Proceeding of the 2004 American Control Conference
Boston, Massachusetts June 30 - July 2, 2004

ACTIVE IDENTIFICATION OF UNKNOWN SYSTEMS: AN INFORMATION THEORETIC APPROACH

M. Baglietto, L. Scardovi, and R. Zoppoli

Abstract— **In this paper, the problem of identifying the value of a set of a priori unknown parameters of a discrete-time dynamic system is faced. To this end, a control law that minimizes the "information gain" about such parameters is required. The Shannon entropy is proposed (and justified) as a suitable information measure. After formulating the general statement of the problem, a viable approximation consisting in formulating a receding horizon problem is proposed. When the measurement channel is linear in the parameters, the optimal solution of the latter is given involving the minimization of a (finite dimensional) cost at each step.**

I. INTRODUCTION

In recent years, great deal of progresses have been made in the area of system identification. However, in the usual approach, the problem consists in estimating the unknown parameters under a giving control policy. It is well known that in general the control policy influences the identification. We pose the following question: What is the control policy by which the Decision Maker (or Controller) gains the maximum information? We shall call this problem *Exploration Problem*.

In statistics, a similar problem is the Optimal Experiment Design (OED), where one has to design an experiment in order to infer about an unknown parameterized system [1], [2]. Also in machine learning a similar problem arises, when one can choose the input patterns to optimize an approximator (Active learning) [3]. In robotics the problem of environment exploration can be formulated as a particular case of our general Problem, and it has been studied from an heuristic point of view [4]. In these last years some researchers have used information theoretic concepts to study control problems (see [5], [6]).

In this paper, we formulate the problem in an information theoretic setting by using the Shannon entropy as a measure of information about a set of unknown parameters. As the problem of maximizing the information gain is almost unsolvable under general hypotheses, we propose a second formulation which results to be more tractable and propose a feedback control law that aims at maximizing the information gain in a receding horizon optimal control setting. We show that, for a particular class of systems,such a control law can be obtained solving on-line a sequence of non-linear constrained minimization problems.

This paper is organized as follows: in Section 2, we define the entropy and the information concepts and then we give some useful properties. In Section 3, we state the exploration problem using two different but correlated formulations. In Section 4 we show how the problem can be reduced, under some hypotheses, to a sequence of non-linear constrained minimization problems. In Section 5 we give a geometrical interpretation to the results when a simple system is considered. Finally experimental results are presented in Section 6.

M. Baglietto, L. Scardovi, and R. Zoppoli are with the Department of Communications, Computer and System Sciences, DIST–University of Genoa, Via Opera Pia 13, 16145 Genova, Italy, {mbaglietto,lucas,rzop}@dist.unige.it

II. PROPERTIES OF THE INFORMATION MEASURE

In this section we will give some definitions and properties of the information measure when the inference is influenced by using a DM acting on the dynamic system.

Let us consider the following state and measurement equation:

$$x_{t+1} = f(x_t, u_t, \xi_t)\,, \quad t = 0, 1, \ldots \tag{1}$$

$$y_t = h(x_t, \eta_t, \theta)\,, \quad t = 0, 1, \ldots \tag{2}$$

where $x_t \in \mathbb{R}^n, y_t \in \mathbb{R}^d, u_t \in U \subseteq \mathbb{R}^m, \xi_t \in \mathbb{R}^p, \theta \in \mathbb{R}^k$. We define the information vector as:

$$I_t = \begin{bmatrix} I_{t-1} \\ u_{t-1} \\ y_t \end{bmatrix}$$

$$I_0 = y_0\ .$$

Here and in the following of the paper, for any time-varying vector v_t, we denote $v_{t_1}^{t_2}$ =col$(v_{t_1}, v_{t_1+1}, \ldots, v_{t_2})$. Let us consider a fixed control law $\pi_0^{T-1} = \{u_0(I_0), u_1(I_1), \ldots\ u_T(I_{T-1})\}$ acting on the system, the consequent measurement vector y_0^T and a prior probability density (shortly, from now on, we shall call it probability) $p_0(\theta)$. In which way is it possible to quantify the information on the parameters the DM have gained with the given informative vector I_T? Two possible choices of information measure are the *Shannon entropy* and the *Kullback divergence*. The Shannon entropy (or simply entropy) of a random variable θ is defined as:

$$H(p(\theta)) \triangleq \int_{\mathbb{R}^z} p(\theta) \ln \frac{1}{p(\theta)} d\theta.$$

The difference between prior and posterior entropies will be addressed as "information gain", and is given by:

$$\mathcal{I}(p(\theta), I_T) = \int_{\mathbb{R}^z} p(\theta|I_T) \ln p(\theta|I_T) d\theta - \int_{\mathbb{R}^z} p(\theta) \ln p(\theta) d\theta\ .$$

0-7803-8335-4/04/$17.00 ©2004 AACC

The Kullback divergence between prior and posterior probability functions is defined as:

$$D(p(\theta|I_T)||p(\theta)) \triangleq \int_{\mathbb{R}^z} p(\theta|I_T) \ln \frac{p(\theta|I_T)}{p(\theta)} d\theta.$$

Since we are interested in the expected value of the information gain given the control policy, we can indistinctly choose one of the above mentioned measures, as stated by the following theorem:

Theorem 1:

$$\mathop{\mathrm{E}}_{y_0^T} D(p(\theta|I_T)||p_0(\theta)) = \mathop{\mathrm{E}}_{y_0^T} \mathcal{I}(p_0(\theta), I_T)$$

Proof:
By using the Bayes formula:

$$\begin{aligned}
&\mathop{\mathrm{E}}_{y_0^T} \mathcal{I}(p_0(\theta), I_T) = \\
&\int_{\mathbb{R}^{d\times(T+1)}} p(y_0^T) \int_{\mathbb{R}^z} p(\theta|I_T) \ln p(\theta|I_T) d\theta - \\
&\int_{\mathbb{R}^z} p(\theta) \ln p(\theta) d\theta dy_0^T = \\
&\int_{\mathbb{R}^{d\times(T+1)}} \int_{\mathbb{R}^z} p(y_0^T, \theta|u_0^{T-1}) \ln \frac{p(y_0^T|\theta, u_0^{T-1}) p(\theta)}{p(y_0^T|u_0^{T-1})} dy_0^T \\
&- \int_{\mathbb{R}^z} \int_{\mathbb{R}^{d\times(T+1)}} p(y_0^T|u_0^{T-1}) p(\theta|y_0^T, u_0^{T-1}) dy_0^T \ln p_0(\theta) d\theta \\
&= \int_{\mathbb{R}^{d\times(T+1)}, \mathbb{R}^z} p(y_0^T, \theta|u_0^{T-1}) \ln \frac{p(y_0^T, \theta|u_0^{T-1})}{p(y_0^T|u_0^{T-1}) p(\theta)} dy_0^T d\theta
\end{aligned}$$

□

Given a sequence of control functions $\pi_t^{t+T-1} \triangleq \{u_t(I_t), u_{t+1}(I_{t+1}), \ldots u_{t+T-1}(I_{t+T-1})\}$ and a probability $p(\theta|I_t)$ we define

$$\begin{aligned}
&\bar{\mathcal{I}}(\pi_t^{t+T-1}, p(\theta|I_t)) \triangleq \mathop{\mathrm{E}}_{y_{t+1}^{t+T}} \{\mathcal{I}_T(p(\theta|I_t), I_{t+T})\} = \\
&\mathop{\mathrm{E}}_{y_{t+1}^{t+T}} \{H(p(\theta|I_t)) - H(p(\theta|I_{t+T}))\}
\end{aligned}$$

to be the expected information gain obtained by applying the control sequence π_t^{t+T-1}. We give now an important property:

Proposition 1:

$$\bar{\mathcal{I}}(\pi_t^{t+T-1}, p(\theta|I_t) \geq 0$$

Proof:

This fact can be proved using the fact that K.D. is a pseudo-metric and then it is always positive see [7]

□

III. Problem formulations

Let us consider the state equation (1) and the measurement equation (2). In the following, we shall state the exploration problem in a general case. The control objective is to gain all the information achievable about the parameters, minimizing a given process cost. A possible, approximate solution, consists in formulating a second problem in which the control objective is to maximize the information in a receding horizon (RH) setting. To fix the ideas let us consider an explicative example. A DM (e.g. a robot) that must explore totally the unknown environment in which it is moving. After choosing the parameterized model of the environment, the DM must move and gain local information by means of its sensors. During the process, a cost is paid (e.g. fuel consumption). For this example the first formulation correspond to explore totally the environment minimizing the total cost. Clearly nobody knows what is the time horizon in which the DM will finish its task. The second formulation correspond to maximize the information the DM achieves given the size of the control window. A constant cost for every time step is considered. Clearly we can see the second technique as an approximate way to solve the general problem when we have a constant cost (minimum time problem). We now give first the general and then the RH formulation for the exploration Problem. In next section we shall give the solution to the second problem for a particular class of systems.

Problem 1: Find an admissible control law $\left(\pi_0^{T-1}\right)^\circ = \{u_0^\circ(I_0), u_1^\circ(I_1), \ldots, u_{T-1}^\circ(I_{T-1})\}$ that minimizes the following:

$$J = \mathop{\mathrm{E}}_{\xi_t, \eta_t, x_0, \theta} \sum_{t=0}^{T-1} g(x_t, \theta, u_t(I_t), \xi_t) + g_T(x_T, \theta) \quad (3)$$

subject to:

$$x_{t+1} = f(x_t, u_t(I_t), \xi_t), \quad t = 0, 1, \ldots \quad (4)$$

$$y_t = h(x_t, \eta_t, \theta), \quad t = 0, 1, \ldots$$

$$e^{H(p(\theta|I_T))} \leq \epsilon \quad (5)$$

$$u_t \in U$$

where the time horizon T is a priori unknown and ϵ is an arbitrary small constant.

□

Constraint (5) states that the final probability density function must belong to the set of functions for which the exponential of the entropy is under a given arbitrarily small constant. It is worth noting that the exponential function is needed because the entropy (if the random variable is continuous) assumes decreasing (even negative) values with the "contraction" of the density function.

From now on, we shall concentrate on minimum time problems, in this case we have $g(\cdot, \cdot, \cdot, \cdot) = 1$. The general hypotheses under which Problem 1 has been formulated make it impossible to solve it exactly. The properties of the entropy given in the previous section suggest us to renounce to solve Problem 1 and to reformulate it in a RH form. To this end, let us consider the following sequence of finite-horizon cost functions:

$$\bar{J}_t = -\bar{\mathcal{I}}(u_t^{t+N-1}, p(\theta|I_t)), \; t = 0, 1, \ldots \quad (6)$$

where

$$u_t^{t+N-1} = \{u_{t,t}, u_{t+1,t}, \ldots, u_{t+N-1,t}\}.$$

We can now formulate the following problem:

Problem 2: Find sequentially, stage after stage, the optimal RH control vectors $u_t^* = \bar{u}_{t,t}(I_t),\ t = 0, 1, \ldots$, where $\bar{u}_t^{t+N-1} = \{\bar{u}_{t,t}, \bar{u}_{t+1,t}, \ldots, \bar{u}_{t+N-1,t}\}$, that minimizes the function $\bar{J}_t,\ t = 0, 1, \ldots$.

□Then, when the system is at stage t, a sequence of N optimal control vectors minimizing (6) is derived but only the first control function is used to generate the control vector $u_t^* = \bar{u}_{t,t}(I_t)$. The same procedure is repeated at stages $t+1, t+2, \ldots$ up to infinity. $\bar{\mathcal{I}}(u_t^{t+N-1}, p(\theta|I_t))$ represents the expected information gain in the control window that starts at t and finishes at $t + N - 1$.

IV. Problem solution

In this section we shall focus on a particular class of systems, i.e.:

$$x_{t+1} = f(x_t, u_t(I_t)) , \quad t = 0, 1, \ldots$$
$$y_t = h(x_t)'\theta + \eta_t , \quad t = 0, 1, \ldots \qquad (7)$$

where η_t is an additive Gaussian noise with density $G(0, \sigma^2)$ and θ is $G(\bar{\theta}, \Sigma_0)$ ($'$ denote the transpose). The state of the system is perfectly measurable, i.e. the DM, at every time step, knows perfectly its state but not the parameters. Then we can state the following:

Theorem 2: For the System (7):

$$\bar{J}_t = -\frac{1}{2}\sum_{i=1}^{N} \ln(1 + \frac{1}{\sigma^2}(h(x_{t+i}))^{'}\Sigma_{t+i}h(x_{t+i})) \qquad (8)$$

Where Σ_t denote the covariance matrix of the Gaussian density of the parameters updated at time t.

Proof:

Using the linearity of the channel and the Gaussian assumption we know that:

$$\Sigma_{t+1}^{-1} = \Sigma_t^{-1} + \frac{1}{\sigma^2}h(x_t)h'(x_t)$$

then:

$$H_t = \frac{k}{2}(1 + 2\pi) + \ln|\Sigma_t|$$

Similarly:

$$\begin{aligned} H_{t+1} \; = \; & \frac{k}{2}(1 + 2\pi) + \ln|\Sigma_{t+1}| = \\ & \frac{k}{2}(1 + 2\pi) - \ln|\Sigma_t^{-1} + \frac{1}{\sigma^2}h(x_t)h'(x_t)| = \\ & \frac{k}{2}(1 + 2\pi) - \ln|\Sigma_t^{-1}|(1 + \frac{1}{\sigma^2}h'(x_t)\Sigma_t h(x_t)) \end{aligned}$$

where we have used the following identity in the scalar case:

$$|A + FF'| = |A||I + F'A^{-1}F|$$

then it follows:

$$\begin{aligned} H_t - H_{t+1} &= \ln|\Sigma_t| + \ln|\Sigma_t^{-1}|(1 + \frac{1}{\sigma^2}h'(x_t)\Sigma_t h(x_t)) \\ &= \ln(1 + \frac{1}{\sigma^2}h'(x_t)\Sigma_t h(x_t)) \,. \end{aligned}$$

If we consider a control window grater then one step then, using the additivity of the entropy functional we have:

$$\bar{J}_t = -\frac{1}{2}\sum_{i=1}^{N} \ln(1 + \frac{1}{\sigma^2}(h(x_{t+i}))^{'}\Sigma_{t+i}h(x_{t+i})) \,.$$

□

The most informative control $u_t^*\ t = 1, 2, \ldots$ is found solving the following constrained non linear minimization problem:

$$\bar{u}_t^{t+N-1} = \arg\min \bar{J}_t \,, \quad t = 1, 2, \ldots \qquad (9)$$

s.t.

$$\bar{u}_{t+i,t} \in U \,, \quad i = 1, 2, \ldots, N-1 \,, \quad t = 1, 2, \ldots \,.$$

The control law will take on the following form:

$$u_t^* = \gamma(x_t, \Sigma_t) \,.$$

Then, under the given hypotheses, the control law can be found by using constrained mathematical programming techniques.

Remark 1 The assumption of linearity of the channel function respect to the parameters is a strength hypothesis. In the case of a non-linear channel function belonging to the class of $\mathcal{C}^1$ functions, we can give an approximate version of (3) through a linearization of the channel function near the estimate $\hat{\theta}$. This procedure leads to the following:

$$\mathcal{I}(u_t, x_t, \hat{\theta}) = \frac{1}{2}\ln(1 + \frac{1}{\sigma^2}\nabla^{'}h(f(x_t, u_t), \hat{\theta})\Sigma_t\nabla h(f(x_t, u_t), \hat{\theta})$$

Remark 2 The result stated in Theorem 1 has been obtained for $N = 1$. If we would want to consider $N > 1$, two possibilities arise. The first one consists in carrying on calculations similar to that in the Proof of Theorem 1. This would lead to find, at each stage t, a sequence of RH control vectors $\bar{u}_t^{t+N-1}$ corresponding to the solution of the open loop minimization of $\bar{J}_t(u_t^{t+N-1}, p(\theta|I_t))$, though the resulting RH strategy $u_t^* = \bar{u}_{t,t}$ would be a feedback one. At each stage a constrained mathematical programming problem must be solved. The second possibility consists in solving exactly Problem 2, thus searching for sequences of feedback control functions $\bar{\pi}_t^{t+N-1},\ t = 0, 1, \ldots$. At each stage t, a finite horizon stochastic optimal control problem would be faced. Solving it exactly (by means of Dynamic Programming) under general hypotheses is unfeasible. One could resort to approximate techniques such as the Extended Ritz Method (see [8]) or the Neuro Dynamic Programming approach (see [9]).

In the particular case of parameters identification problem where the space of the variables are finite and discrete, we are able to calculate the information gain in a straightforward way, because it reduces simply to a summation. As an example, let us consider the problem of exploring a two-dimensional environment by a DM or a team of DMs. The mapping problem consists in constructing a map of the ground by identifying its obstacle-free parts and the parts occupied by obstacles. In order to model the environment, one could choose a discrete formalization dividing the ground into regular squares or *cells*. In [10] and [11], simulation results show that the approach has been successful.

V. A SIMPLE GEOMETRICAL INTERPRETATION

In this section we show some geometrical interpretations of the results shown in the previous section. For the sake of clarity we consider a very simple system and measurement equation and set $N = 1$. Let the system be:

$$x_{t+1} = u_t\,, \quad t = 0, 1, \ldots$$
$$y_t = x_t'\theta + \eta_t\,, \quad t = 0, 1, \ldots$$

s.t.

$$||u_t|| \leq \delta$$

Accordingly with our results, the control objective is to to solve at every control step the following:

$$u_t^\circ = \arg\max (h(x_{t+1}))'\Sigma_{t+1} h(x_{t+1})$$

or equivalently

$$u_t^\circ = \arg\max ||u_t||_{\Sigma_{t+1}} \tag{10}$$

As Σ_{t+1} is a positive definite matrix, the vector u_t which satisfies (10) will have maximum norm $||u_t|| = \delta$. Moreover the direction of u_t will be the direction of the eigen-vector associated to the maximum eigen-value of the matrix Σ_{t+1}. Observing the recursion equation relative to the covariance matrix we see that this choice causes a reduction of the maximum eigen-value, and then contracts a the ellipsoid associated to Σ_{t+2}. Then the most informative control, causes a contraction of the posterior density function, and reduces the uncertainty about the parameters.

VI. EXPERIMENTAL RESULTS

Let us consider the mass-spring system:

$$m\ddot{x} = Fx - kx - \omega\dot{x}$$

The associated state form is:

$$\dot{x} = \begin{bmatrix} 0 & 1 \\ -\frac{k}{m} & -\frac{\omega}{m} \end{bmatrix} x + \begin{bmatrix} 0 \\ \frac{1}{m} \end{bmatrix} u \tag{11}$$

We have considered the discrete time version of (11) with a sample time $T_s = 0.1$. Let us consider the following measurement equation:

$$y_t = [cN(\alpha_1, \gamma_1), cN(\alpha_2, \gamma_2), \ldots, cN(\alpha_m, \gamma_m)]'\theta + \eta$$

where

$$N(\alpha_i, \gamma_i) = \frac{1}{\sqrt{2\pi\gamma_i^2}} e^{\frac{1}{\gamma_i^2}(x^1 - \alpha_i)^2}\,.$$

The system represents a controlled mass-spring system with a sensor giving information about an object with an increasing precision (high signal-noise ratio) with the alignment of the mass with the object (see Fig. 1). The control purpose is to gain the maximum amount of information in a receding horizon setting about the objects (rapresented with 4 unknown parameters). Figures 2, 3, 4 show the position of the system, the control and the increasing information gain respectively, setting $N = 5, m = 1, k = 1, \omega = 1$, $\alpha_1 = 5$, $\alpha_2 = 2$, $\alpha_3 = -5$, $\alpha_2 = -2, \gamma_1 = \gamma_2 = \gamma_3 = \gamma_4 = 4, \sigma = 1, c = 10$. We have compared our results with the information achieved with several sinusoidal input signals (the structure of the problem suggest the function form to be near to the optimal), in Fig. 4 we have compared the maximum informative input sinusoidal signal with the control sequence according to our approach.

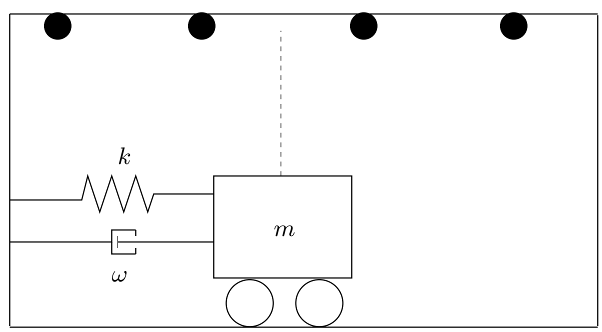

Fig. 1. Mass-spring system

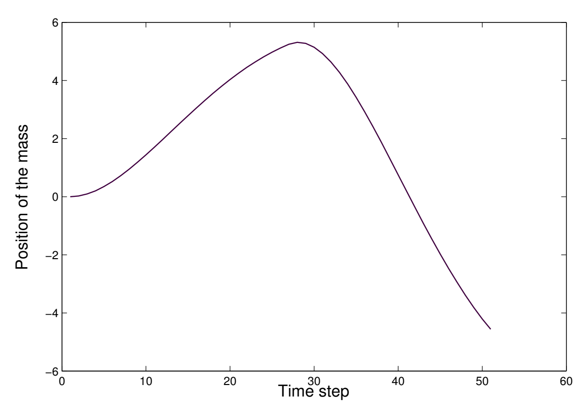

Fig. 2. Evolution of the position of the mass

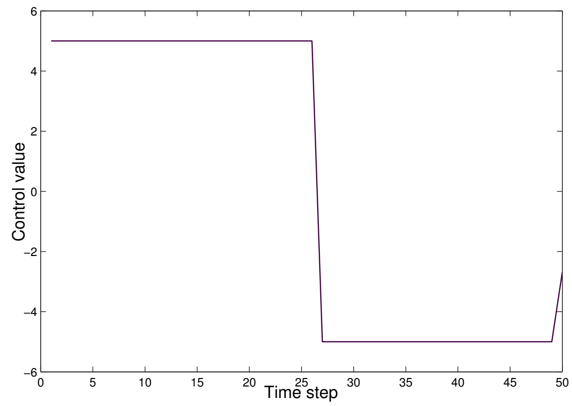

Fig. 3. Evolution of the control values

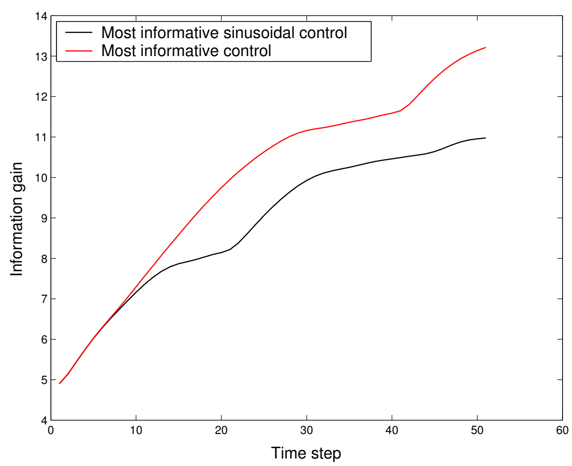

Fig. 4. Information gain comparison

VII. Conclusions

The problem of identifying a set of a priori unknown parameters for discrete time dynamic systems has been faced. To quantify the information gain the Shannon entropy has been proposed, motivated by some of its properties. The identification problem has been formulated in a general setting, where a closed-loop strategy is searched that minimizes a process cost while leading to a "sufficiently large" information gain. The general problem has then been approximated by a receding horizon one, which result to be numerically tractable. In the case when the measurement channel is linear in the parameters, the optimal solution is given involving on-line mathematical programming. The same procedure can be applied in the non linear case as an approximate solution. Numerical results will be given in the final version of the paper in the case of a linear measurement channel as well as in the nonlinear case.

References

[1] V. V. Fedorov, *Theory of optimal experiments*. Academic Press, 1972.

[2] Z. Ghahramani, D. A. Cohn, and M. I. Jordan, "Active learning with statistical models," *J. of Artificial Intelligence Research*, vol. 4, pp. 129–145, 1996.

[3] D. MacKay, "Information-based objective functions for active data selection," *Neural Computation*, vol. 4, pp. 590–604, 1992.

[4] B. Yamauchi, "A frontier based approach for autonomous exploration," in *IEEE International Symposium on Computational Intelligence in Robotics and Automation*, Monterey, CA, 1997, pp. 146–151.

[5] G. N. Saridis, "Entropy formulation of optimal and adaptive control," *IEEE Trans. Automatic Control*, vol. 33, pp. 713–721, 1988.

[6] K. A. Loparo, X. Feng, and Y. Fang, "Optimal state estimation for stochastic systems: an information theoretic approach," *IEEE Trans. Automatic Control*, vol. 42, pp. 771–785, 1997.

[7] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. Wiley, 1991.

[8] R. Zoppoli, M. Sanguineti, and T. Parisini, "Approximating networks and the extended ritz method for the solution of functional optimization problems," *Journal of Optimization Theory and Applications*, vol. 112, pp. 403–439, 2002.

[9] D. P. Bertsekas and J. N. Tsitsiklis, *Neuro-Dynamic Programming*. Athena Scientific, 1996.

[10] M. Baglietto, M. Paolucci, L. Scardovi, and R. Zoppoli, "Information based multi-agent exploration," in *IEEE Third International Workshop on Robot Motion and Control*, Bukowy Dworek, Poland, 2002, pp. 173–179.

[11] ——, "Entropy-based environment exploration and stochastic optimal control," in *42nd IEEE Conference on Decision and Control*, Maui, Hawaii, 2003, pp. 2938–2941.