Optimization-based Estimation of Expected Values with Application to Stochastic Programming

Raphael Chinchilla and João P. Hespanha

***Abstract*— This paper constructs bounds on the expected value of a scalar function of a random vector. The bounds are obtained using an optimization method, which can be computed efficiently using state-of-the-art solvers, and do not require integration or sampling the random vector. This optimization based approach is especially useful in stochastic programming, where the criteria to be minimized takes the form of an expected value. In particular, we minimize the bounds to solve problems of discrete time finite horizon open-loop control with stochastic perturbations and also uncertainty in the system's parameters. We illustrate this application with two numerical examples.**

I. Introduction

This paper addresses the development of efficient methods to estimate the expected value of a function of a random variable and optimize it with respect to a set of (deterministic) parameters that affect the function. The expected value for continuous random variables is given by an integral, and its exact computation is computationally intensive in high dimensions. Hence, an array of techniques has been developed to tackle this problem.

Numerical integration methods such as the trapezoid method or the adaptive Gauss-Kronrod quadrature method [1] provide a numerical solution, but do not scale well with the dimension. Another approach is to use first order approximations, where the expected value of the cost function is substituted by the cost function of the expected value. It is a simple method and when the cost function is convex, it provides a lower bound on the expected value via Jensen's inequality. However, this method provides no formal guarantees for more general optimization criteria.

Another set of methods is based on sampling. Markov Chain Monte Carlo [2], [3] and Monte Carlo integration can be used to estimate the expected value by computing the empirical mean of samples. In the context of stochastic optimization, stochastic search algorithms, like Recursive Least Square and Stochastic Gradient Descent [4], provide algorithms analogous to their deterministic counterpart. However, such algorithms have slow rates of convergence. Methods such as Scenario Approach [5] or Stochastic Average Approximation [6] substitutes minimizing the expected value by minimizing, respectively, the maximum or the average of the cost function in different "scenarios", obtained through sampling the random variables. They require optimizations with large numbers of samples and variables, which can make them slower. For all these methods, the sampling process can be a computational bottleneck. While sampling from unconditioned distributions is generally straightforward, sampling from conditional distributions (*i.e.*, where there is knowledge that a certain informative event has occured) is more complicated. Few conditional distributions have efficient samplers [7]–[10]. In the lack of these, one can use generic samplers such as Metropolis Hastings algorithm and Importance sampling, which can be less effective or more computationally demanding. We refer to [3] for a broader exposition of Monte Carlo methods.

As an alternative, this paper constructs bounds on an expected value obtained without explicit integration or sampling. The starting point is a result that provides a lower and an upper bound on the expected value of a scalar function $V(\cdot)$ of a random vector Ψ in terms of a minimization and maximization, respectively. The optimization penalizes a criterion that consists of a combination of the function $V(\Psi)$, the logarithm of the probability density function (*p.d.f.*) of Ψ and its differential entropy. This result permits estimating the value of the expected value through an optimization (and thus optimality conditions that involve differentiation) rather than through an integration. The results in Section II actually provide a family of bounds parameterized by a scalar parameter ϵ that can be optimized to improve tightness of the bound.

Our optimization-based approach to compute upper/lower bounds to expected values is especially useful in stochastic programming problems, where the criteria to optimize appears in the form of an expected value. In such problems, one can replace the minimization of the expected value by the minimization of our lower or upper bounds, which results in an optimization over a larger space or a min-max problem, respectively. Using `TensCalc` [11], a high-performance numerical solver that combines symbolic computations with interior point methods, such problem can often be solved very accurately with modest computation.

Our bounds on the expected value are used in Section III to solve the problem of optimal Bayesian estimation [12], [13]. The goal is to find an estimator that minimizes the expected value of a loss function given a set of measurements. For this particular problem, we show that the approach outlined above actually leads to the maximum a posteriori estimation for a class of loss functions and probability density functions.

In Section IV the bounds are used to solve two problems in discrete time finite horizon open-loop optimal stochastic control. In both problems, the goal is to optimize a finite horizon criteria that depends on the trajectory of a dynamical

R. Chinchilla and J.P. Hespanha are with the Center for Control, Dynamical Systems, and Computation (CCDC) at the University of California, Santa Barbara, CA. {raphaelchinchilla,hespanha}@ucsb.edu.
This material is based upon work supported by the National Science Foundation under Grants No. ECCS-1608880 and CNS-1329650.

system with stochastic uncertainty in the form of additive perturbations and parametric uncertainties in the dynamics.

The first problem corresponds to a state-feedback scenario, for which the initial state is known, whereas the second problem corresponds to an output feedback scenario, for which we have past noisy output measurements based on which the initial state needs to be inferred.

A criterion similar to the one developed for the output feedback can be found in [14]–[16], where it is used to construct output feedback model predictive controllers. In those references, this criterion was justified on the basis of that it enabled formal stability proofs for the control scheme in a purely deterministic setting. The results in the present paper provide a formal justification for those criteria in a stochastic control setting.

II. Bounds on an Expected Value and Stochastic Programming

The following result provides upper and lower bounds on the expected value of a random variable:

Theorem 1 (Bounds on the expected value): Let $V : \mathbb{R}^N \to \mathbb{R}$ be a scalar valued function, Ψ a random vector taking values in $\mathbb{R}^N$ with probability density function $p_\Psi(\cdot)$ and $\mathbb{E}_\Psi[V(\Psi)]$ the expected value of $V(\Psi)$. For every scalar $\epsilon_1, \epsilon_2 \in \mathbb{R}$:

$$\inf_\psi \; J(\psi, \epsilon_1) \leqslant \mathbb{E}_\Psi[V(\Psi)] \leqslant \sup_\psi \; J(\psi, \epsilon_2) \qquad (1)$$

with $J(\psi, \epsilon) \coloneqq V(\psi) + \epsilon \log(p_\Psi(\psi)) + \epsilon \mathcal{H}_\Psi$ and where

$$\mathcal{H}_\Psi = -\int \log\left(p_\Psi(\psi)\right) p_\Psi(\psi) d\psi = \mathbb{E}_\Psi[-\log\left(p_\Psi(\Psi)\right)]$$

is the differential entropy of Ψ. ■

Proof. Starting with the upper bound, let us define

$$V_{up}(\epsilon) \coloneqq \sup_\psi \; V(\psi) + \epsilon \log(p_\Psi(\psi)).$$

Then, by definition of supremum,

$$\begin{aligned} & V(\psi) + \epsilon \log(p_\Psi(\psi)) \leqslant V_{up}(\epsilon) && \forall \psi, \forall \epsilon \\ \Rightarrow\; & V(\psi) \leqslant V_{up}(\epsilon) - \epsilon \log(p_\Psi(\psi)) && \forall \psi, \forall \epsilon \\ \Rightarrow\; & V(\psi)\, p_\Psi(\psi) \leqslant \left(V_{up}(\epsilon) - \epsilon \log(p_\Psi(\psi))\right) p_\Psi(\psi) && \forall \psi, \forall \epsilon \end{aligned}$$

and integrating in both sides yields

$$\mathbb{E}_\Psi[V(\Psi)] \leqslant V_{up}(\epsilon) + \epsilon \mathcal{H}_\Psi \qquad \forall \epsilon.$$

The lower bound can be obtained analogously □

Remark 1 (Conditional expected value): In the setting of Theorem 1 and given a particular realization y from another random vector $\mathcal{Y}$, an almost identical deduction can be done for conditional expected values. In this case, $J(\psi, \epsilon) \coloneqq V(\psi) + \epsilon \log\big(p_{\Psi|\mathcal{Y}}(\psi \mid y)\big) + \epsilon \tilde{\mathcal{H}}_{\Psi|\mathcal{Y}}(y)$, where

$$\tilde{\mathcal{H}}_{\Psi|\mathcal{Y}}(y) \coloneqq \int -\log\left(p_{\Psi|\mathcal{Y}}(\psi \mid y)\right) p_{\Psi|\mathcal{Y}}(\psi \mid y) d\psi$$

is a pseudo conditional differential entropy. ■

Remark 2 (Discrete case): Analogous bounds hold for discrete random variables when substituting *p.d.f.* and differential entropy by probability mass function and entropy. ■

Optimizing over ϵ: Theorem 1 defines families of upper and lower bounds parameterized by ϵ. The tightest bounds of these families are obtained by optimizing over ϵ, *i.e.*,

$$\sup_\epsilon \inf_\psi J(\psi, \epsilon) \leqslant E_\Psi[V(\Psi)] \leqslant \inf_\epsilon \sup_\psi J(\psi, \epsilon). \qquad (2)$$

As ϵ is a scalar, optimizing with respect to it is simple, and one could use methods such as bisection, the golden-section search [17] or a linear search. Now we construct a useful stopping criterion for the optimization over ϵ.

We will focus our exposition on the upper bound, as the lower bound has equivalent properties. We define the function $U(\epsilon) \coloneqq \sup_\psi \; J(\psi, \epsilon)$ which is convex [18, §5.1.2] and $\epsilon^* \coloneqq \min\{\arg\inf_\epsilon \; U(\epsilon)\}$.

Assumption 1: Consider the following conditions

i) There exists an $\epsilon \in \mathbb{R}$ such that $\sup_\psi \; J(\psi, \epsilon) < \infty$.
ii) $\sup_\psi \log(p_\Psi(\psi)) < \infty$.
iii) The functions $V(\cdot)$ and $p_\Psi(\cdot)$ are continuously differentiable, and $\forall \epsilon$ such that $U(\epsilon)$ is finite, the supremum takes place at a unique point, defining a continuously differentiable function $\psi^*(\epsilon) \coloneqq \arg\sup_\psi \; J(\psi, \epsilon)$.

Lemma 1: Let $U(\epsilon)$ be finite. If $\log(p_\Psi(\psi^*(\epsilon))) + \mathcal{H}_\Psi \geqslant 0$, then $\epsilon \geqslant \epsilon^*$. Otherwise, $\epsilon < \epsilon^*$. ■

Proof. In view of item (iii) of Assumption 1, taking the derivative of $U(\epsilon)$, one obtains

$$\begin{aligned} \frac{\mathrm{d}U(\epsilon)}{\mathrm{d}\epsilon} &= \frac{\mathrm{d}V(\psi^*(\epsilon)) + \epsilon \log(p_\Psi(\psi^*(\epsilon))) + \epsilon \mathcal{H}_\Psi}{\mathrm{d}\epsilon} \\ &= \frac{\mathrm{d}\psi^*(\epsilon)}{\mathrm{d}\epsilon}' \left.\frac{\mathrm{d}V(\psi) + \epsilon \log(p_\Psi(\psi))}{\mathrm{d}\psi}\right|_{\psi = \psi^*(\epsilon)} \\ &\qquad + \log(p_\Psi(\psi^*(\epsilon))) + \mathcal{H}_\Psi. \end{aligned}$$

As $\psi^*(\epsilon)$ is determined in an open set, the gradient of $J(\psi, \epsilon)$ is zero at that point, *i.e.*

$$\left.\frac{\mathrm{d}J(\psi, \epsilon)}{\mathrm{d}\psi}\right|_{\psi = \psi^*(\epsilon)} = \left.\frac{\mathrm{d}V(\psi) + \epsilon \log(p_\Psi(\psi))}{\mathrm{d}\psi}\right|_{\psi = \psi^*(\epsilon)} = \mathbf{0}.$$

Therefore $\mathrm{d}U(\epsilon)/\mathrm{d}\epsilon = \log(p_\Psi(\psi^*(\epsilon))) + \mathcal{H}_\Psi$. As $U(\epsilon)$ is convex, one can use the sign of its derivative to determine the relation of ϵ to ϵ^*, which finishes the proof. □

Proposition 1 (Linear error on the upper bound): Let ϵ_{low} and ϵ_{up} such that $\epsilon_{low} \leqslant \epsilon^* \leqslant \epsilon_{up}$, obtained using Lemma 1. Then $U(\epsilon_{up}) - U(\epsilon^*) \leqslant (\epsilon_{up} - \epsilon_{low})\, C$, where C is a finite non-negative constant. ■

Proof. Using the definition of $U(\epsilon)$ we obtain

$$\begin{aligned} U(\epsilon_{up}) &= \sup_\psi \; V(\psi) + (\epsilon_{up} + \epsilon^* - \epsilon^*)(\log(p_\Psi(\psi)) + \mathcal{H}_\Psi) \\ &\leqslant \sup_\psi \; [V(\psi) + \epsilon^* \log(p_\Psi(\psi)) + \epsilon^* \mathcal{H}_\Psi] \\ &\qquad + (\epsilon_{up} - \epsilon^*) \sup_\psi \; [\log(p_\Psi(\psi)) + \mathcal{H}_\Psi] \\ &\leqslant U(\epsilon^*) + (\epsilon_{up} - \epsilon_{low}) \sup_\psi \; [\log(p_\Psi(\psi)) + \mathcal{H}_\Psi]; \end{aligned}$$

from which the results follows with $C \coloneqq \sup_\psi \left[\log(p_\Psi(\psi)) + \mathcal{H}_\Psi\right]$, which is finite in view of item (ii) in Assumption 1 and non-negative as $\mathcal{H}_\Psi = -\mathbb{E}_\Psi[\log(p_\Psi(\Psi))] \Rightarrow \sup_\psi \left[\log(p_\Psi(\psi)) + \mathcal{H}_\Psi\right] \geqslant 0$. □

Notice that Proposition 1 implies that if $U(\epsilon) = +\infty$, then $\epsilon < \epsilon^*$ (item (i) in Assumption 1 implies $U(\epsilon^*) < +\infty$).

We illustrate how Proposition 1 can be used as a stopping criteria for the bisection algorithm.

Algorithm 1 Bisection algorithm

Require: a tolerance $\mathcal{T} > 0$, $\epsilon_{low} \leqslant \epsilon^* \leqslant \epsilon_{up}$

```
C ← sup_ψ log(p_Ψ(ψ)) + H_Ψ
k ← 0
while (ε_up − ε_low) C > T do
    ε̄ ← ½(ε_up + ε_low)
    k ← k + 1
    if U(ε̄) = +∞ or log(p_Ψ(ψ*(ε̄))) + H_Ψ < 0 then
        ε_low ← ε̄
    else
        ε_up ← ε̄
    end if
end while
    return ε_up and k
```

Proposition 2: We denote $\{\epsilon_{low}^{(0)}, \epsilon_{up}^{(0)}\}$ and $\{\epsilon_{low}^{(k)}, \epsilon_{up}^{(k)}\}$ as the initial and final values of $\{\epsilon_{low}, \epsilon_{up}\}$. Algorithm 1 converges to ϵ^* and the error obtained by using $\epsilon_{up}^{(k)}$ in lieu of ϵ^* satisfies $U\big(\epsilon_{up}^{(k)}\big) - U\big(\epsilon^*\big) \leqslant (0.5)^k \, (\epsilon_{up}^{(0)} - \epsilon_{low}^{(0)})\, C \leqslant \mathcal{T}$. ■

Proof. The bisection algorithm converges as $\epsilon^* \in [\epsilon_{low}^{(0)}, \epsilon_{up}^{(0)}]$ and as we can use Lemma 1 to classify $\bar{\epsilon}$ [17]. As each step reduces $(\epsilon_{up} - \epsilon_{low})$ by half, $(\epsilon_{up}^{(k)} - \epsilon_{low}^{(k)}) = (0.5)^k \, (\epsilon_{up}^{(0)} - \epsilon_{low}^{(0)})$. Applying Proposition 1 finishes the proof. □

Stochastic Programming: Let $V : \mathbb{R}^N \times \mathbb{R}^M \to \mathbb{R}_{\geqslant 0}$ be a scalar function, $u \in \mathbb{R}^M$ a variable and Ψ a random vector taking values in $\mathbb{R}^N$. Suppose one wants to solve

$$u^* \coloneqq \arg\min_u \; \mathbb{E}_\Psi[V(\Psi, u)], \tag{3}$$

where we assume that Ψ does not depend on u. Let us define

$$\mathcal{J}(\psi, \epsilon, u) \coloneqq V(\psi, u) + \epsilon \log(p_\Psi(\psi)) + \epsilon \mathcal{H}_\Psi.$$

Theorem 2 (Bounds on Stochastic Programming): Let u° be the solution to the optimization problem

$$u^\circ \coloneqq \arg\min_u \inf_\epsilon \sup_\psi \; \mathcal{J}(\psi, \epsilon, u) \tag{4}$$

The following bounds hold:

$$\begin{aligned} &\sup_\epsilon \min_u \inf_\psi \; \mathcal{J}(\psi, \epsilon, u) && (5a)\\ &\leqslant \mathbb{E}_\Psi[V(\Psi, u^*)] && (5b)\\ &\leqslant \mathbb{E}_\Psi[V(\Psi, u^\circ)] && (5c)\\ &\leqslant \inf_\epsilon \sup_\psi \; \mathcal{J}(\psi, \epsilon, u^\circ). && (5d)\end{aligned}$$

■

Proof. The proof follows from applying the bounds obtained in Theorem 1. □

Theorem 2 provides the formal justification to use the solution u° from (4) in lieu of the actual optimum u^* from (3) by giving performance guarantees. On one hand, the expected value using u° in (5c) will not exceed the optimal expected value in (5b) by more than the upper bound in (5d). On the other hand, the control u^* can never do better than the lower bound in (5a).

Notice that since $\mathcal{H}_\Psi$ does not depend on the optimization variables, one can obtain its value prior to the optimizations.

Remark 3: In the case where one can sample from Ψ, one could estimate (5c) using Monte Carlo integration [3], providing a better confidence interval for (5b). This would still require fewer samples than to solve (5b) and determine u^* using sample based stochastic optimization methods. ■

III. Optimal Bayesian Estimation

A common Bayesian formulation [12], [13] for estimating an unknown parameter Θ based on a set of measurements $\mathcal{Y}$ consists on finding the estimate ϕ that minimizes the conditional expectation of a semi-positive loss function $L(\Theta, \phi)$ given the measurements:

$$\phi^* = \arg\min_\phi \mathbb{E}_{\Theta|\mathcal{Y}}[L(\Theta, \phi)]. \tag{6}$$

Motivated by Theorem 2, one could replace the conditional expectation in the right-hand side of (6) and solve instead:

$$\phi^\circ \coloneqq \arg\min_{\phi,\epsilon} \sup_\theta \; L(\theta, \phi) + \epsilon \log\big(p_{\Theta|\mathcal{Y}}(\theta \mid y)\big) + \epsilon \mathcal{H}_{\Theta|\mathcal{Y}}(y). \tag{7}$$

Such estimate ϕ° has a strong relationship with the maximum a posteriori estimation, as we state in the next proposition.

Proposition 3 (Relationship with maximum a posteriori): Suppose that for $\epsilon = \epsilon^*$ the $\min$ and the $\sup$ in the definition of (7) commute. For any semi-positive loss function such that $L(\theta, \phi) = 0 \Leftrightarrow \theta = \phi$, the estimate ϕ° corresponds to the maximum a posteriori estimate of θ. ■

Proof. If the $\min$ and $\sup$ commute, then

$$\begin{aligned} &\min_\phi \sup_\theta \; L(\theta, \phi) + \epsilon^* \log\big(p_{\Theta|\mathcal{Y}}(\theta \mid y)\big) + \epsilon^* \mathcal{H}_{\Theta|\mathcal{Y}}(y)\\ &= \sup_\theta \min_\phi \; L(\theta, \phi) + \epsilon^* \log\big(p_{\Theta|\mathcal{Y}}(\theta \mid y)\big) + \epsilon^* \mathcal{H}_{\Theta|\mathcal{Y}}(y).\end{aligned}$$

As $L(\theta, \phi)$ is minimized by setting $\phi = \theta$, the minimizer will necessarily pick $\phi^\circ(\theta) = \theta$, and we conclude that

$$\begin{aligned} &\sup_\theta \min_\phi \; L(\theta, \phi) + \epsilon^* \log\big(p_{\Theta|\mathcal{Y}}(\theta \mid y)\big) + \epsilon^* \mathcal{H}_{\Theta|\mathcal{Y}}(y)\\ &= \sup_\theta \; \epsilon^* \log\big(p_{\Theta|\mathcal{Y}}(\theta \mid y)\big) + \epsilon^* \mathcal{H}_{\Theta|\mathcal{Y}}(y),\end{aligned}$$

achieved at the maximum a posteriori estimate of θ. □

IV. Open-loop Optimal Stochastic Control

This section addresses the use of Theorem 2 to solve discrete time finite horizon open-loop optimal stochastic control problem. The key challenge is to determine the probability density function (*p.d.f.*) and differential entropy of the states.

A. State-feedback

Consider the dynamical system

$$\mathcal{X}_{t+1} = f(\mathcal{X}_t, u_t, \Theta) + \mathcal{D}_t, \tag{8}$$

where $\mathcal{X}_t$ takes values in $\mathbb{R}^{N_x}$ and is the state of the system, Θ is a random vector representing unknown parameters, $u_t \in \mathbb{R}^{N_u}$ is the control signal and $\mathcal{D}_t$ is a zero mean random vector taking values in $\mathbb{R}^{N_x}$ called disturbance. The random parameter Θ and the random process $\mathcal{D}_t$ are all independent and have *p.d.f.* $p_\Theta(\cdot)$ and $p_{\mathcal{D}_t}(\cdot)$ and differential entropies $\mathcal{H}_\Theta$ and $\mathcal{H}_{\mathcal{D}_t}$, respectively.

For simplicity, we assume full knowledge of the initial state $\mathcal{X}_0$, but the results could be generalized to the case where we only know its distribution. Given a time horizon T, a cost function $V(\cdot)$ that depends on the unknown parameter θ, on the sequence of states $\mathcal{X}_{0:T} := \{\mathcal{X}_0, \mathcal{X}_1, \dots, \mathcal{X}_T\}$ and control inputs $u_{0:T-1} := \{u_0, u_1, \dots, u_{T-1}\}$ and an admissible set $\mathcal{U}$ for $u_{0:T-1}$, our goal is to solve the finite horizon optimal control problem

$$u^*_{0:T-1} := \arg\min_{u_{0:T-1}} \mathbb{E}_{\mathcal{X}_{1:T},\Theta}[V(\mathcal{X}_{0:T}, \Theta, u_{0:T-1})]$$
$$\text{s.t} \quad u_{0:T-1} \in \mathcal{U},$$

For this problem, the function $\mathcal{J}(\cdot)$ in Theorem 2 is

$$\mathcal{J}(x_{0:T}, \theta, \epsilon, u_{0:T-1}) = V(x_{0:T}, \theta, u_{0:T-1}) + \epsilon p_{\mathcal{X}_{1:T},\Theta}(x_{0:T}, \theta) + \epsilon \mathcal{H}_{\mathcal{X}_{1:T},\Theta}.$$

Proposition 4: In $\mathcal{J}(x_{0,T}, \theta, \epsilon, u_{0:T-1})$'s expression,

$$p_{\mathcal{X}_{1:T},\Theta}(x_{0:T}, \theta) = p_\Theta(\theta) \prod_{t=0}^{T-1} p_{\mathcal{D}_t}(x_{t+1} - f(x_t, u_t, \theta)),$$

and $\mathcal{H}_{\mathcal{X}_{1:T},\Theta} = \mathcal{H}_\Theta + \sum_{t=0}^{T-1} \mathcal{H}_{\mathcal{D}_t}$. ■

Proof. Using the Markov Chain property of stochastic dynamic systems

$$p_{\mathcal{X}_{1:T},\Theta}(x_{0:T}, \theta) = p_\Theta(\theta) \prod_{t=0}^{T-1} p_{\mathcal{X}_{t+1}|\mathcal{X}_t,\Theta}(x_{t+1}, | \, x_t, \theta).$$

Using the property of change of variable of probability density functions,

$$\begin{aligned} &p_{\mathcal{X}_{t+1},|\mathcal{X}_t,\Theta}(x_{t+1}, | \, x_t, \theta) \\ &= \left|\det\left(\frac{\mathrm{d}(x_{t+1} - f(x_t, u_t, \theta))}{\mathrm{d}x_{t+1}}\right)\right| p_{\mathcal{D}_t}(x_{t+1} - f(x_t, u_t, \theta)) \\ &= p_{\mathcal{D}_t}(x_{t+1} - f(x_t, u_t, \theta)). \end{aligned}$$

For the differential entropy, we use the chain rule

$$\mathcal{H}_{\mathcal{X}_{1:T},\Theta} = \mathcal{H}_\Theta + \sum_{t=0}^{T-1} \mathcal{H}_{\mathcal{X}_{t+1}|\mathcal{X}_t,\Theta}.$$

Calculating $\mathcal{H}_{\mathcal{X}_{t+1}|\mathcal{X}_t,\Theta}$:

$$\begin{aligned} &-\int p_{\mathcal{X}_{t+1}\mathcal{X}_t,\Theta}(x_{t+1}, x_t, \theta) \\ &\qquad \log\big(p_{\mathcal{X}_{t+1}|\mathcal{X}_t,\Theta}(x_{t+1} \mid x_t, \theta)\big) dx_{t+1} dx_t d\theta \\ &= -\int p_{\mathcal{X}_{t+1}|\mathcal{X}_t,\Theta}(x_{t+1} \mid x_t, \theta) p_{\mathcal{X}_t}(x_t) p_\Theta(\theta) \\ &\qquad \log\big(p_{\mathcal{X}_{t+1}|\mathcal{X}_t,\Theta}(x_{t+1} \mid x_t, \theta)\big) dx_{t+1} dx_t d\theta \\ &= -\int p_{\mathcal{D}_t}(x_{t+1} - f(x_t, u_t, \theta)) p_{\mathcal{X}_t}(x_t) p_\Theta(\theta) \\ &\qquad \log(p_{\mathcal{D}_t}(x_{t+1} - f(x_t, u_t, \theta))) dx_{t+1} dx_t d\theta \\ &= \int \mathcal{H}_{\mathcal{D}_t} p_{\mathcal{X}_t}(x_t) p_\Theta(\theta) dx_t d\theta = \mathcal{H}_{\mathcal{D}_t}. \end{aligned}$$
□

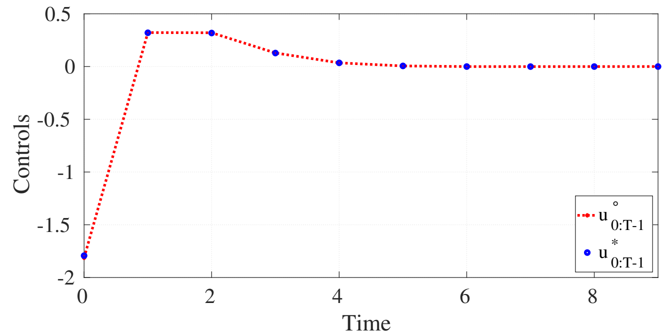


Fig. 1. Results for Example 1 comparing $u^\circ_{0:T-1}$ and the analytical solution $u^*_{0:T-1}$ obtained solving (10). The solutions are so similar that $u^*_{0:T-1}$ had to be represented with dots otherwise $u^\circ_{0:T-1}$ was hidden.

Example 1: We illustrate the use of Theorem 2 to open-loop control with state-feedback with the following example, which was selected so that we could compute the solution analytically and compare it with the control obtained using Theorem 2. Consider a linear time invariant system given by

$$\mathcal{X}_{t+1} = A\,\mathcal{X}_t + B\,u_t + \mathcal{D}_t, \tag{9}$$

where $\mathcal{D}_t$ is an independent zero mean Gaussian random vector and a quadratic cost function $V(\mathcal{X}_{0:T}, u_{0:T-1}) = \sum_{t=0}^{T-1}(\mathcal{X}_t' Q \mathcal{X}_t + u_t' R u_t) + \mathcal{X}_T' F \mathcal{X}_T$. The expected value has an analytical expression given by

$$\begin{aligned} &\mathbb{E}_{\mathcal{X}_{1:T}}[V(\mathcal{X}_{0:T}, u_{0:T-1})] \\ &= \sum_{t=0}^{T-1}(\bar{x}_t' Q \bar{x}_t + u_t' R u_t) + \bar{x}_T' F \bar{x}_T + \sum_{t=0}^{T-1} \mathrm{tr}\big(Q\bar{S}_t\big) + \mathrm{tr}\big(F\bar{S}_T\big), \end{aligned} \tag{10}$$

where $\bar{x}_t$ is the state of the nominal system (*i.e.*, with no perturbations) and $\bar{S}_t$ is the covariance matrix of $\mathcal{X}_t$.

We chose a horizon $T = 10$, and dynamics

$$A = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}, \; B = \begin{bmatrix} 0 \\ 1 \end{bmatrix},$$

Figure 1 shows the controls $u^\circ_{0:T-1}$ given by Theorem 2 as well as the true optimum, obtained by solving (10). They turn out to be essentially the same.

B. Output-feedback

Considered the system in (8) complemented by an observation equation:

$$\begin{aligned} \mathcal{X}_{t+1} &= f(\mathcal{X}_t, u_t, \Theta) + \mathcal{D}_t \\ \mathcal{Y}_t &= h(\mathcal{X}_t, \theta) + \mathcal{N}_t, \end{aligned} \tag{11}$$

where $\mathcal{Y}_t$ is the observation vector and $\mathcal{N}_t$ is a zero mean random vector called noise, both taking values in $\mathbb{R}^{N_y}$. The random processes $\mathcal{N}_t$ is independent of $\mathcal{D}_t$ and Θ and has *p.d.f.* $p_{\mathcal{N}_t}(\cdot)$.

We assume given a past window of size K with a set of controls $u_{-K:-1} \coloneqq \{u_{-K}, u_{-K+1}, \dots, u_{-1}\}$ applied before time $t = 0$ and outputs $y_{-K:0} \coloneqq \{y_{-K}, y_{-K+1}, \dots, y_0\}$ a realization of $\mathcal{Y}_{-K:0}$, observed before time 0. Given a time horizon T, a function $V(\cdot)$ that depends on the unknown parameter Θ, on the sequence of states $\mathcal{X}_{0:T} \coloneqq \{\mathcal{X}_0, \mathcal{X}_1, \dots, \mathcal{X}_T\}$ and control inputs $u_{0:T-1} \coloneqq \{u_0, u_1, \dots, u_{T-1}\}$ from time 0 to time $T-1$ and an admissible set $\mathcal{U}$ for $u_{0:T-1}$, our goal is to solve the finite horizon optimal control problem

$$\begin{aligned} u^*_{0:T-1} &\coloneqq \arg\min \mathbb{E}_{\mathcal{X}_{-K:T},\Theta|\mathcal{Y}_{-K:0}}[V(\mathcal{X}_{0:T}, \Theta, u_{0:T-1})] \\ &\text{s.t} \quad u_{0:T-1} \in \mathcal{U}, \end{aligned}$$

Different from the problem in §IV-A, the expected value is now also taken with respect to past values of $\mathcal{X}_t$ and is conditioned to the observations.

For this problem, the function $\mathcal{J}(\cdot)$ in Theorem 2 is

$$\begin{aligned} \mathcal{J}(x_{-K:T}, \theta, \epsilon, u_{0:T-1}) &= V(x_{0:T}, \theta, u_{0:T-1}) \\ &+ \epsilon p_{\mathcal{X}_{-K:T},\Theta|\mathcal{Y}_{-K:0}}(x_{-K:T}, \theta \mid y_{-K:0}) \\ &+ \epsilon \tilde{\mathcal{H}}_{\mathcal{X}_{-K:T},\Theta|\mathcal{Y}_{-K:0}}(y_{-K:0}) \end{aligned}$$

Proposition 5: In $\mathcal{J}(x_{-K:T}, \theta, \epsilon, u_{0:T-1})$'s expression,

$$p_{\mathcal{X}_{-K:T},\Theta|\mathcal{Y}_{-K:0}}(x_{-K:T}, \theta \mid y_{-K:0}) = \prod_{t=-K}^{0} p_{\mathcal{N}_t}(y_t - h(x_t, \theta))$$

$$p_{\mathcal{X}_{-K}}(x_{-K}) \prod_{t=-K}^{T-1} p_{\mathcal{D}_t}(x_{t+1} - f(x_t, u_t, \theta)) p_\Theta(\theta) / p_{\mathcal{Y}_{-K:0}}(y_{-K:0})$$

and

$$\tilde{\mathcal{H}}_{\mathcal{X}_{-K:T},\Theta|\mathcal{Y}_{-K:0}}(y_{-K:0}) = \sum_{t=0}^{T-1} \mathcal{H}_{\mathcal{D}_t} + \tilde{\mathcal{H}}_{\mathcal{X}_{-K:0},\Theta|\mathcal{Y}_{-K:0}}(y_{-K:0}),$$

which does not depend on $u_{0:T-1}$. ■

Proof. Using Bayes Theorem we can determine the *p.d.f.*

$$p_{\mathcal{X}_{-K:T},\Theta|\mathcal{Y}_{-K:0}}(x_{-K:T}, \theta \mid y_{-K:0}) = \frac{p_{\mathcal{Y}_{-K:0}|\mathcal{X}_{-K:T},\Theta}(y_{-K:0} \mid x_{-K:T}, \theta) p_{\mathcal{X}_{-K:T}|\Theta}(x_{-K:T} \mid \theta) p_\Theta(\theta)}{p_{\mathcal{Y}_{-K:0}}(y_{-K:0})}.$$

Proving that

$$\begin{aligned} &p_{\mathcal{X}_{-K:T}|\Theta}(x_{-K:T} \mid \theta) \\ &= p_{\mathcal{X}_{-K}}(x_{-K}) \prod_{t=-K}^{T-1} p_{\mathcal{D}_t}(x_{t+1} - f(x_t, u_t, \theta)) \end{aligned}$$

is analogous to what was done in the proof in Proposition 4. As the observations are conditionally independent,

$$p_{\mathcal{Y}_{-K:0}|\mathcal{X}_{-K:0},\Theta}(y_{-K:0} \mid x_{-K:0}, \theta) = \prod_{t=-K}^{0} p_{\mathcal{Y}_t|\mathcal{X}_t,\Theta}(y_t \mid x_t, \theta),$$

from which a change of variable gives the result

$$p_{\mathcal{Y}_t|\mathcal{X}_t,\Theta}(y_t \mid x_t, \theta) = p_{\mathcal{N}_t}(y_t - h(x_t, \theta)).$$

Therefore, the *p.d.f.* of $\mathcal{X}_{-K:T}$ can be determined from the *p.d.f.* of $\mathcal{X}_{-K}$, Θ, $\mathcal{D}_t$ and $\mathcal{N}_t$.

The second part of the proof is to show that $\tilde{\mathcal{H}}_{\mathcal{X}_{-K:T},\Theta|\mathcal{Y}_{-K:0}}(y_{-K:0})$ does not depend on $u_{0:T-1}$. The key element is to separate $\tilde{\mathcal{H}}_{\mathcal{X}_{-K:T},\Theta|\mathcal{Y}_{-K:0}}(y_{-K:0})$ into past and future values and show that the future values depend only on the differential entropy of the disturbance $\mathcal{H}_{\mathcal{D}_t}$.

$$\begin{aligned} &\tilde{\mathcal{H}}_{\mathcal{X}_{-K:T},\Theta|\mathcal{Y}_{-K:0}}(y_{-K:0}) \\ &= -\int p_{\mathcal{X}_{-K:T},\Theta|\mathcal{Y}_{-K:0}}(x_{-K:T}, \theta \mid y_{-K:0}) \\ &\qquad \log\big(p_{\mathcal{X}_{-K:T},\Theta|\mathcal{Y}_{-K:0}}(x_{-K:T}, \theta \mid y_{-K:0})\big) dx_{-K:T} d\theta \\ &= -\int p_{\mathcal{X}_{-K:T},\Theta|\mathcal{Y}_{-K:0}}(x_{-K:T}, \theta \mid y_{-K:0}) \\ &\qquad \log\big(p_{\mathcal{X}_{1:T},\Theta|\mathcal{X}_0}(x_{1:T}, \theta \mid x_0)\big) dx_{-K:T} d\theta \\ &\qquad - \int p_{\mathcal{X}_{-K:T},\Theta|\mathcal{Y}_{-K:0}}(x_{-K:T}, \theta \mid y_{-K:0}) \\ &\qquad \log\big(p_{\mathcal{X}_{-K:0},\Theta|\mathcal{Y}_{-K:0}}(x_{-K:0}, \theta \mid y_{-K:0})\big) dx_{-K:T} d\theta \\ &= -\int p_{\mathcal{X}_{1:T},\Theta|\mathcal{X}_0}(x_{1:T}, \theta \mid x_0) \\ &\qquad \log\big(p_{\mathcal{X}_{1:T},\Theta|\mathcal{X}_0}(x_{1:T}, \theta \mid x_0)\big) dx_{1:T} d\theta \\ &\qquad - \int p_{\mathcal{X}_{-K:0},\Theta|\mathcal{Y}_{-K:0}}(x_{-K:0}, \theta \mid y_{-K:0}) \\ &\qquad \log\big(p_{\mathcal{X}_{-K:0},\Theta|\mathcal{Y}_{-K:0}}(x_{-K:0}, \theta \mid y_{-K:0})\big) dx_{-K:0} d\theta \\ &= \sum_{t=0}^{T-1} \mathcal{H}_{\mathcal{D}_t} + \tilde{\mathcal{H}}_{\mathcal{X}_{-K:0},\Theta|\mathcal{Y}_{-K:0}}(y_{-K:0}), \end{aligned}$$

where the last line follows from the expression of $\mathcal{H}_{\mathcal{X}_{1:T}}$ deduced in Proposition 4. From that proof, $\sum_{t=0}^{T-1} \mathcal{H}_{\mathcal{D}_t}$ does not depend on $u_{0:T-1}$ and, by causality, neither does $\tilde{\mathcal{H}}_{\mathcal{X}_{-K:0},\Theta|\mathcal{Y}_{-K:0}}(y_{-K:0})$. Therefore, $\tilde{\mathcal{H}}_{\mathcal{X}_{-K:T},\Theta|\mathcal{Y}_{-K:0}}(y_{-K:0})$ does not depend on $u_{0:T-1}$. □

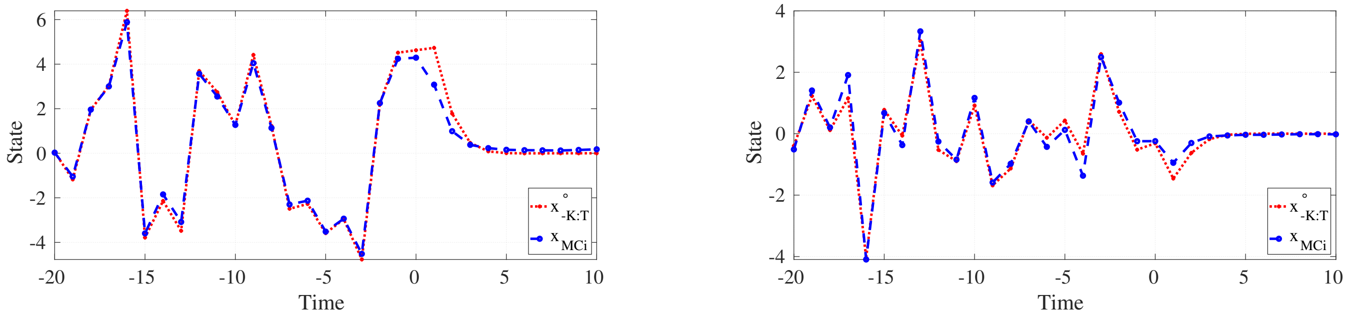
Remark 4: Solving the upper bound optimization from Theorem 2 provides both a control $u^\circ_{0:T-1}$ and a corresponding state trajectory $x^\circ_{-K:T}$. When the min and the sup commute, one can interpret $x^\circ_{-K:0}$ as a state estimate of the past values and $x^\circ_{1:T}$ as a state estimate of the future values for the control $u^\circ_{0:T-1}$. In the absence of a control objective, this would correspond to a maximum a posteriori estimate, as we saw in Section III. ■

Example 2: We illustrate the use of Theorem 2 to open-loop control with output-feedback with the following example which, to the best of our knowledge, cannot be solved in closed form. Consider a linear system

$$\begin{aligned} \mathcal{X}_{t+1} &= A\,\mathcal{X}_t + B\,u_t + \mathcal{D}_t \\ \mathcal{Y}_t &= C\,\mathcal{X}_t + \mathcal{N}_t, \end{aligned} \tag{12}$$

where $\mathcal{D}_t$ and $\mathcal{N}_t$ are independent zero mean Gaussian processes. The system is time-invariant, C is an identity matrix, but the matrices A and B are unknown stochastic parameters such that we have the following priors on them

$$A = \begin{bmatrix} \mathcal{N}(1,1) & \mathcal{N}(1,1) \\ 0 & \mathcal{N}(1,1) \end{bmatrix} \; B = \begin{bmatrix} 0 \\ \mathcal{N}(1,1) \end{bmatrix},$$



(a) First component of the state.

(b) Second component of the state.

Fig. 2. Results from Example 2. We compare the state trajectories for $x^{\circ}_{-K:T}$, obtained solving the min-max and x_{MCi}, obtained with Monte Carlo integration for the control $u^{\circ}_{0:T-1}$. We can see how both are similar, reinforcing that $x^{\circ}_{-K:T}$ can be seen as an estimation.

where $\mathcal{N}(\mu, \sigma^2)$ is a Gaussian random variable with mean μ and variance σ^2. We chose a quadratic cost, a future horizon $T = 10$ and a past horizon $K = 20$.

The problem was solved on MATLAB® using `TensCalc` [11], which compiles an efficient optimization code based on symbolic computation and interior point methods. We obtained the controls $u^{\circ}_{0:T-1}$.

We now apply Theorem 2. The optimal expected value (that we do not know) is larger than 22 according to (5a), and smaller than 4.5×10^5 (5d). In light of Remark 3, we can refine the confidence interval on the optimal expected value by using Monte Carlo integration to estimate the conditional expected value $\mathbb{E}_{\mathcal{X}_{-K:T},\Theta|\mathcal{Y}_{-K:0}}\left[V(\mathcal{X}_{0:T}, \Theta, u^{\circ}_{0:T-1})\right]$, *i.e.* equation (5c), which is equal to 250.

Figure 2 compares the trajectories from $x^{\circ}_{-K:T}$, obtained solving the min-max upper bound optimization of Theorem 2, and the estimated expected value of the trajectory for the control $u^{\circ}_{0:T-1}$, obtained using Monte Carlo integration. As we can see, they match closely, reinforcing the interpretation that $x^{\circ}_{-K:T}$ can be seen as an estimation of the states trajectory (see Remark 4). Furthermore, although we do not have any guarantee on the optimality of $u^{\circ}_{0:T-1}$, it does drive the expected value of the trajectory towards the origin.

V. Conclusion and Future Work

We presented a method to determine lower and upper bounds on the expected value of a scalar function of a random vector and how it can be used in stochastic programming. The bounds are computed through an optimization requiring only the probability density function of the underlying random vector and its differential entropy. These bounds can be applied to the problem of Bayesian estimation and how they relate to maximum a posteriori estimation.They can also be used to solve optimization arising in open-loop control of finite horizon stochastic dynamical systems with either state or output feedback. We were able to efficiently compute the control using solvers generated from `TensCalc`.

Directions for future research include further investigations on under each conditions the bounds are finite and on the possibility of using a convex combination of the lower and upper bound to solve the stochastic programming problem. In the context of optimal control, extension of our bounds to stochastic model predictive control or infinite horizon optimal stochastic control could also be investigated.

References

[1] L. F. Shampine, "Vectorized adaptive quadrature in matlab," *Journal of Computational and Applied Mathematics*, vol. 211, no. 2, pp. 131–140, 2008.

[2] C. Robert, *The Bayesian choice: from decision-theoretic foundations to computational implementation*. Springer Science & Business Media, 2007.

[3] C. Robert and G. Casella, *Monte Carlo statistical methods*. Springer Science & Business Media, 2013.

[4] J. C. Spall, *Introduction to stochastic search and optimization: estimation, simulation, and control*, vol. 65. John Wiley & Sons, 2005.

[5] M. C. Campi and S. Garatti, *Introduction to the Scenario Approach*. Society for Industrial and Applied Mathematics.

[6] A. Shapiro, D. Dentcheva, and A. Ruszczynski, *Lectures on Stochastic Programming: Modeling and Theory*. Society for Industrial and Applied Mathematics, second edition ed.

[7] F. Orieux, O. Féron, and J.-F. Giovannelli, "Sampling high-dimensional gaussian distributions for general linear inverse problems," *IEEE Signal Processing Letters*, vol. 19, no. 5, pp. 251–254, 2012.

[8] F. Orieux, E. Sepulveda, V. Loriette, B. Dubertret, and J.-C. Olivo-Marin, "Bayesian estimation for optimized structured illumination microscopy," *IEEE Transactions on image processing*, vol. 21, no. 2, pp. 601–614, 2012.

[9] J.-F. Giovannelli, "Unsupervised bayesian convex deconvolution based on a field with an explicit partition function," *IEEE Transactions on Image Processing*, vol. 17, no. 1, pp. 16–26, 2008.

[10] T. Park and G. Casella, "The bayesian lasso," *Journal of the American Statistical Association*, vol. 103, no. 482, pp. 681–686, 2008.

[11] J. P. Hespanha, "TensCalc: A toolbox to generate fast code to solve nonlinear constrained minimizations and compute Nash equilibria.," tech. rep., Center for Control, Dynamical Systems and Computation; University of California, Santa Barbara, https://github.com/hespanha/tenscalc, 2017.

[12] S. M. Kay, "Fundamentals of statistical signal processing, volume i: Estimation theory (v. 1)," *PTR Prentice-Hall, Englewood Cliffs*, 1993.

[13] E. L. Lehmann and G. Casella, *Theory of point estimation*. Springer Science & Business Media, 2006.

[14] D. A. Copp and J. P. Hespanha, "Simultaneous nonlinear model predictive control and state estimation," *Automatica*, vol. 77, pp. 143–154, 2017.

[15] D. Copp and J. Hespanha, "Addressing adaptation and learning in the context of model predictive control with moving-horizon estimation," in *Control of Complex Systems*, pp. 187–209, Elsevier, 2016.

[16] D. A. Copp and J. P. Hespanha, "Nonlinear output-feedback model predictive control with moving horizon estimation," in *53rd IEEE Conference on Decision and Control*, pp. 3511–3517, IEEE, 2014.

[17] E. K. Chong and S. H. Zak, *An introduction to optimization*, vol. 76. John Wiley & Sons, 2013.

[18] S. Boyd and L. Vandenberghe, *Convex optimization*. Cambridge university press, 2004.