  $\mathcal{C}$  strictly, and similarly that there exists a strictly feasible point in each  $\mathcal{E}^{(i)}$ . Then we have that  $\mathbf{x} \in \mathbb{R}^n$  satisfies the constraint (17) with  $\mathcal{U} = \mathcal{C} \times \mathcal{E}^{(1)} \times \dots \times \mathcal{E}^{(k)}$  if and only if there exist  $\boldsymbol{\zeta} \in \mathbb{R}^{\ell_{\mathcal{C}}}$ ,  $\boldsymbol{\theta}^{(i)} \in \mathbb{R}^{\ell_i}$  for  $i = 1, \dots, k$  such that:*

$$\mathbf{g}'\boldsymbol{\zeta} \leq 1, \boldsymbol{\zeta} \in \mathcal{K}_{\mathcal{C}}^{\star}, \mathfrak{H}_{\mathcal{C}}^{\dagger}(\boldsymbol{\zeta}) = \mathbf{0}, \text{ and for } i = 1, \dots, k, \\ \mathfrak{F}^{(i)\dagger}(\boldsymbol{\theta}^{(i)}) + \mathbf{x} = \mathbf{0}, \mathfrak{H}^{(i)\dagger}(\boldsymbol{\theta}^{(i)}) = \mathbf{0}, \mathbf{g}^{(i)'}\boldsymbol{\theta}^{(i)} \leq \log(-[\mathfrak{F}_{\mathcal{C}}^{\dagger}(\boldsymbol{\zeta})]_i), \boldsymbol{\theta}^{(i)} \in \mathcal{K}^{(i)\star}.$$

*Note* Here  $\mathfrak{F}^{\dagger}$ ,  $\mathfrak{F}^{(i)\dagger}$ , and  $\mathfrak{H}^{\dagger}$  represent the adjoints of the operators  $\mathfrak{F}$ ,  $\mathfrak{F}^{(i)}$ , and  $\mathfrak{H}$  respectively. The cones  $\mathcal{K}_{\mathcal{C}}^{\star}$  and  $\mathcal{K}^{(i)\star}$  denote the duals of the cones  $\mathcal{K}_{\mathcal{C}}$  and  $\mathcal{K}^{(i)}$  respectively. The assumption that there exist points that strictly satisfy the constraints specifying  $\mathcal{C}$  and those specifying each  $\mathcal{E}^{(i)}$  allows us to appeal to strong duality in deriving our result [48]. We give a concrete illustration of the power of this result in the sequel as well as an application to hitting-time estimation in dynamical systems in Sect. 3.3.

*Proof* The constraint (17) can be reformulated as follows for uncertainty sets  $\mathcal{U}$  that are decomposable according to (18):

$$\exists \mathbf{y} \in \mathbb{R}^k \text{ s.t.} \quad \sup_{\mathbf{c} \in \mathcal{C}} \mathbf{y}'\mathbf{c} \leq 1 \\ \sup_{\mathbf{a}^{(i)} \in \mathcal{E}^{(i)}} \mathbf{a}^{(i)'}\mathbf{x} \leq \log(\mathbf{y}_i), \quad i = 1, \dots, k.$$

This restatement is possible because the set  $\mathcal{C}$  is a subset of the nonnegative orthant  $\mathbb{R}_{+}^k$ , and because the uncertainty sets  $\mathcal{E}^{(i)}$  are decoupled (from  $\mathcal{C}$  and from each other) and are therefore independent of each other. The first expression,  $\sup_{\mathbf{c} \in \mathcal{C}} \mathbf{y}'\mathbf{c} \leq 1$ , is a universal quantification for all  $\mathbf{c} \in \mathcal{C}$ . In order to convert this universal quantifier to an existential quantifier, we appeal to convex duality as is commonly done in the theory of robust optimization [3, 5]. Specifically, by noting that  $\mathcal{C}$  has a conic representation and by appealing to conic duality [7], we have that:

$$\forall \mathbf{c} \in \mathcal{C}, \mathbf{y}'\mathbf{c} \leq 1 \\ \Updownarrow \\ \exists \boldsymbol{\zeta} \in \mathbb{R}^{\ell} \text{ s.t. } \mathfrak{F}_{\mathcal{C}}^{\dagger}(\boldsymbol{\zeta}) + \mathbf{y} \leq \mathbf{0}, \mathfrak{H}_{\mathcal{C}}^{\dagger}(\boldsymbol{\zeta}) = \mathbf{0}, \boldsymbol{\zeta} \in \mathcal{K}_{\mathcal{C}}^{\star}, \mathbf{g}'\boldsymbol{\zeta} \leq 1. \quad (19)$$

Similarly, we have that:

$$\forall \mathbf{a}^{(i)} \in \mathcal{E}^{(i)}, \mathbf{a}^{(i)'}\mathbf{x} \leq \log(\mathbf{y}_i) \\ \Updownarrow \\ \exists \boldsymbol{\theta}^{(i)} \in \mathbb{R}^{\ell_i} \text{ s.t. } \mathfrak{F}^{(i)\dagger}(\boldsymbol{\theta}^{(i)}) + \mathbf{x} = \mathbf{0}, \mathfrak{H}^{(i)\dagger}(\boldsymbol{\theta}^{(i)}) = \mathbf{0}, \\ \boldsymbol{\theta}^{(i)} \in \mathcal{K}^{(i)\star}, \mathbf{g}^{(i)'}\boldsymbol{\theta}^{(i)} \leq \log(\mathbf{y}_i). \quad (20)$$

The assumptions on strict feasibility are required to derive (19) and (20). Combining these results and eliminating  $\mathbf{y}$ , we have the desired conclusion.  $\square$

In summary, Proposition 3 gives an extended formulation for (17) with additional variables. In a similar spirit to Proposition 2, if the sets  $\mathcal{C}$  and  $\mathcal{E}^{(1)}, \dots, \mathcal{E}^{(k)}$  are representable via LP or SOCP (i.e., the cones  $\mathcal{K}_{\mathcal{C}}$  and  $\mathcal{K}^{(i)}$  are orthants, second-order cones, or relative entropy cones of suitable dimensions), then the inequality (17) can be represented via REP. In other words, robust GPs in which the uncertainty

sets are decomposable according to (18) and are tractably specified as polyhedral or ellipsoidal sets (or more generally, sets that are tractably represented via relative entropy inequalities) can be reformulated exactly and efficiently via REPs. (As before, robust GPs in which the cones  $\mathcal{K}_{\mathcal{C}}$  and  $\mathcal{K}^{(i)}$  are semidefinite cones are not directly reformulated as REPs, but can nonetheless be solved efficiently by combining barrier penalties for the relative entropy cone and the semidefinite cone.)

In contrast to previous results on robust GP, note that the form of uncertainty for which we obtain an efficient and exact REP reformulation is significantly more general than that considered in [5], in which  $\mathcal{C}$  can only be a restricted kind of ellipsoidal set and each  $\mathcal{E}^{(i)}$  must be a singleton set (no uncertainty). On the other hand, Hsiung et al. [36] consider robust GPs with polyhedral or ellipsoidal uncertainty that may be coupled across  $\mathcal{E}^{(i)}$ , but their reformulation is inexact. A distinction of our approach relative to those in [5] and in [36] is that our reformulation is a REP, while those described in [5] and in [36] are GPs. In particular, note that some of the inequalities described in the reformulation in Proposition 3 involve a combination of an affine term and a logarithmic term; such inequalities cannot be represented via GPs, and it is the additional expressive power provided by REPs that enables the efficient reformulation and solution of the general class of robust GPs considered here.

*Example* To illustrate these distinctions concretely, consider a specific instance of a robust GP constraint (17) in which the uncertainty set  $\mathcal{U}$  is decomposable according to (18) as  $\mathcal{U} = \mathcal{C} \times \mathcal{E}^{(1)} \times \cdots \times \mathcal{E}^{(k)}$ . Further, suppose  $\mathcal{C} \subset \mathbb{R}^k$  is a convex set of the form:

$$\mathcal{C} = \{\mathbf{c} \in \mathbb{R}^k \mid \mathbf{c} \succeq \mathbf{0}, \mathbf{G} + \mathfrak{F}_{\mathcal{C}}(\mathbf{c}) \succeq \mathbf{0}\}$$

for  $\mathbf{G} \in \mathbb{S}^{\ell}$  and a linear operator  $\mathfrak{F}_{\mathcal{C}} : \mathbb{R}^k \rightarrow \mathbb{S}^{\ell}$ , and suppose also that each  $\mathcal{E}^{(i)} \subset \mathbb{R}^n$  for  $i = 1, \dots, k$  is a set of the form:

$$\mathcal{E}^{(i)} = \{\mathbf{q} \in \mathbb{R}^n \mid \mathbf{G}^{(i)} + \mathfrak{F}^{(i)}(\mathbf{q}) \succeq \mathbf{0}\}$$

for  $\mathbf{G}^{(i)} \in \mathbb{S}^{\ell}$  and linear operators  $\mathfrak{F}^{(i)} : \mathbb{R}^n \rightarrow \mathbb{S}^{\ell}$ . With these uncertainty sets, we have from Proposition 3 that

$$\begin{aligned} \sup_{\mathbf{c} \in \mathcal{C}, \mathbf{a}^{(i)} \in \mathcal{E}^{(i)}} \sum_{i=1}^k \mathbf{c}_i \exp(\mathbf{a}^{(i)'} \mathbf{x}) &\leq 1 \\ \Leftrightarrow \\ \exists \mathbf{Z} \in \mathbb{S}^{\ell} \text{ and } \boldsymbol{\Theta}^{(i)} \in \mathbb{S}^{\ell}, i = 1, \dots, k \text{ s.t.} \\ \text{Tr}(\mathbf{G}\mathbf{Z}) &\leq 1, \mathbf{Z} \succeq \mathbf{0}, \text{ and for each } i = 1, \dots, k, \\ \mathfrak{F}^{(i)\dagger}(\boldsymbol{\Theta}^{(i)}) + \mathbf{x} &= \mathbf{0}, \text{Tr}(\mathbf{G}^{(i)} \boldsymbol{\Theta}^{(i)}) \leq \log(-[\mathfrak{F}^{\dagger}(\mathbf{Z})]_i), \boldsymbol{\Theta}^{(i)} \succeq \mathbf{0}. \end{aligned}$$

Note that the uncertainty sets  $\mathcal{C}$  and  $\mathcal{E}^{(i)}$  are SDP-representable sets. The corresponding robust GP inequality (17) cannot be handled by the previous approaches [5, 36] described in Sect. 3.1, but it can be efficiently reformulated via semidefinite and relative entropy inequalities.

### 3.3 Application to hitting-time estimation in dynamical systems

Consider a linear dynamical system with state-space equations as follows:

$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t), \quad (21)$$

where the state  $\mathbf{x}(t) \in \mathbb{R}^n$  for  $t \geq 0$ . We assume throughout this section that the transition matrix  $\mathbf{A} \in \mathbb{R}^{n \times n}$  is diagonal; otherwise, one can always change to the appropriate *modal coordinates* given by the eigenvectors of  $\mathbf{A}$  (assuming  $\mathbf{A}$  is diagonalizable). The diagonal entries of  $\mathbf{A}$  are called the *modes* of the system. Suppose that the parameters of the dynamical system can take on a range of values with  $\mathbf{A} \in \mathcal{A}$  and  $\mathbf{x}(0) \in \mathcal{X}_{\text{initial}}$ ; the set  $\mathcal{A}$  specifies the set of modes and the set  $\mathcal{X}_{\text{initial}} \subseteq \mathbb{R}^n$  specifies the set of initial conditions. Suppose also that we are given a *target set*  $\mathcal{X}_{\text{target}} \subseteq \mathbb{R}^n$ , and we wish to find the smallest time required for the system to reach a state in  $\mathcal{X}_{\text{target}}$  from an arbitrary initial state in  $\mathcal{X}_{\text{initial}}$ . Formally, we define the *worst-case hitting-time* of the dynamical system (21) to be:

$$\tau(\mathcal{X}_{\text{initial}}, \mathcal{X}_{\text{target}}, \mathcal{A}) \triangleq \inf \left\{ t \in \mathbb{R}_+ \mid \forall \mathbf{x}(0) \in \mathcal{X}_{\text{initial}}, \mathbf{A} \in \mathcal{A} \text{ we have } \exp(\mathbf{A}t) \mathbf{x}(0) \in \mathcal{X}_{\text{target}} \right\}. \quad (22)$$

Indeed, for an initial state  $\mathbf{x}(0)$ , the state of the system (21) at time  $t$  is given by  $\mathbf{x}(t) = \exp(\mathbf{A}t) \mathbf{x}(0)$ ; consequently, the quantity  $\tau(\mathcal{X}_{\text{initial}}, \mathcal{X}_{\text{target}}, \mathcal{A})$  represents the amount of time that the worst-case trajectory of the system, taken over all initial conditions  $\mathcal{X}_{\text{initial}}$  and mode values  $\mathcal{A}$ , requires to enter the target set  $\mathcal{X}_{\text{target}}$ . An assumption underlying this definition is that the target set  $\mathcal{X}_{\text{target}}$  is *absorbing* so that once a trajectory enters  $\mathcal{X}_{\text{target}}$  it remains in  $\mathcal{X}_{\text{target}}$  for all subsequent time.

Hitting-times are of interest in system analysis and verification [47]. As an example, suppose that a system has the property that  $\tau(\mathcal{X}_{\text{initial}}, \mathcal{X}_{\text{target}}, \mathcal{A}) = \infty$ ; this provides a proof that from certain initial states in  $\mathcal{X}_{\text{initial}}$ , the system never enters the target region  $\mathcal{X}_{\text{target}}$ . On the other hand, if the hitting-time  $\tau(\mathcal{X}_{\text{initial}}, \mathcal{X}_{\text{target}}, \mathcal{A}) = 0$ , we have a certificate that  $\mathcal{X}_{\text{target}} \subseteq \mathcal{X}_{\text{initial}}$ . While verification of linear systems has been extensively studied via Lyapunov and barrier certificate methods, the study of hitting-times has received relatively little attention, with a few exceptions such as [56]. In particular, the approaches in [56] can lead to loose bounds as the worst-case hitting-time is computed based on box-constrained outer approximations of  $\mathcal{X}_{\text{initial}}$  and  $\mathcal{X}_{\text{target}}$ .

For a particular class of dynamical systems, we show next that the hitting-time can be computed *exactly* by solving an REP. Specifically, we make the following assumptions regarding the structure of the dynamical system (21):

- The set of modes is given by

$$\mathcal{A} = \{ \mathbf{A} \in \mathbb{R}^{n \times n} \mid \mathbf{A} \text{ diagonal with } \mathbf{A}_{j,j} \in [\ell_j, u_j] \ \forall j = 1, \dots, n \}$$

with  $\ell_j, u_j \leq 0 \ \forall j = 1, \dots, n$ .



- The set of initial states  $\mathcal{X}_{\text{initial}} \subseteq \mathbb{R}_+^n$  is given by a convex set with a tractable representation via affine and conic constraints. In particular, as in Proposition 3,  $\mathcal{X}_{\text{initial}}$  is specified as follows:

$$\mathcal{X}_{\text{initial}} = \{\mathbf{x} \in \mathbb{R}_+^n \mid \exists \mathbf{z} \in \mathbb{R}^m \text{ s.t. } \mathbf{g} + \mathfrak{F}(\mathbf{x}) + \mathfrak{H}(\mathbf{z}) \in \mathcal{K}\}. \quad (23)$$

Here  $\mathbf{g} \in \mathbb{R}^\ell$ , the maps  $\mathfrak{F} : \mathbb{R}^n \rightarrow \mathbb{R}^\ell$  and  $\mathfrak{H} : \mathbb{R}^m \rightarrow \mathbb{R}^\ell$  are linear, and the cone  $\mathcal{K} \subset \mathbb{R}^\ell$  is efficiently described.

- The target set  $\mathcal{X}_{\text{target}} \subset \mathbb{R}^n$  is representable as the intersection of finitely many half spaces:

$$\mathcal{X}_{\text{target}} = \{\mathbf{x} \in \mathbb{R}_+^n \mid \mathbf{c}^{(i)'} \mathbf{x} \leq \mathbf{d}_i, i = 1, \dots, k\}.$$

with  $\mathbf{c}^{(i)} \in \mathbb{R}_+^n$ ,  $i = 1, \dots, k$ , and  $\mathbf{d} \in \mathbb{R}_+^k$ .

The first condition on the modes of the dynamical system and the subsequent conditions on the structure of the initial and target states ensure that the target state is absorbing. Indeed, as each  $\mathbf{c}^{(i)} \in \mathbb{R}_+^n$ , each  $\mathbf{A}_{j,j} \leq 0$ , and  $\mathcal{X}_{\text{initial}} \subset \mathbb{R}_+^n$ , one can check for each  $i = 1, \dots, k$  and any  $\mathbf{x}(0) \in \mathcal{X}_{\text{initial}}$  that  $\sum_{j=1}^n \mathbf{c}_j^{(i)} \mathbf{x}_j(0) \exp(\mathbf{A}_{j,j} t_0) \leq \mathbf{d}_i$  for some  $t_0 \in \mathbb{R}_+$  implies that  $\sum_{j=1}^n \mathbf{c}_j^{(i)} \mathbf{x}_j(0) \exp(\mathbf{A}_{j,j} t) \leq \mathbf{d}_i$  for all  $t \geq t_0$ . Under these conditions, the worst-case hitting-time  $\tau(\mathcal{X}_{\text{initial}}, \mathcal{X}_{\text{target}}, \mathcal{A})$  can be computed exactly as follows:

$$\begin{aligned} \tau(\mathcal{X}_{\text{initial}}, \mathcal{X}_{\text{target}}, \mathcal{A}) = \inf_{t \geq 0} \quad & t \\ \text{s.t.} \quad & \sup_{\substack{\mathbf{x}(0) \in \mathcal{X}_{\text{initial}} \\ \mathbf{A} \in \mathcal{A}}} \sum_{j=1}^n \mathbf{c}_j^{(i)} \mathbf{x}_j(0) \exp(\mathbf{A}_{j,j} t) \leq \mathbf{d}_i, \forall i. \end{aligned} \quad (24)$$

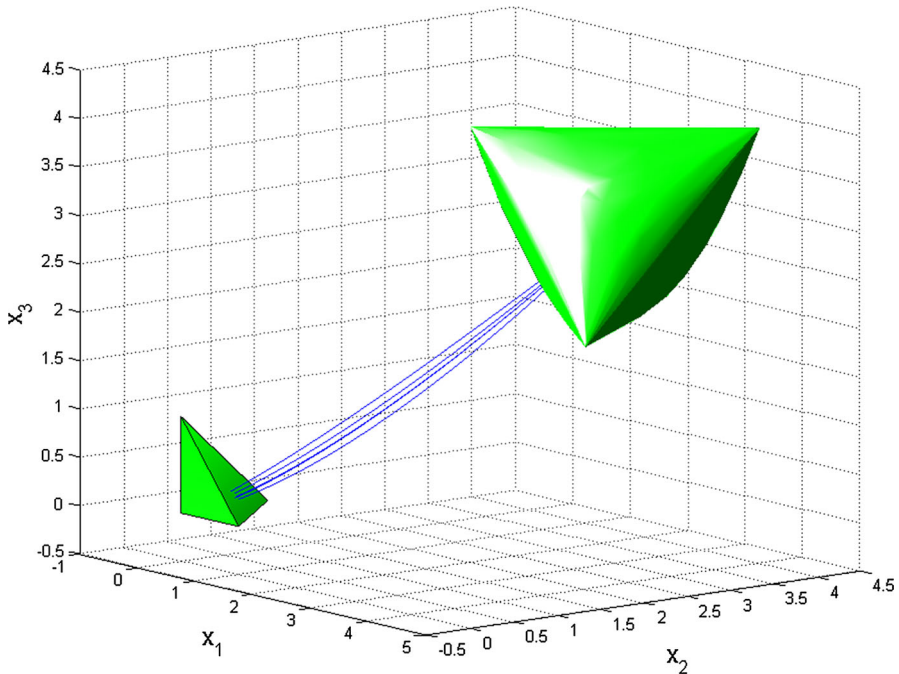
Each constraint here is a robust GP inequality of the form (17) with the uncertainty set being decomposable according to (18). Consequently, the hitting-time computation problem for the particular family of dynamical systems we consider can be reformulated as an REP [possibly with additional conic constraints depending on the cone  $\mathcal{K}$  in (23)] by appealing to Proposition 3.

*Example* Consider a dynamical system with three modes that each take on values in certain ranges as follows:

$$\begin{aligned} \mathcal{A} = \left\{ \mathbf{A} \in \mathbb{R}^{3 \times 3} \mid \mathbf{A} \text{ diagonal, } \mathbf{A}_{1,1} \in [-0.4, -0.45], \right. \\ \left. \mathbf{A}_{2,2} \in [-0.5, -0.6], \mathbf{A}_{3,3} \in [-0.6, -0.7] \right\}. \end{aligned}$$

The set of initial states is a shifted *elliptope* that is contained within the nonnegative orthant:

$$\mathcal{X}_{\text{initial}} = \left\{ \mathbf{x}(0) \in \mathbb{R}_+^3 : \begin{bmatrix} 1 & \mathbf{x}_1(0) - 3 & \mathbf{x}_2(0) - 3 \\ \mathbf{x}_1(0) - 3 & 1 & \mathbf{x}_3(0) - 3 \\ \mathbf{x}_2(0) - 3 & \mathbf{x}_3(0) - 3 & 1 \end{bmatrix} \succeq 0 \right\}.$$



**Fig. 1** Some sample trajectories of a linear system from  $\mathcal{X}_{\text{initial}}$  to  $\mathcal{X}_{\text{target}}$  for the dynamical system described in the example in Sect. 3.3. The (shifted) ellipsope specifies the set of feasible starting points, while the tetrahedron specifies the target set. The system consists of three modes

Finally, we let the target region be a tetrahedron:

$$\mathcal{X}_{\text{target}} = \left\{ \mathbf{x} \in \mathbb{R}_+^3 \mid \mathbf{1}'\mathbf{x} \leq 1 \right\}.$$

For these system attributes, a few sample trajectories are shown in Fig. 1. We solve the REP (24) and obtain that  $\tau(\mathcal{X}_{\text{initial}}, \mathcal{X}_{\text{target}}, \mathcal{A}) = 7.6253$  is the worst-case hitting-time.

## 4 Quantum relative entropy optimization

In this section we discuss optimization problems involving the quantum relative entropy function. Both the classical and the quantum relative entropy cones can be viewed as limiting cases of more general families of cones known as power cones; we describe this relationship in Sect. 4.1. Next we survey applications involving the solution of Von-Neumann entropy optimization problems, which constitute an important special class of quantum relative entropy optimization problems. In Sect. 4.3 we discuss a method to obtain bounds on certain capacities associated to quantum communication channels via Von-Neumann entropy optimization. Finally, in Sect. 4.4 we consider a matrix analog of a robust GP and the utility of the quantum relative entropy function in providing lower bounds on the optimal value of this problem; this stylized

problem provides a natural setting to examine the similarities and differences between the classical and the quantum relative entropy functions.

#### 4.1 Relation to power cones

The relative entropy cone  $\mathcal{RE}_1 \subset \mathbb{R}^3$  can be viewed as a limiting case of a more general family of cones known as power cones [24]. Specifically, for each  $\alpha \in [0, 1]$ , the following map is a convex function on  $\mathbb{R}_+ \times \mathbb{R}_+$ :

$$(\nu, \lambda) \mapsto -\nu^\alpha \lambda^{1-\alpha}. \quad (25)$$

As this map is positively homogenous, its epigraph gives the following *power cone* in  $\mathbb{R}^3$ :

$$\mathcal{P}^\alpha = \{(\nu, \lambda, \delta) \in \mathbb{R}_+ \times \mathbb{R}_+ \times \mathbb{R} \mid -\nu^\alpha \lambda^{1-\alpha} \leq \delta\}. \quad (26)$$

Power cones in higher dimensions are obtained by taking Cartesian products of  $\mathcal{P}^\alpha$ . The relative entropy function can be obtained as a suitable limit of the functions defined in (25):

$$\lim_{\alpha \rightarrow 1} \frac{1}{1-\alpha} [-\nu^\alpha \lambda^{1-\alpha} + \nu] = \nu \log \left( \frac{\nu}{\lambda} \right).$$

Hence the relative entropy cone  $\mathcal{RE}_1$  can be viewed as a kind of “extremal” power cone.

One can also define a matrix analog of the power cones  $\mathcal{P}^\alpha$  based on the following theorem due to Lieb [40]. In analogy to the classical case, the quantum relative entropy cone can also be viewed as a limiting case of these quantum power cones.

**Theorem 2** [40] *For each fixed  $\mathbf{X} \in \mathbb{R}^{d \times d}$  and  $\alpha \in [0, 1]$ , the following function is convex on  $\mathbb{S}_+^d \times \mathbb{S}_+^d$ :*

$$(\mathbf{N}, \mathbf{A}) \mapsto -\text{Tr}(\mathbf{X}' \mathbf{N}^\alpha \mathbf{X} \mathbf{A}^{1-\alpha}).$$

This theorem has significant consequences in various mathematical and statistical aspects of quantum mechanics. As with the map defined in (25), the function considered in this theorem is also positively homogenous, and consequently its epigraph defines a convex cone analogous to  $\mathcal{P}^\alpha$ :

$$\mathcal{Q}\mathcal{P}^\alpha = \{(\mathbf{N}, \mathbf{A}, \delta) \in \mathbb{S}_+^d \times \mathbb{S}_+^d \times \mathbb{R} \mid -\text{Tr}(\mathbf{X}' \mathbf{N}^\alpha \mathbf{X} \mathbf{A}^{1-\alpha}) \leq \delta\}. \quad (27)$$

As before, the quantum relative entropy function (7) may be viewed as a limiting case as follows:

$$D(\mathbf{N}, \mathbf{A}) = \lim_{\alpha \rightarrow 1} \frac{1}{1-\alpha} [-\text{Tr}(\mathbf{N}^\alpha \mathbf{A}^{1-\alpha}) + \text{Tr}(\mathbf{N})].$$

Indeed, this perspective of the quantum relative entropy function offers a proof of its joint convexity with respect to both arguments, as the function  $-\text{Tr}(\mathbf{N}^\alpha \mathbf{A}^{1-\alpha}) + \text{Tr}(\mathbf{N})$  is jointly convex with respect to  $(\mathbf{N}, \mathbf{A})$  for each  $\alpha \in [0, 1]$ .

## 4.2 Survey of applications of Von-Neumann entropy optimization

An important subclass of QREPs are optimization problems involving the Von-Neumann entropy function, which is obtained by restricting the second argument of the negative quantum relative entropy function to be the identity matrix:

$$\begin{aligned} H(\mathbf{N}) &= -D(\mathbf{N}, \mathbf{I}) \\ &= -\text{Tr}[\mathbf{N} \log(\mathbf{N})]. \end{aligned} \quad (28)$$

Here  $\mathbf{N}$  is a positive semidefinite matrix and  $\mathbf{I}$  is the identity matrix of appropriate dimension. In quantum information theory, the Von-Neumann entropy is typically considered for positive semidefinite matrices with trace equal to one (i.e., quantum density matrices), but the function is concave over the full cone of positive semidefinite matrices.

Based on the next proposition, Von-Neumann entropy optimization problems are essentially REPs with additional semidefinite constraints:

**Proposition 4** [7] *Let  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  be a convex function that is invariant under permutation of its argument, and let  $g : \mathbb{S}^n \rightarrow \mathbb{R}$  be the convex function defined as  $g(\mathbf{N}) = f(\lambda(\mathbf{N}))$ . Here  $\lambda(\mathbf{N})$  refers to the list of eigenvalues of  $\mathbf{N}$ . Then we have that*

$$\begin{aligned} g(\mathbf{N}) \leq t \\ \Updownarrow \\ \exists \mathbf{x} \in \mathbb{R}^n \text{ s.t. } f(\mathbf{x}) \leq t \\ \mathbf{x}_1 \geq \cdots \geq \mathbf{x}_n \\ s_r(\mathbf{N}) \leq \mathbf{x}_1 + \cdots + \mathbf{x}_r, \quad r = 1, \dots, n-1 \\ \text{Tr}(\mathbf{N}) = \mathbf{x}_1 + \cdots + \mathbf{x}_n, \end{aligned}$$

where  $s_r(\mathbf{N})$  is the sum of the  $r$  largest eigenvalues of  $\mathbf{N}$ .

The function  $s_r$  is SDP-representable [7]. Further, we observe that the Von-Neumann entropy function  $H(\mathbf{N})$  is concave and invariant under conjugation of the argument  $\mathbf{N}$  by any orthogonal matrix. That is, it is a concave and permutation-invariant function of the eigenvalues of  $\mathbf{N}$ :

$$H(\mathbf{N}) = -d(\lambda(\mathbf{N}), \mathbf{1}),$$

where  $\mathbf{1}$  is the all-ones vector. Consequently, convex optimization problems involving linear constraints as well as constraints on the Von-Neumann entropy function  $H(\mathbf{N})$  are REPs with additional semidefinite constraints.

Several previously proposed methods for solving problems in varied application domains can be viewed as special cases of Von-Neumann entropy optimization problems, involving the maximization of the Von-Neumann entropy of a positive semidefinite matrix subject to affine constraints on the matrix. We briefly survey these methods here.

*Quantum state tomography* Quantum state tomography arises in the characterization of optical systems and in quantum computation [26]. The goal is to reconstruct a quantum state specified by a density matrix given partial information about the state. Such information is typically provided via *measurements* of the state, which can be viewed as linear functionals of the underlying density matrix. As measurements can frequently be expensive to obtain, tomography is a type of inverse problem in which one must choose among the many density matrices that satisfy the limited measurement information that is acquired. One proposal for tomography, based on the original work of Jaynes [37] on the maximum entropy principle, is to find the Von-Neumann-entropy-maximizing density matrix among all density matrices consistent with the measurements [26]—the rationale behind this approach is that the entropy-maximizing matrix makes the least assumptions about the quantum state beyond the constraints imposed by the acquired measurements. Using this method, the optimal reconstructed quantum state can be computed as the solution of a Von-Neumann entropy optimization problem.

*Equilibrium densities in statistical physics* In statistical mechanics, a basic objective is to investigate the properties of a system at equilibrium given information about macroscopic attributes of the system such as energy, number of particles, volume, and pressure. In mathematical terms, the situation is quite similar to the previous setting with quantum state tomography—specifically, the system is described by a density matrix (as in quantum state tomography, this matrix is symmetric positive semidefinite with unit trace), and the macroscopic properties can be characterized as constraints on linear functionals of this density matrix. The Massieu–Planck extremum principle in statistical physics states that the system at equilibrium is given by the Von-Neumann-entropy-maximizing density matrix among all density matrices that satisfy the specified constraints [50]. As before, this equilibrium density can be computed efficiently via Von-Neumann entropy optimization.

*Kernel learning* A commonly encountered problem in machine learning is to measure similarity or affinity between entities that may not belong to a Euclidean space, e.g., text documents. A first step in such settings is to specify coordinates for the entities in a linear vector space after performing some nonlinear mapping, which is followed by a computation of pairwise similarities in the linear space. Kernel methods approach this problem implicitly by working directly with inner-products between pairs of entities, thus combining the nonlinear mapping and the similarity computation in a single step. This approach leads naturally to the question of learning a good kernel matrix, which is a symmetric positive semidefinite matrix specifying similarities between pairs of entities. Kulis et al. [39] propose to minimize the quantum relative entropy between a kernel  $\mathbf{M} \in \mathbb{S}_+^n$  (the decision variable) and a given reference kernel  $\mathbf{M}_0 \in \mathbb{S}_+^n$  (specifying prior knowledge about the affinities among the  $n$  entities):

$$D(\mathbf{M}, \mathbf{M}_0) = -H(\mathbf{M}) - \text{Tr}[\mathbf{M} \log(\mathbf{M}_0)]. \quad (29)$$

This minimization is subject to the decision variable  $\mathbf{M}$  being positive definite as well as constraints on  $\mathbf{M}$  that impose bounds on similarities between pairs of entities in

the linear space. We refer the reader to [39] for more details, and for the virtues of minimizing the quantum relative entropy as a method for learning kernel matrices. In the context of the present paper, the optimization problems considered in [39] can be viewed as Von-Neumann entropy maximization problems subject to linear constraints, as the second argument in the quantum relative entropy function (29) is fixed.

### 4.3 Approximating quantum channel capacity via Von-Neumann entropy optimization

In this section we describe a Von-Neumann entropy optimization approach to bounding the capacity of quantum communication channels. A quantum channel transmits quantum information, and every such channel has an associated *quantum capacity* that characterizes the maximum rate at which quantum information can be reliably communicated across the channel. We refer the reader to the literature for an overview of this topic [35, 45, 49, 51]. The input and the output of a quantum channel are quantum states that are specified by quantum density matrices. Formally, a quantum channel is characterized by a linear operator  $\mathcal{L} : \mathbb{S}^n \rightarrow \mathbb{S}^k$  that maps density matrices to density matrices; the dimensions of the input and output density matrices may, in general, be different. For the linear operator  $\mathcal{L}$  to specify a valid quantum channel, it must map positive semidefinite matrices in  $\mathbb{S}^n$  to positive semidefinite matrices in  $\mathbb{S}^k$  and it must be trace-preserving. These properties ensure that density matrices are mapped to density matrices. In addition,  $\mathcal{L}$  must satisfy a property known as complete-positivity, which is required by the postulates of quantum mechanics—see [45] for more details. Any linear map  $\mathcal{L} : \mathbb{S}^n \rightarrow \mathbb{S}^k$  that satisfies all these properties can be expressed via matrices<sup>2</sup>  $\mathbf{A}^{(j)} \in \mathbb{R}^{k \times n}$  as follows [33]:

$$\mathcal{L}(\rho) = \sum_j \mathbf{A}^{(j)} \rho \mathbf{A}^{(j)'} \quad \text{where} \quad \sum_j \mathbf{A}^{(j)'} \mathbf{A}^{(j)} = \mathbf{I}.$$

Letting  $S^n$  denote the  $n$ -sphere, the capacity of the channel specified by  $\mathcal{L} : \mathbb{S}^n \rightarrow \mathbb{S}^k$  is defined as:

$$C(\mathcal{L}) \triangleq \sup_{\substack{\mathbf{v}^{(i)} \in S^n \\ p_i \geq 0, \sum_i p_i = 1}} H \left[ \mathcal{L} \left( \sum_i p_i \mathbf{v}^{(i)} \mathbf{v}^{(i)'} \right) \right] - \sum_i p_i H \left[ \mathcal{L} \left( \mathbf{v}^{(i)} \mathbf{v}^{(i)'} \right) \right]. \quad (30)$$

Here  $H$  is the Von-Neumann entropy function (28) and the number of states  $\mathbf{v}^{(i)}$  is part of the optimization.

We note that there are several kinds of capacities associated with a quantum channel, depending on the protocol employed for encoding and decoding states transmitted across the channel; the version considered here is the one most commonly investigated, and we refer the reader to Shor's survey [51] for details of the other types of capacities.

<sup>2</sup> In full generality, density matrices are trace-one, positive semidefinite *Hermitian* matrices, and  $\mathbf{A}^{(j)} \in \mathbb{C}^{k \times n}$ . As with SDPs, the Von-Neumann entropy optimization framework can handle linear matrix inequalities on Hermitian matrices, but we stick with the real case for simplicity.

The quantity  $C(\mathcal{L})$  (30) on which we focus here is called the  $C_{1,\infty}$  quantum capacity—roughly speaking, it is the capacity of a quantum channel in which the sender cannot couple inputs across multiple uses of the channel, while the receiver can jointly decode messages received over multiple uses of the channel.

*Shor's approach for lower bounds via LP* In [51] Shor describes a procedure based on LP to provide lower bounds on  $C(\mathcal{L})$ . As the first step of this method, one fixes a finite set of states  $\{\mathbf{v}^{(i)}\}_{i=1}^m$  with each  $\mathbf{v}^{(i)} \in S^n$  and a density matrix  $\rho \in \mathbb{S}^n$ , so that  $\rho$  is in the convex hull  $\text{conv}(\{\mathbf{v}^{(i)}\mathbf{v}^{(i)'}\}_{i=1}^m)$ . With these quantities fixed, one can obtain a lower bound on  $C(\mathcal{L})$  by solving the following LP:

$$C\left(\mathcal{L}, \left\{\mathbf{v}^{(i)}\right\}_{i=1}^m, \rho\right) = \sup_{\substack{p \in \mathbb{R}_+^m, \sum_i p_i = 1 \\ \rho = \sum_i p_i \mathbf{v}^{(i)} \mathbf{v}^{(i)'}}} H[\mathcal{L}(\rho)] - \sum_i p_i H\left[\mathcal{L}\left(\mathbf{v}^{(i)} \mathbf{v}^{(i)'}\right)\right] \quad (31)$$

In [51] Shor also suggests local heuristics to search for better sets of states and density matrices to improve the lower bound (31). It is clear that  $C\left(\mathcal{L}, \left\{\mathbf{v}^{(i)}\right\}_{i=1}^m, \rho\right)$  as computed in (31) is a lower bound on  $C(\mathcal{L})$ . Indeed, one can check that:

$$C(\mathcal{L}) = \sup_{\mathbf{v}^{(i)} \in S^n} \sup_{\rho \in \text{conv}\{\mathbf{v}^{(i)} \mathbf{v}^{(i)'}\}} C\left(\mathcal{L}, \left\{\mathbf{v}^{(i)}\right\}, \rho\right). \quad (32)$$

Here the number of states is not explicitly denoted, and it is also a part of the optimization.

*Improved lower bounds via Von-Neumann entropy optimization* In the computation of the lower bound by Shor's approach (31), the density matrix  $\rho$  remains fixed and the optimization is over decompositions of  $\rho$  as a convex combination of elements from the set  $\{\mathbf{v}^{(i)} \mathbf{v}^{(i)'}\}_{i=1}^m$ . We describe an improvement upon Shor's lower bound (31) by further optimizing over the set of density matrices  $\rho$  in the convex hull  $\text{conv}(\{\mathbf{v}^{(i)} \mathbf{v}^{(i)'}\}_{i=1}^m)$ . Our improved lower bound entails the solution of a Von-Neumann entropy optimization problem. Specifically, we observe that for a fixed set of states  $\{\mathbf{v}^{(i)}\}_{i=1}^m$  the following optimization problem involves Von-Neumann entropy maximization subject to affine constraints:

$$\begin{aligned} C\left(\mathcal{L}, \left\{\mathbf{v}^{(i)}\right\}_{i=1}^m\right) &= \sup_{\rho \in \text{conv}\{\mathbf{v}^{(i)} \mathbf{v}^{(i)'}\}_{i=1}^m} C\left(\mathcal{L}, \left\{\mathbf{v}^{(i)}\right\}_{i=1}^m, \rho\right) \\ &= \sup_{p_i \geq 0, \sum_{i=1}^m p_i = 1} H\left[\mathcal{L}\left(\sum_{i=1}^m p_i \mathbf{v}^{(i)} \mathbf{v}^{(i)'}\right)\right] \\ &\quad - \sum_{i=1}^m p_i H\left[\mathcal{L}\left(\mathbf{v}^{(i)} \mathbf{v}^{(i)'}\right)\right]. \end{aligned} \quad (33)$$

In particular, the optimization over density matrices  $\rho \in \text{conv}\{\mathbf{v}^{(i)}\mathbf{v}^{(i)'}\}_{i=1}^m$  in (32) can be folded into the computation of (31) at the expense of solving a Von-Neumann entropy maximization problem instead of an LP. It is easily seen from (30), (32), and (33) that for a fixed set of states  $\{\mathbf{v}^{(i)}\}_{i=1}^m$ :

$$C(\mathcal{L}) \geq C\left(\mathcal{L}, \left\{\mathbf{v}^{(i)}\right\}_{i=1}^m\right) \geq C\left(\mathcal{L}, \left\{\mathbf{v}^{(i)}\right\}_{i=1}^m, \rho\right).$$

For a finite set of states  $\{\mathbf{v}^{(i)}\}_{i=1}^m$ , the quantity  $C(\mathcal{L}, \{\mathbf{v}^{(i)}\}_{i=1}^m)$  in (33) is referred to as the *classical-to-quantum capacity* with respect to the states  $\{\mathbf{v}^{(i)}\}_{i=1}^m$  [35, 49]; the reason for this name is that  $C(\mathcal{L}, \{\mathbf{v}^{(i)}\}_{i=1}^m)$  is the capacity of a quantum channel in which the input is “classical” as it is restricted to be a convex combination of the finite set  $\{\mathbf{v}^{(i)}\mathbf{v}^{(i)'}\}_{i=1}^m$ . One can improve upon the bound  $C(\mathcal{L}, \{\mathbf{v}^{(i)}\}_{i=1}^m)$  by progressively adding more states to the collection  $\{\mathbf{v}^{(i)}\}_{i=1}^m$ . It is of interest to develop techniques to accomplish this in a principled manner.

In summary, in Shor’s approach one fixes a set of states  $\{\mathbf{v}^{(i)}\}_{i=1}^m$  and a density matrix  $\rho \in \text{conv}(\{\mathbf{v}^{(i)}\mathbf{v}^{(i)'}\}_{i=1}^m)$ , and one obtains a lower bound on the  $C_{1,\infty}$  quantum capacity  $C(\mathcal{L})$  (30) by optimizing over decompositions of  $\rho$  in terms of convex combinations of elements of the set  $\{\mathbf{v}^{(i)}\mathbf{v}^{(i)'}\}_{i=1}^m$ . In contrast, in our method we fix a set of states  $\{\mathbf{v}^{(i)}\}_{i=1}^m$ , and we optimize simultaneously over density matrices in  $\text{conv}(\{\mathbf{v}^{(i)}\mathbf{v}^{(i)'}\}_{i=1}^m)$  as well as decompositions of these matrices. Shor’s method involves the solution of an LP, while the improved bound using our approach comes at the cost of solving a Von-Neumann entropy optimization problem. The question of computing the  $C_{1,\infty}$  capacity  $C(\mathcal{L})$  (30) exactly in a tractable fashion remains open.

#### 4.3.1 Numerical examples

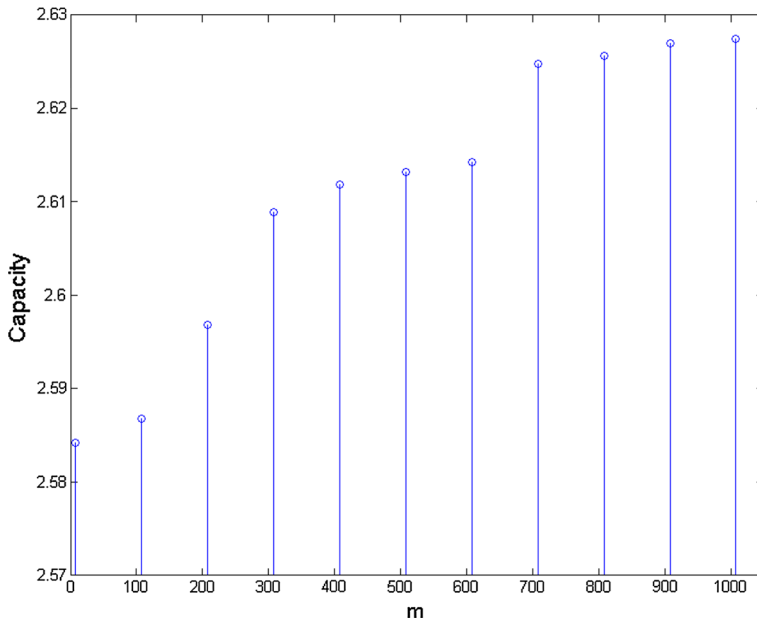
We consider a quantum channel given by a linear operator  $\mathcal{L} : \mathbb{S}^8 \rightarrow \mathbb{S}^8$  with:

$$\mathcal{L}(\rho) = \mathbf{A}^{(1)}\rho\mathbf{A}^{(1)'} + \epsilon^2\mathbf{A}^{(2)}\rho\mathbf{A}^{(2)'} + \mathbf{A}^{(3)}\rho\mathbf{A}^{(3)'}, \quad (34)$$

where  $\mathbf{A}^{(1)}$  is a random  $8 \times 8$  diagonal matrix (entries chosen to be i.i.d. standard Gaussian),  $\mathbf{A}^{(2)}$  is a random  $8 \times 8$  matrix (entries chosen to be i.i.d. standard Gaussian),  $\epsilon \in [0, 1]$ , and  $\mathbf{A}^{(3)}$  is the symmetric square root of  $\mathbf{I} - \mathbf{A}^{(1)'}\mathbf{A}^{(1)} - \epsilon^2\mathbf{A}^{(2)'}\mathbf{A}^{(2)}$ . We scale  $\mathbf{A}^{(1)}$  and  $\mathbf{A}^{(2)}$  suitably so that  $\mathbf{I} - \mathbf{A}^{(1)'}\mathbf{A}^{(1)} - \epsilon^2\mathbf{A}^{(2)'}\mathbf{A}^{(2)}$  is positive definite for all  $\epsilon \in [0, 1]$ . For this quantum channel, we describe the results of two numerical experiments.

In the first experiment, we compute the classical-to-quantum capacity of this channel with the input states  $\mathbf{v}^{(i)} \in \mathbb{S}^8$  for  $i = 1, \dots, 8$  being the 8 standard basis vectors in  $\mathbb{R}^8$ . In other words, the input density  $\rho \in \mathbb{S}_+^8$  is given by a diagonal matrix of unit trace. We add unit vectors in  $\mathbb{R}^8$  at random to this collection, and we plot the increase in Fig. 2 in the corresponding classical-to-quantum channel capacity (33)—in each case, the capacity is evaluated at  $\epsilon = 0.5$ . In this manner, Von-Neumann entropy optimization problems can be used to provide successively tighter lower bounds for the capacity of a quantum channel.



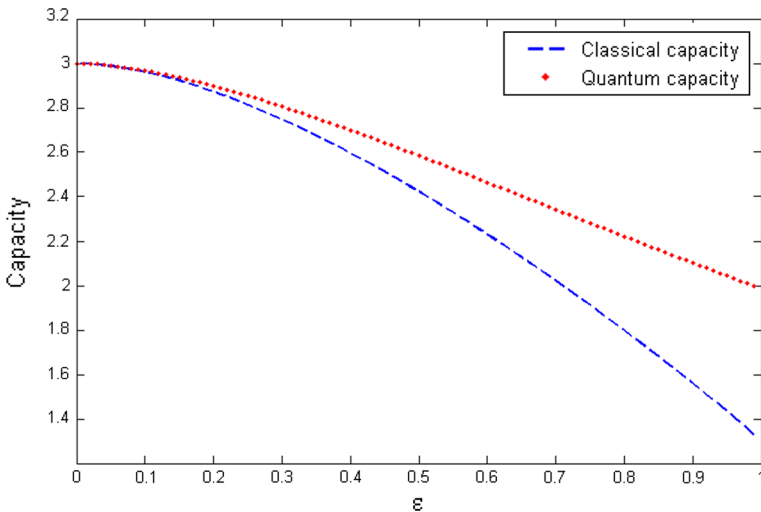


**Fig. 2** A sequence of increasingly tighter lower bounds on the quantum capacity of the channel specified by (34), obtained by computing classical-to-quantum capacities with increasingly larger collections of input states

A quantum channel can also be restricted in a natural manner to a classical channel. In the next experiment, we compare the capacity of such a purely classical channel induced by the quantum channel (34) to a classical-to-quantum capacity of the channel (34). In both cases we consider input states  $\mathbf{v}^{(i)} \in \mathbb{S}^8$ ,  $i = 1, \dots, 8$  given by the 8 standard basis vectors in  $\mathbb{R}^8$ . With these input states, the classical-to-quantum capacity of the channel (34) is computed via Von-Neumann entropy optimization by solving (33); in other words, this capacity corresponds to a setting in which the input is classical (i.e., a diagonal density matrix) while the output is in general a quantum state (i.e., specified by an arbitrary density matrix). To obtain a purely classical channel from (34) in which both the inputs and the outputs are classical, let  $\mathfrak{P}_{\text{diag}} : \mathbb{S}^8 \rightarrow \mathbb{S}^8$  denote an operator that projects a matrix onto the space of diagonal matrices (i.e., zeros out the off-diagonal entries), and consider the following linear map:

$$\mathcal{L}_{\text{classical}}(\rho) \triangleq \mathfrak{P}_{\text{diag}}(\mathcal{L}(\rho)). \quad (35)$$

For diagonal input density matrices  $\rho$  (i.e., classical inputs), this map specifies a classical channel induced by the quantum channel (34), as the output is also restricted to be a diagonal density matrix (i.e., a classical output). Figure 3 shows two plots for  $\epsilon$  ranging from 0 to 1 of both the classical-to-quantum capacity of (34) and the classical capacity of the induced classical channel (35). Note that for  $\epsilon = 0$  the output of the operator (34) is diagonal if the input is given by a diagonal density matrix. Therefore, the two curves coincide at  $\epsilon = 0$  in Fig. 3. For larger values of  $\epsilon$ , the classical-to-



**Fig. 3** Comparison of a classical-to-quantum capacity of the quantum channel specified by (34) and the classical capacity of a classical channel induced by the quantum channel given by (34)

quantum channel is greater than the capacity of the induced classical channel, thus demonstrating the improvement in capacity as one goes from a classical-to-classical communication protocol to a classical-to-quantum protocol. (The  $C_{1,\infty}$  capacity  $C(\mathfrak{L})$  (34) of the channel (33) is in general even larger than the classical-to-quantum capacity computed here.)

#### 4.4 QREP bounds for a robust matrix GP

Moving beyond the family of Von-Neumann entropy optimization problems, which are effectively just REPs with additional semidefinite constraints, we discuss QREPs that exploit the full expressive power of the quantum relative entropy function. In contrast to REPs, general QREPs are optimization problems involving non-commuting variables. This non-commutativity leads to some important distinctions between the classical and the quantum settings, and we present these differences in the context of a stylized application.

We investigate the problem of obtaining bounds on the optimal value of a matrix analog of a robust GP. Specifically, given a collection of matrices  $\mathbf{B}^{(j)} \in \mathbb{S}^k$ ,  $j = 1, \dots, n$  and a convex set  $\mathcal{M} \subset \mathbb{S}_+^k$ , consider the following function of  $\mathcal{M}$  and the  $\mathbf{B}^{(j)}$ 's:

$$F(\mathcal{M}; \{\mathbf{B}^{(j)}\}_{j=1}^n) = \sup_{\mathbf{C} \in \mathcal{M}} \left[ \inf_{\substack{\mathbf{x} \in \mathbb{R}^n \\ \mathbf{Y} \in \mathbb{S}^k \\ \mathbf{Z} \in \mathbb{S}_+^k}} \text{Tr}(\mathbf{C}\mathbf{Z}) \text{ s.t. } \mathbf{Y} \preceq \log(\mathbf{Z}), \sum_{j=1}^n \mathbf{B}^{(j)} \mathbf{x}_j = \mathbf{Y} \right]. \quad (36)$$

In the inner optimization problem, the constraint set specified by the inequality  $\mathbf{Y} \preceq \log(\mathbf{Z})$  is a convex set as the matrix logarithm function is operator concave. The inner problem is a non-commutative analog of an unconstrained GP (3); the set  $\mathcal{M}$ , over which the outer supremum is taken, specifies “coefficient uncertainty.” Hence, the nested optimization problem (36) is a matrix equivalent of a robust unconstrained GP with coefficient uncertainty. To see this connection more precisely, consider the analogous problem over vectors for a convex set  $\mathcal{C} \subset \mathbb{R}_+^n$  and a collection  $\{\mathbf{b}^{(j)}\}_{j=1}^n \subset \mathbb{R}^k$ :

$$f(\mathcal{C}; \{\mathbf{b}^{(j)}\}_{j=1}^n) = \sup_{\mathbf{c} \in \mathcal{C}} \inf_{\substack{\mathbf{x} \in \mathbb{R}^n \\ \mathbf{y} \in \mathbb{R}^k \\ \mathbf{z} \in \mathbb{R}_+^n}} \mathbf{c}'\mathbf{z} \text{ s.t. } \mathbf{y} \leq \log(\mathbf{z}), \sum_{j=1}^n \mathbf{b}^{(j)}\mathbf{x}_j = \mathbf{y}. \quad (37)$$

The inner optimization problem here is simply an unconstrained GP, and the set  $\mathcal{C}$  specifies coefficient uncertainty. The reason for this somewhat non-standard description—an equivalent, more familiar specification of (37) via sums of exponentials as in (3) would be  $\sup_{\mathbf{c} \in \mathcal{C}} \inf_{\mathbf{x} \in \mathbb{R}^n} \sum_{i=1}^k \mathbf{c}_i \exp(\mathbf{a}^{(i)'}\mathbf{x})$ , where  $\mathbf{a}^{(i)}$  is the  $i$ 'th row of the  $k \times n$  matrix consisting of the  $\mathbf{b}^{(j)}$ 's as columns—is to make a more transparent connection between the matrix case (36) and the vector case (37).

Our method for obtaining bounds on  $F(\mathcal{M}; \{\mathbf{B}^{(j)}\}_{j=1}^n)$  is based on a relationship between the constraint  $\mathbf{Y} \preceq \log(\mathbf{Z})$  and the quantum relative entropy function via convex conjugacy. To begin with, we describe the relationship between the vector constraint  $\mathbf{y} \leq \log(\mathbf{z})$  and classical relative entropy. Consider the following characteristic function for  $\mathbf{y} \in \mathbb{R}^k$  and  $\mathbf{z} \in -\mathbb{R}_+^n$ :

$$\chi_{\text{aff-log}}(\mathbf{y}, \mathbf{z}) = \begin{cases} 0, & \text{if } \mathbf{y} \leq \log(-\mathbf{z}) \\ \infty, & \text{otherwise.} \end{cases} \quad (38)$$

We then have the following result:

**Lemma 1** Let  $\chi_{\text{aff-log}}(\mathbf{y}, \mathbf{z}) : \mathbb{R}^k \times -\mathbb{R}_+^n \rightarrow \mathbb{R}$  denote the characteristic function defined in (38). Then the convex conjugate  $\chi_{\text{aff-log}}^*(\mathbf{v}, \boldsymbol{\lambda}) = d(\mathbf{v}, e\boldsymbol{\lambda})$  with domain  $(\mathbf{v}, \boldsymbol{\lambda}) \in \mathbb{R}_+^k \times \mathbb{R}_+^n$ , where  $e$  is Euler's constant.

The proof of this lemma follows from a straightforward calculation. Based on this result and by computing the dual of the inner convex program in (37), we have that the function  $f(\mathcal{C}; \{\mathbf{b}^{(j)}\}_{j=1}^n)$  can be computed as the optimal value of an REP:

$$f(\mathcal{C}; \{\mathbf{b}^{(j)}\}_{j=1}^n) = \sup_{\mathbf{c} \in \mathcal{C}, \mathbf{v} \in \mathbb{R}_+^k} -d(\mathbf{v}, e\mathbf{c}) \text{ s.t. } \mathbf{b}^{(j)'}\mathbf{v} = 0, j = 1, \dots, n. \quad (39)$$

Moving to the matrix case, consider the natural matrix analog of the characteristic function (38) for  $\mathbf{Y} \in \mathbb{S}^k$  and  $\mathbf{Z} \in -\mathbb{S}_+^n$ :

$$\chi_{\text{mat-aff-log}}(\mathbf{Y}, \mathbf{Z}) = \begin{cases} 0, & \text{if } \mathbf{Y} \preceq \log(-\mathbf{Z}) \\ \infty, & \text{otherwise.} \end{cases} \quad (40)$$

As the matrix logarithm is operator concave, this characteristic function is also a convex function. However, its relationship to the quantum relative entropy function is somewhat more complicated, as described in the following proposition:

**Proposition 5** *Let  $\chi_{\text{mat-aff-log}}(\mathbf{Y}, \mathbf{Z}) : \mathbb{S}^k \times -\mathbb{S}_+^k \rightarrow \mathbb{R}$  denote the characteristic function defined in (40). Then the convex conjugate  $\chi_{\text{mat-aff-log}}^*(\mathbf{N}, \mathbf{A}) \leq D(\mathbf{N}, e\mathbf{A})$ , with equality holding if the matrices  $\mathbf{N}$  and  $\mathbf{A}$  are simultaneously diagonalizable. Here  $(\mathbf{N}, \mathbf{A}) \in \mathbb{S}_+^k \times \mathbb{S}_+^k$ , and  $e$  is again Euler's constant.*

*Proof* The convex conjugate of  $\chi_{\text{mat-aff-log}}(\mathbf{Y}, \mathbf{Z})$  can be simplified as follows:

$$\begin{aligned} \chi_{\text{mat-aff-log}}^*(\mathbf{N}, \mathbf{A}) &= \sup_{\substack{\mathbf{Y} \in \mathbb{S}^k \\ \mathbf{Z} \in -\mathbb{S}_+^k}} \text{Tr}(\mathbf{N}\mathbf{Y}) + \text{Tr}(\mathbf{A}\mathbf{Z}) \text{ s.t. } \mathbf{Y} \preceq \log(-\mathbf{Z}) \\ &= \sup_{\mathbf{Z} \in -\mathbb{S}_+^k} \text{Tr}[\mathbf{N} \log(-\mathbf{Z})] + \text{Tr}(\mathbf{A}\mathbf{Z}) \\ &\quad [\text{set } \mathbf{W} = \log(-\mathbf{Z})] \\ &= \sup_{\mathbf{W} \in \mathbb{S}^k} \text{Tr}(\mathbf{N}\mathbf{W}) - \text{Tr}[\mathbf{A} \exp(\mathbf{W})] \\ &= \sup_{\mathbf{W} \in \mathbb{S}^k} \text{Tr}(\mathbf{N}\mathbf{W}) - \text{Tr}[\exp(\log(\mathbf{A})) \exp(\mathbf{W})]. \end{aligned} \quad (41)$$

In order to simplify further, we appeal to the Golden–Thompson inequality [25, 52], which states that:

$$\text{Tr}[\exp(\mathbf{A} + \mathbf{B})] \leq \text{Tr}[\exp(\mathbf{A}) \exp(\mathbf{B})], \quad (42)$$

for Hermitian matrices  $\mathbf{A}, \mathbf{B}$ . Equality holds here if the matrices  $\mathbf{A}$  and  $\mathbf{B}$  commute. Consequently, we can bound  $\chi_{\text{mat-aff-log}}^*(\mathbf{N}, \mathbf{A})$  as:

$$\begin{aligned} \chi_{\text{mat-aff-log}}^*(\mathbf{N}, \mathbf{A}) &\stackrel{(i)}{\leq} \sup_{\mathbf{W} \in \mathbb{S}^k} \text{Tr}(\mathbf{N}\mathbf{W}) - \text{Tr}[\exp(\log(\mathbf{A}) + \mathbf{W})] \\ &\quad [\text{set } \mathbf{U} = \exp(\log(\mathbf{A}) + \mathbf{W})] \\ &= \sup_{\mathbf{U} \in \mathbb{S}_+^k} \text{Tr}[\mathbf{N} \log(\mathbf{U})] - \text{Tr}[\mathbf{N} \log(\mathbf{A})] - \text{Tr}(\mathbf{U}) \\ &= \left[ \sup_{\mathbf{U} \in \mathbb{S}_+^k} \text{Tr}[\mathbf{N} \log(\mathbf{U})] - \text{Tr}(\mathbf{U}) \right] - \text{Tr}[\mathbf{N} \log(\mathbf{A})] \\ &\stackrel{(ii)}{=} \text{Tr}[\mathbf{N} \log(\mathbf{N})] - \text{Tr}(\mathbf{N}) - \text{Tr}[\mathbf{N} \log(\mathbf{A})] = D(\mathbf{N}, e\mathbf{A}). \end{aligned}$$

Here (i) follows from the Golden–Thompson inequality (42), and the equality (ii) follows from the fact that the optimal  $\mathbf{U}$  in the previous line is  $\mathbf{N}$ . Consequently, one can check that if  $\mathbf{N}$  and  $\mathbf{A}$  are simultaneously diagonalizable, then the inequality (i) becomes an equality.  $\square$

Thus, the non-commutativity underlying the quantum relative entropy function in contrast to the classical case results in  $D(\mathbf{N}, e\mathbf{A})$  only being an upper bound (in general) on the convex conjugate  $\chi_{\text{mat-aff-log}}^*(\mathbf{N}, \mathbf{A})$ . From the perspective of this result,

Lemma 1 follows from the observation that the relative entropy between two nonnegative vectors can be viewed as the quantum relative entropy between two diagonal positive semidefinite matrices (which are, trivially, simultaneously diagonalizable).

Based on Proposition 5 and again appealing to convex duality, the function  $F(\mathcal{M}; \{\mathbf{B}^{(j)}\}_{j=1}^n)$  can be bounded below by solving a QREP as follows:

$$F(\mathcal{M}; \{\mathbf{B}^{(j)}\}_{j=1}^n) \geq \sup_{\substack{\mathbf{C} \in \mathcal{M} \\ \mathbf{N} \in \mathbb{S}_+^k}} -D(\mathbf{N}, e\mathbf{C}) \quad \text{s.t. } \text{Tr}(\mathbf{B}^{(j)}\mathbf{N}) = 0, \quad j = 1, \dots, n. \quad (43)$$

The quantity on the right-hand-side can be computed, for example, via projected coordinate descent. If the matrices  $\{\mathbf{B}^{(j)}\}_{j=1}^n \cup \{\mathbf{C}\}$  are simultaneously diagonalizable for each  $\mathbf{C} \in \mathcal{M}$ , then the QREP lower bound (43) is equal to  $F(\mathcal{M}; \{\mathbf{B}^{(j)}\}_{j=1}^n)$ . In summary, we have that REPs are useful for computing  $f(\mathcal{C}; \{\mathbf{b}^{(j)}\}_{j=1}^n)$  exactly, while QREPs only provide a lower bound via (43) of  $F(\mathcal{M}; \{\mathbf{B}^{(j)}\}_{j=1}^n)$  in general.

## 5 Further directions

There are several avenues for further research that arise from this paper. It is of interest to develop efficient numerical methods to solve REPs and QREPs in order to scale to large problem instances. Such massive size problems are especially prevalent in data analysis tasks, and are of interest in settings such as kernel learning. On a related note, there exists a vast literature on exploiting the structure of a particular problem instance of an SDP or a GP—e.g., sparsity in the problem parameters—which can result in significant computational speedups in practice in the solution of these problems. A similar set of techniques would be useful and relevant in all of the applications described in this paper.

We also seek a deeper understanding of the expressive power of REPs and QREPs, i.e., of conditions under which convex sets can be tractably represented via REPs and QREPs as well as obstructions to efficient representations in these frameworks (in the spirit of similar results that have recently been obtained for SDPs [9, 27, 34]). Such an investigation would be useful in identifying problems in statistics and information theory that may be amenable to solution via tractable convex optimization techniques.

**Acknowledgments** The authors would like to thank Pablo Parrilo and Yong-Sheng Soh for helpful conversations, and Leonard Schulman for pointers to the literature on Von-Neumann entropy. Venkat Chandrasekaran was supported in part by National Science Foundation Career award CCF-1350590 and Air Force Office of Scientific Research grant FA9550-14-1-0098.

## 6 Appendix

The second-order cone  $L_2 \subset \mathbb{R}^3$  from (5) can be written as:

$$L_2 = \left\{ (\mathbf{x}, y) \in \mathbb{R}^2 \times \mathbb{R} \mid \begin{pmatrix} y - \mathbf{x}_1 & \mathbf{x}_2 \\ \mathbf{x}_2 & y + \mathbf{x}_1 \end{pmatrix} \succeq \mathbf{0} \right\}.$$

Combining this reformulation with the next result gives us the description (6).

**Proposition 6** *We have that  $\begin{pmatrix} a & c \\ c & b \end{pmatrix} \in \mathbb{S}_+^2$  if and only if there exists  $\nu \in \mathbb{R}_+$  such that:*

$$\begin{aligned} \nu \log\left(\frac{\nu}{a}\right) + \nu \log\left(\frac{\nu}{b}\right) - 2\nu &\leq 2c \\ \nu \log\left(\frac{\nu}{a}\right) + \nu \log\left(\frac{\nu}{b}\right) - 2\nu &\leq -2c \\ a, b &\in \mathbb{R}_+. \end{aligned}$$

*Proof* We have that  $\begin{pmatrix} a & c \\ c & b \end{pmatrix} \in \mathbb{S}_+^2$  if and only if  $a\bar{\mathbf{z}}_1^2 + b\bar{\mathbf{z}}_2^2 + 2c\bar{\mathbf{z}}_1\bar{\mathbf{z}}_2 \geq 0$ ,  $\forall \bar{\mathbf{z}} \in \mathbb{R}^2$ . This latter condition can in turn be rewritten to obtain the following reformulation:

$$\begin{pmatrix} a & c \\ c & b \end{pmatrix} \in \mathbb{S}_+^2 \Leftrightarrow a\mathbf{z}_1^2 + b\mathbf{z}_2^2 + 2c\mathbf{z}_1\mathbf{z}_2 \geq 0 \text{ and } a\mathbf{z}_1^2 + b\mathbf{z}_2^2 - 2c\mathbf{z}_1\mathbf{z}_2 \geq 0 \forall \mathbf{z} \in \mathbb{R}_+^2. \quad (44)$$

Each of these inequalities with universal quantifiers can be reformulated by appealing to GP duality. Specifically, based on the change of variables  $\mathbf{w}_i \leftarrow \log(\mathbf{z}_i)$ ,  $i = 1, 2$ , which is commonly employed in the GP literature [11, 20], we have from (44) that  $\begin{pmatrix} a & c \\ c & b \end{pmatrix} \in \mathbb{S}_+^2$  if and only if:

$$\inf_{\mathbf{w} \in \mathbb{R}^2} a \exp\{\mathbf{w}_1 - \mathbf{w}_2\} + b \exp\{\mathbf{w}_2 - \mathbf{w}_1\} \geq \max\{2c, -2c\}. \quad (45)$$

As the optimization problem on the left-hand-side is a GP for  $a, b \in \mathbb{R}_+$ , we can appeal to convex duality to conclude that

$$\inf_{\mathbf{w} \in \mathbb{R}^2} a \exp\{\mathbf{w}_1 - \mathbf{w}_2\} + b \exp\{\mathbf{w}_2 - \mathbf{w}_1\} = \sup_{\nu \in \mathbb{R}_+} -\nu \log\left(\frac{\nu}{a}\right) - \nu \log\left(\frac{\nu}{b}\right) + 2\nu.$$

Combining this result with (45), we have that  $\begin{pmatrix} a & c \\ c & b \end{pmatrix} \in \mathbb{S}_+^2$  if and only if:

$$a, b \in \mathbb{R}_+ \text{ and } \exists \nu \in \mathbb{R}_+ \text{ s.t. } -\nu \log\left(\frac{\nu}{a}\right) - \nu \log\left(\frac{\nu}{b}\right) + 2\nu \geq \max\{2c, -2c\}. \quad (46)$$

This concludes the proof.  $\square$

## References

1. Bapat, R.B., Beg, M.I.: Order statistics for nonidentically distributed variables and permanents. *Sankhya Indian J. Stat. A* **51**, 79–93 (1989)
2. Barvinok, A.I.: Computing mixed discriminants, mixed volumes, and permanents. *Discrete Comput. Geom.* **18**, 205–237 (1997)
3. Ben-Tal, A., El Ghaoui, L., Nemirovski, A.: *Robust Optimization*. Princeton University Press, Princeton (2009)
4. Ben-Tal, A., Nemirovski, A.: Optimal design of engineering structures. *Optima*. **47**, 4–8 (1995)
5. Ben-Tal, A., Nemirovski, A.: Robust convex optimization. *Math. Oper. Res.* **23**, 769–805 (1998)
6. Ben-Tal, A., Nemirovski, A.: On polyhedral approximations of the second-order cone. *Math. Oper. Res.* **26**, 193–205 (2001)

7. Ben-Tal, A., Nemirovskii, A.: Lectures on Modern Convex Optimization. Society for Industrial and Applied Mathematics, Philadelphia (2001)
8. Betke, U.: Mixed volumes of polytopes. *Arch. Math.* **58**, 388–391 (1992)
9. Blekherman, G., Parrilo, P., Thomas, R.: Semidefinite Optimization and Convex Algebraic Geometry. Society for Industrial and Applied Mathematics, Philadelphia (2013)
10. Boyd, S., Kim, S.J., Patil, D., Horowitz, M.: Digital circuit optimization via geometric programming. *Oper. Res.* **53**, 899–932 (2005)
11. Boyd, S., Kim, S.J., Vandenberghe, L., Hassibi, A.: A tutorial on geometric programming. *Optim. Eng.* **8**, 67–127 (2007)
12. Chandrasekaran, V., Shah, P.: Conic geometric programming. In: Proceedings of the Conference on Information Sciences and Systems (2014)
13. Chandrasekaran, V., Shah, P.: Relative entropy relaxations for signomial optimization. *SIAM J. Optim.* (2014)
14. Chiang, M.: Geometric programming for communication systems. *Found. Trends Commun. Inf. Theory* **2**, 1–154 (2005)
15. Chiang, M., Boyd, S.: Geometric programming duals of channel capacity and rate distortion. *IEEE Trans. Inf. Theory* **50**, 245–258 (2004)
16. Cover, T., Thomas, J.: Elements of Information Theory. Wiley, New York (2006)
17. Cox, D.R.: The regression analysis of binary sequences. *J. R. Stat. Soc.* **20**, 215–242 (1958)
18. Dinkel, J.J., Kochenberger, G.A., Wong, S.N.: Entropy maximization and geometric programming. *Environ. Plan. A* **9**, 419–427 (1977)
19. Drew, J.H., Johnson, C.R.: The maximum permanent of a 3-by-3 positive semidefinite matrix, given the eigenvalues. *Linear Multilinear Algebra* **25**, 243–251 (1989)
20. Duffin, R.J., Peterson, E.L., Zener, C.M.: Geometric Programming: Theory and Application. Wiley, New York (1967)
21. Egorychev, G.P.: Proof of the Van der Waerden conjecture for permanents (english translation; original in russian). *Sib. Math. J.* **22**, 854–859 (1981)
22. El Ghaoui, L., Lebret, H.: Robust solutions to least-squares problems with uncertain data. *SIAM J. Matrix Anal. Appl.* **18**, 1035–1064 (1997)
23. Falikman, D.I.: Proof of the Van der Waerden conjecture regarding the permanent of a doubly stochastic matrix (english translation; original in russian). *Math. Notes* **29**, 475–479 (1981)
24. Glineur, F.: An extended conic formulation for geometric optimization. *Found. Comput. Decis. Sci.* **25**, 161–174 (2000)
25. Golden, S.: Lower bounds for the Helmholtz function. *Phys. Rev. Ser. II* **137**, B1127–B1128 (1965)
26. Goncalves, D.S., Lavor, C., Gomes-Ruggiero, M.A., Cesrio, A.T., Vianna, R.O., Maciel, T.O.: Quantum state tomography with incomplete data: maximum entropy and variational quantum tomography. *Phys. Rev. A* **87** (2013)
27. Gouveia, J., Parrilo, P., Thomas, R.: Lifts of convex sets and cone factorizations. *Math. Oper. Res.* **38**, 248–264 (2013)
28. Grone, R., Johnson, C.R., Eduardo, S.A., Wolkowicz, H.: A note on maximizing the permanent of a positive definite hermitian matrix, given the eigenvalues. *Linear Multilinear Algebra* **19**, 389–393 (1986)
29. Gurvits, L.: Van der Waerden/Schrijver-Valiant like conjectures and stable (aka hyperbolic) homogeneous polynomials: one theorem for all. *Electron. J. Comb.* **15** (2008)
30. Gurvits, L., Samorodnitsky, A.: A deterministic algorithm for approximating the mixed discriminant and mixed volume, and a combinatorial corollary. *Discrete Comput. Geom.* **27**, 531–550 (2002)
31. Han, S., Preciado, V.M., Nowzari, C., Pappas, G.J.: Data-Driven Network Resource Allocation for Controlling Spreading Processes. *IEEE Trans. Netw. Sci. Eng.* **2**(4), 127–38 (2015)
32. Hastie, T., Tibshirani, R., Friedman, J.: The Elements of Statistical Learning. Springer, Berlin (2008)
33. Hellwig, K., Krauss, K.: Operations and measurements II. *Commun. Math. Phys.* **16**, 142–147 (1970)
34. Helton, J.W., Nie, J.: Sufficient and necessary conditions for semidefinite representability of convex hulls and sets. *SIAM J. Optim.* **20**, 759–791 (2009)
35. Holevo, A.S.: The capacity of the quantum channel with general signal states. *IEEE Trans. Inf. Theory* **44**, 269–273 (1998)
36. Hsiung, K.L., Kim, S.J., Boyd, S.: Tractable approximate robust geometric programming. *Optim. Eng.* **9**, 95–118 (2008)
37. Jaynes, E.T.: Information theory and statistical mechanics. *Phys. Rev. Ser. II* **106**, 620–630 (1957)

38. Jerrum, M., Sinclair, A., Vigoda, E.: A polynomial-time approximation algorithm for the permanent of a matrix with non-negative entries. *J. ACM* **51**, 671–697 (2004)
39. Kulis, B., Sustik, M., Dhillon, I.: Low-rank kernel learning with Bregman matrix divergences. *J. Mach. Learn. Res.* **10**, 341–376 (2009)
40. Lieb, E.: Convex trace functions and the Wigner–Yanase–Dyson conjecture. *Adv. Math.* **11**, 267–288 (1973)
41. Linial, N., Samorodnitsky, A., Wigderson, A.: A deterministic strongly polynomial algorithm for matrix scaling and approximate permanents. *Combinatorica* **20**, 545–568 (2000)
42. Lobo, M., Vandenberghe, L., Boyd, S., Lebret, H.: Applications of second-order cone programming. *Linear Algebra Appl.* **284**, 193–228 (1998)
43. Minc, H.: *Permanents*. Cambridge University Press, Cambridge (1984)
44. Nesterov, Y., Nemirovski, A.: *Interior-Point Polynomial Algorithms in Convex Programming*. Society of Industrial and Applied Mathematics, Philadelphia (1994)
45. Nielsen, M., Chuang, I.: *Quantum Computation and Quantum Information*. Cambridge University Press, Cambridge (2011)
46. Potchinkov, A.W., Reemsten, R.M.: The design of FIR filters in the complex plane by convex optimization. *Signal Process.* **46**, 127–146 (1995)
47. Prajna, S., Jadbabaie, A.: Safety Verification of Hybrid Systems Using Barrier Certificates. In: *Hybrid Systems: Computation and Control*, pp. 477–492. Springer (2004)
48. Rockafellar, R.T.: *Convex Analysis*. Princeton University Press, Princeton (1970)
49. Schumacher, B., Westmoreland, M.D.: Sending classical information via noisy quantum channels. *Phys. Rev. A* **56**, 131–138 (1997)
50. Scott, C.H., Jefferson, T.R.: Trace optimization problems and generalized geometric programming. *J. Math. Anal. Appl.* **58**, 373–377 (1977)
51. Shor, P.W.: Capacities of quantum channels and how to find them. *Math. Program. B* **97**, 311–335 (2003)
52. Thompson, C.J.: Inequality with applications in statistical mechanics. *J. Math. Phys.* **6**, 1812–1813 (1965)
53. Valiant, L.: The complexity of computing the permanent. *Theor. Comput. Sci.* **8**, 189–201 (1979)
54. Vandenberghe, L., Boyd, S., Wu, S.: Determinant maximization with linear matrix inequality constraints. *SIAM J. Matrix Anal. Appl.* **19**, 499–533 (1998)
55. Wall, T., Greening, D., Woolsey, R.: Solving complex chemical equilibria using a geometric programming based technique. *Oper. Res.* **34**, 345–355 (1986)
56. Yazarel, H., Pappas, G.: Geometric programming relaxations for linear system reachability. In: *Proceedings of the American Control Conference* (2004)