Generalized maximum entropy estimation

Tobias Sutter   TOBIAS.SUTTER@EPFL.CH
Risk Analytics and Optimization Chair
EPFL, Switzerland

David Sutter   SUTTEDAV@PHYS.ETHZ.CH
Institute for Theoretical Physics
ETH Zurich, Switzerland

Peyman Mohajerin Esfahani   P.MohajerinEsfahani@tudelft.nl
Delft Center for Systems and Control
TU Delft, The Netherlands

John Lygeros   LYGEROS@CONTROL.EE.ETHZ.CH
Department of Electrical Engineering and Information Technology
ETH Zurich, Switzerland

Editor: Benjamin Recht

Abstract

We consider the problem of estimating a probability distribution that maximizes the entropy while satisfying a finite number of moment constraints, possibly corrupted by noise. Based on duality of convex programming, we present a novel approximation scheme using a smoothed fast gradient method that is equipped with explicit bounds on the approximation error. We further demonstrate how the presented scheme can be used for approximating the chemical master equation through the zero-information moment closure method, and for an approximate dynamic programming approach in the context of constrained Markov decision processes with uncountable state and action spaces.

Keywords: Entropy maximization, convex optimization, relative entropy minimization, fast gradient method, approximate dynamic programming

1. Introduction

This article investigates the problem of estimating an unknown probability distribution given a finite number of observed moments that might be corrupted by noise. Given that the observed moments are consistent, i.e., there exists a probability distribution that satisfies all the moment constraints, the problem is underdetermined and has infinitely many solutions. This raises the question of which solution to choose. A natural choice would be to pick the one with the highest entropy, called the *MaxEnt distribution*. The main reason why the MaxEnt distribution is a natural choice is due to a concentration phenomenon described by Jaynes (2003):

> "*If the information incorporated into the maximum-entropy analysis includes all the constraints actually operating in the random experiment, then the distribution predicted by maximum entropy is overwhelmingly the most likely to be observed experimentally.*"

See Jaynes (2003); Grünwald (2008) for a rigorous statement. This maximum entropy estimation problem subject to moment constraints, also known as the *principle of maximum entropy*, is applicable to large classes of problems in natural and social sciences — in particular in economics, see Golan (2008) for a comprehensive survey. Furthermore it has important applications in approximation methods to dynamical objects, such as in systems biology, where MaxEnt distributions are key objects in the so-called *moment closure method* to approximate the chemical master equation Smadbeck and Kaznessis (2013), or more recently in the context of approximating dynamic programming Mohajerin Esfahani et al. (2018) where MaxEnt distributions act as a regularizer, leading to computationally more efficient optimization programs.

Their operational significance motivates the study of numerical methods to compute MaxEnt distributions, which are the solutions of an infinite-dimensional convex optimization problem and as such computationally intractable in general. Since it was shown that the MaxEnt distribution subject to a finite number of moment constraints (if it exists) belongs to the exponential family of distributions Csiszár (1975), its computation can be reduced to solving a system of nonlinear equations, whose dimension is equal to the number of moment constraints Mead and Papanicolaou (1984). Furthermore, the system of nonlinear equations involves evaluating integrals over the support set $\mathbb{K}$ that are computationally difficult in general. Even if $\mathbb{K}$ is finite, finding the MaxEnt distribution is not straightforward, since solving a system of nonlinear equations can be computationally demanding.

In this article, we present a new approximation scheme to minimize the relative entropy subject to noisy moment constraints. This is a generalization of the introduced maximum entropy problem and extends the principle of maximum entropy to the so-called *principle of minimum discriminating information* Kullback (1959). We show that its dual problem exhibits a particular favorable structure that allows us to apply Nesterov's smoothing method Nesterov (2005) and hence tackle the presented problem using a fast gradient method obtaining process convergence properties, unlike Lasserre (2009).

Computing the MaxEnt distribution has applications in randomized rounding and the design of approximation algorithms. More precisely, it has been shown how to improve the approximation ratio for the symmetric and the asymmetric traveling salesman problem via MaxEnt distributions Asadpour et al. (2010); Gharan et al. (2011). Often, it is important to efficiently compute the MaxEnt distribution. For example, the zero-information moment closure method Smadbeck and Kaznessis (2013) (see Section 6), a recent approximate dynamic programming method for constrained Markov decision processes (see Section 7), as well as the approximation of the channel capacity of a large class of memoryless channels Sutter et al. (2015) deal with iterative algorithms that require the numerical computation of the MaxEnt distribution in each iteration step.

Related results. Before comparing the approach presented in this article with existing methods we provide a brief digression on the moment problem. Consider a one-dimensional moment problem formulated as follows: Given a set $\mathbb{K} \subset \mathbb{R}$ and a sequence $(y_i)_{i \in \mathbb{N}} \subset \mathbb{R}$ of moments, does there exist a measure μ supported on $\mathbb{K}$ such that

$$y_i = \int_{\mathbb{K}} x^i \mu(\mathrm{d}x) \quad \text{for all } i \in \mathbb{N} ?$$
(1)

2

For $\mathbb{K} = \mathbb{R}$ and $\mathbb{K} = [a, b]$ with $-\infty < a < b < \infty$ the above moment problem is known as the *Hamburger moment problem* and *Hausdorff moment problem*, respectively. If the moment sequence is finite, the problem is called a *truncated moment problem*. In both full and truncated cases, a measure μ that satisfies (1), is called a *representing measure* of the sequence $(y_i)_{i \in \mathbb{N}}$. If a representing measure is unique, it is said to be *determined by its moments*. From the Stone-Weierstrass theorem it followes directly that every non-truncated representing measure with compact support is determined by its moments. In the Hamburger moment problem, given a representing measure μ for a moment sequence $(y_i)_{i \in \mathbb{N}}$, a sufficient condition for μ being determined by its moments is the so-called *Carleman condition*, i.e., $\sum_{i=1}^{\infty} y_{2i}^{-1/2i} = \infty$. Roughly speaking this says that the moments should not grow too fast, see Akhiezer (1965) for further details. For the Hamburger and the Hausdorff moment problem, there are necessary and sufficient conditions for the existence of a representing measure for a given moment sequence $(y_i)_{i \in \mathbb{N}}$ in both the full as well as the truncated setting, that exploit the rich algebraic connection with Hankel matrices see (Lasserre, 2009, Theorems 3.2, 3.3, 3.4).

In (Lasserre, 2009, Section 12.3) it is shown that the maximum entropy subject to finite moment constraints can be approximated by using duality of convex programming. The problem can be reduced to an unconstrained finite-dimensional convex optimization problem and an approximation hierarchy of its gradient and Hessian in terms of two single semidefinite programs involving two linear matrix inequalities is presented. The desired accuracy is controlled by the size of the linear matrix inequalities. The method seems to be powerful in practice, however a rate of convergence has not been proven. Furthermore, it is not clear how the method extends to the case of uncertain moment constraints. In a finite dimensional setting, Dudik et al. (2007) presents a treatment of the maximum entropy principle with generalized regularization measures, that as a special case contains the setting presented here. However, convergence rates of algorithms presented are not known and again it is not clear how the method extends to the case of uncertain moment constraints. The discrete case, where the support set $\mathbb{K}$ is discrete, has been studied in more detail in the past. It has been shown the the maximum entropy problem in the discrete case has a succinct description that is polynomial-size in the input and can be efficiently computed Singh and Vishnoi (2014); Straszak and Vishnoi (2017). Furthermore it was shown that the maximum entropy problem is equivalent to the counting problem Singh and Vishnoi (2014).

Structure. The layout of this paper is as follows: In Section 2 we formally introduce the problem setting. Our results on an approximation scheme in a continuous setting are reported in Section 3. In Section 4, we show how these results simplify in the finite-dimensional case. Section 5 discusses the gradient approximation that is the dominant step of the proposed approximation method from a computational perspective. The theoretical results are applied in Section 6 to the zero-information moment closure method and in Section 7 to constrained Markov decision processes. We conclude in Section 8 with a summary of our work and comment on possible subjects of further research.

Notation. The logarithm with basis 2 and e is denoted by $\log(\cdot)$ and $\ln(\cdot)$, respectively. We define the standard $n-$simplex as $\Delta_n := \{x \in \mathbb{R}^n : x \geq 0, \sum_{i=1}^{n} x_i = 1\}$. For a probability mass function $p \in \Delta_n$ we denote its entropy by $H(p) := \sum_{i=1}^{n} -p_i \log p_i$. Let

$B(y, r) := \{x \in \mathbb{R}^n : \|x - y\|_2 \le r\}$ denote the ball with radius r centered at y. Throughout this article, measurability always refers to Borel measurability. For a probability density p supported on a measurable set $B \subset \mathbb{R}$ we denote the differential entropy by $h(p) := -\int_B p(x) \log p(x) \mathrm{d}x$. For $\mathbb{A} \subset \mathbb{R}$ and $1 \le p \le \infty$, let $\mathrm{L}^p(\mathbb{A})$ denote the space of L^p-functions on the measure space $(\mathbb{A}, \mathcal{B}(\mathbb{A}), \mathrm{d}x)$, where $\mathcal{B}(\mathbb{A})$ denotes the Borel σ-algebra and $\mathrm{d}x$ the Lebesgue measure. Let $\mathbb{X}$ be a compact metric space, equipped with its Borel σ-field $\mathcal{B}(\cdot)$. The space of all probability measures on $(\mathbb{X}, \mathcal{B}(\mathbb{X}))$ will be denoted by $\mathcal{P}(\mathbb{X})$. The *relative entropy* (or Kullback-Leibler divergence) between any two probability measures $\mu, \nu \in \mathcal{P}(\mathbb{X})$ is defined by

$$\mathsf{D}(\mu\|\nu) := \begin{cases} \int_{\mathbb{X}} \log\left(\frac{\mathrm{d}\mu}{\mathrm{d}\nu}\right) \mathrm{d}\mu, & \text{if } \mu \ll \nu \\ +\infty, & \text{otherwise}, \end{cases}$$

where $\ll$ denotes absolute continuity of measures, and $\frac{\mathrm{d}\mu}{\mathrm{d}\nu}$ is the Radon-Nikodym derivative. The relative entropy is non-negative, and is equal to zero if and only if $\mu \equiv \nu$. Let $\mathbb{X}$ be restricted to a compact metric space and let us consider the pair of vector spaces $(\mathbb{M}(\mathbb{X}), \mathbb{B}(\mathbb{X}))$ where $\mathbb{M}(\mathbb{X})$ denotes the space of finite signed measures on $\mathcal{B}(\mathbb{X})$ and $\mathbb{B}(\mathbb{X})$ is the Banach space of bounded measurable functions on $\mathbb{X}$ with respect to the sup-norm and consider the bilinear form

$$\langle \mu, f \rangle := \int_{\mathbb{X}} f(x) \mu(\mathrm{d}x).$$

This induces the total variation norm as the dual norm on $\mathbb{M}(\mathbb{X})$, since by (Hernández-Lerma and Lasserre, 1999, p.2)

$$\|\mu\|_* = \sup_{\|f\|_\infty \le 1} \langle \mu, f \rangle = \|\mu\|_{\mathsf{TV}},$$

making $\mathbb{M}(\mathbb{X})$ a Banach space. In the light of (Hernández-Lerma and Lasserre, 1999, p. 206) this is a dual pair of Banach spaces; we refer to (Anderson and Nash, 1987, Section 3) for the details of the definition of dual pairs. The Lipschitz norm is defined as $\|u\|_L := \sup_{x, x' \in \mathbb{X}}\{|u(x)|, \frac{|u(x) - u(x')|}{\|x - x'\|_\infty}\}$ and $\mathcal{L}(\mathbb{X})$ denotes the space of Lipschitz functions on $\mathbb{X}$.

2. Problem Statement

Let $\mathbb{K} \subset \mathbb{R}$ be compact and consider the scenario where a probability measure $\mu \in \mathcal{P}(\mathbb{K})$ is unknown and only observed via the following measurement model

$$y_i = \langle \mu, x^i \rangle + u_i, \quad u_i \in \mathcal{U}_i \quad \text{for } i = 1, \ldots, M, \tag{2}$$

where u_i represents the uncertainty of the obtained data point y_i and $\mathcal{U}_i \subset \mathbb{R}$ is compact, convex and $0 \in \mathcal{U}_i$ for all $i = 1, \ldots, M$. Given the data $(y_i)_{i=1}^M \subset \mathbb{R}$, the goal is to estimate a probability measure μ that is consistent with the measurement model (2). This problem (given that M is finite) is underdetermined and has infinitely many solutions. Among all possible solutions for (2), we aim to find the solution that maximizes the entropy. Define the set $T := \times_{i=1}^M \{y_i - u : u \in \mathcal{U}_i\} \subset \mathbb{R}^M$ and the linear operator $\mathcal{A} : \mathbb{M}(\mathbb{K}) \to \mathbb{R}^M$ by

$$(\mathcal{A}\mu)_i := \langle \mu, x^i \rangle = \int_{\mathbb{K}} x^i \mu(\mathrm{d}x) \quad \text{for all} \quad i = 1, \ldots, M.$$

The operator norm is defined as $\|\mathcal{A}\| := \sup_{\|\mu\|_{\mathsf{TV}}=1,\|y\|_2=1} \langle \mathcal{A}\mu, y \rangle$. Note that due to the compactness of $\mathbb{K}$ the operator norm is bounded, see Lemma 4 for a formal statement. The adjoint operator to $\mathcal{A}$ is given by $\mathcal{A}^* : \mathbb{R}^M \to \mathbb{B}(\mathbb{K})$, where $\mathcal{A}^* z(x) := \sum_{i=1}^{M} z_i x^i$; note that the domain and image spaces of the adjoint operator are well defined as $(\mathbb{B}(\mathbb{K}), \mathbb{M}(\mathbb{K}))$ is a topological dual pairs and the operator $\mathcal{A}$ is bounded (Hernández-Lerma and Lasserre, 1999, Proposition 12.2.5).

Given a reference measure $\nu \in \mathcal{P}(\mathbb{K})$, the problem of minimizing the relative entropy subject to moment constraints (2) can be formally described by

$$J^\star = \min_{\mu \in \mathcal{P}(\mathbb{K})} \left\{ \mathsf{D}(\mu\|\nu) \; : \; \mathcal{A}\mu \in T \right\}. \tag{3}$$

We note that the reference measure $\nu \in \mathcal{P}(\mathbb{K})$ is always fixed a priori. A typical choice for ν is the uniform measure over $\mathbb{K}$.

Proposition 1 (Existence & uniqueness of (3)**)** *The optimization problem* (3) *attains an optimal feasible solution that is unique.*

Proof The variational representation of the relative entropy (Boucheron et al., 2013, Corollary 4.15) implies that the mapping $\mu \mapsto \mathsf{D}(\mu\|\nu)$ is lower-semicontinuous Luenberger (1969). Note also that the space of probability measures on $\mathbb{K}$ is compact (Aliprantis and Border, 2007, Theorem 15.11). Moreover, since the linear operator $\mathcal{A}$ is bounded, it is continuous. As a result, the feasible set of problem (3) is compact and hence the optimization problem attains an optimal solution. Finally, the strict convexity of the relative entropy Csiszár (1975) ensures uniqueness of the optimizer. ∎

Note that if $\mathcal{U}_i = \{0\}$ for all $i = 1, \ldots, M$, i.e., there is no uncertainty in the measurement model (2), Proposition 1 reduces to a known result Csiszár (1975). Consider the special case where the reference measure ν is the uniform measure on $\mathbb{K}$ and let p denote the Radon-Nikodym derivative $\frac{\mathrm{d}\mu}{\mathrm{d}\nu}$ (whose existence can be assumed without loss of generality). Since $\mathcal{A}$ is weakly continuous and the differential entropy is known to be weakly lower semi-continuous Boucheron et al. (2013), we can restrict attention to a (weakly) dense subset of the feasible set and hence assume without loss of generality that $p \in \mathrm{L}^1(\mathbb{K})$. Problem (3) then reduces to

$$\max_{p \in \mathrm{L}^1(\mathbb{K})} \left\{ h(p) \; : \; \int_{\mathbb{K}} p(x)\mathrm{d}x = 1, \; \int_{\mathbb{K}} x^i p(x)\mathrm{d}x \in T_i, \; \forall i = 1, \ldots, M \right\}. \tag{4}$$

Problem (4) is a generalized maximum entropy estimation problem that, in case $\mathcal{U}_i = \{0\}$ for all $i = 1, \ldots, M$, simplifies to the standard entropy maximization problem subject to M moment constraints. In this article, we present a new approach to solve (3) that is based on its dual formulation. It turns out that the dual problem of (3) has a particular structure that allows us to apply Nesterov's smoothing method Nesterov (2005) to accelerate convergence. Furthermore, we will show how an ε-optimal solution to (3) can be reconstructed. This is done by solving the dual problem of (3). To achieve additional feasibility guarantees for the ε-optimal solution to (3), we introduce and a second smoothing step that is motivated by Devolder et al. (2012). The problem of entropy maximization subject to uncertain moment constraints (4) can be seen as a special case of (3).

3. Relative entropy minimization

We start by recalling that an unconstrained minimization of the relative entropy with an additional linear term in the cost admits a closed form solution. Let $c \in \mathbb{B}(\mathbb{K})$, $\nu \in \mathcal{P}(\mathbb{K})$ and consider the optimization problem

$$\min_{\mu \in \mathcal{P}(\mathbb{K})} \left\{ \mathsf{D}(\mu \| \nu) - \langle \mu, c \rangle \right\}. \tag{5}$$

Lemma 2 (Gibbs distribution) *The unique optimizer to problem* (5) *is given by the Gibbs distribution, i.e.,*

$$\mu^\star(\mathrm{d}x) = \frac{2^{c(x)} \nu(\mathrm{d}x)}{\int_{\mathbb{K}} 2^{c(x)} \nu(\mathrm{d}x)} \quad \textit{for } x \in \mathbb{K},$$

which leads to the optimal value of $-\log \int_{\mathbb{K}} 2^{c(x)} \nu(\mathrm{d}x)$.

Proof The result is standard and follows from Csiszár (1975) or alternatively by (Sutter et al., 2015, Lemma 3.10). ∎

Let $\mathbb{R}^M \ni z \mapsto \sigma_T(z) := \max_{x \in T} \langle x, z \rangle \in \mathbb{R}$ denote the support function of T, which is continuous since T is compact (Rockafellar, 1997, Corollary 13.2.2). The primal-dual pair of problem (3) can be stated as

$$(\text{primal program}): \quad J^\star = \min_{\mu \in \mathcal{P}(\mathbb{K})} \left\{ \mathsf{D}(\mu \| \nu) + \sup_{z \in \mathbb{R}^M} \left\{ \langle \mathcal{A}\mu, z \rangle - \sigma_T(z) \right\} \right\} \tag{6}$$

$$(\text{dual program}): \quad J_{\mathsf{D}}^\star = \sup_{z \in \mathbb{R}^M} \left\{ -\sigma_T(z) + \min_{\mu \in \mathcal{P}(\mathbb{K})} \left\{ \mathsf{D}(\mu \| \nu) + \langle \mathcal{A}\mu, z \rangle \right\} \right\}, \tag{7}$$

where the dual function is given by

$$F(z) = -\sigma_T(z) + \min_{\mu \in \mathcal{P}(\mathbb{K})} \left\{ \mathsf{D}(\mu \| \nu) + \langle \mathcal{A}\mu, z \rangle \right\}. \tag{8}$$

Note that the primal program (6) is an infinite-dimensional convex optimization problem. The key idea of our analysis is driven by Lemma 2 indicating that the dual function, that involves a minimization running over an infinite-dimensional space, is analytically available. As such, the dual problem becomes an unconstrained finite-dimensional convex optimization problem, which is amenable to first-order methods.

Lemma 3 (Zero duality gap) *There is no duality gap between the primal program* (6) *and its dual* (7), *i.e.,* $J^\star = J_{\mathsf{D}}^\star$. *Moreover, if there exists* $\bar{\mu} \in \mathcal{P}(\mathbb{K})$ *such that* $\mathcal{A}\bar{\mu} \in \mathsf{int}(T)$, *then the set of optimal dual variables in* (7) *is compact.*

Proof Recall that the relative entropy is known to be lower semicontinuous and convex in the first argument, which can be seen as a direct consequence of the duality relation for the relative entropy (Boucheron et al., 2013, Corollary 4.15). Hence, the desired zero duality gap follows by Sion's minimax theorem (Sion, 1958, Theorem 4.2). The compactness of the

set of dual optimizers is due to (Bertsekas, 2009, Proposition 5.3.1). ■

Because the dual function (8) turns out to be non-smooth, in the absence of any additional structure, the efficiency estimate of a black-box first-order method is of order $O(1/\varepsilon^2)$, where ε is the desired absolute additive accuracy of the approximate solution in function value Nesterov (2004). We show, however, that the generalized entropy maximization problem (6) has a certain structure that allows us to deploy the recent developments in Nesterov (2005) for approximating non-smooth problems by smooth ones, leading to an efficiency estimate of order $O(1/\varepsilon)$. This, together with the low complexity of each iteration step in the approximation scheme, offers a numerical method that has an attractive computational complexity. In the spirit of Nesterov (2005); Devolder et al. (2012), we introduce a smoothing parameter $\eta := (\eta_1, \eta_2) \in \mathbb{R}^2_{>0}$ and consider a smooth approximation of the dual function

$$F_\eta(z) := -\max_{x \in T} \left\{ \langle x, z \rangle - \frac{\eta_1}{2} \|x\|_2^2 \right\} + \min_{\mu \in \mathcal{P}(\mathbb{K})} \left\{ \mathsf{D}(\mu\|\nu) + \langle \mathcal{A}\mu, z \rangle \right\} - \frac{\eta_2}{2} \|z\|_2^2 , \qquad (9)$$

with respective optimizers denoted by $x_z^\star$ and $\mu_z^\star$. Consider the projection operator $\pi_T : \mathbb{R}^m \to \mathbb{R}$, $\pi_T(y) = \arg\min_{x \in T} \|x - y\|_2^2$. It is straightforward to see that the optimizer $x_z^\star$ is given by

$$x_z^\star = \arg\min_{x \in T} \|x - \eta_1^{-1} z\|_2^2 = \pi_T \left(\eta_1^{-1} z \right).$$

Hence, the complexity of computing $x_z^\star$ is determined by the projection operator onto T; for simple enough cases (e.g., 2-norm balls, hybercubes) the solution is analytically available, while for more general cases (e.g., simplex, 1-norm balls) it can be computed at relatively low computational effort, see (Richter, 2012, Section 5.4) for a comprehensive survey. The optimizer $\mu_z^\star$ according to Lemma 2 is given by

$$\mu_z^\star(B) = \frac{\int_B 2^{-\mathcal{A}^\star z(x)} \nu(\mathrm{d}x)}{\int_{\mathbb{K}} 2^{-\mathcal{A}^\star z(x)} \nu(\mathrm{d}x)}, \quad \text{for all } B \in \mathcal{B}(\mathbb{K}).$$

Lemma 4 (Lipschitz gradient) *The dual function F_η defined in (9) is η_2-strongly concave and differentiable. Its gradient $\nabla F_\eta(z) = -x_z^\star + \mathcal{A}\mu_z^\star - \eta_2 z$ is Lipschitz continuous with Lipschitz constant $\frac{1}{\eta_1} + \left(\sum_{i=1}^M B^i \right)^2 + \eta_2$ and $B := \max\{|x| \; : \; x \in \mathbb{K}\}$.*

Proof The proof follows along the lines of (Nesterov, 2005, Theorem 1) and in particular by recalling that the relative entropy (in the first argument) is strongly convex with convexity parameter one and Pinsker's inequality, that says that for any $\mu \in \mathcal{P}(\mathbb{K})$ we have

$$\|\mu - \nu\|_{\mathsf{TV}} \le \sqrt{2\mathsf{D}(\mu\|\nu)} . \qquad (10)$$

Moreover, we use the bound

$$
\begin{aligned}
\|\mathcal{A}\| &= \sup_{\lambda \in \mathbb{R}^M,\, \mu \in \mathcal{P}(\mathbb{K})} \left\{ \langle \mathcal{A}\mu, \lambda \rangle \ : \ \|\lambda\|_2 = 1, \ \|\mu\|_{\mathsf{TV}} = 1 \right\} \\
&\leq \sup_{\lambda \in \mathbb{R}^M,\, \mu \in \mathcal{P}(\mathbb{K})} \left\{ \|\mathcal{A}\mu\|_2 \, \|\lambda\|_2 \ : \ \|\lambda\|_2 = 1, \ \|\mu\|_{\mathsf{TV}} = 1 \right\} \qquad (11) \\
&\leq \sup_{\mu \in \mathcal{P}(\mathbb{K})} \left\{ \|\mathcal{A}\mu\|_1 \ : \ \|\mu\|_{\mathsf{TV}} = 1 \right\} \\
&= \sup_{\mu \in \mathcal{P}(\mathbb{K})} \left\{ \sum_{i=1}^{M} \left| \int_{\mathbb{K}} x^i \mu(\mathrm{d}x) \right| \ : \ \|\mu\|_{\mathsf{TV}} = 1 \right\} \\
&\leq \sum_{i=1}^{M} B^i \,,
\end{aligned}
$$

where (11) is due to the Cauchy-Schwarz inequality. ∎

Note that F_η is η_2-strongly concave and according to Lemma 4 its gradient is Lipschitz continuous with constant $L(\eta) := \frac{1}{\eta_1} + \|\mathcal{A}\|^2 + \eta_2$. We finally consider the approximate dual program given by

$$
\text{(smoothed dual program)}: \quad J_\eta^\star = \sup_{z \in \mathbb{R}^M} F_\eta(z) \,. \qquad (12)
$$

It turns out that (12) belongs to a favorable class of smooth and strongly convex optimization problems that can be solved by a fast gradient method given in Algorithm 1 (see Nesterov (2004)) with an efficiency estimate of the order $O(1/\sqrt{\varepsilon})$.

Algorithm 1: Optimal scheme for smooth & strongly convex optimization

---

Choose $w_0 = y_0 \in \mathbb{R}^M$ and $\eta \in \mathbb{R}^2_{>0}$

For $k \geq 0$ do* **Step 1:** Set $y_{k+1} = w_k + \frac{1}{L(\eta)} \nabla F_\eta(w_k)$

  Step 2: Compute $w_{k+1} = y_{k+1} + \frac{\sqrt{L(\eta)} - \sqrt{\eta_2}}{\sqrt{L(\eta)} + \sqrt{\eta_2}} (y_{k+1} - y_k)$

[*The stopping criterion is explained in Remark 7]

Under an additional regularity assumption, solving the smoothed dual problem (12) provides an estimate of the primal and dual variables of the original non-smooth problems (6) and (7), respectively, as summarized in the next theorem (Theorem 5). The main computational difficulty of the presented method lies in the gradient evaluation ∇F_η. We refer to Section 5, for a detailed discussion on this subject.

Assumption 1 (Slater point) *There exits a strictly feasible solution to (3), i.e., $\mu_0 \in \mathcal{P}(\mathbb{K})$ such that $\mathcal{A}\mu_0 \in T$ and $\delta := \min_{y \in T^c} \|\mathcal{A}\mu_0 - y\|_2 > 0$.*

Note that finding a Slater point μ_0 such that Assumption 1 holds, in general can be difficult. In Remark 8 we present a constructive way of finding such an interior point. Given Assumption 1, for $\varepsilon > 0$ define

$$C := \mathsf{D}\big(\mu_0 \| \nu\big), \qquad D := \frac{1}{2} \max_{x \in T} \|x\|_2, \qquad \eta_1(\varepsilon) := \frac{\varepsilon}{4D}, \qquad \eta_2(\varepsilon) := \frac{\varepsilon \delta^2}{2C^2}$$

$$N_1(\varepsilon) := 2 \left(\sqrt{\frac{8DC^2}{\varepsilon^2 \delta^2} + \frac{2\|\mathcal{A}\|^2 C^2}{\varepsilon \delta^2} + 1} \right) \ln \left(\frac{10(\varepsilon + 2C)}{\varepsilon} \right) \tag{13}$$

$$N_2(\varepsilon) := 2 \left(\sqrt{\frac{8DC^2}{\varepsilon^2 \delta^2} + \frac{2\|\mathcal{A}\|^2 C^2}{\varepsilon \delta^2} + 1} \right) \ln \left(\frac{C}{\varepsilon \delta (2 - \sqrt{3})} \sqrt{4 \left(\frac{4D}{\varepsilon} + \|\mathcal{A}\|^2 + \frac{\varepsilon \delta^2}{2C^2} \right) \left(C + \frac{\varepsilon}{2} \right)} \right).$$

Theorem 5 (Convergence rate) *Given Assumption 1 and (13), let $\varepsilon > 0$ and $N(\varepsilon) := \lceil \max\{N_1(\varepsilon), N_2(\varepsilon)\} \rceil$. Then, $N(\varepsilon)$ iterations of Algorithm 1 produce approximate solutions to the problems (7) and (6) given by*

$$\hat{z}_{k,\eta} := y_k \quad and \quad \hat{\mu}_{k,\eta}(B) := \frac{\int_B 2^{-\mathcal{A}^* \hat{z}_{k,\eta}(x)} \nu(\mathrm{d}x)}{\int_{\mathbb{K}} 2^{-\mathcal{A}^* \hat{z}_{k,\eta}(x)} \nu(\mathrm{d}x)}, \quad for\ all\ B \in \mathcal{B}(\mathbb{K}), \tag{14}$$

which satisfy

dual ε-optimality: $\qquad\qquad 0 \le J^\star - F(\hat{z}_{k(\varepsilon)}) \le \varepsilon$ $\qquad\qquad$ (15a)

primal ε-optimality: $\qquad\qquad |\mathsf{D}\big(\hat{\mu}_{k(\varepsilon)} \| \nu\big) - J^\star| \le 2(1 + 2\sqrt{3})\varepsilon$ $\qquad$ (15b)

primal ε-feasibility: $\qquad\qquad d(\mathcal{A}\hat{\mu}_{k(\varepsilon)}, T) \le \dfrac{2\varepsilon\delta}{C}$, $\qquad\qquad$ (15c)

where $d(\cdot, T)$ denotes the distance to the set T, i.e., $d(x,T) := \min_{y \in T} \|x - y\|_2$.

In some applications, Assumption 1 does not hold, as for example in the classical case where $\mathcal{U}_i = \{0\}$ for all $i = 1, \ldots, M$, i.e., there is no uncertainty in the measurement model (2). Moreover, in other cases satisfying Assumption 1 using the construction described in Remark 8 might be computationally expensive. Interestingly, Algorithm 1 can be run irrespective of whether Assumption 1 holds or not, i.e. for any choice of C and δ. While explicit error bounds of Theorem 5 as well as the a-posteriori error bound discussed below do not hold anymore, the asymptotic convergence is not affected.

Proof Using Assumption 1, note that the constant defined as

$$\iota := \frac{\mathsf{D}\big(\mu_0 \| \nu\big) - \min_{\mu \in \mathcal{P}(\mathbb{K})} \mathsf{D}(\mu \| \nu)}{\min_{y \in T^c} \|\mathcal{A}\mu_0 - y\|_2} = \frac{C}{\delta}$$

can be shown to be an upper bound for the optimal dual multiplier (Nedić and Ozdaglar, 2008, Lemma 1), i.e., $\|z^\star\|_2 \le \iota$. The dual function can be bounded from above by C, since weak duality ensures $F(z) \le J^\star \le \mathsf{D}\big(\mu_0 \| \nu\big) = C$ for all $z \in \mathbb{R}^M$. Moreover, if we recall the preparatory Lemmas 3 and 4, we are finally in the setting such that the presented error bounds can be derived from Devolder et al. (2012), see Appendix A for a detailed derivation. ∎

Theorem 5 directly implies that we need at most $O(\frac{1}{\varepsilon}\log\frac{1}{\varepsilon})$ iterations of Algorithm 1 to achieve ε-optimality of primal and dual solutions as well as ε-feasible primal variable. Note that Theorem 5 provides an explicit bound on the so-called *a-priori errors*, together with approximate optimizer of the primal (6) and dual (7) problem. The latter allows us to derive an *a-posteriori error* depending on the approximate optimizers, which is often significantly smaller than the a-priori error.

Corollary 6 (Posterior error estimation) *Given Assumption 1, the approximate primal and dual variables $\hat\mu$ and $\hat z$ given by (14), satisfy the following a-posteriori error bound*

$$F(\hat z) \leq J^\star \leq \mathsf{D}(\hat\mu\|\nu) + \frac{C}{\delta}d(\mathcal{A}\hat\mu, T),$$

where $d(\cdot, T)$ denotes the distance to the set T, i.e., $d(x, T) := \inf_{y\in T}\|x - y\|_2$.

Proof The two key ingredients of the proof are Theorem 5 and the Lipschitz continuity of the so-called perturbation function of convex programming. Let $z^\star$ denote the dual optimizer to (7). We introduce the perturbed program as

$$
\begin{aligned}
J^\star(\varepsilon) &= \min_{\mu\in\mathcal{P}(\mathbb{K})}\left\{\mathsf{D}(\mu\|\nu) \,:\, d(\mathcal{A}\mu, T) \leq \varepsilon\right\} \\
&= \min_{\mu\in\mathcal{P}(\mathbb{K})}\mathsf{D}(\mu\|\nu) + \sup_{\lambda\geq 0}\inf_{y\in T}\lambda\|\mathcal{A}\mu - y\| - \lambda\varepsilon \\
&= \sup_{\lambda\geq 0}-\lambda\varepsilon + \inf_{\substack{\mu\in\mathcal{P}(\mathbb{K}) \\ y\in T}}\sup_{\|z\|_2\leq\lambda}\langle\mathcal{A}\mu - y, z\rangle + \mathsf{D}(\mu\|\nu) \qquad\text{(16a)} \\
&= \sup_{\substack{\lambda\geq 0 \\ \|z\|_2\leq\lambda}}-\lambda\varepsilon + \inf_{\substack{\mu\in\mathcal{P}(\mathbb{K}) \\ y\in T}}\langle\mathcal{A}\mu - y, z\rangle + \mathsf{D}(\mu\|\nu) \qquad\text{(16b)} \\
&\geq -\|z^\star\|_2\,\varepsilon + \inf_{\substack{\mu\in\mathcal{P}(\mathbb{K}) \\ y\in T}}\langle\mathcal{A}\mu - y, z^\star\rangle + \mathsf{D}(\mu\|\nu) \\
&= -\|z^\star\|_2\,\varepsilon + J^\star.
\end{aligned}
$$

Equation (16a) uses the strong duality property that follows by the existence of a Slater point that is due to the definition of the set T, see Section 2. Step (16b) follows by Sion's minimax theorem (Sion, 1958, Theorem 4.2). Hence, we have shown that the perturbation function is Lipschitz continuous with constant $\|z^\star\|_2$. Finally, recalling $\|z^\star\|_2 \leq \frac{C}{\delta}$, established in the proof of Theorem 5 completes the proof. ∎

Remark 7 (Stopping criterion of Algorithm 1) *There are two alternatives for defining a stopping criterion for Algorithm 1. Choose desired accuracy $\varepsilon > 0$.*

(i) a-priori stopping criterion*: Theorem 5 provides the required number of iterations $N(\varepsilon)$ to ensure an ε-close solution.*

(ii) a-posteriori stopping criterion*: Choose the smoothing parameter η as in (13). Fix a (small) number of iterations ℓ that are run using Algorithm 1. Compute the a-posteriori error $\mathsf{D}(\hat\mu\|\nu) + \frac{C}{\delta}d(\mathcal{A}\hat\mu, T) - F(\hat z)$ according to Corollary 6 and if it is smaller than If ε terminate the algorithm. Otherwise continue with another ℓ iterations.*

Remark 8 (Slater point computation) *To compute the respective constants in Assumption 1, we need to construct a strictly feasible point for (3). For this purpose, we consider a polynomial density of degree r defined as $p_r(\alpha, x) := \sum_{i=0}^{r-1} \alpha_i x^i$. For notational simplicity we assume that the support set is the unit interval ($\mathbb{K} = [0,1]$), such that the moments induced by the polynomial density are given by*

$$\langle p_r(\alpha, x), x^i \rangle = \int_0^1 \sum_{j=0}^{r-1} \alpha_j x^{j+i} \mathrm{d}x = \sum_{j=0}^{r-1} \frac{\alpha_j}{j+i+1},$$

for $i = 0, \ldots, M$. Consider $\beta \in \mathbb{R}^{M+1}$, where $\beta_1 = 1$ and $\beta_i = y_{i-1}$ for $i = 2, \ldots, M+1$. Hence, the feasibility requirement of (3) can be expressed as the linear constraint $A\alpha = \beta$, where $A \in \mathbb{R}^{(M+1) \times r}$, $\alpha \in \mathbb{R}^r$, $\beta \in \mathbb{R}^{M+1}$ and $A_{i,j} = \frac{1}{i+j-1}$ and finding a strictly feasible solution reduces to the following feasibility problem

$$\begin{cases} \max\limits_{\alpha \in \mathbb{R}^r} & const \\ \mathrm{s.\,t.} & A\alpha = \beta \\ & p_r(\alpha, x) \geq 0 \quad \forall x \in [0,1], \end{cases} \tag{17}$$

where p_r is a polynomial function in x of degree r with coefficients α. The second constraint of the program (17) (i.e., $p_r(\alpha, x) \geq 0 \; \forall x \in [0,1]$)[1] can be equivalently reformulated as linear matrix inequalities of dimension $\lceil \frac{r}{2} \rceil$, using a standard result in polynomial optimization, see (Lasserre, 2009, Chapter 2) for details. We note that for small degree r, the set of feasible solutions to problem (17) may be empty, however, by choosing r large enough and assuming that the moments can be induced by a continuous density, problem (17) becomes feasible. Moreover, if $0 \in \mathrm{int}(T)$ the Slater point leads to a $\delta > 0$ in Assumption 1.

Example 1 (Density estimation) *We are given the first 3 moments of an unknown probability measure defined on $\mathbb{K} = [0,1]$ as[2]*

$$y := \left(\frac{1 - \ln 2}{\ln 2}, \frac{\ln 4 - 1}{\ln 4}, \frac{5 - \ln 64}{\ln 64} \right) \approx (0.44, \; 0.28, \; 0.20).$$

The uncertainty set in the measurement model (2) is assumed to be $\mathcal{U}_i = [-u, u]$ for all $i = 1, \ldots, 3$. A Slater point is constructed using the method described in Remark 8, where $r = 5$ is enough for the problem (17) to be feasible, leading to the constant $C = 0.0288$. The Slater point is depicted in Figure 1 and its differential entropy can be numerically computed as -0.0288.

We consider two simulations for two different uncertainty sets (namely, $u = 0.01$ and $u = 0.005$). The underlying maximum entropy problem (4) is solved using Algorithm 1. The respective features of the a-priori guarantees by Theorem 5 as well as the a-posteriori guarantees (upper and lower bounds) by Corollary 6 are reported in Table 1. Recall that $\hat{\mu}_{k(\varepsilon)}$ denotes the approximate primal variable after k-iterations of Algorithm 1 as defined in Theorem 5 and that $d(\mathcal{A}\hat{\mu}_{k(\varepsilon)}, T)$ (resp. $\frac{2\varepsilon\delta}{C}$) represent the a-posteriori (resp. a-priori)

---

1. In a multi-dimensional setting one has to consider a tightening (i.e., $p_r(\alpha, x) > 0 \; \forall x \in [0,1]^n$).
2. The considered moments are actually induced by the probability density $f(x) := (\ln 2 \, (1+x))^{-1}$. We, however, do not use this information at any point of this example.

feasibility guarantees. It can be seen in Table 1 that increasing the uncertainty set $\mathcal{U}$ leads to a higher entropy, where the uniform density clearly has the highest entropy. This is also intuitively expected since enlarging the uncertainty set is equivalent to relaxing the moment constraints in the respective maximum entropy problem. The corresponding densities are graphically visualized in Figure 1.

Table 1: Some specific simulation points of Example 1.

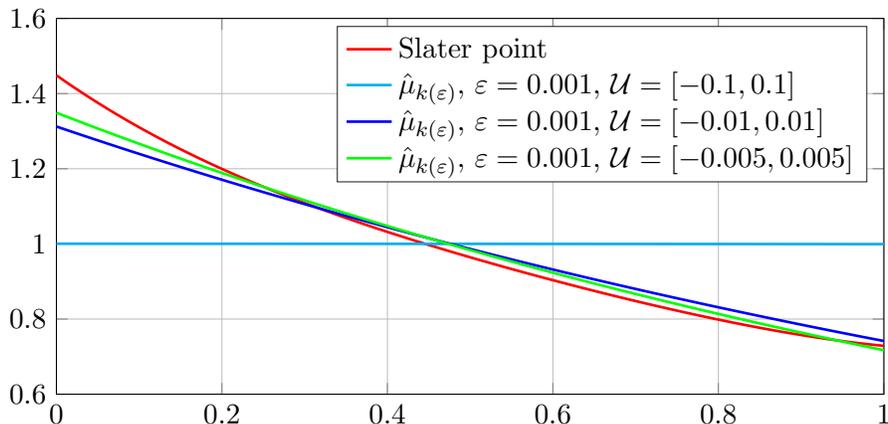|	$\mathcal{U} = [-0.01, 0.01]$				$\mathcal{U} = [-0.005, 0.005]$			
a-priori error ε	1	0.1	0.01	0.001	1	0.1	0.01	0.001
J_{UB}	-0.0174	-0.0189	-0.0194	-0.0194	-0.0223	-0.0236	-0.0237	-0.0238
J_{LB}	-0.0220	-0.0279	-0.0204	-0.0195	-0.0263	-0.0298	-0.0244	-0.0238
iterations $k(\varepsilon)$	99	551	5606	74423	232	1241	12170	157865
$d(\mathcal{A}\hat{\mu}_{k(\varepsilon)}, T)$	0.0008	0.0036	0.0005	0	0	0.001	0.0001	0
$\frac{2\varepsilon\delta}{C}$	0.69	0.069	0.0069	0.00069	0.35	0.035	0.0035	0.00035
runtime $[\mathrm{s}]$[3]	1.4	1.4	2.3	12.9	1.4	1.5	3.3	26.1

Figure 1: Maximum entropy densities obtained by Algorithm 1 for two different uncertainty sets. As a reference, the Slater point density, that was computed as described in Remark 8 is depicted in red.

4. Finite-dimensional case

We consider the finite-dimensional case where $\mathbb{K} = \{1, \ldots, N\}$ and hence we optimize in (3) over the probability simplex $\mathcal{P}(\mathbb{K}) = \Delta_N$. One substantial simplification, when restricting to the finite-dimensional setting, is that the Shannon entropy is non-negative and bounded from above (by $\log N$). Therefore, we can substantially weaken Assumption 1 to the following assumption.

---

3. Runtime includes Slater point computation. Simulations were run with Matlab on a laptop with a 2.2 GHz Intel Core i7 processor.

Assumption 2 (Regularity)

(i) *There exists $\delta > 0$ such that $B(0, \delta) \subset \{x - \mathcal{A}\mu \; : \; \mu \in \Delta_N, x \in T\}$.*

(ii) *The reference measure $\nu \in \Delta_N$ has full support, i.e., $\min\limits_{1 \le i \le N} \nu_i > 0$.*

Consider the the definitions given in (13) with $C := \max\limits_{1 \le i \le N} \log \frac{1}{\nu_i}$, then the following finite-dimensional equivalent to Theorem 5 holds.

Corollary 9 (A-priori error bound) *Given Assumption 2, $C := \max\limits_{1 \le i \le N} \log \frac{1}{\nu_i}$ and the definitions (13), let $\varepsilon > 0$ and $N(\varepsilon) := \lceil \{N_1(\varepsilon), N_2(\varepsilon)\} \rceil$. Then, $N(\varepsilon)$ iterations of Algorithm 1 produce the approximate solutions to the problems (7) and (6), given by*

$$\hat{z}_{k(\varepsilon)} := y_{k(\varepsilon)} \quad and \quad \hat{\mu}_{k(\varepsilon)}(B) := \frac{\sum_{i \in B} 2^{-\left(\mathcal{A}^* \hat{z}_{k(\varepsilon)}\right)_i} \nu_i}{\sum_{i=1}^{N} 2^{-\left(\mathcal{A}^* \hat{z}_{k(\varepsilon)}\right)_i} \nu_i} \quad for \ all \ B \subset \{1, 2, \dots, N\}, \quad (18)$$

which satisfy

$$dual \ \varepsilon\text{-}optimality: \qquad\qquad 0 \le F(\hat{z}_{k(\varepsilon)}) - J^\star \le \varepsilon \qquad\qquad (19a)$$

$$primal \ \varepsilon\text{-}optimality: \qquad\qquad |\mathsf{D}(\hat{\mu}_{k(\varepsilon)} \| \nu) - J^\star| \le 2(1 + 2\sqrt{3})\varepsilon \qquad (19b)$$

$$primal \ \varepsilon\text{-}feasibility: \qquad\qquad d(\mathcal{A}\hat{\mu}_{k(\varepsilon)}, T) \le \frac{2\varepsilon\delta}{C}, \qquad\qquad (19c)$$

where $d(\cdot, T)$ denotes the distance to the set T, i.e., $d(x, T) := \min_{y \in T} \|x - y\|_2$.

Proof Under Assumption 2 the dual optimal solutions in (7) are bounded by

$$\|z^\star\| \le \frac{1}{r} \max_{1 \le i \le N} \log \frac{1}{\nu_i}. \qquad\qquad (20)$$

This bound on the dual optimizer follows along the lines of (Devolder et al., 2012, Theorem 6.1). The presented error bounds can then be derived along the lines of Theorem 5. ∎

In addition to the explicit error bound provided by Corollary 9, the a-posteriori upper and lower bounds presented in Corollary 6 directly apply to the finite-dimensional setting as well.

5. Gradient Approximation

The computationally demanding element for Algorithm 1 is the evaluation of the gradient $\nabla F_\eta(\cdot)$ given in Lemma 4. In particular, Theorem 5 and Corollary 9 assume that this gradient is known exactly. While this is not restrictive if, for example, $\mathbb{K}$ is a finite set, in general, $\nabla F_\eta(\cdot)$ involves an integration that can only be computed approximately. In particular if we consider a multi-dimensional setting (i.e., $\mathbb{K} \subset \mathbb{R}^d$), the evaluation of the gradient $\nabla F_\eta(\cdot)$ represents a multi-dimensional integration problem. This gives rise to the question of how the fast gradient method (and also Theorem 5) behaves in a case of inexact first-order information. Roughly speaking, the fast gradient method Algorithm 1,

while being more efficient than the classical gradient method (if applicable), is less robust when dealing with inexact gradients Devolder et al. (2014). Therefore, depending on the computational complexity of the gradient, one may consider the possibility of replacing Algorithm 1 with a classical gradient method. A detailed mathematical analysis of this tradeoff is a topic of further research, and we refer the interested readers to Devolder et al. (2014) for further details in this regard.

In this section we discuss two numerical methods to approximate this gradient. Note that in Lemma 4, given that T is simple enough the optimizer $x_z^\star$ is analytically available, so what remains is to compute $\mathcal{A}\mu_z^\star$, that according to Lemma 2 is given by

$$(\mathcal{A}\mu_z^\star)_i = \frac{\int_\mathbb{K} x^i 2^{-\mathcal{A}^* z(x)} \nu(\mathrm{d}x)}{\int_\mathbb{K} 2^{-\mathcal{A}^* z(x)} \nu(\mathrm{d}x)} \quad \text{for all } i = 1, \dots, M\,. \tag{21}$$

Semidefinite programming. Due to the specific structure of the considered problem, (21) represents an integration of exponentials of polynomials for which an efficient approximation in terms of two single semidefinite programs (SDPs) involving two linear matrix inequalities has been derived, where the desired accuracy is controlled by the size of the linear matrix inequalities constraints, see Bertsimas et al. (2008); Lasserre (2009) for a comprehensive study and for the construction of those SDPs. While the mentioned hierarchy of SDPs provides a certificate of optimality (hat is easy to evaluate and asymptotic convergence (in the size of the SDPs), a convergence rate that explicitly quantifies the size of the SDPs required for a desired accuracy is unknown. In practice, the hierarchy often converges in few iteration steps, which however, depends on the problem and is not known a priori.

Quasi-Monte Carlo. The most popular methods for integration problems of the from (21) are Monte Caro (MC) schemes, see Robert and Casella (2004) for a comprehensive summary. The main advantage of MC methods is that the root-mean-square error of the approximation converges to 0 with a rate of $O(N^{-1/2})$ that is independent of the dimension, where N are the number of samples used. In practise, this convergence often is too slow. Under mild assumptions on the integrand, the MC methods can be significantly improved with a more recent technique known as Quasi-Monte Carlo (QMC) methods. QMC methods can reach a convergence rate arbitrarily close to $O(N^{-1})$ with a constant not depending on the dimension of the problem. We would like to refer the reader to Dick et al. (2013); Sloan and Woźniakowski (1998); Kuo and Sloan (2005); Niederreiter (2010); Thély (2017) for a detailed discussion about the theory of QMC methods.

Remark 10 (Computational stability) *The evaluation of the gradient in Lemma 4 involves the term $\mathcal{A}\mu_z^\star$, where $\mu_z^\star$ is the optimizer of the second term in (9). By invoking Lemma 2 and the definition of the operator $\mathcal{A}$, the gradient evaluation reduces to*

$$(\mathcal{A}\mu_z^\star)_i = \frac{\int_\mathbb{K} x^i 2^{-\sum_{j=1}^M z_j x^j} \mathrm{d}x}{\int_\mathbb{K} 2^{-\sum_{j=1}^M z_j x^j} \mathrm{d}x} \quad \text{for } i = 1, \dots, M\,. \tag{22}$$

Note that a straightforward computation of the gradient via (22) is numerically difficult. To alleviate this difficulty, we follow the suggestion of (Nesterov, 2005, p. 148) which we briefly

elaborate here. Consider the functions $f(z, x) := -\sum_{j=1}^{M} z_j x^j$, $\bar{f}(z) := \max_{x \in \mathbb{K}} f(z, x)$ and $g(z, x) := f(z, x) - \bar{f}(z)$. Note, that $g(z, x) \geq 0$ for all $(z, x) \in \mathbb{R}^M \times \mathbb{R}$. One can show that

$$(\mathcal{A}\mu_z^\star)_i = \frac{\int_{\mathbb{K}} 2^{g(z,x)} \frac{\partial}{\partial z_i} g(z, x) \mathrm{d}x}{\int_{\mathbb{K}} 2^{g(z,x)} \mathrm{d}x} + \frac{\partial}{\partial z_i} \bar{f}(z) \quad \text{for } i = 1, \ldots, M \,,$$

which can be computed with significantly smaller numerical error than (22) *as the numerical exponent are always negative, leading to values always being smaller than* 1.

6. Zero-information moment closure method

In chemistry, physics, systems biology and related fields, stochastic chemical reactions are described by the *chemical master equation* (CME), that is a special case of the Chapman-Kolmogorov equation as applied to Markov processes van Kampen (1981); Wilkinson (2006). These equations are usually infinite-dimensional and analytical solutions are generally impossible. Hence, effort has been directed toward developing of a variety of numerical schemes for efficient approximation of the CME, such as stochastic simulation techniques (SSA) Gillespie (1976). In practical cases, one is often interested in the first few moments of the number of molecules involved in the chemical reactions. This motivated the development of approximation methods to those low-order moments without having to solve the underlying infinite-dimensional CME. One such approximation method is the so-called *moment closure method* Gillespie (2009), that briefly described works as follows: First the CME is recast in terms of moments as a linear ODE of the form

$$\frac{\mathrm{d}}{\mathrm{d}t}\mu(t) = A\mu(t) + B\zeta(t) \,, \tag{23}$$

where $\mu(t)$ denotes the moments up to order M at time t and $\zeta(t)$ is an infinite vector describing the contains moments of order $M + 1$ or higher. In general ζ can be an infinite vector, but for most of the standard chemical reactions considered in, e.g., systems biology it turns out that only a finite number of higher order moments affect the evolution of the first M moments. Indeed, if the chemical system involves only the so-called zeroth and first order reactions the vector ζ has dimension zero (reduces to a constant affine term), whereas if the system also involves second order reactions then ζ also contains some moments of order $M + 1$ only. It is widely speculated that reactions up to second order are sufficient to realistically model most systems of interest in chemistry and biology Gillespie (2007); Gillespie et al. (2013). The matrix A and the linear operator B (that may potentially be infinite-dimensional) can be found analytically from the CME. The ODE (23), however, is intractable due to its higher order moments dependence. The approximation step is introduced by a so-called closure function

$$\zeta = \varphi(\mu) \,,$$

where the higher-order moments are approximated as a function of the lower-oder moments, see Singh and Hespanha (2006, 2007). A closure function that has recently attracted interest is known as the *zero-information* closure function (of order M) Smadbeck and Kaznessis (2013), and is given by

$$(\varphi(\mu))_i = \langle p_\mu^\star, x^{M+i} \rangle, \text{ for } i = 1, 2, \ldots \,, \tag{24}$$

where $p_\mu^\star$ denotes the maximizer to the problem (4), where $T = \times_{i=1}^M \left\{ \langle \mu, x^i \rangle - u \; : \; u \in \mathcal{U}_i \right\}$, where $\mathcal{U}_i = [-\kappa, \kappa] \subset \mathbb{R}$ for all i and for a given $\kappa > 0$, that acts as a regularizer, in the sense of Assumption 1. This approximation reduces the infinite-dimensional ODE (23) to a finite-dimensional ODE

$$\frac{\mathrm{d}}{\mathrm{d}t}\mu(t) = A\mu(t) + B\varphi(\mu(t)). \tag{25}$$

To numerically solve (25) it is crucial to have an efficient evaluation of the closure function φ. In the zero-information closure scheme this is given by to the entropy maximization problem (4) and as such can be addressed using Algorithm 1.

To illustrate this point, we consider a reversible dimerisation reaction where two monomers ($\mathcal{M}$) combine in a second-order monomolecular reaction to form a dimer ($\mathcal{D}$); the reverse reaction is first order and involves the decomposition of the dimer into the two monomers. This gives rise to the chemical reaction system

$$
\begin{aligned}
2\mathcal{M} &\xrightarrow{k_1} \mathcal{D} \\
\mathcal{D} &\xrightarrow{k_2} 2\mathcal{M} ,
\end{aligned}
\tag{26}
$$

with reaction rate constants $k_1, k_2 > 0$. Note that the system as described has a single degree of freedom since $\mathcal{M} = 2\mathcal{D}_0 - 2\mathcal{D} + \mathcal{M}_0$, Where $\mathcal{M}$ denotes the count of the monomers, $\mathcal{D}$ the count of dimers, and $\mathcal{M}_0$ and $\mathcal{D}_0$ the corresponding initial conditions. Therefore, the matrices can be reduced to include only the moments of one component as a simplification and as such the zero-information closure function (24) consists of solving a one-dimensional entropy maximization problem such as given by (4), where the support are the natural numbers (upper bounded by $\mathcal{M}_0 + \mathcal{D}_0$ and hence compact). For illustrative purposes, let us look at a second order closure scheme, where the corresponding moment vectors are defined as $\mu = (1, \langle \mathcal{M} \rangle, \langle \mathcal{M}^2 \rangle)^\top \in \mathbb{R}^3$ and $\zeta = \langle \mathcal{M}^3 \rangle \in \mathbb{R}$ and the corresponding matrices are given by

$$
A = \begin{pmatrix} 0 & 0 & 0 \\ k_2 S_0 & 2k_1 - k_2 & -2k_1 \\ 2k_2 S_0 & 2k_2(S_0 - 1) - 4k_1 & 8k_1 - 2k_2 \end{pmatrix}, \quad B = \begin{pmatrix} 0 \\ 0 \\ -4k_1 \end{pmatrix},
$$

where $S_0 = \mathcal{M}_0 + 2\mathcal{D}_0$. The simulation results, Figure 2, show the time trajectory for the average and the second moment of the number of $\mathcal{M}$ molecules in the reversible dimerization model (26), as calculated for the zero information closure (25) using Algorithm 1, for a second-order closure as well as a third-order closure. To solve the ODE (25) we use an explicit Runge-Kutta (4,5) formula (ode45) built into MATLAB. The results are compared to the average of 10^6 SSA Wilkinson (2006) trajectories. It can be seen how increasing the order of the closure method improves the approximation accuracy.

7. Approximate dynamic programming for constrained Markov decision processes

In this section, we show that problem (3) naturally appears in the context of approximate dynamic programming, which is at the heart of reinforcement learning (see Bertsekas and Tsitsiklis (1996); Bertsekas (1995); Recht (2018) and references therein). We consider
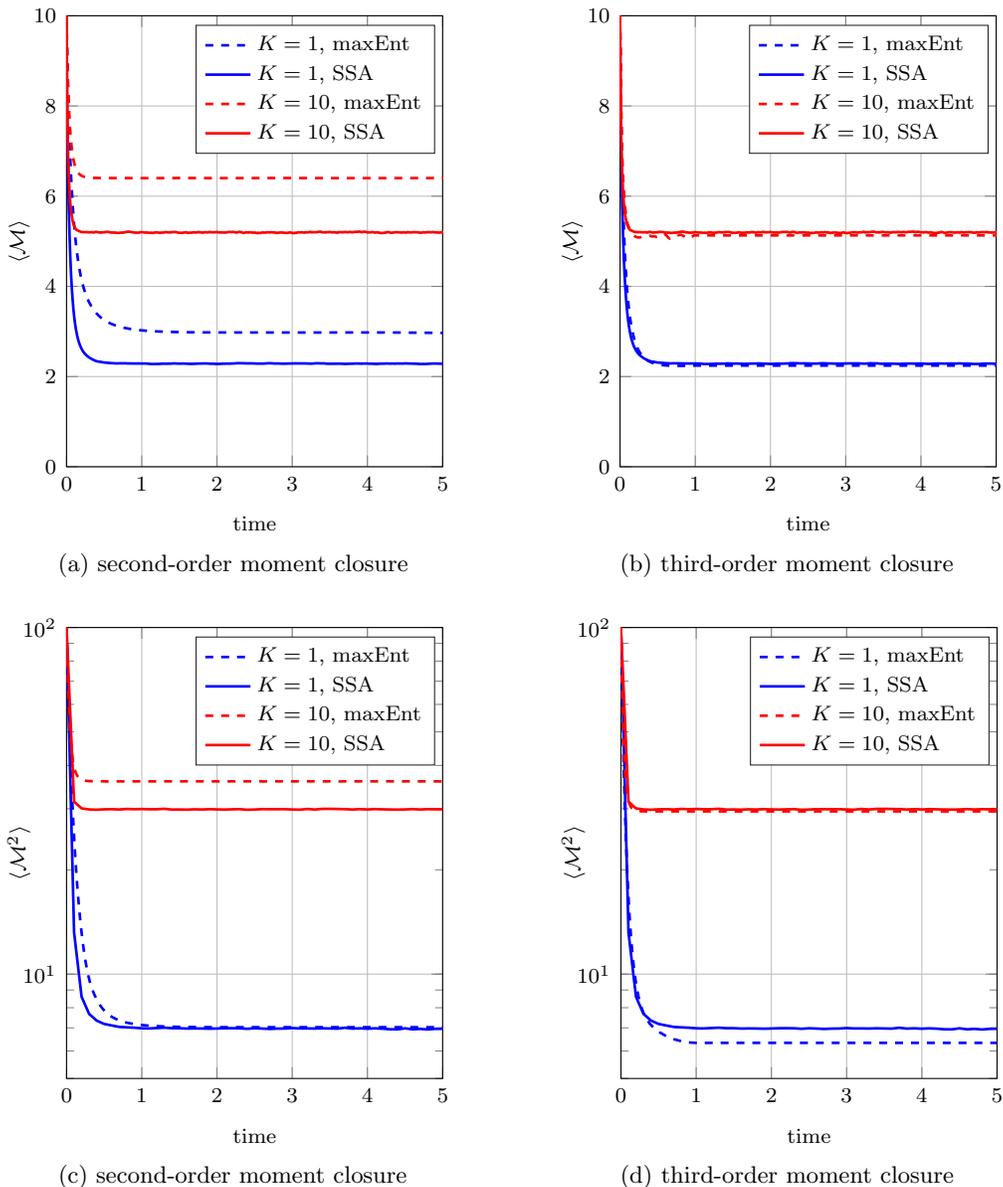
Figure 2: Reversible dimerization system (26) with reaction constants $K = k_2/k_1$: Comparison of the zero-information moment closure method (25), solved using Algorithm 1 and the average of 10^6 SSA trajectories. The initial conditions are $\mathcal{M}_0 = 10$ and $\mathcal{D}_0 = 0$ and the regularization term $\kappa = 0.01$.

constrained Markov decision processes (MDPs) that form an important class of stochastic control problems with applications in many areas; see Piunovskiy (1997); Altman (1999) and the comprehensive bibliography therein. A look at these references shows that most of the literature concerns constrained MDPs where the state and action spaces are either finite or countable. Inspired by the recent work Mohajerin Esfahani et al. (2018), we show here

that the entropy maximization problem (3) is a key element to approximate constrained MDPs on general, possibly uncountable, state and action spaces.

7.1. Constrained MDP problem formulation

Consider a discrete-time constrained MDP $\big(S, A, \{A(s) : s \in S\}, Q, c, d, \kappa\big)$, where S (resp. A) is a metric space called *state space* (resp. *action space*) and for each $s \in S$ the measurable set $A(s) \subseteq A$ denotes the set of *feasible actions* when the system is in state $s \in S$. The *transition law* is a stochastic kernel Q on S given the feasible state-action pairs in $\mathbb{K} := \{(s, a) : s \in S, a \in A(s)\}$. A stochastic kernel acts on real-valued measurable functions u from the left as $Qu(s, a) := \int_S u(s')Q(\mathrm{d}s'|s, a)$, for all $(s, a) \in \mathbb{K}$, and on probability measures μ on $\mathbb{K}$ from the right as $\mu Q(B) := \int_K Q(B|s, a)\mu\big(\mathrm{d}(s, a)\big)$, for all $B \in \mathcal{B}(S)$. The (measurable) function $c : \mathbb{K} \to \mathbb{R}_+$ denotes the so-called *one-stage cost function*, and the (measurable) function $d : \mathbb{K} \to \mathbb{R}^q$ the one-stage constraint cost along with the preassigned budget level $\kappa \in \mathbb{R}^q$.

In short, the MDP model described above may read as follows: When the system at the state $s \in S$ deploys the action $a \in A(s)$, it incurs the one-stage cost and constraint $c(s, a)$ and $d(s, a)$, respectively, and subsequently lands in the next state whose distribution is supported on S and described via $Q(\cdot|s, a)$. We consider the expected long-run average cost criteria[4], i.e., for any measurable function $\psi : \mathbb{K} \to \mathbb{R}^p$ with $p \in \{1, q\}$, we define

$$J_\psi(\pi, \nu) := \limsup_{n \to \infty} \frac{1}{n} \mathbb{E}_\nu^\pi \left(\sum_{t=0}^{n-1} \psi(s_t, a_t) \right).$$

Using the defintion above, the central object of this section is the optimization problem

$$J^\star := \begin{cases} \inf_{\nu, \pi} & J_c(\pi, \nu) \\ \text{s.t.} & J_d(\pi, \nu) \leq \kappa \\ & \nu \in \mathcal{P}(S), \ \pi \in \Pi, \end{cases} \tag{27}$$

where Π denotes the set of all control policies. We refer the reader to Hernández-Lerma and Lasserre (1996); Hernández-Lerma et al. (2003); Altman (1999) for a detailed mathematical treatment of this setting. It is well known that the MDP problem (27) can be stated equivalently as an infinite-dimensional linear program and its corresponding dual

$$J_P^\star := \begin{cases} \inf_{\mu} & \langle \mu, c \rangle \\ \text{s.t.} & \mu(B \times A) = \mu Q(B) \quad \forall B \in \mathbb{B}(S) \\ & \langle \mu, d \rangle \leq \kappa \\ & \mu \in \mathcal{P}(\mathbb{K}), \end{cases} \tag{28a}$$

$$J_D^\star := \begin{cases} \sup_{u, \rho, \gamma} & \rho - \gamma^\top \kappa \\ \text{s.t.} & \rho + u(x) - Qu(x, a) \leq c(x, a) + \gamma^\top d(x, a) \quad \forall(x, a) \in \mathbb{K} \\ & \langle \mu, d \rangle \leq \kappa \\ & u \in, \rho \in \mathbb{R}, \gamma \in \mathbb{R}_+^q. \end{cases} \tag{28b}$$

---

4. We refer the interested reader to Mohajerin Esfahani et al. (2018) for extension to the discounted cost.

The following regularity assumption is required in order to ensure that the solutions are well posed and that equivalence between (27) and the LPs (28a), (28b) holds.

Assumption 3 (Control model) *We stipulate that*

(i) *the set of feasible state-action pairs is the unit hypercube $\mathbb{K} = [0,1]^{\dim(S \times A)}$;*

(ii) *the transition law Q is Lipschitz continuous, i.e., there exists $L_Q > 0$ such that for all $k, k' \in \mathbb{K}$ and all continuous functions u*

$$|Qu(k) - Qu(k')| \leq L_Q \|u\|_\infty \|k - k'\|_{\ell_\infty};$$

(iii) *the cost function c is non-negative and Lipschitz continuous on $\mathbb{K}$ and d is continuous on $\mathbb{K}$.*

Under this assumption, strong duality between the linear programs (28a) and (28b) holds (i.e., the supremum and infimum are attained and $J_P^\star = J_D^\star$). Moreover, the LP formulation is equivalent to the original problem (27) in the sense that $J^\star = J_P^\star = J_D^\star$, see (Hernández-Lerma et al., 2003, Theorem 5.2).

Finding exact solutions to either (28a) or (28b) generally is impossible as the linear programs are infinite dimensional. This challenge has given rise to a wealth of approximation schemes in the literature under the names of approximate dynamic programming. Typically, one restricts decision space in (28b) to a finite dimensional subspace spanned by basis functions $\{u_i\}_{i=1}^n \subset \mathcal{L}(S)$ denoted by $\mathbb{U}_n := \{\sum_{i=1}^n \alpha_i u_i \ : \ \|\alpha\|_2 \leq \theta\}$. Motivated by de Farias and Roy (2004); Mohajerin Esfahani et al. (2018) we then approximate the solution $J^\star$ by

$$J_{P,n}^\star := \begin{cases} \inf_\mu & \langle \mu, c \rangle + \theta \|\mathcal{T}_n \mu - e\|_2 \\ \text{s.t.} & \langle \mu, d \rangle \leq \kappa \\ & \mu \in \mathcal{P}(\mathbb{K}), \end{cases} \tag{29}$$

where the operator $\mathcal{T}_n : \mathcal{P}(\mathbb{K}) \to \mathbb{R}^{n+1}$ is defined as $(\mathcal{T}_n \mu)_1 = -1, (\mathcal{T}_n \mu)_{i+1} = \langle Qu_i - u_i, \mu \rangle$, $i = 1, \ldots, n$ and $e := (-1, 0, \ldots, 0) \in \mathbb{R}^{n+1}$. The optimization problem (29) can be solved with an accelerated first-order method provided by Algorithm 2, stated below, where each iteration step involves solving a problem an entropy maximization problem of the form (3). We define

$$c_{\alpha,\zeta} := \frac{1}{\zeta}\left(c - \alpha_0 + \sum_{i=1}^n \alpha_i (Qu_i - u_i)\right), \qquad \mathcal{Y} := \{\mu \in \mathcal{P}(\mathbb{K}) \ : \ \langle \mu, d \rangle \leq \kappa\},$$

$$\mathbb{T}(q, \alpha) := (\alpha - q)\min\{1, \theta\|q - \alpha\|_2^{-1}\}, \quad y_\zeta^\star(\alpha) := \arg\min_{y \in \mathcal{Y}} \{\mathsf{D}(y \| \lambda) + \langle y, c_{\alpha,\zeta} \rangle\}. \tag{30}$$

Steps 2 and 3 of Algorithm 2 are simple arithmetic operations. Step 1 relies on the solution to the optimization problem (30) that can be reduced to (3). To see this, note that the additional linear term in the objective function can be directly integrated into the analysis of Section 3. We provide the explicit construction in Section 7.2 for an inventory management system. We are interested in establishing bounds on the quality of the approximate solution $J_{n,\zeta}^{(k)}$ obtained by Algorithm 2. The results of (Mohajerin Esfahani et al., 2018, Theorem 3.3 and 5.3) allow us to obtain the following bound.

19

Algorithm 2:   Approximate dynamic programming scheme

Input: $n, k \in \mathbb{N}$, $\zeta, \theta > 0$, and $w^{(0)} \in \mathbb{R}^{n+1}$ such that $\|w^{(0)}\|_2 \leq \theta$

For $0 \leq \ell \leq k$ **do**   **Step 1:**   Define $r^{(\ell)} := \frac{\zeta}{4n}(e - \mathcal{T}_n \, \mathrm{y}^\star_\zeta(w^{(\ell)}))$

                                     Step 2:   Let $z^{(\ell)} := \mathbb{T}\left(\sum_{j=0}^{\ell} \frac{j+1}{2} r^{(j)}, 0\right)$ and $\beta^{(\ell)} = \mathbb{T}\left(r^{(\ell)}, w^{(\ell)}\right)$

                                       Step 3:   Set $w^{(\ell+1)} = \frac{2}{\ell+3} z^{(\ell)} + \frac{\ell+1}{\ell+3} \beta^{(\ell)}$

Output: $J_{n,\zeta}^{(k)} := \langle c, \hat{y}_\zeta \rangle + \theta \|\mathcal{T}_n \hat{y}_\zeta - e\|_2$ with $\hat{y}_\zeta := \sum_{j=0}^{k} \frac{2(j+1)}{(k+1)(k+2)} \, \mathrm{y}^\star_\zeta(w^{(j)})$

Theorem 11 (Approximation error) *Under Assumption 3, Algorithm 2 provides an approximation to* (27) *with the following error bound*

$$|J_{n,\zeta}^{(k)} - J^\star| \leq (1 + \max\{L_Q, 1\})\|u^\star - \Pi_{\mathbb{U}_n}(u^\star)\|_L + \left(\frac{4n^2\theta}{k^2\zeta} + \dim(\mathbb{K})\zeta \max\{\log\left(\beta/\zeta\right), 1\}\right),$$

where $\beta := \frac{e}{\dim(\mathbb{K})}(\theta\sqrt{n}(\max\{L_Q, 1\} + 1) + \|c\|_L)$.

Note that the bound depends on the projection residual $\|u^\star - \Pi_{\mathbb{U}_n}(u^\star)\|_L$, where the projection mapping is defined as $\Pi_{\mathbb{U}_n}(u^\star) := \arg\min_{u \in \mathbb{U}_n} \|u^\star - u\|_2$, the parameters of the problem (notably the dimensions of the state and action spaces, the stage cost, and the Lipschitz constant of the kernel L_Q), and the design choices of the algorithm (the number of basis functions n, norm bound θ and ζ).

Remark 12 (Tuning parameters)

(i) *The residual error* $\|u^\star - \Pi_{\mathbb{U}_n}(u^\star)\|_L$ *can be approximated by leveraging results from the literature on universal function approximation. Prior information about the value function* $u^\star$ *may offer explicit quantitative bounds. For instance, for MDP under Assumption 3 we know that* $u^\star$ *is Lipschitz continuous. For appropriate choice of basis functions, we can therefore ensure a convergence rate of* $n^{-1/\dim(S)}$, *see for instance Farouki (2012) for polynomials and Olver (2009) for the Fourier basis functions.*

(ii) *The regularization parameter* θ *has to be chosen such that* $\theta > \|c\|_L$. *An optimal choice for* θ *and* ζ *is described in (Mohajerin Esfahani et al., 2018, Remark 4.6 and Theorem 5.3).*

7.2. Inventory management system

Consider an inventory model in which the state variable describes the stock level s_t at the beginning of period t. The control or action variable a_t at t is the quantity ordered and immediately supplied at the beginning of period t, and the "disturbance" or "exogenous" variable ξ_t is the demand during that period. We assume ξ_t to be i.i.d. random variables following an exponential distribution with parameter λ. The system equation Hernández-Lerma and Lasserre (1996) is

$$s_{t+1} = \max\{0, s_t + a_t - \xi_t\} =: (s_t + a_t - \xi_t)_+, \quad t = 0, 1, 2, \ldots. \tag{31}$$

We assume that the system has a finite capacity C. Therefore $S = A = [0, C]$ and since the current stock plus the amount ordered cannot exceed the system's capacity, the set of feasible actions is $A(s) = [0, C - s]$ for every $s \in S$. Suppose we wish to maximize an expected profit for operating the system, we might take the net profit at stage t to be

$$r(s_t, a_t, \xi_t) := v \min\{s_t + a_t, \xi_t\} - p a_t - h(s_t + a_t), \tag{32}$$

which is of the form "profit = sales - production cost - holding cost". In (32), v, p and h are positive numbers denoting unit sale price, unit production cost, and unit holding cost, respectively. To write the cost (32) in the form of our control model (27), we define

$$
\begin{aligned}
c(s, a) :=\ & \mathbb{E}[-r(s_t, a_t, \xi_t) | s_t = s, a_t = a] \\
=\ & -v(s + a)e^{-\lambda(s+a)} - \frac{v}{\lambda}\left(1 - e^{-\lambda(s+a)}\left(\lambda(s + a) + 1\right)\right) + pa + h(s + a).
\end{aligned}
\tag{33}
$$

Note that non-negativity of the cost can be ensured by subtracting the term $2vC$ from (32). We assume that there are regulatory constraints on the required stock level s_t, for example to avoid the risk of running into a shortage of a certain critical product. For simplicity, let us assume that the regulator enforces constraints on the long-term first and second moments of the stock s_t in the following sense

$$\limsup_{n\to\infty} \frac{1}{n} \mathbb{E}_\nu^\pi \left(\sum_{t=0}^{n-1} s_t\right) \geq \ell_1, \quad \text{and} \quad \limsup_{n\to\infty} \frac{1}{n} \mathbb{E}_\nu^\pi \left(\sum_{t=0}^{n-1} s_t^2\right) \leq \ell_2, \tag{34}$$

for given $\ell_1, \ell_2 \in \mathbb{R}_+$, where we assume that $\ell_1^2 < \ell_2$[5]. To express it in the form of our control model (27), we define $d_1(s, a) := -s$, $d_2(s, a) := s^2$, and $\kappa := (-\ell_1, \ell_2) \in \mathbb{R}^2$. From the described assumptions on the constants ℓ_1 and ℓ_2, it can be directly seen that the Slater point assumption described in Lemma 3 holds (consider the set $T := \{x \in \mathbb{R}^2 \ : \ x_1 \geq \ell_1, x_2 \leq \ell_2, x_1^2 \leq x_2\}$).

In the following, we will describe how Algorithm 2 and the respective error bounds given by Theorem 11 can be applied to the inventory management system. In particular we show how the computationally demanding part of Algorithm 2, given by (30), is directly addressed by the methodology presented in Section 3. To fulfill Assumption 3, we equivalently reformulate the above problem using the dynamics

$$s_{t+1} = (\min\{C, s_t + a_t\} - \xi_t)_+, \quad t = 0, 1, 2, \ldots, \tag{35}$$

where the admissible actions set is now the state-independent set $A = [0, C]$. Finally by normalizing the state and action variables through the definitions $\tilde{s}_t := \frac{s_t}{C}$ and $\tilde{a}_t := \frac{a_t}{C}$, we have

$$\tilde{s}_{t+1} = \left(\min\{1, \tilde{s}_t + \tilde{a}_t\} - \frac{\xi_t}{C}\right)_+, \quad t = 0, 1, 2, \ldots. \tag{36}$$

Furthermore, it can be seen directly using Leibnitz' rule that the transition law is Lipschitz continuous and $L_Q \leq \sqrt{2}C\lambda$. It remains to argue how (30) can be addressed by the

---

5. Note that the enforced constraints (34) imply the upper and lower bounds $\ell_1 \leq \limsup_{n\to\infty} \frac{1}{n}\mathbb{E}_\nu^\pi \left(\sum_{t=0}^{n-1} s_t\right) \leq \sqrt{\ell_2}$ and $\ell_1^2 \leq \limsup_{n\to\infty} \frac{1}{n}\mathbb{E}_\nu^\pi \left(\sum_{t=0}^{n-1} s_t^2\right) \leq \ell_2$.

methodology presented in Section 3. We introduce the linear operator $\mathcal{A} : \mathcal{P}(\mathbb{K}) \to \mathbb{R}^2$, defined by $(\mathcal{A}\mu)_i := \langle \mu, d_i \rangle$ for $i = 1, 2$. Then, the optimization problem (30) can be expressed as

$$\text{(primal program)}: \quad J^\star = \min_{\mu \in \mathcal{P}(\mathbb{K})} \left\{ \mathsf{D}(\mu \| \nu) - \langle \mu, c_{\alpha,\zeta} \rangle + \sup_{z \in \mathbb{R}^M} \left\{ \langle \mathcal{A}\mu, z \rangle - \sigma_T(z) \right\} \right\}, \quad (37)$$

$$\text{(dual program)}: \quad J_{\mathsf{D}}^\star = \sup_{z \in \mathbb{R}^M} \left\{ -\sigma_T(z) + \min_{\mu \in \mathcal{P}(\mathbb{K})} \left\{ \mathsf{D}(\mu \| \nu) + \langle \mu, \mathcal{A}^* z - c_{\alpha,\zeta} \rangle \right\} \right\}, \quad (38)$$

which apart from the additional linear term are in the form of (6) and (7). Due to our assumptions on ℓ_1 and ℓ_2 described above, there exists a strictly feasible solution to (30), i.e., $\mu_0 \in \mathcal{P}(\mathbb{K})$ such that $\mathcal{A}\mu_0 \in T$ and $\delta := \min_{y \in T^c} \|\mathcal{A}\mu_0 - y\|_2 > 0$. Hence, the method from Section 3 is applicable.

Numerical Simulation. For a given set of model parameters ($C = 1$, $\lambda = \frac{1}{2}$, $v = 1$, $p = \frac{1}{2}$, $h = \frac{1}{10}$) we consider four different scenarios:

(i) Unconstrained case (i.e., $\ell_1 = 0$ and $\ell_2 = 1$);

(ii) $\ell_1 = 0.5$ and $\ell_2 = 0.4$;

(iii) $\ell_1 = 0.5$ and $\ell_2 = 0.3$, which is strictly more constrained than scenario (ii);

(iv) $\ell_1 = 0.1$ and $\ell_2 = 0.1$.

For each scenario we run Algorithm 2 and plot the value of the resulting approximation $J_{n,\zeta}^{(k)}$ as a function of the number of iterations in Figure 3. In each iteration of Algorithm 2 the variable (30) is required, which is computed with Algorithm 1. The following simulation parameters were used:

Algorithm 1: $\eta_1 = \eta_2 = 10^{-3}$, 1500 iterations;

Algorithm 2: $\zeta = 10^{-1.5}$, $\theta = 3$, $n = 10$, Fourier basis $u_{2i-1}(s) = \frac{C}{2i\pi} \cos\left(\frac{2i\pi s}{C}\right)$ and $u_{2i}(s) = \frac{C}{2i\pi} \sin\left(\frac{2i\pi s}{C}\right)$ for $i = 1, \ldots, \frac{n}{2}$.

As shown in Figure 3, the value of $J_{n,\zeta}^{(k)}$ converges at around 1000 iterations of Algorithm 2.[6] The fact that scenario (ii) is a (strict) relaxation in terms of the constraints compared to scenario (iii) is visualized by the numerical simulation as the expected profit of scenario (ii) is continuously higher compared to scenario (iii). Figure 3 also indicates that in Scenario (iv) the constraints are the most restrictive.

8. Conclusion and future work

We presented an approximation scheme to a generalization of the classical problem of estimating a density via a maximum entropy criterion, given some moment constraints. The key idea used is to apply smoothing techniques to the non-smooth dual function of the entropy maximization problem, that enables us to solve the dual problem efficiently with fast

---

6. 1000 iterations of Algorithm 2 took around 5.5 hours with Matlab on a laptop with a 2.2 GHz Intel Core i7 processor.
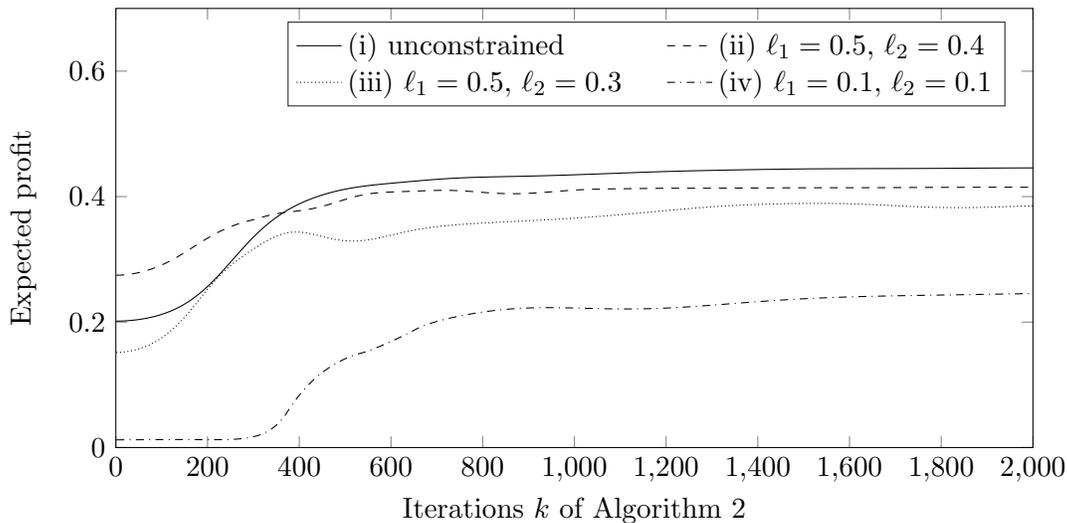
Figure 3: The expected profit of the inventory system, approximated by $-J_{n,\zeta}^{(k)}$, resulting from Algorithm 2 is displayed for four different scenarios (i)-(iv) representing four different constraints on the inventory system.

gradient methods. Due to the favorable structure of the considered entropy maximization problem, we provide explicit error bounds on the approximation error as well as a-posteriori error estimates.

The proposed method requires one to evaluate the gradient (21) in every iteration step, which, as highlighted in Section 5, in the infinite-dimensional setting involves an integral. As such the method used to compute those integrals has to be included to the complexity of the proposed algorithm and, in higher dimensions, may become is the dominant factor. Therefore, it would be interesting to investigate this integration step in more detail. Two approaches, one based on semidefinite programming and another invoking Quasi-Monte Carlo integration techniques, are briefly sketched. What remains open is to quantify the accuracy required in the gradient approximations, which could be done along the lines of Devolder et al. (2014). Another potential direction, would be to test the proposed numerical method in the context of approximating the channel capacity of a large class of memoryless channels Sutter et al. (2015), as mentioned in the introduction.

Finally it should be mentioned that the approximation scheme proposed in this article can be further generalized to quantum mechanical entropies. In this setup probability mass functions are replaced by density matrices (i.e., positive semidefinite matrices, whose trace is equal to one). The von Neumann entropy of such a density matrix ρ is defined by $H(\rho) := -\mathrm{tr}(\rho \log \rho)$, which reduces to the (Shannon) entropy in case the density matrix ρ is diagonal. Also the relative entropy can be generalized to the quantum setup Umegaki (1962) and general treatment of our approximation scheme, its analysis can be lifted to the this (strictly) more general framework. As demonstrated in Sutter et al. (2016), (quantum) entropy maximization problems can be used to efficiently approximate the classical capacity of quantum channels.

23

Appendix A. Detailed proof of Theorem 5

Given the bounds $\|z^\star\|_2 \leq \frac{C}{\delta}$ and $F(z) \leq J^\star \leq \mathsf{D}(\mu_0 \| \nu) = C$ for all $z \in \mathbb{R}^M$ as argued above, we show how the error bounds of Theorem 5 follow from Devolder et al. (2012). The dual ε-optimality (15a) is derived from (Devolder et al., 2012, Equation (7.8)), that in our setting states

$$0 \leq J^\star - F(\hat{z}_k) \leq \frac{3\varepsilon}{4} + 5 \left(F(z^\star) - F(0) + \frac{\varepsilon}{2} \right) e^{-\frac{k}{2}\sqrt{\frac{\eta_2}{L(\eta)}}}. \tag{39}$$

Using the parameters as defined in Theorem 5, i.e.,

$$C := \mathsf{D}(\mu_0 \| \nu), \qquad D := \frac{1}{2} \max_{x \in T} \|x\|_2, \qquad \eta_1(\varepsilon) := \frac{\varepsilon}{4D}, \qquad \eta_2(\varepsilon) := \frac{\varepsilon \delta^2}{2C^2}, \tag{40}$$

the fact that $F(z^\star) - F(0) = F(z^\star) \leq C$, and the Lipschitz constant $L(\eta) = \frac{1}{\eta_1} + \|\mathcal{A}\|^2 + \eta_2$ derived in Lemma 4, inequality (39) ensures $0 \leq J^\star - F(\hat{z}_k) \leq \varepsilon$ if

$$k \geq N_1(\varepsilon) := 2 \left(\sqrt{\frac{8DC^2}{\varepsilon^2 \delta^2} + \frac{2\|\mathcal{A}\|^2 C^2}{\varepsilon \delta^2} + 1} \right) \ln \left(\frac{10(\varepsilon + 2C)}{\varepsilon} \right).$$

The primal ε-optimality bound (15b) following the derivations in Devolder et al. (2012) is implied if k is chosen large enough such that $\|\nabla F_\eta(z_k)\|_2 \leq \frac{2\varepsilon\delta}{C}$. According to (Devolder et al., 2012, Equation (7.11)) the gradient can be bounded by

$$\|\nabla F_\eta(z_k)\|_2 \leq \sqrt{4L(\eta) \left(F(z^\star) - F(0) + \frac{\varepsilon}{2} \right)} e^{-\frac{k}{2}\sqrt{\frac{\eta_2}{L(\eta)}}} + 2\sqrt{3}\eta_2 \frac{C}{\delta}. \tag{41}$$

Again using the parameters and constants as we did for the dual ε-optimality bound we find that $\|\nabla F_\eta(z_k)\|_2 \leq \frac{2\varepsilon\delta}{C}$ if

$$k \geq N_2(\varepsilon) := 2 \left(\sqrt{\frac{8DC^2}{\varepsilon^2 \delta^2} + \frac{2\|\mathcal{A}\|^2 C^2}{\varepsilon \delta^2} + 1} \right) \ln \left(\frac{C}{\varepsilon \delta(2 - \sqrt{3})} \sqrt{4 \left(\frac{4D}{\varepsilon} + \|\mathcal{A}\|^2 + \frac{\varepsilon \delta^2}{2C^2} \right) \left(C + \frac{\varepsilon}{2} \right)} \right).$$

The primal ε-feasiblility (15c) finally directly follows from (Devolder et al., 2012, Equation (7.14)) and the bound $\|z^\star\|_2 \leq \frac{C}{\delta}$.

Acknowledgments

References

N. I. Akhiezer. *The classical moment problem and some related questions in analysis.* Translated by N. Kemmer. Hafner Publishing Co., New York, 1965.

Charalambos D. Aliprantis and Kim C. Border. *Infinite Dimensional Analysis: A Hitchhiker's Guide.* Springer, 2007. ISBN 9783540326960. doi: 10.1007/3-540-29587-9.

Eitan Altman. *Constrained Markov decision processes.* Stochastic Modeling. Chapman & Hall/CRC, Boca Raton, FL, 1999. ISBN 0-8493-0382-6.

Edward J. Anderson and Peter Nash. *Linear programming in infinite-dimensional spaces: theory and applications.* Wiley-Interscience Series in Discrete Mathematics and Optimization. Wiley, 1987. ISBN 9780471912507.

Arash Asadpour, Michel X. Goemans, Aleksander M?dry, Shayan Oveis Gharan, and Amin Saberi. *An $O(\log n/\log\log n)$-approximation Algorithm for the Asymmetric Traveling Salesman Problem*, pages 379–389. 2010. doi: 10.1137/1.9781611973075.32. URL `https://epubs.siam.org/doi/abs/10.1137/1.9781611973075.32`.

D. Bertsekas. *Dynamic Programming and Optimal Control.* Athena Scientific, 1995.

D. Bertsekas and J. Tsitsiklis. *Neuro-Dynamic Programming.* Athena Scientific, 1996.

Dimitri P. Bertsekas. *Convex Optimization Theory.* Athena Scientific optimization and computation series. Athena Scientific, 2009. ISBN 9781886529311.

Dimitris Bertsimas, Xuan Vinh Doan, and Jean Lasserre. Approximating integrals of multivariate exponentials: a moment approach. *Oper. Res. Lett.*, 36(2):205–210, 2008. ISSN 0167-6377. doi: 10.1016/j.orl.2007.07.002. URL `http://dx.doi.org/10.1016/j.orl.2007.07.002`.

Stéphane Boucheron, Gábor Lugosi, and Pascal Massart. *Concentration inequalities: A nonasymptotic theory of independence.* Oxford University Press, 2013. ISBN 978-0-19-953525-5. doi: 10.1093/acprof:oso/9780199535255.001.0001. URL `http://dx.doi.org/10.1093/acprof:oso/9780199535255.001.0001`.

I. Csiszár. *I*-divergence geometry of probability distributions and minimization problems. *Annals of Probability*, 3:146–158, 1975. doi: 10.1214/aop/1176996454.

D.P. de Farias and B. Van Roy. On Constraint Sampling in the Linear Programming Approach to Approximate Dynamic Programming. *Mathematics of Operations Research*, 29(3):462–478, 2004.

Olivier Devolder, Francois Glineur, and Yurii Nesterov. Double smoothing technique for large-scale linearly constrained convex optimization. *SIAM Journal on Optimization*, 22 (2):702–727, 2012. doi: 10.1137/110826102.

Olivier Devolder, François Glineur, and Yurii Nesterov. First-order methods of smooth convex optimization with inexact oracle. *Math. Program.*, 146(1-2, Ser. A):37–75, 2014.

ISSN 0025-5610. doi: 10.1007/s10107-013-0677-5. URL https://doi.org/10.1007/s10107-013-0677-5.

Josef Dick, Frances Y. Kuo, and Ian H. Sloan. High-dimensional integration: The Quasi-Monte Carlo way. *Acta Numerica*, 22:133–288, 5 2013. ISSN 1474-0508. doi: 10.1017/S0962492913000044. URL http://journals.cambridge.org/article_S0962492913000044.

Miroslav Dudik, Steven J. Phillips, and Robert E. Schapire. Maximum entropy density estimation with generalized regularization and an application to species distribution modeling. *Journal of Machine Learning Research*, 8:1217–1260, 2007. ISSN 1532-4435.

Rida T. Farouki. The Bernstein polynomial basis: A centennial retrospective. *Computer Aided Geometric Design*, 29(6):379 – 419, 2012.

Shayan Oveis Gharan, Amin Saberi, and Mohit Singh. A randomized rounding approach to the traveling salesman problem. In *Proceedings of the 2011 IEEE 52Nd Annual Symposium on Foundations of Computer Science*, FOCS 11, pages 550–559, Washington, DC, USA, 2011. IEEE Computer Society. ISBN 978-0-7695-4571-4. doi: 10.1109/FOCS.2011.80. URL http://dx.doi.org/10.1109/FOCS.2011.80.

C. S. Gillespie. Moment-closure approximations for mass-action models. *IET Systems Biology*, 3(1):52–58, January 2009. ISSN 1751-8849. doi: 10.1049/iet-syb:20070031.

Daniel T. Gillespie. A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. *J. Computational Phys.*, 22(4):403–434, 1976. ISSN 0021-9991. doi: 10.1016/0021-9991(76)90041-3.

Daniel T. Gillespie. Stochastic simulation of chemical kinetics. *Annual Review of Physical Chemistry*, 58(1):35–55, 2007. doi: 10.1146/annurev.physchem.58.032806.104637. URL https://doi.org/10.1146/annurev.physchem.58.032806.104637. PMID: 17037977.

Daniel T Gillespie, Andreas Hellander, and Linda R Petzold. Perspective: Stochastic algorithms for chemical kinetics. *The Journal of Chemical Physics*, 138(17):170901, 2013. doi: 10.1063/1.4801941.

Amos Golan. Information and entropy econometrics a review and synthesis. *Foundations and Trends in Econometrics*, 2:1–145, 2008. ISSN 1551-3076. doi: 10.1561/0800000004. URL http://dx.doi.org/10.1561/0800000004.

Peter Grünwald. Entropy concentration and the empirical coding game. *Statistica Neerlandica*, 62(3):374–392, 2008. ISSN 1467-9574. doi: 10.1111/j.1467-9574.2008.00401.x. URL http://dx.doi.org/10.1111/j.1467-9574.2008.00401.x.

O. Hernández-Lerma and J.B. Lasserre. *Discrete-Time Markov Control Processes: Basic Optimality Criteria*. Applications of Mathematics Series. Springer, 1996. ISBN 9780387945798.

O. Hernández-Lerma and J.B. Lasserre. *Further topics on discrete-time Markov control processes*. Applications of Mathematics Series. Springer, 1999. ISBN 9780387986944. doi: 10.1007/978-1-4612-0561-6.

Onésimo Hernández-Lerma, Juan González-Hernández, and Raquiel R. López-Martínez. Constrained average cost Markov control processes in Borel spaces. *SIAM J. Control Optim.*, 42(2):442–468, 2003. ISSN 0363-0129. doi: 10.1137/S0363012999361627. URL https://doi.org/10.1137/S0363012999361627.

E. T. Jaynes. *Probability theory: The logic of science*. Cambridge University Press, Cambridge, 2003. ISBN 0-521-59271-2. doi: 10.1017/CBO9780511790423. URL http://dx.doi.org/10.1017/CBO9780511790423.

Solomon Kullback. *Information theory and statistics*. John Wiley and Sons, Inc., New York; Chapman and Hall, Ltd., London, 1959.

Frances Y Kuo and Ian H Sloan. Lifting the curse of dimensionality. *Notices of the AMS*, 52(11):1320–1328, 2005.

Jean B. Lasserre. *Moments, Positive Polynomials and Their Applications*. Imperial College Press optimization series. Imperial College Press, 2009. ISBN 9781848164468. doi: 10. 1142/9781848164468_0001.

David G. Luenberger. *Optimization by Vector Space Methods*. John Wiley & Sons, Inc., New York-London-Sydney, 1969.

Lawrence R. Mead and N. Papanicolaou. Maximum entropy in the problem of moments. *J. Math. Phys.*, 25(8):2404–2417, 1984. ISSN 0022-2488. doi: 10.1063/1.526446. URL http://dx.doi.org/10.1063/1.526446.

Peyman Mohajerin Esfahani, Tobias Sutter, Daniel Kuhn, and John Lygeros. From infinite to finite programs: explicit error bounds with applications to approximate dynamic programming. *SIAM J. Optim.*, 28(3):1968–1998, 2018. ISSN 1052-6234. doi: 10.1137/17M1133087. URL https://doi.org/10.1137/17M1133087.

Angelia Nedić and Asuman Ozdaglar. Approximate primal solutions and rate analysis for dual subgradient methods. *SIAM Journal on Optimization*, 19(4):1757–1780, 2008.

Yurii Nesterov. *Introductory Lectures on Convex Optimization: A Basic Course*. Applied Optimization. Springer, 2004. ISBN 9781402075537. doi: 10.1007/978-1-4419-8853-9.

Yurii Nesterov. Smooth minimization of non-smooth functions. *Mathematical Programming*, 103(1):127–152, 2005. ISSN 0025-5610. doi: 10.1007/s10107-004-0552-5. URL http://dx.doi.org/10.1007/s10107-004-0552-5.

Harald Niederreiter. Quasi-Monte Carlo methods. *Encyclopedia of quantitative finance*, 2010.

Sheehan Olver. On the convergence rate of a modified Fourier series. *Mathematics of Computation*, 78(267):1629–1645, 2009.

A. B. Piunovskiy. *Optimal control of random sequences in problems with constraints*, volume 410 of *Mathematics and its Applications*. Kluwer Academic Publishers, Dordrecht, 1997. ISBN 0-7923-4571-1. doi: 10.1007/978-94-011-5508-3. URL `https://doi.org/10.1007/978-94-011-5508-3`. With a preface by V. B. Kolmanovskii and A. N. Shiryaev.

Benjamin Recht. A Tour of Reinforcement Learning: The View from Continuous Control. *arXiv e-prints*, art. arXiv:1806.09460, June 2018.

Stefan Richter. Computational complexity certification of gradient methods for real-time model predictive control. *PhD thesis, ETH Zurich*, 2012. doi: 10.3929/ethz-a-007587480.

C. P. Robert and G. Casella. *Monte Carlo Statistical Methods*. Springer, 2 edition, 2004.

R. Tyrrell Rockafellar. *Convex analysis*. Princeton Landmarks in Mathematics and Physics Series. Princeton University Press, 1997. ISBN 9780691015866.

A. Singh and J. P. Hespanha. Lognormal moment closures for biochemical reactions. In *Proceedings of the 45th IEEE Conference on Decision and Control*, pages 2063–2068, Dec 2006. doi: 10.1109/CDC.2006.376994.

Abhyudai Singh and João Pedro Hespanha. A derivative matching approach to moment closure for the stochastic logistic model. *Bull. Math. Biol.*, 69(6):1909–1925, 2007. ISSN 0092-8240. doi: 10.1007/s11538-007-9198-9. URL `http://dx.doi.org/10.1007/s11538-007-9198-9`.

Mohit Singh and Nisheeth K. Vishnoi. Entropy, optimization and counting. In *Proceedings of the Forty-sixth Annual ACM Symposium on Theory of Computing*, STOC '14, pages 50–59, New York, NY, USA, 2014. ACM. ISBN 978-1-4503-2710-7. doi: 10.1145/2591796.2591803. URL `http://doi.acm.org/10.1145/2591796.2591803`.

Maurice Sion. On general minimax theorems. *Pacific Journal of Mathematics*, 8:171–176, 1958. ISSN 0030-8730.

Ian H Sloan and Henryk Woźniakowski. When are Quasi-Monte Carlo algorithms efficient for high dimensional integrals? *Journal of Complexity*, 14(1):1–33, 1998.

Patrick Smadbeck and Yiannis N. Kaznessis. A closure scheme for chemical master equations. *Proceedings of the National Academy of Sciences*, 110(35):14261–14265, 2013. doi: 10.1073/pnas.1306481110. URL `http://www.pnas.org/content/110/35/14261.abstract`.

Damian Straszak and Nisheeth K. Vishnoi. Computing maximum entropy distributions everywhere, 2017. available at `arXiv:1711.02036`.

D. Sutter, T. Sutter, P. Mohajerin Esfahani, and R. Renner. Efficient approximation of quantum channel capacities. *IEEE Transactions on Information Theory*, 62(1):578–598, Jan 2016. ISSN 0018-9448. doi: 10.1109/TIT.2015.2503755.

T. Sutter, D. Sutter, P. Mohajerin Esfahani, and J. Lygeros. Efficient approximation of channel capacities. *IEEE Transactions on Information Theory*, 61(4):1649–1666, April 2015. ISSN 0018-9448. doi: 10.1109/TIT.2015.2401002.

Maxime Thély. Stochastic programming with Quasi-Monte Carlo methods. Master's thesis, ETH Zurich, Switzerland, 2017.

H. Umegaki. Conditional expectation in an operator algebra. *Kodai Mathematical Seminar Reports*, 14:59–85, 1962. doi: 10.2996/kmj/1138844604.

N. G. van Kampen. *Stochastic processes in physics and chemistry*, volume 888 of *Lecture Notes in Mathematics*. North-Holland Publishing Co., Amsterdam-New York, 1981. ISBN 0-444-86200-5.

Darren James Wilkinson. *Stochastic modelling for systems biology*. Chapman & Hall/CRC Mathematical and Computational Biology Series. Chapman & Hall/CRC, Boca Raton, FL, 2006. ISBN 978-1-58488-540-5; 1-58488-540-8.