

Dual Particle Output Feedback Control based on Lyapunov drifts for nonlinear systems

Emilien Flayac, Karim Dahia, Bruno Hérissé, and Frédéric Jean

Abstract—This paper presents a dual receding horizon output feedback controller for a general non linear stochastic system with imperfect information. The novelty of this controller is that stabilization is treated, inside the optimization problem, as a negative drift constraint on the control that is taken from the theory of stability of Markov chains. The dual effect is then created by maximizing information over the stabilizing controls which makes the global algorithm easier to tune than our previous algorithm. We use a particle filter for state estimation to handle nonlinearities and multimodality. The performance of this method is demonstrated on the challenging problem of terrain aided navigation.

I. INTRODUCTION

Stochastic model predictive control (SMPC) is a widespread technique to deal with control problems where the state is subject to stochastic disturbances, a hard constraint on the input and sometimes soft constraints on the state. Its principle is to solve a finite horizon stochastic optimal control problem and to apply the optimal control in a receding horizon way. Although, when only partial and noisy information on the state is available through some observations, classic SMPC combined with a state estimator may lead to overcautious controls or even destabilizing ones. It is due to the fact that, in general, the control influences the observations and then state estimation. It is called the dual effect property of the control.

Stochastic optimal control problems with imperfect information are much harder than their counterpart in the perfect information case. The main reason is that to be optimal with partial information, one needs to anticipate the information that will be available, in addition to guiding the system in a standard way. Consequently, the optimal control possesses the property of dual effect but as they are computationally intractable, suboptimal outputfeedback control laws are computed instead, with the idea to keep the guiding property and the dual effect property. Consequently, they are called dual controllers.

Usually, dual controllers are either implicit or explicit. In implicit dual control, one tries to approximate the Bellman equation of the problem. Implicit dual control is not treated in this paper. See [1] for a review on implicit dual control and [2] for an example of use of particle filter in implicit dual control. On the contrary, in explicit dual control, one modifies the original problem to incorporate an explicit excitation of the system to maintain the dual effect. Explicit dual effect can be included in the optimization problem as a constraint or in the cost. Including it as a constraint generally leads to controllers with persistent excitation, constraints on information or approximation with scenario trees. Including it in the cost

leads to integrated experiment design, where a measure of information is added to the original cost. In [3], a persistent excitation controller is presented but it is supposed that the system is linearizable which is not a suitable assumption in our framework. Controllers with constraints on the information may lead to infeasibility problems that are also hard to anticipate in nonlinear cases. In [4], this infeasibility problem is addressed but for deterministic dynamics. Scenario tree methods are computationally demanding and usually use Kalman filters inside the optimization problem, like in [5], [6] which are not adapted to multimodal cases.

In integrated experiment design, a new term based on the Fisher information matrix (FIM) is added to the original cost as classically done in optimal design [7]. However, if the original cost to minimize conditions the stability of the system, then the resulting trade off may destabilize the system.

That's why, in this paper, we propose an output feedback dual SMPC based on integrated experiment design for information probing and on a Lyapunov constraint for nonlinear system on the control to ensure stability. Output feedback is obtained by setting the initial condition of the optimization problem as particles from a particle filter that is able to deal with state estimation for arbitrary systems.

The paper is organized as follows. Section II defines some important probabilistic notations. Section III recalls the bases of SMPC with imperfect information and describes our approach. Section IV recalls the principle of particle filtering and section V describes the particular optimization problem that is solved in a receding horizon manner. In [8] and [9], stochastic output feedback stability is proved but a separation principle was assumed. Actually, proving stochastic output feedback stability in the general case, when observability depends on the control, and when the only efficient estimators are particle filters, has never been proved to the best of our knowledge. However, we demonstrate the efficiency of our method on a challenging application in section VI.

II. NOTATION

Let (Ω, \mathcal{F}, P) be a probability space. In the following, random variables refer to \mathcal{F} -measurable functions defined on Ω . For $i \in \mathbb{N}$, $\mathcal{B}(\mathbb{R}^i)$ denotes the set of Borel sets of \mathbb{R}^i . For a random variable X and a probability distribution, $X \sim p$ means that p is the probability law of X . $P(\cdot|\cdot)$ and $E(\cdot|\cdot)$ denotes conditional probability and expectation. P_x and E_x denotes especially probability and expectation conditionally

to $X_0 = x$, for $x \in \mathbb{R}^i$. P_p and E_p denotes probability and expectation conditionally to $X_0 \sim p$.

III. PROBLEM SETUP

A. Description of the system

We consider a discrete-time controlled stochastic dynamical system $(X_k)_{k \in \mathbb{N}}$ valued in \mathbb{R}^{n_x} described by the following equation, $\forall k \in \mathbb{N}$:

$$\begin{aligned} X_{k+1} &= f(X_k, U_k, \xi_k), \\ X_0 &\sim p_0, \end{aligned} \quad (1)$$

where:

- p_0 is a probability density.
- $(U_k)_{k \in \mathbb{N}}$ is the control process valued in \mathbb{R}^{n_u} .
- $(\xi_k)_{k \in \mathbb{N}}$ are i.i.d. random variables valued in \mathbb{R}^{n_ξ} distributed according to p_ξ . For each $k \in \mathbb{N}$, ξ_k represents an external disturbance on the dynamics.
- $f : \mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \times \mathbb{R}^{n_\xi} \rightarrow \mathbb{R}^{n_x}$ is measurable.

We also assume that the state of the system is only available through some observations represented by a stochastic process $(Y_k)_{k \in \mathbb{N}}$ valued in \mathbb{R}^{n_y} which verifies, $\forall k \in \mathbb{N}$:

$$Y_k = h(X_k, \eta_k), \quad (2)$$

where:

- $(\eta_k)_{k \in \mathbb{N}}$ are i.i.d. random variables valued in \mathbb{R}^{n_η} distributed according to p_η . For each $k \in \mathbb{N}$, η_k represents an external disturbance on the observations.
- $h : \mathbb{R}^{n_x} \times \mathbb{R}^{n_\eta} \rightarrow \mathbb{R}^{n_y}$ is measurable.

For $k \in \mathbb{N}$, we define the *information vector* I_k as follow:

$$I_k = (Y_0, U_0, \dots, Y_{k-1}, U_{k-1}, Y_k), \quad (3)$$

or equivalently, for $k \in \mathbb{N}$:

$$\begin{aligned} I_0 &= Y_0, \\ I_{k+1} &= (I_k, U_k, Y_{k+1}). \end{aligned} \quad (4)$$

From I_k , one can derive two important quantities in stochastic control with imperfect information that are the conditional distribution of X_k given I_k denoted by μ_k and the conditional distribution of X_{k+i} given $(I_k, U_k, \dots, U_{k+i-1})$ for any $i \in \mathbb{N}^*$, denoted by $\mu_{k+i|k}$.

Moreover, we denote by K the Markov kernel defined by equation (1) and we assume that the conditional distribution defined by equation (2) has a density with respect to the Lebesgue measure such that there exists a likelihood function denoted by ρ . Therefore, for $k \in \mathbb{N}$, $A \in \mathcal{B}(\mathbb{R}^{n_x})$ and $B \in \mathcal{B}(\mathbb{R}^{n_y})$:

$$\begin{aligned} P(X_{k+1} \in A | X_k = x_k, U_k = u_k) &= K(A, x_k, u_k), \\ P(Y_k \in B | X_k = x_k) &= \int_B \rho(y_k, x_k) dy_k. \end{aligned}$$

μ_0 is supposed to be known. Thus, $\forall (k, i) \in \mathbb{N}^2$, and $\forall A \in \mathcal{B}(\mathbb{R}^{n_x})$, μ_k and $\mu_{k+i|k}$ verify the following nonlinear

filtering equations

$$\mu_{k+1|k}(A) = \int_{\mathbb{R}^{n_x}} K(A, x_k, U_k) \mu_k(dx_k), \quad (5)$$

$$\mu_{k+1|k+1}(A) = \frac{\int_A \rho(Y_{k+1}, x_{k+1}) \mu_{k+1|k}(dx_{k+1})}{\int_{\mathbb{R}^{n_x}} \rho(Y_{k+1}, x_{k+1}) \mu_{k+1|k}(dx_{k+1})}, \quad (6)$$

$$\mu_{k+i+1|k}(A) = \int_{\mathbb{R}^{n_x}} K(A, x_{k+i}, U_{k+i}) \mu_{k+i|k}(dx_{k+i}). \quad (7)$$

One can sum up equations (5) to (7), in the following way, $\forall (k, i) \in \mathbb{N}^2$:

$$\mu_{k+1} = F(\mu_k, Y_{k+1}, U_k), \quad (8)$$

$$\mu_{k+i+1|k} = G(\mu_{k+i|k}, U_{k+i}), \quad (9)$$

A control policy at time k is then a map, denoted by π_k , that maps a conditional distribution μ_k to a control U_k . We denote a sequence of control policies by:

$$\pi_{i:j} := (\pi_i, \dots, \pi_j), \text{ for } i \leq j.$$

B. Dual stochastic model predictive control

Stochastic model predictive control is a widely used method for the design of a controller in the presence of possibly unbounded disturbances in nonlinear dynamics as the one described in equation (1). It consists in solving a finite-horizon discrete stochastic optimal problem, to only apply the first control policy of the optimal sequence and to solve the problem again starting from the new state of the system. See [10] for a review on general stochastic MPC.

A MPC scheme is generally defined by the following features:

- A time horizon, $T \in \mathbb{N}^*$
- A family of set of constraints on the control, $U_i \subset \mathbb{R}^{n_u}$, $\forall i = 0, \dots, T-1$.
- $\forall i = 0, \dots, T-1$, an instantaneous cost $g_i: \mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \times \mathbb{R}^{n_\xi} \rightarrow \mathbb{R}$ and a final cost $g_T: \mathbb{R}^{n_x} \rightarrow \mathbb{R}$.

The choice of the objective functions g_i and g_T and of the control constraints U_i is a matter of modelling. Indeed, in receding horizon control, one often has an economic cost to minimize that comes from practical considerations, like a price or a fuel consumption, for example. Moreover, the control must be designed to attain some target in the state space, it is the guiding problem. There are two classical ways to address this issue in MPC:

- Adding a new term to the cost to enforce stability in some sens. In this case, the general cost is decomposed in the following way, for $i = 0, \dots, T$:

$$g_i = g_i^{\text{stab}} + g_i^{\text{eco}}, \quad (10)$$

Thus, g_i and g_T realize a compromise between convergence and economic costs. For instance, in the LQG case, $g_i(x, u, \xi) = x^T M_x x + u^T M_u u$ where M_x and M_u are positive definite matrices. The first term drives the state of the system to zero and the second term penalizes high controls, that can be expensive. The compromise is dealt with by tuning the matrices M_x

and M_u .

- Adding a drift constraint on the first control, U_0 , that enforces decreasing of some Lyapunov-like function, during the first time step only, such that, for $i = 0, \dots, T$:

$$g_i = g_i^{\text{eco}}, \quad (11)$$

a negative drift condition on U_0 .

Actually, since only U_0 is applied on the system, the Lyapunov function decreases along the whole trajectory and then stability is obtained. It is also known as Lyapunov Economic MPC, see [11] for a review in the deterministic setting. In the stochastic setting, it has been applied with output feedback for continuous-time nonlinear systems in [9] and for a discrete-time linear system with bounded controls in [8].

In the presence of partial information represented by equation (2), the theoretical receding horizon problem to solve is more complex than in the perfect information case, as the new state of the system is μ_k [12], and it evolves according to equation (8). Thus, the problem to solve at each time k denoted by (P_C^k) can generally be written as follow:

$$\begin{aligned} \min_{\pi_{0:T-1}} \quad & E_{\mu_k} \left[\sum_{i=0}^{T-1} E[g_i(X_i, U_i, \xi_i) | I_i] + E[g_T(X_T) | I_T] \right] \\ \text{s.t.} \quad & \tilde{\mu}_{i+1} = F(\tilde{\mu}_i, Y_{i+1}, U_i), \\ & U_i = \pi_i(\tilde{\mu}_i), \\ & U_i \in \mathbb{U}_i, \forall i = 0, \dots, T-1, \\ & \tilde{\mu}_0 = \mu_k. \end{aligned}$$

In (P_C^k) , we make appear explicitly the conditional expectation w.r.t. the available information I_i to emphasize the fact that the information vectors, and notably the observations, are propagated from k to $k+T$ and that they must depend on the control. Therefore, as recalled in [1] and [13], one can see from the Dynamic Programming Principle that the optimal control policies of (P_C^k) generally influence the future observations. The control is then said to have the *dual effect* property [14]. As optimal policies are intractable in the general nonlinear case due to the curse of dimensionality, suboptimal policies are constructed instead. To mimic the properties of the optimal policies, suboptimal policies are designed to have two roles:

- controlling the system in the standard way defined by equation (10) or (11).
- actively probing information to improve the quality of the observations i.e. keeping the dual effect property.

Such policies are also called dual controllers. In this article, we focus on explicit dual controllers that take information into account inside a new cost. This technique is called integrated experiment design. In fact, the new cost denoted by g_i^{ex} realizes a trade off between the original costs g_i and g_T and a measure of information denoted by g_i^{info} such that, for $i = 0, \dots, T$:

$$g_i^{\text{ex}} = g_i + g_i^{\text{info}}, \quad (12)$$

Integrated experiment design is, a priori, adapted to our

framework. However, an important issue appears, when, additionally, the guiding objective is incorporated in the cost as described in equation (10). Consequently, (12) represents the sum of three terms that usually contain weights that must be tuned. By construction, the weights affect the convergence of the system to the guiding goal which makes their tuning difficult. Indeed, if too much importance is given to g_i^{info} then probing will be efficient but the system may get stuck far from the target. On the contrary, if too much importance is given to g_i^{stab} then the probing effect will not be sufficient and output feedback performance may be poor. It is difficult to know which case will occur a priori depending of the value of the weights because the optimal costs are not known at first sight.

That's why, in this paper, we present a new output-feedback explicit dual controller based on the minimization of the Fisher information matrix subject to a negative mean drift condition for nonlinear discrete-time systems coupled with a particle filter for state estimation. In other words, we have chosen to model our problem with equations (11) and (12).

Furthermore, in explicit dual control, one does not have to propagate the information inside the dynamics with equation (8) anymore because it is dealt with in a different way. It is sufficient to propagate $\mu_{i|0}$ with equation (9) which is much simpler. Besides, one wants to solve in practice a finite dimensional optimization problem so one looks for control values and not policies anymore. Then, the general optimization problem associated with our method, denoted by (P_{ex}^k) , can be written as follows:

$$\begin{aligned} \min_{u_{0:T-1}} \quad & E_{\mu_k} \left[E[\sum_{i=0}^{T-1} g_i^{\text{ex}}(X_i, u_i, \xi_i) + g_T^{\text{ex}}(X_T) | I_0] \right] \\ \text{s.t.} \quad & \tilde{\mu}_{i+1|0} = G(\tilde{\mu}_{i|0}, u_i), \\ & u_i \in \mathbb{U}_i, \forall i = 0, \dots, T-1, \\ & \tilde{\mu}_0 = \mu_k. \\ & \text{negative drift condition on } u_0, \end{aligned}$$

IV. STATE ESTIMATION

From the definition of the (P_C^k) and (P_{ex}^k) , it is clear that dual SMPC requires the computation of μ_k . However, in the general nonlinear case, μ_k cannot be computed explicitly so approximations are needed. Kalman filters are widespread and easy-to-compute approximations of the posterior distribution but they may fail in the presence of high nonlinearities and multimodality. That's why we use particle filters that are known to handle these difficulties at the price of a higher computational cost. A particle filter approximates the posterior distribution μ_k by a set of N particles, $(x_k^l)_{l=1, \dots, N}$ valued in \mathbb{R}^{n_x} , associated with nonnegative and normalized weights $(\omega_k^l)_{l=1, \dots, N}$. The approximate distribution, denoted by, μ_k^N , is then defined as follow:

$$\mu_k^N = \sum_{l=1}^N \omega_k^l \delta_{x_k^l}, \quad (13)$$

Like a Kalman filter, a particle filter can be computed recursively following two steps: prediction and correction. During the prediction step, the particles are propagated using an importance distribution that is often chosen as the Markov kernel from the dynamics, K . During the correction step, the weights are updated thanks to the last observation and the particles are resampled from the updated weights. Actually, resampling is not needed at each time step and criteria have been developed to decide whether or not resampling is needed. These steps defines a particle filter with adaptive resampling and are summed up in Algorithm 1.

Furthermore, we define the conditional expectation of the state, denoted by \hat{x}_k w.r.t. the information I_k and its particle approximation, the empirical mean of the filter \hat{x}_k^N i.e:

$$\hat{x}_k = E[X_k | I_k],$$

$$\hat{x}_k^N = \frac{1}{N} \sum_{l=1}^N \omega_k^l x_k^l.$$

These will be useful in the definition of the finite horizon optimization problem.

Algorithm 1 Particle filter with adaptive resampling

- 1: Create a sample of N particles x_0^l according to the law μ_0 and initialize the weights $\omega_0^{(l)}$ with $\frac{1}{N}$
- 2: **for** $i = 0, 1, 2 \dots$ **do**
- 3: **Prediction:**
- 4: Given a control u_i and a set of particles $(x_i^l)_{l=1, \dots, N}$, compute the predicted particles by drawing samples from K i.e.

$$x_{i+1|i}^l \sim K(dx_{i+1|i}, x_i^l, u_i), \text{ for } l = 1, \dots, N.$$

- 5: **Correction:**
- 6: Get the new observation Y_{i+1}
- 7: Compute the updated weights $(\tilde{\omega}_i^l)_{l=1, \dots, N}$ thanks to the likelihood function ρ :

$$\tilde{\omega}_i^l = \omega_i^l * \rho(Y_{i+1}, x_{i+1|i}^l)$$

- 8: **if Resampling then**
 - 9: Draw the *a posteriori* particles $(x_{i+1}^l)_{l=1, \dots, N}$ from the set $(x_{i+1|i}^l)_{l=1, \dots, N}$ and $(\tilde{\omega}_i^l)_{l=1, \dots, N}$ using a resampling technique and set $\omega_{i+1}^l = \frac{1}{N}$
 - 10: **else**
 - 11: Set $x_{i+1}^l = x_{i+1|i}^l$ and $\omega_{i+1}^l = \tilde{\omega}_i^l / \sum_{l=1}^N \tilde{\omega}_i^l$
 - 12: **end if**
 - 13: **end for**
-

V. CONTROL POLICY

The goal of this section is to present our new explicit dual receding horizon control scheme. The main idea is to prioritize the guiding goal by adding a stabilizing constraint on the first control of the finite horizon optimization problem. This constraint consists in forcing the decreasing of a Lyapunov-like function during the first time step starting from the current estimator the state. It is equivalent to

stabilizing an estimator of the state. The success of such a technique depends highly on the estimation error. Indeed, if the estimation error is high or even diverges, driving the estimator to the target does not imply driving the true state to the target. It is a well known problem in the deterministic setting [15]. In [9] and [8], output feedback stability is proved but it was assumed that the estimation error could controlled a priori uniformly on the control. This assumption is not true anymore in dual control because the observations, and by consequence the estimation error, do depend on the control and can make the estimator diverge. To solve this problem, in this paper, we have chosen to obtain the lowest estimation error w.r.t. the Fisher information matrix over the stabilizing controls.

A. Fisher information matrix

Intuitively, in our context, the Fisher information matrix at time k , denoted by J_k , quantifies the information about the state X_k that is contained in the vector of observations (Y_0, \dots, Y_k) . More formally, its inverse, the Cramer-Rao bound is a lower bound on the covariance of the estimation error for any unbiased estimator. Thus, the FIM is often used as an empirical measure of information that is to be maximized. It is very popular in Optimal Design [7] and is also classic in explicit dual control. It was shown in [16] that J_k could be computed recursively which is very interesting in practice. For the sake of simplicity of the recursive formulation of the FIM, we assume that the disturbances are additive and gaussian in the dynamics and the observation equations i.e. $\forall k \in \mathbb{N}$:

$$f(X_k, U_k, \xi_k) = f^{det}(X_k, U_k) + \xi_k,$$

$$h(X_k, \eta_k) = h^{det}(X_k) + \eta_k,$$

$$p_\xi = \mathcal{N}(0, Q),$$

$$p_\eta = \mathcal{N}(0, R),$$

where Q and R are covariance matrices, f^{det} and h^{det} are differentiable w.r.t. the state. Thus, the FIM satisfies the following recursive equation, $\forall k \in \mathbb{N}^*$:

$$J_0 = P_0^{-1},$$

$$J_{k-1}^+ = J_{k-1} + E [F_{k-1}^T Q^{-1} F_{k-1}], \quad (14)$$

$$J_k = E [H_k R_k^{-1} H_k^T] -$$

$$Q^{-1} E [F_{k-1}] (J_{k-1}^+)^{-1} E [F_{k-1}]^T Q^{-1},$$

where P_0 is the covariance matrix associated with the initial distribution p_0 , $F_{k-1} = \nabla_x f^{det}(X_{k-1}, U_{k-1})$ and $H_k = \nabla h^{det}(X_k)$.

B. Foster-Lyapunov drift in case of perfect information

Perfect information is met when full knowledge of the state X_k is available. In that case, the control is computed from state feedback control policies i.e. measurable maps, denoted by α_k , that maps a state X_k to a control U_k . Since equation (1) is time homogeneous, it is sufficient to consider a fixed control policy α . Thus, for any, α , one can define the

corresponding closed loop system as follows, $\forall k \in \mathbb{N}$:

$$\begin{aligned} X_{k+1} &= f(X_k, \alpha(X_k), \xi_k), \\ X_0 &\sim p_0. \end{aligned} \quad (15)$$

A state feedback control policy is said to be *admissible* if $\forall x \in \mathbb{R}^n, \alpha(x) \in \mathbb{U}_0$. Therefore, equation (15) defines also a time-homogeneous Markov chain whose stability can be studied via the classical theory of negative drifts conditions discussed in [17] and recalled in [18]. In proposition 1, we focus on *geometric drifts conditions* that are closely related to Lyapunov conditions for exponential stability for continuous-time processes.

Proposition 1: Suppose that there exist $b > 0$ and $\lambda_{min} \in [0, 1[$, a measurable function $V : \mathbb{R}^{n_x} \rightarrow [0, +\infty[$, a compact set $C \subset \mathbb{R}^{n_x}$ and an admissible state feedback control policy α such that $E[V(f(x, \alpha(x), \xi_0))] \leq \lambda_{min}V(x)$, $\forall x \notin C$ and $\sup_{x \in C} E_x[V(X_1)] = b$

Then, $\forall \lambda \in [\lambda_{min}, 1[$,

$$E_x[V(X_k)] \leq \lambda^k V(x) + b(1 - \lambda)^{-1}, \forall k \in \mathbb{N}, \forall x \in \mathbb{R}^n,$$

where X_k is computed with equation (15).

Proposition 1 is a slightly different reformulation of proposition 1 in [18], but its proof follows from the one in [18]. In particular, we consider that the parameter λ can be chosen arbitrarily in $[\lambda_{min}, 1[$. In practice, λ is a parameter to tune that determines the convergence speed of the system. Moreover, as explained in [18], if proposition 1 is verified for a norm-like function V then for $r > 0$, $P_x(\|X_k\| > r)$ decreases as the inverse of V so the distribution of state concentrates itself around 0. That's why, in the rest of the paper, we suppose that our guiding goal is to drive the system (1) to 0. To do that, we also suppose that the assumptions in proposition 1 are fulfilled and notably that the system (1) can be stabilized with perfect information with some admissible state feedback control policy α .

C. Receding horizon policy

The purpose of our MPC scheme is to compute control that explicitly look for information by minimizing some functions of the FIM, g_i^{fish} and g_T^{fish} , over the controls that stabilize a state estimator. This is guaranteed by the drift condition taken from proposition 1 and, applied to \hat{x}_k only when $\hat{x}_k \notin C$. Thus, the stochastic optimization problem to solve, denoted by, (P_{CFL}^k) , can be written as follow, $\forall \lambda \in [\lambda_{min}, 1[$:

$$\begin{aligned} \min_{u_0 \dots u_{T-1}} E_{\mu_k} \left[E[\sum_{i=0}^{T-1} g_i^{\text{ex}}(X_i, u_i, \xi_i) + g_T^{\text{ex}}(X_T) | I_0] \right] \\ \text{s.t.} \quad \begin{aligned} \tilde{\mu}_{i+1|0} &= G(\tilde{\mu}_{i|0}, u_i), \\ u_i &\in \mathbb{U}_i, \forall i = 0, \dots, T-1, \\ \tilde{\mu}_0 &= \mu_k, \\ E_{\hat{x}_k}[V(f(\hat{x}_k, u_0, \xi_0))] &\leq \lambda V(\hat{x}_k), \text{ when } \hat{x}_k \notin C \end{aligned} \end{aligned}$$

where:

$$g_i^{\text{ex}}(X_k, U_k, \xi_k) = g_i^{\text{eco}}(X_k, U_k, \xi_k) + g_i^{\text{fish}}(J_k), \quad (16)$$

$$g_T^{\text{ex}}(X_T) = g_T^{\text{eco}}(X_T) + g_T^{\text{fish}}(J_T), \quad (17)$$

It is important to notice that the admissibility of the drift

constraint is guaranteed by the existence of the admissible state feedback stabilizing α .

A good approximation technique to approach (P_{CFL}^k) in practice is the scenario approach [19]. It appears that, in output feedback MPC, there is a synergy between the scenario approach and particle filtering. In fact, the initial condition for each independent scenario is chosen as a particle from the current set of particles. It improves global performance compared to a similar technique involving a Kalman filter in which, the initials conditions are always drawn according to a unimodal law. This method was already used in [20] and [13]. The approximation, denoted by $(P_{CFL}^{k,N})$, can be defined, $\forall \lambda \in [\lambda_{min}, 1[$, by:

$$\begin{aligned} \min_{u_0 \dots u_{T-1}} \sum_{l=1}^{N_s} \omega_k^l \left(\sum_{i=0}^{T-1} g_i^{\text{ex}}(X_i^l, u_i, \xi_i^l) + g_T^{\text{ex}}(X_T^l) \right) \\ \text{s.t.} \quad \begin{aligned} X_{i+1}^l &= f(X_i^l, u_i, \xi_i^l), \\ u_i &\in \mathbb{U}_i, \\ X_0^l &= x_k^l, \forall l = 1, \dots, N_s, \forall i = 0, \dots, T-1, \end{aligned} \\ \frac{1}{N_{dr}} \sum_{\ell=1}^{N_{dr}} V(f(\hat{x}_k^N, u_0, \xi_0^\ell)) \leq \lambda V(\hat{x}_k^N), \text{ when } \hat{x}_k^N \notin C, \end{aligned}$$

where:

- $N_s < N$ is the number of scenarios considered. It is supposed to be less than the number of particles for computational reasons, so N_s particles must be extracted from the original set.
- $N_{dr} \in \mathbb{N}^*$ is the size of the sample used to approximate the expectation in the drift constraint.
- $(\xi_i^l)_{i=0, \dots, T-1, l=1, \dots, N_s}$ and $(\xi_0^\ell)_{\ell=1, \dots, N_{dr}}$ are i.i.d. random variables sampled from p_ξ .

Additionally, when $\hat{x}_k^N \in C$ the drift condition is not necessarily feasible so we decided to apply the stabilizing policy α to \hat{x}_k^N , such that $U_k = \alpha(\hat{x}_k^N)$, as it is done in classical certainty equivalence controllers. This means that if \hat{x}_k^N enters C there is no probing anymore. It is not a problem as one can see C as the target set.

D. Global algorithm

The complete Output feedback algorithm is summarized in Algorithm 2. This method has two main advantages compared to the one we presented in [13]:

- First, In [13], both the stability properties and the information were incorporated as two terms of the cost. Therefore, the weights between the terms were difficult to tune and, especially in receding horizon control where they deeply influence the convergence of the system. In particular, it was compulsory to decrease the weights on g_i^{fish} and g_T^{fish} with time otherwise the system converged to a point that was far from the target. Moreover, tuning the decreasing of this weight was also complicated a priori. In our new method, the stability properties are much less influenced by the cost because of the drift condition which is a constraint. The most important parameter to tune is λ , and, in principle, it

influences only the convergence speed of the system and not its qualitative properties of stability.

- Secondly, from a numerical point of view, as stated in [9], the stability properties are contained in one constraint and are easier to achieve in practice. Indeed, the convergence of the system depends little on the quality of the solution of the optimization problem and much more on its admissibility which is easy to obtain with classical solvers.

Algorithm 2 Fisher/Lyapunov Output Feedback Control

- 1: Create a sample of N particles x_0^l according to the law μ_0 and initialize the weights ω_0^l with $\frac{1}{N}$.
 - 2: **for** $k = 0, 1, 2, \dots$ **do**
 - 3: **if** $\hat{x}_k^N \notin C$ **then**
 - 4: Solve (\tilde{P}_{CF}^k) starting from the particles x_k^l and the weights ω_k^l .
 - 5: Get an optimal sequence $(u_0^*, \dots, u_{T-1}^*)$.
 - 6: Set $U_k = u_0^*$.
 - 7: **else**
 - 8: Set $U_k = \alpha(\hat{x}_k^N)$
 - 9: **end if**
 - 10: Compute the *a posteriori* particles x_{k+1}^l and weights ω_{k+1}^l given U_k according to Algorithm 1
 - 11: **end for**
-

VI. APPLICATION AND NUMERICAL RESULTS

A. Description of the application

Algorithm 2 has been applied to the guidance and localization of a drone by terrain-based navigation. Our first goal is to drive a drone in a 3D space from an uncertain initial condition X_0 to a compact set centered around 0. If the original target is not 0 then a translation can be made to center the problem around 0. In the Cartesian coordinates, we assume that the dynamics of the drone can be described as follow:

- the state is composed of 3 positions and 3 speeds: $X_k = (x_k, y_k, z_k, v_k^x, v_k^y, v_k^z)^T$ and the control of 3 accelerations $U_k = (u_k^x, u_k^y, u_k^z)^T \cdot (x_k, y_k)$ represents a horizontal position and z_k an altitude.
- its dynamics (1) is linear with bounded controls such that, for $k \in \mathbb{N}$:

$$X_{k+1} = AX_k + BU_k + \xi_k, \quad (18)$$

$$\|U_k\| \leq U_{max}, \quad (19)$$

where $U_{max} > 0$ and $A \in \mathbb{R}^{n_x \times n_x}$ and $B \in \mathbb{R}^{n_x \times n_u}$ correspond to a discrete-time second order system with damping on the speed. In this setting, classical linear controllers do not lead to admissible control almost surely so a nonlinear feedback controller must be designed.

As described in section V, the guiding problem is addressed with a drift constraint for a particular drift that depends on the dynamics. We have used the drift proposed

in proposition 8 of [18] such that $V(X) = e^{\|X\|}$ and C is a ball centered around 0 whose radius depends only on the disturbances of the dynamics To guarantee the existence of an admissible stabilizing state feedback policy, α , the maximum control U_{max} must be sufficiently high. See [18] for the precise definition of α , U_{max} and C . Although it applies only to orthonormal dynamics, we suppose that A is *neutrally stable* in the sens that, after a change of coordinates, it can be decomposed in a Shur-stable block and an orthonormal block. The drift is then considered only on the last block as the first one already ensures stability.

We assume that the dynamics has a relatively simple form because the main difficulty of this application is the nature of the observations. Indeed, we suppose that the speed is a measure of the difference between the altitude of the drone, z_k , and the altitude of the corresponding vertical point on the ground. We also suppose that the ground is represented by a map, $h^{map} : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ that maps a horizontal position (x, y) to the corresponding height of the terrain. In practice, h^{map} is determined by a smooth interpolation of data points so it is highly nonlinear. Therefore, in this case, $h^{det}(X_k) = (z_k - h^{map}(x_k, y_k), v_k^x, v_k^y, v_k^z)^T$. Thus our second goal is to estimate the state of the system. It is not very restrictive to suppose that the whole speed vector is measured because the speed on its own is not enough to reconstruct the whole state so the altitude measure is still required. The use of particle filters is totally justified by the particular definition of h^{det} . Actually, the conditional distribution, μ_k is multimodal and its modes are closely related to the level set of the map of the terrain, h^{map} so Kalman filters cannot accurately deal with this problem.

Moreover, it appears very naturally that dual control is required in this application. Indeed, the quality of the observations depends on the control and more precisely on the area that is flied over by the drone. Let us assume that the drone flies over a flat area with constant altitude then one measurement of height matches a whole horizontal area and the estimation error on (x_k, y_k) is of the order of magnitude of the size of the area which can be very large. On the contrary, if the drone flies over a rough terrain, then one measurement of height correspond to a smaller area on the ground and the estimation error is reduced. Our last goal is then to probe information by avoiding to fly over uninformative areas on the ground. It is dealt with by selecting the controls that minimizes some loss function of the FIM inspired from Optimal Design, among the controls that verifies the drift constraint.

B. Numerical results

Figure 1 represents the horizontal projection of a trajectory obtained by algorithm 2 with the terrain map in its background. The scenario has been chosen such that, if the system goes in straight line to the target then it flies over a flat area and then the estimation error will be high. We can see that in figure 1 that, as expected, the system makes a detour to avoid the flat area, so that the particles tighten

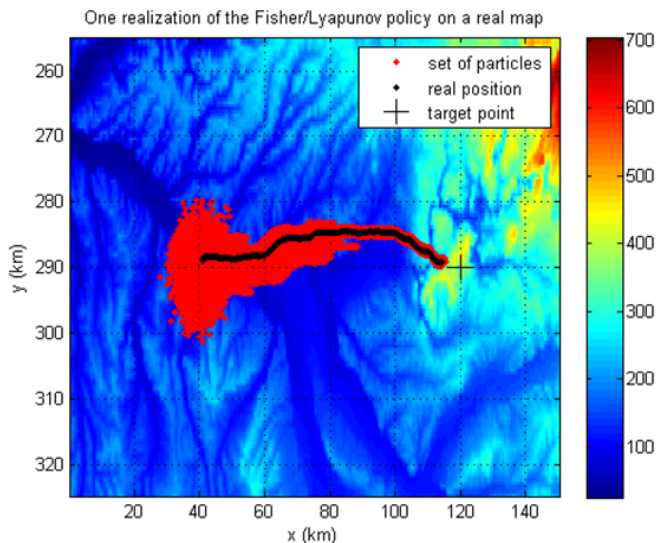


Fig. 1. Plot of one trajectory obtained by Fisher/Lyapunov particle control and of the particles from the particle filter

around the true position, and finally reaches the target. In this run, the algorithm has been stopped when the target enters the compact C .

Figure 2 and 3 shows RMSE respectively in horizontal position and altitude after 30 Monte Carlo simulations of three different policies: a policy with no information probing, the policy described in this article and the policy described in [13] where stability was forced by the minimization of the distance to the target. The setup in figure 1 is less favourable for estimation than the one in figures 2 and 3 that's why the error diminishes faster in figures 2 and 3 than what appears in figure 1.

It is clear that the horizontal estimation error is reduced with information probing. It is actually due to the fact that the filter diverges in 30% of the cases without the FIM and 15% with the FIM. Our new method also seems slightly better than the previous one. Moreover, in [13], the crucial parameters that regulated the convergence of the system were actually the weights on the FIM and on the distance to the target. In particular, the drone could get stuck over a mountain if the weight on the FIM was too high because mountains are very informative in this application. Therefore, it was compulsory to make this weight decrease over time or tune it dynamically to give more importance to the distance to the target and force the system to go to it eventually. We claim that this kind of complicated tuning is not necessary with our new method because the system cannot get stuck due to the drift constraint that compel it to go toward the target set at each time step. We only choose the convergence *speed* of the drone through the parameter λ of proposition 1 and, in practice, it affects only marginally the qualitative property of stability.

The simulations were run in MATLAB and the optimization problems were solved using the modelling language AMPL and the solver Ipopt. Besides, compared to [13],

thanks to user-defined functions in AMPL and a library in C for interpolation, we have managed to deal efficiently with the interpolation of a real map in AMPL that is necessary for the computation of the FIM inside the optimization problem ($P_{CFL}^{k,N}$).

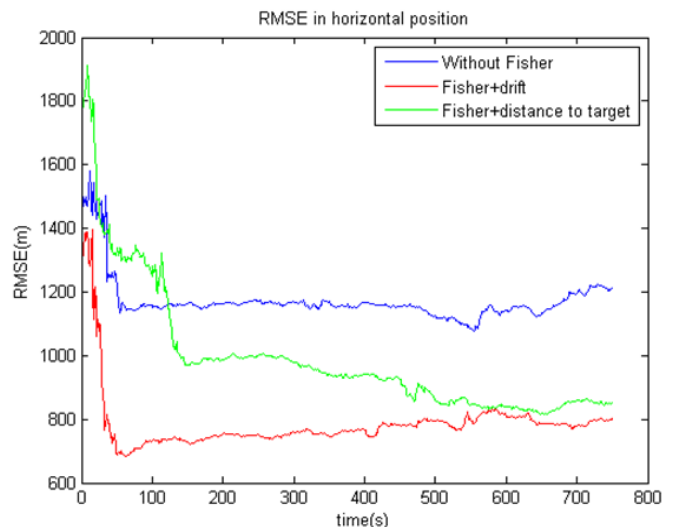


Fig. 2. Plot of the RMSE in horizontal position for 3 policies: without the FIM (blue), with the FIM and the drift (red), with the FIM and the distance to the target (green)

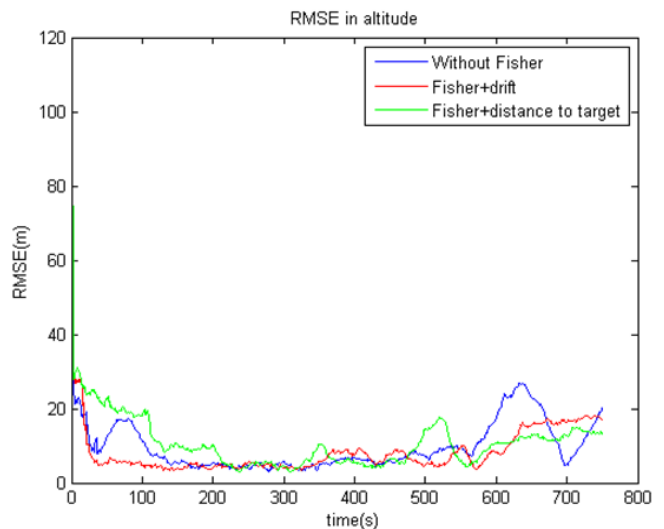


Fig. 3. Plot of the RMSE in altitude for 3 policies: without the FIM (blue), with the FIM and the drift (red), with the FIM and the distance to the target (green)

VII. CONCLUSION

In this paper, we have presented a new explicit dual output feedback stochastic MPC for nonlinear systems. Its principle is to choose at each time step the controls that maximizes the information over the control that forces the mean decreasing of some Lyapunov-like function for discrete-time nonlinear

stochastic systems. In particular it does not involve penalization of the guiding goal in the cost which is a classical feature of MPC. Output feedback is obtained by coupling the resolution of an optimization problem with a particle filter. The method is applied to terrain aided navigation and appears to be easier to tune than one of our previous methods.

REFERENCES

- [1] Ali Mesbah. Stochastic model predictive control with active uncertainty learning: A Survey on dual control. *Annual Reviews in Control*, November 2017.
- [2] David S. Bayard and Alan Schumitzky. Implicit dual control based on particle filtering and forward dynamic programming. *International Journal of Adaptive Control and Signal Processing*, 2008.
- [3] Giancarlo Marafioti, Robert R. Bitmead, and Morten Hovd. Persistently exciting model predictive control: Persistently exciting model predictive control. *International Journal of Adaptive Control and Signal Processing*, 28(6):536–552, June 2014.
- [4] D. Telen, B. Houska, M. Vallerio, F. Logist, and J. Van Impe. A study of integrated experiment design for NMPC applied to the Droop model. *Chemical Engineering Science*, 160:370–383, March 2017.
- [5] Kristian G. Hanssen and Bjarne Foss. Scenario Based Implicit Dual Model. In *5th IFAC Conference on Nonlinear Model Predictive Control NMPC 2015*, volume 48, pages 416–241, Seville, Spain, September 2015. Elsevier.
- [6] Sankaranarayanan Subramanian, Sergio Lucia, and Sebastian Engell. Economic Multi-stage Output Feedback NMPC using the Unscented Kalman Filter. *IFAC-PapersOnLine*, 48(8):38–43, 2015.
- [7] Valerii V. Fedorov and Peter Hackl. *Model-oriented design of experiments*, volume 125. Springer Science & Business Media, 2012.
- [8] Peter Hokayem, Eugenio Cinquemani, Debasish Chatterjee, Federico Ramponi, and John Lygeros. Stochastic receding horizon control with output feedback and bounded controls. *Automatica*, 48(1):77–88, 2012.
- [9] Tyler Homer and Prashant Mhaskar. Output-Feedback Lyapunov-based Predictive Control of Stochastic Nonlinear Systems. *IEEE Transactions on Automatic Control*, pages 1–1, 2017.
- [10] Ali Mesbah. Stochastic Model Predictive Control: An Overview and Perspectives for Future Research. *IEEE Control Systems*, 36(6):30–44, December 2016.
- [11] Matthew Ellis, Helen Durand, and Panagiotis D. Christofides. A tutorial review of economic model predictive control methods. *Journal of Process Control*, 24(8):1156–1178, August 2014.
- [12] Dimitri P. Bertsekas. *Dynamic programming and optimal control. Vol. 1: [...]*. Number 3 in Athena scientific optimization and computation series. Athena Scientific, Belmont, Mass, 3. ed edition, 2005. OCLC: 835987011.
- [13] Emilien Flayac, Karim Dahia, Bruno Hérisse, and Frédéric Jean. Nonlinear Fisher Particle Output Feedback Control and its application to Terrain Aided Navigation. In *56th IEEE Conference on Decision and Control*, Melbourne, December 2017.
- [14] Yaakov Bar-Shalom and Edison Tse. Dual Effect, Certainty Equivalence, and Separation in Stochastic Control. *IEEE Transactions on Automatic Control*, 19(5):494–500, 1974.
- [15] Vincent Andrieu and Laurent Praly. A unifying point of view on output feedback designs for global asymptotic stabilization. *Automatica*, 45(8):1789–1798, August 2009.
- [16] Petr Tichavsky, Carlos H. Muravchik, and Arye Nehorai. Posterior Cramr-Rao bounds for discrete-time nonlinear filtering. *IEEE Transactions on signal processing*, 46(5):1386–1396, 1998.
- [17] Sean P. Meyn and Richard L. Tweedie. *Markov Chains and Stochastic stability*. Cambridge University Press, Cambridge, 2nd edition, 2009.
- [18] Debasish Chatterjee and John Lygeros. On stability and performance of stochastic predictive control techniques. *IEEE Transactions on Automatic Control*, 60(2):509–514, 2015.
- [19] Giuseppe C. Calafiore and Marco C. Campi. The Scenario Approach to Robust Control Design. *IEEE Transactions on Automatic Control*, 51(5):742–753, May 2006.
- [20] Martin A. Sehr and Robert R. Bitmead. Particle Model Predictive Control: Tractable Stochastic Nonlinear Output-Feedback MPC. *arXiv preprint arXiv:1612.00505*, 2016.