A Neural Approach to Active Estimation of Nonlinear Systems

M. Baglietto, D. Garassino, L. Scardovi, L. Zanchi, and R. Zoppoli

Abstract— **In this paper, we consider the problem of actively identifying the state of a stochastic dynamic system over a finite horizon. We formalize this Problem as a Stochastic Optimal Control one, in which the minimization of a suitable *uncertainty measure* is performed. To this end, the use of the Renyi Entropy is proposed and motivated. A neural control scheme, based on the application of the Extended Ritz Method and on the use of a Gaussian Sum Filter, is then presented. Simulation results show the effectiveness of the approach.**

I. INTRODUCTION

In this paper, we address the problem of *actively estimating* the state of a stochastic dynamic system over a finite horizon (FH). By *active estimation* we intend the problem of finding a feedback control law that aims at "maximizing the amount of information" on the state of the system. In particular we formalize the problem in the framework of stochastic optimal control, where the cost to be minimized is quantified by a suitable uncertainty measure.

We assume that the classical linear quadratic (LQ) hypotheses are not met. Indeed, if these hypotheses are satisfied, the well known separation principle affirms that the choice of the control law does not affect the estimation process, that is, any control law is "equally informative."

In statistics, a similar problem is the Optimal Experiment Design (OED), in which one has to design an experiment in order to infer about an unknown parameterized system [1]. Also in machine learning a similar problem arises when one can choose the input patterns in the training process (active learning) [2], [3]. In robotics, the problem of environment exploration can be formulated as a particular case of the Active Identification Problem, and it has been studied from an heuristic point of view [4], [5], [6]. In these last years some researchers have used information theoretic concepts to study control problems (see [7], [8]). In [9], [10] the problem of actively identifying a set of unknown parameters in a linear regression setting has been addressed.

In this paper, we formulate the problem in an information theoretic setting by using the *Renyi entropy* as a measure of information about the state of the system. This choice is motivated by the the possibility of deriving a closed-form expression for the amount of information, thus avoiding to resort to computationally demanding nonlinear programming techniques.

As is well known, solving a FH stochastic optimal control problem requires the knowledge of the conditional probability density function $p(x_k|I_k)$ $k = 0, 1, \ldots, N-1$, where x_k is the state vector of the controlled plant and I_k is the information vector consisting of all the measures taken by the controller up to stage k and of all the control actions

performed up to stage $k-1$ (we assume that all stochastic vectors are mutually independent). Dynamic programming could be an effective tool to be applied, at least in principle. This technique, however, entails the recursive computation of the state conditional probability. Unfortunately, such computation can be accomplished analytically only in very few cases, typically, under the classical LQG hypotheses. The conditional probability function is explicitly needed to calculate the measure of uncertainty (which plays the role of the cost function), but no analytical expression will be available in general. This leads us to look for a suitable approximation. To this end, a Gaussian Sum Filtering approach is adopted [11].

The approximating technique adopted in this paper to solve the resulting optimal control problem, consists in assigning a given structure to the control laws. The latter are given a fixed structure, where a fixed number of parameters have to be determined in order to minimize the cost function. Multilayer feedforward neural networks have been chosen, motivated by their good approximating properties. Actually, this family of neural networks is characterized by the ability of approximating nonlinear functions (in our case, the optimal control functions) by using a number of parameters that may be surprisingly smaller than the one required by traditional expansions, like the polynomial and trigonometric ones (this applies to a class of functions to be approximated, characterized by suitable smoothness assumptions). Such a property, proved by Barron [12], should explain the successful experimental results achieved by feedforward neural networks in solving many application problems. Constraining the control functions to take on a fixed structure enables us to reduce the problem of finding the optimal control law (which is a functional optimization problem) to a nonlinear programming one. Such a technique has been used successfully to solve non-LQG deterministic and stochastic optimal control problems (in both finite, infinite, and receding horizon cases (see [13] and references within). In the lines of [14], the technique proposed in this paper to optimize the free parameters relies on the so-called stochastic gradient algorithm (see [15]). Here the possibility of differentiating the expression of the Renyi entropy and that of the Gaussian sum filter plays a central role.

II. PROBLEM FORMULATION

Let us consider a discrete-time stochastic nonlinear system given by

$$x_{k+1} = f_{k+1}(x_k, u_k) + \xi_k \quad k = 0, \dots, N-1 \quad (1a)$$
$$y_k = h_k(x_k) + \eta_k \quad\quad k = 0, \dots, N \quad (1b)$$

where $x_k \in \mathbb{R}^n$, $y_k \in \mathbb{R}^m$ and $u_k \in \mathbb{R}^p$ are the state vector, the measurement vector and the control vector respectively and where $\xi_k \in \mathbb{R}^n$ and $\eta_k \in \mathbb{R}^m$ are two independent white noise processes. The initial state x_0 is known in probability according to the initial density function $p(x_0)$.

Let us define the information vector, by which the Decision Maker (DM) makes its decisions, as $I_k = col(y_0, \dots, y_k, u_0, \dots, u_{k-1})$, $k = 1, \dots, N$; $I_0 = y_0$. Then the DM's control functions take on the form

$$u_k = \gamma_k(I_k), \quad\quad k = 0, 1, \dots, N-1 . \quad (2)$$

While controlling the system, a *process cost* $g_k(x_k, u_k)$ is incurred at any stage k. The final process cost is denoted by $g_N(x_N)$. Our objective is to control the system in N decisional stages in order to gain the "maximum amount of information" on the state vector.

It is possible to formalize the above statement in the framework of a *stochastic optimal control problem*, by adding to the cost function to be minimized a suitable term quantifying the uncertainty in the knowledge of the state vector. We shall call the resulting Problem an *Active Estimation Problem* (AEP).

Problem 1 (AEP): Find a sequence of control functions $\{u_0^\circ = \gamma_0^\circ(I_0), \dots, u_{N-1}^\circ = \gamma_{N-1}^\circ(I_{N-1})\}$ that minimizes the expected value of the cost functional

$$J = \rho \left[\sum_{k=0}^{N-1} g_k(x_k, \gamma_k(I_k)) + g_N(x_N) \right] + (1-\rho)U(I_N) \quad (3)$$

subject to the system equation (1a) and the measurement equation (1b). The scalar $\rho \in [0, 1]$ regulates the trade-off between the process cost and the term $U(I_N)$ that denotes a suitable uncertainty measure penalizing the lack of knowledge on x_N at $k = N$.

To address the problem we are dealing with, two concepts are important: the concept of *information* and the concept of *sufficient statistic*, which will be the subjects of the next section.

III. INFORMATION MEASURES AND SUFFICIENT STATISTIC

In this paper we shall concentrate on two information measures: the *differential Shannon Entropy* and the *differential Renyi entropy* [16]. Let us recall their respective definitions.

Definition 1: The differential Shannon entropy $H(p)$ of a continuous random variable $x \in \mathcal{X}$ is defined by

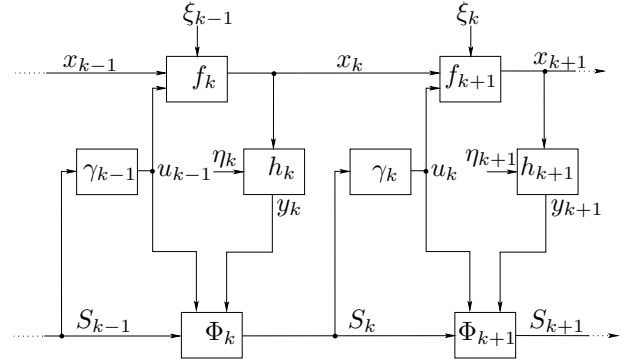$$H(p) \triangleq - \int_{\mathcal{X}} p(x) \log p(x) dx. \quad (4)$$

Fig. 1.   Sufficient statistic propagation according to Bayes rule.

Definition 2: The differential Renyi entropy $H_r(p)$ of a continuous random variable $x \in \mathcal{X}$ is defined by

$$H_r(p) \triangleq \frac{1}{1-r} \log_2 \int_{\mathcal{X}} p^r(x) dx \quad (5)$$

where $r \in \mathbb{R}^+ \setminus \{1\}$.

It can be proven (see for example [16]) that

$$\lim_{r \to 1} H_r(p) = H(p).$$

If we set $r = 2$, then (5) is called *quadratic Renyi entropy*. It is worth noting that the above measures are scalar functionals of the probability density function $p(x)$. We shall address the problem of choosing one among these measures later. Given a stochastic optimal control problem, the information on the state of the system is contained in the conditional probability $p(x_k|I_k)$. It is possible to show that this is indeed a *sufficient statistic* for the problem and its importance is both from the analytical and the conceptual point of view [17]. The conditioned probability $p(x_k|I_k)$ can be generated recursively by the Bayes law and can be viewed as the state of a controlled discrete-time dynamic system,

$$p(x_{k+1}|I_{k+1}) = \Phi(p(x_k|I_k), u_k, y_{k+1}), \quad k = 0, 1, \dots, N-1$$

where $\Phi(\cdot, \cdot, \cdot)$ represents the recursive Bayes updating law

$$\Phi(p(x_k|I_k), u_k, y_{k+1}) = \frac{p(y_{k+1}|x_{k+1})p(x_{k+1}|I_k, u_k)}{p(y_{k+1}|I_k, u_k)} \quad (6)$$

where

$$p(x_{k+1}|I_k, u_k) = \int_{R^n} p(x_{k+1}|x_k, u_k)p(x_k|I_k) dx_k,$$

$$p(y_{k+1}|I_k, u_k) = \int_{R^n} p(y_{k+1}|x_{k+1})p(x_{k+1}|I_k, u_k) dx_{k+1}.$$

In Fig. 1 a block diagram showing the evolution of the conditioned probability density function according to (6) is presented.

Unfortunately an explicit form for (6) is not available in general. An important exception is when the system is linear and the random variables are normal distributed: in this case the conditioned probability is also normal and (6) takes on the form of the well known *Kalman Filter* equations.

Since we are addressing the problem under general assumptions and the conditioned probability function is needed to calculate the measure of uncertainty, we must resort to an approximation of the conditional probability.

A possible approach is to approximate the conditional probability density functions $p(x_k|I_k)$ by means of suitable fixed-structured functions, in which a finite dimensional parameter vector $\hat{S}_k$ have to be fixed in order to approximate the true function. The initial p.d.f. $p(x_0)$ can be approximated by optimizing the corresponding parameter vector $\hat{S}_{-1}$ with respect to some suitable loss function.

An updating rule to propagate over time the parameters characterizing such a representation is then needed. Given the parameters vector $\hat{S}_k$ (characterizing $p(x_k|I_k)$), the applied control u_k and the measure y_{k+1}, such a rule will have the following structure

$$\hat{S}_0 = \hat{\Phi}_0(\hat{S}_{-1}, y_0)$$
$$\hat{S}_{k+1} = \hat{\Phi}_{k+1}(\hat{S}_k, u_k, y_{k+1}), \quad k = 0, \ldots, N-1. \quad (7)$$

The description of one possible technique is presented in the next section.

A. Gaussian Sum Filtering

It is a known result that any probability density function $p(x)$ can be approximated with a Gaussian sum representation

$$\hat{p}(x) = \sum_{i=1}^{z} \alpha_i \mathcal{N}(x - \mu_i, P_i) \quad (8)$$
$$\sum_{i=1}^{z} \alpha_i = 1, \quad \alpha_i \geq 0 \ \ \forall i$$

so that $\hat{p}$ converges uniformly to any $p(x)$ as the number of *basis functions* z increases [11], [18]. By using this parametric representation to approximate $p(x_k|I_k)$, we obtain

$$\hat{p}(x_k, \hat{S}_k) = \sum_{i=1}^{z_k} \alpha_{ki} \mathcal{N}(x_k - \mu_{ki}, P_{ki}) \quad (9)$$

were the *approximate sufficient statistic* is given by $\hat{S}_k = \{\alpha_{ki}, \mu_{ki}, P_{ki}, \ i = 1, \ldots, z_k\}$, $k = 0, \ldots, N$.

It is possible to show that if a probability density function $p(x_k|I_{k-1})$ has a Gaussian sum representation, also $p(x_k|I_k)$ and consequently $p(x_{k+1}|I_k)$ admit the same representation and that the accuracy improves as $P_{ki} \to 0$.

Each Gaussian term in the sum may be propagated independently using an Extended Kalman Filter (EKF) and the result is then normalized by equating the zero moment of each Gaussian distribution [19]. This corresponds to the so called *Gaussian Sum Filter* (GSF).

Let

$$\hat{p}(x_k|I_k) = \sum_{i=1}^{z_k} \alpha_{ki} \mathcal{N}(x_k - \mu_{ki}, P_{ki}) \quad (10)$$
$$\hat{p}(x_{k+1}|I_k) = \sum_{i=1}^{z_{k+1}} \alpha'_{(k+1)i} \mathcal{N}(x_{k+1} - \mu'_{(k+1)i}, P'_{(k+1)i}) \quad (11)$$

be the Gaussian sum representation of the ex-post and ex-ante probability density function respectively and be $p(x_0|I_{-1})$ the approximation of $p(x_0)$ according to (8).

Two cases must be taken in account to obtain (11) from (10). The first one is faced when the system noise covariance Q_k is comparable to that of the terms of the Gaussian sum P_{ki}. Then the prediction step is performed as follows:

$$z_{k+1} = z_k \ ,$$
$$F_{(k+1)i} = \left. \frac{\partial f_{(k+1)}}{\partial x_k} \right|_{\mu_{ki}, u_k} \ ,$$
$$\mu'_{(k+1)i} = f_{(k+1)}(\mu_{ki}, u_k) \ ,$$
$$P'_{(k+1)i} = F_{(k+1)i} P_{ki} F_{(k+1)i}^T + Q_k \ ,$$
$$\alpha'_{(k+1)i} = \alpha_{ki} \ .$$

Following the same approach as in EKF, the innovation step is given by:

$$H_{(k+1)i} = \left. \frac{\partial h_{k+1}}{\partial x_{k+1}} \right|_{\mu'_{(k+1)i}} \ ,$$
$$T_{(k+1)i} = H_{(k+1)i} P'_{(k+1)i} H_{(k+1)i}^T + R_{k+1} \ ,$$
$$K_{(k+1)i} = P'_{(k+1)i} H_{(k+1)i}^T T_{(k+1)i}^{-1} \ ,$$
$$\beta_{(k+1)i} = \mathcal{N}\left(y_{k+1} - h_{k+1}(\mu'_{(k+1)i}), T_{(k+1)i}\right) \ ,$$
$$\mu_{(k+1)i} = \mu'_{(k+1)i} + K_{ki}\left[y_{k+1} - h_k(\mu'_{(k+1)i})\right] \ ,$$
$$P_{(k+1)i} = P'_{(k+1)i} - K_{(k+1)i} H_{(k+1)i} P'_{(k+1)i} \ ,$$
$$\alpha_{(k+1)i} = \frac{\alpha'_{(k+1)i} \beta_{(k+1)i}}{\sum_{h=1}^{z_{k+1}} \alpha'_{(k+1)h} \beta_{(k+1)h}} \ .$$

The second case corresponds to the situation when the covariance Q_k is large if compared to P_{ki}. In this case it may be necessary to introduce a Gaussian sum representation for $p(\xi_k)$ to prevent all approximating Gaussian collapsing into a single term and the GSF reducing to a single EKF [11].

In any case if the elements of the approximating covariance overcome a fixed threshold depending on the initial approximation, it may be necessary to re-approximate the probability density according to (9) [20].

It can be shown that if $z_k \to \infty$, $P_{ki} \to 0$ and $P'_{ki} \to 0$, the Gaussian sum algorithm gives the exact evolution of the a posteriori density function of $p(x_k|I_k)$ (see [20], [11]).

B. Information content of a Gaussian sum

Shannon entropy of a Gaussian mixture presents the limitation to be not calculable in closed form. On the contrary, the Rènyi quadratic entropy can be calculated in closed form as stated by the following result

Theorem 1: Let x be a random variable such that

$$p(x) = \sum_{i=1}^{z} \alpha_i \mathcal{N}(x - \mu_i, P_i)$$

Then the quadratic Renyi entropy $H_2(x)$ can be expressed in closed-form by the expression

$$H_2(x) = -\log_2\left[\alpha^T C \alpha\right]$$

where

$$\alpha = [\alpha_1, \alpha_2, \ldots, \alpha_z]^T$$

and C is the symmetric matrix of elements

$$c_{ji} = \mathcal{N}\left(\mu_j - \mu_i, P_j + P_i\right)$$

Proof: by declaring beforehand the following result valid for the product of two Gaussians

$$\mathcal{N}\left(x - \mu_i, P_i\right)\mathcal{N}\left(x - \mu_j, P_j\right) = c_{ij}\mathcal{N}\left(x - \mu_{ij}, P_{ij}\right)$$

where

$$
\begin{aligned}
c_{ij} = c_{ji} &= \mathcal{N}\left(\mu_i - \mu_j, P_i + P_j\right), \\
\mu_{ij} &= \left(P_i^{-1} + P_j^{-1}\right)^{-1}\left(P_i^{-1}\mu_i + P_j^{-1}\mu_j\right), \\
P_{ij} &= \left(P_i^{-1} + P_j^{-1}\right)^{-1},
\end{aligned}
$$

it is possible to calculate $H_2(p)$ as follows:

$$
\begin{aligned}
H_2(x) &= -\log_2 \int_{R^n} \left[\sum_{i=1}^{z} \alpha_i \mathcal{N}(X - \mu_i, P_i)\right]^2 dX \\
&= -\log_2 \left[\sum_{i=1}^{z}\sum_{j=1}^{z} \alpha_i c_{ij} \alpha_j\right] = -\log_2 \left[\alpha^T C \alpha\right].
\end{aligned}
$$

$\square$

IV. APPROXIMATE PROBLEM SOLUTION

In the previous section we have shown how to obtain an approximate representation of the probability density functions, and, in particular, of $p(\boldsymbol{x}_N | \boldsymbol{I}_N)$ as Gaussian sums and how it is possible to evaluate analytically the corresponding information content. Consequently (3) can be approximated as

$$
\begin{aligned}
\hat{J}\left(\hat{S}_{-1}, x_0, \xi, \eta,\right) &= \\
= \rho \left[\sum_{k=0}^{N-1} g_k(x_k, \bar{\gamma}(\hat{S}_k)) + g_N(x_N)\right] &+ (1 - \rho)H_2(\hat{S}_N)
\end{aligned}
$$

$$(12)$$

where the control laws here take on the form $u_k = \bar{\gamma}_k(\hat{S}_k)$. Even assuming the control laws to depend on the approximate statistic, the AEP is not easy to solve: in fact the control laws are functions to be determined by minimizing the cost (12) as in a functional optimization problem. Following the lines of [13], by exploiting again (as done in approximating $p(x_k | I_k)$) the properties of parameterized structures, we approximate the control functions that are the unknowns in our functional optimization problem, by approximate a fixed structure of the form

$$u_k = \hat{\gamma}(\hat{S}_k, w_k) \tag{13}$$

where w_k is the set of parameters of the chosen representation and $\hat{\gamma}$ is a nonlinear approximator. This technique takes on the name of *Extended Ritz Method* [13].

This allows us to give a more tractable (though approximate) definition for Problem 1. For the sake of notational compactness, let us define

$$\xi \triangleq \{\xi_0, \ldots, \xi_{N-1}\}, \quad \eta \triangleq \{\eta_0, \ldots, \eta_N\},$$

$$w \triangleq col[w_k, \ k = 1, \ldots, N-1].$$

We define an Active Neural Estimation Problem the following

Problem 2 (ANEP): Let $u_k = \hat{\gamma}(\hat{S}_k, w_k)$ and let $\hat{S}_{-1}$ be the parameters of the Gaussian sum representation of $p(x_0)$. Find the sequence of the optimal parameter vectors

$$w^\circ = \left\{w_0^\circ, \ldots, w_k^\circ, \ldots, w_{N-1}^\circ\right\}$$

that minimizes the cost functional

$$\bar{J} = \underset{x_0, \xi, \eta}{\mathrm{E}} \quad \hat{J}\left(\hat{S}_{-1}, x_0, \xi, \eta, w\right) \tag{14}$$

where in (12) the control functions $\bar{\gamma}(\hat{S}_k)$ are substituted by $\hat{\gamma}(\hat{S}_k, w_k)$.

The functional Problem 1 has then been reduced to a *nonlinear programming problem*, which can be solved by the application of the gradient method.

Let h be a generic step of the algorithm. Then the solution of problem (2) may be calculated by applying the following updating rule

$$w^{h+1} = w^h - t^h \nabla_w^h \bar{J}, \quad h = 0, 1, \ldots$$

where $\nabla_w^h \bar{J}$ denotes $\nabla_w \bar{J}|_{w^h}$. The evaluation of (14) and of its gradient are in general difficult to achieve but impossible. The application of the stochastic gradient method [15] allows us to overcome this new problem. The new updating rule becomes

$$w^{h+1} = w^h - t^h \nabla_w^h \hat{J}, \quad h = 0, 1, \ldots$$

where t^h is the step-size of the algorithm and $\nabla_w^h J$ is evaluated at each step in a particular realization of x_0, ξ and η.

For the sake of simplicity, let us consider the process cost depending only on the control vectors, that is

$$J = \rho \ G(u_0, \ldots, u_{N-1}) + (1 - \rho) \ H_2(\hat{S}_N)$$

$$G(u_0, \ldots, u_{N-1}) = \sum_{k=0}^{N-1} g_k(u_k)$$

Finally, let us briefly describe the learning mechanism which includes the *back-propagation* (BP) algorithm through which $\nabla_w^h J$ is evaluated. The training algorithm consists at each training iteration of two main steps that are iterated up to convergence.

In the first one, called *forward pass*, the random variables are generated according to their distribution and the system is made evolve up to the last stage. More specifically, the following steps are performed: 1) the initial state x_0^h and the random vectors ξ^h and η^h are randomly generated; 2)

the optimal neural control vectorsare generated by means of the approximate control functions (13) and the approximate sufficient statistics are propagated up to the last stage N. Then we obtain

$$J_h \triangleq J\left(\hat{S}_{-1}, x_0^h, \xi^h, \eta^h, w^h\right)$$

In the second pass in the iteration, called *backward pass*, the gradient of the cost with respect of the parameter vectors w_k of each net, denoted with $\nabla_{w_k}^h J$, is recursively evaluated from the last stage ($k = N$) to the initial one. The initialization of the algorithm is given by defining the following quantities:

$$\psi_N \triangleq \frac{\partial J_h}{\partial \hat{S}_N} = (1 - \rho)\frac{\partial H_2}{\partial \hat{S}_N} \ ,$$
$$\lambda_N \triangleq \frac{\partial J_h}{\partial x_N} = \rho\frac{\partial G}{\partial x_N} + (1-\rho)\frac{\partial H_2}{\partial x_N} = (1-\rho)\frac{\partial H_2}{\partial x_N}$$
$$= (1-\rho)\frac{\partial H_2}{\partial \hat{S}_N}\frac{\partial \hat{\Phi}_N}{\partial y_N}\frac{\partial h_N}{\partial x_N} = \psi_N \frac{\partial \hat{\Phi}_N}{\partial y_N}\frac{\partial h_N}{\partial x_N} \ .$$

Then, for $k = N - 1, \ldots, 0$, the following quantities are generated recursively:

$$\frac{\partial J_h}{\partial u_k} = \rho\frac{\partial G}{\partial u_k} + (1-\rho)\frac{\partial H_2}{\partial u_k} \ ,$$
$$\frac{\partial H_2}{\partial u_k} = \frac{\partial H_2}{\partial \hat{S}_{k+1}}\frac{\partial \hat{\Phi}_{k+1}}{\partial u_k} + \frac{\partial H_2}{\partial x_{k+1}}\frac{\partial f_{k+1}}{\partial u_k} \ ,$$
$$\frac{\partial J_h}{\partial u_k} = \rho\frac{\partial G}{\partial u_k} + \psi_{k+1}\frac{\partial \hat{\Phi}_{k+1}}{\partial u_k} + \lambda_{k+1}\frac{\partial f_{k+1}}{\partial u_k} \ ,$$
$$\nabla_{w_k}^h J = \frac{\partial J_h}{\partial u_k}\frac{\partial \hat{\gamma}}{\partial w_k}\bigg|_{w_k^h} \ ,$$
$$\left(\frac{\partial J_h}{\partial \hat{S}_k}\right)_\gamma \triangleq \frac{\partial J_h}{\partial u_k}\frac{\partial \hat{\gamma}}{\partial \hat{S}_k} \ ,$$
$$\psi_k = \frac{\partial J_h}{\partial \hat{S}_k} = \psi_{k+1}\frac{\partial \hat{\Phi}_{k+1}}{\partial \hat{S}_k} + \left(\frac{\partial J_h}{\partial \hat{S}_k}\right)_\gamma \ ,$$
$$\lambda_k = \frac{\partial J_h}{\partial x_k} = \frac{\partial J_h}{\partial x_{k+1}}\frac{\partial f_{k+1}}{\partial x_k} + \frac{\partial J_h}{\partial \hat{S}_k}\frac{\partial \hat{\Phi}_k}{\partial y_k}\frac{\partial h_k}{\partial x_k}$$
$$= \lambda_{k+1}\frac{\partial f_{k+1}}{\partial x_k} + \psi_k\frac{\partial \hat{\Phi}_k}{\partial y_k}\frac{\partial h_k}{\partial x_k} \ .$$

The quantities $\partial \hat{\Phi}_{k+1}/\partial \hat{S}_k$, $\partial \hat{\Phi}_k/\partial y_k$, $\partial \hat{\Phi}_{k+1}/\partial u_k$ are obtained by differentiating (7), while $\partial \hat{\gamma}/\partial w_k$ and $\partial \hat{\gamma}/\partial \hat{S}_k$ are evaluated by applying the BP algorithm to each net [13].

V. SIMULATION RESULTS

In this section we present two applications of the AEP. The proposed examples show the effectiveness of the active neural control in a contest generally considered as difficult, that is in the presence of:

- hard nonlinearities;
- non Gaussian random variables;
- non stationary processes.

In the following examples the process cost has been considered as a quadratic form given by

$$G(u_0, \ldots, u_{N-1}) = \sum_{k=0}^{N-1} u_k \Gamma u_k$$

where $\Gamma > 0$ is a matrix introduced to limit the control vectors' module. In the following examples we have chosen one hidden layer neural network made up of 20 neurons as fixed structures.

A. Localization

Let us consider an observer O, who is allowed to move along a circumference of given radius r. The observer task is to determine his angular position, not known with certainty, using a fixed point P as an absolute reference. Figure 2 gives a geometric representation of the problem. The angle β is the angular position of the observer and θ is the measure available to him. The mathematical formulation of
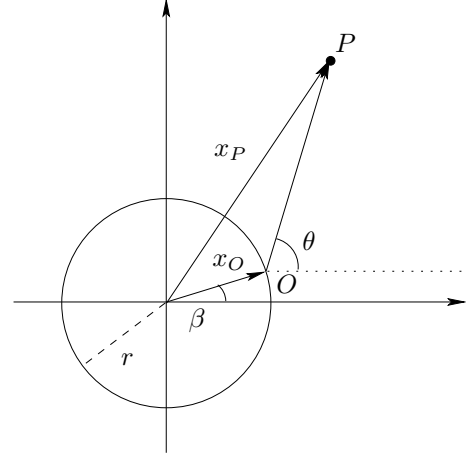

Fig. 2.   Geometric representation of the problem.

the problem is the following:

$$\beta_{k+1} = \beta_k + u_k + \xi_k$$
$$\theta_k = \tan^{-1}\left(\frac{y_P - r\,sen\beta_k}{x_P - r\,cos\beta_k}\right) + \eta_k$$
$$k = 0, 1, \ldots$$

(15)

where x_P and y_P are the cartesian coordinates of P and β_0 is assumed to be uniformly distributed in interval $[0, 2\pi]$. In the training of the neural controllers we considered a time horizon $N = 18$ steps.

The evolution of the Renyi quadratic entropy obtained by the active neural control is represented in Fig. 3 as the average in each step over 20.000 simulation runs and it is compared with those obtained by the application of both random and heuristic controls. Heuristic controls have been chosen according to the idea that the most valuable states to obtain an information gain are those for which we have a large measure variation.
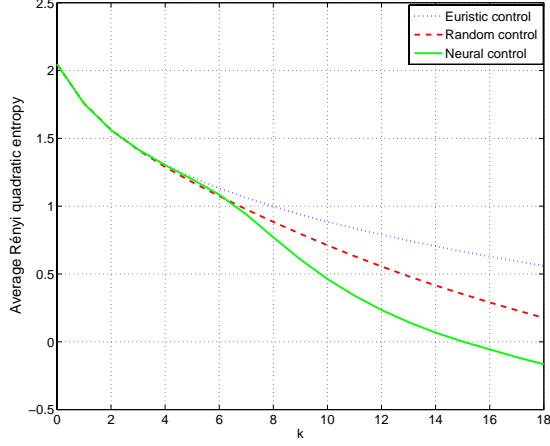
Fig. 3. Entropy mean evolution along N steps

Fig. 5. An example of density evolution.

An example of the trajectories followed by the system is depicted in Figure 4 as a polar plot, where the angle is the state β_k and the radial distance increases with time. The dotted line shows the two angles for which the measure channel presents an extremum.
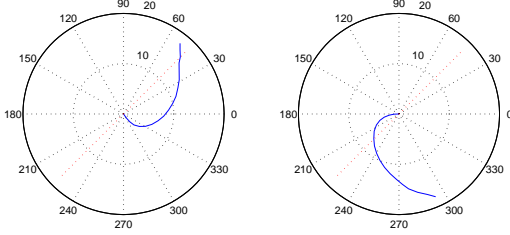

Fig. 4. Two example of trajectories.

An example of evolution of the probability density function is represented in Figure 5. As we can see from a starting uniform probability density the neural control functions succeeded in controlling the system reduceing the uncertainty on the knowledge of the state.
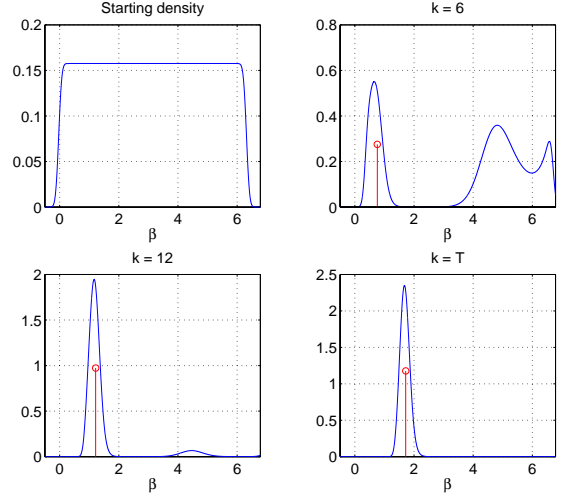
B. Bearing Only Motion Planning (BOMP)
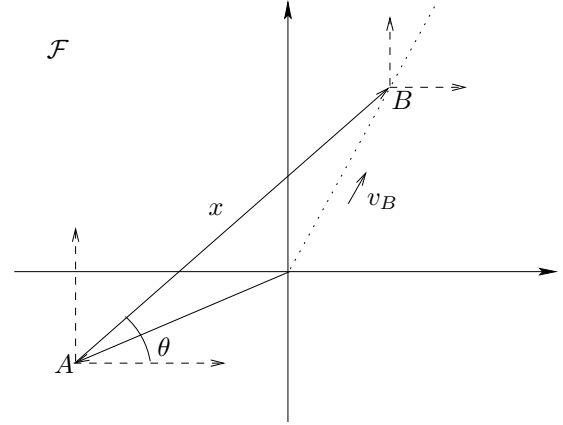
Let us consider two generic geometric points A and B in an absolute reference frame $\mathcal{F}$. Point B, whose absolute position x_B is known with certainty, leaves the origin of $\mathcal{F}$ according to a rectilinear uniform motion denoted by v_B which can be directed anywhere. Point A, instead, whose absolute position x_A is known with uncertainty, has to follow the most appropriate trajectory to establish as well as possible his relative position $x = x_B - x_A$. Such a situation is depicted in Fig. 6: A can only measure the angle θ.

The mathematical formulation of the problem is the

Fig. 6. Geometric representation of the problem.

following:

$$x_{B(k+1)} = x_{Bk} + cv_B$$
$$x_{A(k+1)} = x_{Ak} + u_k - \xi_k$$
$$y_k = \tan^{-1}\left(\frac{x_{k2}}{x_{k1}}\right) + \eta_k$$
$$k = 0, 1, \ldots$$

or, equivalently, by considering the system state as the relative position between the two points $x = x_B - x_A$

$$x_{k+1} = x_k - u_k + cv_B + \xi_k$$
$$y_k = \tan^{-1}\left(\frac{x_{k2}}{x_{k1}}\right) + \eta_k$$
$$k = 0, 1, \ldots$$

where the term $cv_B + \xi_k$ can be regarded as a $WN\left(cv_B, \Sigma_\xi\right)$ acting on the system.

Fig. 7, 8 and 9 are the analogous for the BOMP of Fig. 3, 4 and 5 for the localization and show how our neural control policy is satisfying in terms of information gain. In Fig. 8, in particular are represented for clearness the separate

trajectories followed by the two points instead of the state of the system. Point B is the one which starts from the origin of $\mathcal{F}$. It must be noticed that good behaviours have been
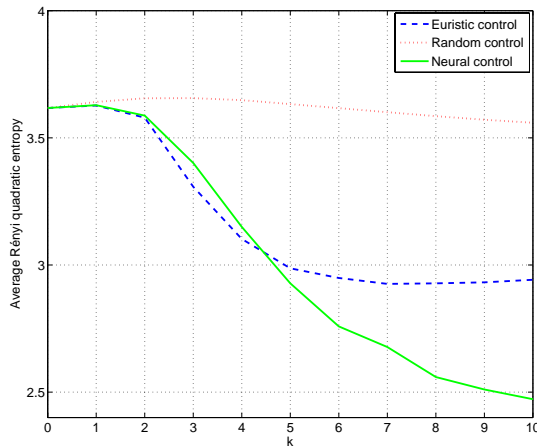

Fig. 7. Entropy mean evolution along N steps

obtained even if B does not move according to the rectilinear uniform motions for which the neural networks has been trained.
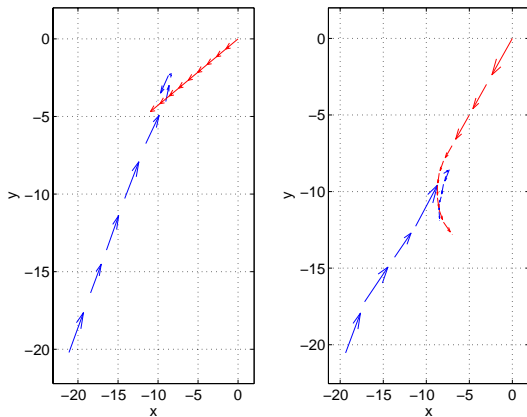

Fig. 8. An example of trajectories.

REFERENCES

[1] V. V. Fedorov, *Theory of optimal experiments*. Academic Press, 1972.
[2] D. MacKay, "Information-based objective functions for active data selection," *Neural Computation*, vol. 4, pp. 590–604, 1992.
[3] 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.
[4] 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.
[5] ——, "Entropy-based environment exploration and stochastic optimal control," in *42nd IEEE Conference on Decision and Control*, Maui, Hawaii, 2003, pp. 2938–2941.
[6] 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.
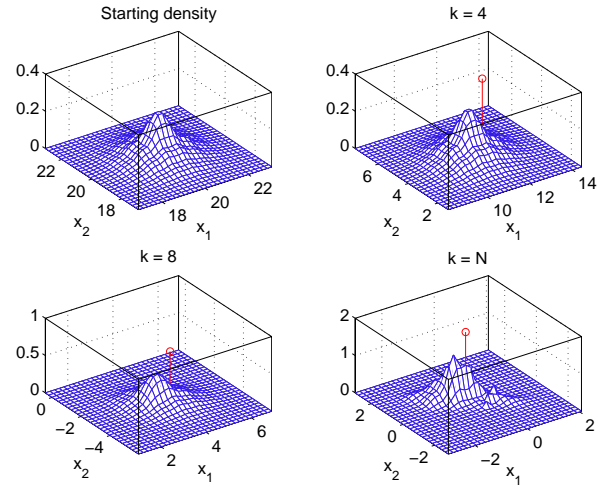
Fig. 9. Example of density evolution.

[7] G. N. Saridis, "Entropy formulation of optimal and adaptive control," *IEEE Trans. Automatic Control*, vol. 33, pp. 713–721, 1988.
[8] 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.
[9] L. Scardovi, M. Baglietto, G. Cannata, and R. Zoppoli, "Active identification of unknown systems," in *16th IFAC World Congress*, Prague, 2005, to appear.
[10] M. Baglietto, L. Scardovi, and R. Zoppoli, "Active identification of unknown systems: an information theoretic approach," in *42nd IEEE American Control Conference*, Boston, 2004, pp. 3826–3830.
[11] D. L. Alspach and H. W. Sorenson, "Nonlinear bayesian estimation using gaussian sum approximation," *IEEE Trans. Automatic Control*, vol. 17, pp. –, 1972.
[12] A. R. Barron, "Universal approximation bounds for superpositions of a sigmoidal function," *IEEE Transactions on Information Theory, vol. 39, no. 3*, 1993.
[13] 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.
[14] T. Parisini and R. Zoppoli, "Neural networks for feedback feedforward nonlinear control systems," *IEEE Trans. on Neural Networks*, vol. 5, pp. 436–449, 1994.
[15] H. J. Kushner and G. G. Yin, *Stochastic Approximation Algorithms and Applications*. Springer–Verlag, New York, 1997.
[16] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. Wiley, 1991.
[17] D. P. Bertsekas, *Dynamic Programming and Optimal Control*. Athena Scientific, 2001.
[18] F. Girosi, "Regularization theory, radial basis functions and networks," in *From Statistics to Neural Networks. Theory and Pattern Recognition Applications*, J. H. Friedman, V. Cherkassky, and H. Wechsler, Eds. Springer-Verlag, Berlin, 1993.
[19] K. Ito and K. Xiong, "Gaussian filters for nonlinear filtering problems," *IEEE Trans. Automatic Control*, vol. 45, pp. –, 2000.
[20] B. Anderson and J. Moore, *Optimal Filtering*. Prentice-Hall, 1991.