STATIONARY FULLY PROBABILISTIC CONTROL DESIGN

Tatiana V. Guy
Adaptive Systems Department
Institute of Information Theory and Automation
P.O. Box 18, 182 08 Prague 8, Czech Republic

Miroslav Kárný
Adaptive Systems Department
Institute of Information Theory and Automation
P.O. Box 18, 182 08 Prague 8, Czech Republic

Keywords: Stochastic control design, stationary control, fully probabilistic design, stationary state space models.

Abstract: Stochastic control design chooses the controller that makes the closed-loop behavior as close as possible to the desired one. The fully probabilistic design describes both the closed-loop and its desired behavior in probabilistic terms and uses the Kullback-Leibler divergence as their proximity measure. Such a design provides explicit minimizer, which opens a way for a simpler approximations of analytically infeasible cases. The current formulations are oriented towards finite-horizon design. Consequently, the optimal strategy is non-stationary one. This paper provides infinite-horizon problem formulation and solution. It leads to a stationary strategy whose approximation is much easier.

1 INTRODUCTION

Stochastic control design (Kushner, 1971) chooses control law that makes the closed-loop behaviour of the controlled system as close as possible to the desired behaviour. To review the numerous existing solution methods and major restrictions of their application, a survey (Lee and Lee 2004) can be advised.

In a wider context, the stochastic control design can be viewed as a specific case of Bayesian dynamic decision making (Berger, 1985), which minimises the expected value of a loss function expressing control aim. Probabilistic description of both: the closed-loop behaviour of the controlled system and the desired behaviour makes a ground for fully probabilistic design (FPD) of stochastic control. This design (Kárný, 1996; Kárný et al., 2003; Kárný et al., 2005; Kárný and Guy, 2004) selects randomised control laws that make the *entire* joint distribution of variables describing closed-loop behaviour as close as possible to their desired distribution. The paper considers asymptotic version of FPD. It suits to the situations when control horizon is large enough and leads to simplified design applicable to a wider set of control problems than the non-stationary, finite-horizon version.

The next section introduces necessary notions and notations. Section 3 recalls the FPD in the most general state-space setting and provides the extension of its solution for the growing horizon. This is the main result of the paper. Section 4 concludes the paper by discussion about the practical significance of the result obtained.

2 PRELIMINARIES

In the paper, $\equiv$ stands for the equality by definition; X^* denotes a set of all values of X; $\mathring{X}$ means cardinality of a finite set X^*; $X(t)$ stands for the sequence $(X_1, \ldots, X_t)$, $f(\cdot|\cdot)$ denotes probability density function (pdf); t labels discrete-time moments, $t \in t^* \equiv \{1, \ldots, \mathring{t}\}$; $\mathring{t}$ is a given control horizon that can grow up to the infinity; $d_t = (y_t, u_t)$ is the finite-dimensional data record at the discrete time t, consisting of the observed system output y_t and of the optional system input u_t; x_t stands for the finite-dimensional unobserved system state at time t.

Arguments distinguish respective pdfs and no formal distinction is made between a random variable, its realisation and an argument of a pdf. Integrals encountered are multivariate and definite with integration domains coinciding with those of integrands.

The FPD exploits the *Kullback-Leibler (KL) divergence*, an information entropy measure, (Kullback and Leibler, 1951)

V. Guy T. and Kárný M. (2005).
STATIONARY FULLY PROBABILISTIC CONTROL DESIGN.
In *Proceedings of the Second International Conference on Informatics in Control, Automation and Robotics*, pages 109-112
DOI: 10.5220/0001183101090112

$$\mathcal{D}\left(f||\tilde{f}\right) \equiv \int f(X) \ln\left(\frac{f(X)}{\tilde{f}(X)}\right) dX. \quad (1)$$

It measures proximity of pdfs $f, \tilde{f}$ acting on a set X^* and has the following key property

$$\mathcal{D}(f||\tilde{f}) \geq 0,\ \mathcal{D}(f||\tilde{f}) = 0 \text{ iff } f = \tilde{f} \quad (2)$$

almost everywhere on X^*.

The joint pdf $f(d(\mathring{t}), x(\mathring{t})|x_0, d(0)) f(x_0|d(0)) = f(d(\mathring{t}), x(\mathring{t})|x_0) f(x_0)$ of all random considered variables is the most complete probabilistic description of the closed-loop behaviour. The variable x_t concerns to have character of closed-loop state with x_0 is an initial uncertain state. $d(0)$ stands for the prior information serving for the choice of the input at time $t = 1$. Further on, $d(0)$ is considered implicitly only. The chain rule for pdfs (Peterka, 1981) implies the following decomposition of the joint pdf

$$\begin{aligned} & f(d(\mathring{t}), x(\mathring{t})|x_0) \\ & = f(x_0) \prod_{t \in t^*} f(y_t|u_t, d(t-1), x(t)) \quad (3) \\ & \times f(x_t|u_t, d(t-1), x(t-1)) f(u_t|d(t-1), x(t-1)). \end{aligned}$$

The chosen decomposition (3) distinguishes:

observation model $f(y_t|u_t, d(t-1), x(t))$;

time evolution model $f(x_t|u_t, d(t-1), x(t-1))$;

randomized controller law $f(u_t|d(t-1), x(t-1))$.

As the variable x_t represent closed-loop state, all above models do not depend on its history, i.e $x(t) = x_t$. Moreover, assumed *admissible controllers* generate the system input u_t using at most the historical data observed $d(t-1)$ and cannot use the unobserved states $x(t-1)$. Besides, the addressed stationary design requires all functions to be time independent. To gain this, observed data $d(t-1)$, growing with time t, has to enter the models via a fixed-dimensional *observable state* ϕ_{t-1}. Thus, the introduced closed-loop description (3) reduces to:

$$\begin{aligned} & f(d(\mathring{t}), x(\mathring{t})|x_0) \quad (4) \\ & = \prod_{t \in t^*} f(y_t|\psi_t, x_t) f(x_t|\psi_t, x_{t-1}) f(u_t|d(t-1)), \end{aligned}$$

where the *regression vector* $\psi_t \equiv [u_t, \phi_{t-1}]$.

3 FULLY PROBABILISTIC DESIGN

The control aim and constraints are quantified by the so-called *ideal pdf* that defines the desired joint distribution of the closed-loop variables considered. It is constructed in the way analogous to (4) with the user-specified factors are marked by the superscript ^{I}

$$\begin{aligned} & {}^I f(d(\mathring{t}), x(\mathring{t})\,|\,x_0)\, {}^I f(x_0) = \prod_{t \in t^*} {}^I f(y_t|\psi_t, x_t) \\ & \times {}^I f(x_t|\psi_t, x_{t-1})\, {}^I f(u_t|d(t-1)) f(x_0). \end{aligned}$$

The pdfs ${}^I f(y_t|\psi_t, x_t),\ {}^I f(x_t|\psi_t, x_{t-1})$ describe the ideal models of observation and state evolution and ${}^I f(u_t|d(t-1))$ the ideal control law. The prior pdf on possible initial states x_0^* cannot be influenced by the optimized control strategy so that it is left to its fate, i.e. ${}^I f(x_0) = f(x_0)$.

The formulation of FPD is straightforward: **find an admissible control strategy minimizing the KL divergence** $\mathcal{D}\left(f(d(\mathring{t}), x(\mathring{t}))||\, {}^I f(d(\mathring{t}), x(\mathring{t}))\right)$.

Solution of the FPD requires the solution of stochastic filtering problem in the closed-loop.

Proposition 1 (Filtering in the closed-loop) *Let the prior pdf* $f(x_0)$ *be given. Then, the pdf* $f(x_t|d(t))$, *determining the* state estimate, *and the pdf* $f(x_t|u_t, d(t-1))$, *determining the* state prediction, *evolve according to the coupled equations*

<u>*Time updating*</u>

$$f(x_t|u_t, d(t-1)) = \int f(x_t|\psi_t, x_{t-1}) f(x_{t-1}|d(t-1)) dx_{t-1}$$

<u>*Data updating*</u>

$$f(x_t|d(t)) = \frac{f(y_t|\psi_t, x_t) f(x_t|u_t, d(t-1))}{\underbrace{\int f(y_t|\psi_t, x_t) f(x_t|u_t, d(t-1))\, dx_t}_{f(y_t|u_t, d(t-1))}}.$$

The stochastic filtering does not depend on the used admissible control strategy $\{f(u_t|d(t-1))\}_{t \in t^*}$ *but only on the generated inputs.*

Let the time-invariant state estimate ${}^s f(x_t|V_t) = f(x_t|d(t))$ *exist, where* V_t *is a finite-dimensional statistic. Then, this function solves the equation*

$$\begin{aligned} & {}^s f(x_t|V_t) = \quad (5) \\ & \frac{f(y_t|\psi_t, x_t) \int f(x_t|\psi_t, x_{t-1})\, {}^s f(x_{t-1}|V_{t-1})\, dx_{t-1}}{\underbrace{\int f(y_t|\psi_t, x_t) f(x_t|\psi_t, x_{t-1})\, {}^s f(x_{t-1}|V_{t-1})\, dx_{t-1}\, dx_t}_{f(y_t|\psi_t, V_{t-1})}} \end{aligned}$$

Proof: See, e.g. (Kárný and Guy, 2004). □

Example 1 (Stationary Kalman filter) *Let the time evolution model be* $f(x_t|\psi_t, x_{t-1}) = \mathcal{N}_{x_t}(Ax_{t-1} + Bu_t, R)$, *with* $\mathcal{N}_z(\hat{z}, w)$ *denoting normal pdf of* z *having expectation* $\hat{z}$ *and covariance matrix* w. *Let also the observation model be normal* $f(y_t|\psi_t, x_t) = \mathcal{N}_{y_t}(Cx_t + Du_t, r)$. *Assuming that matrices* A, B, C, D, R, r *are known, the stationary* *estimate* ${}^{s}f(x_t|V_t \equiv (\hat{x}_t, P)) = \mathcal{N}_{x_t}(\hat{x}_t, P)$. *The expectation and covariance matrix of this estimate fulfill the equations coinciding with stationary Kalman filter (Meditch, 1969)*

$\hat{x}_t = A\hat{x}_{t-1} + Bu_t + K(y_t - C\hat{x}_{t-1} - Du_t)$, *with* $K = PC'r^{-1}$ *and* $P^{-1} = C'r^{-1}C + (APA' + R)^{-1}$.

Respecting the aim of the paper, the solution of FPD for the stationary state estimate is written.

Proposition 2 (Solution of FPD) *Let the state estimate reached its stationary form* ${}^{s}f(x_t|V_t)$. *Then, the optimal admissible control strategy in FPD sense is the randomized one given by the pdfs* ${}^{o}f(u_t|\phi_{t-1}, V_{t-1})$

$$= {}^{I}f(u_t|\phi_{t-1})\frac{\exp[-\omega(\psi_t, V_{t-1})]}{\gamma(\phi_{t-1}, V_{t-1})}, \quad t \in t^*, \quad (6)$$

$$\gamma(\phi_{t-1}, V_{t-1}) \equiv \int {}^{I}f(u_t|\phi_{t-1})\exp[-\omega(\psi_t, V_{t-1})]\, du_t.$$

Starting with $\gamma(\phi_{\mathring{t}}, V_{\mathring{t}}) \equiv 1$, *the functions* $\omega(\psi_t, V_{t-1})$ *are generated recursively in the backward manner for* $t = \mathring{t}, \mathring{t}-1, \ldots, 1$, *as follows*

$$\omega(\psi_t, V_{t-1}) \equiv \int \Omega(\psi_t, x_{t-1})\, {}^{s}f(x_{t-1}|V_{t-1}) dx_{t-1}.$$

${}^{s}f(x_t|V_t)$ *is updated according to Proposition 1 and*

$$\Omega(\psi_t, x_{t-1}) \equiv \int f(y_t|\psi_t, x_t) f(x_t|\psi_t, x_{t-1}) \times \ln\left(\frac{f(y_t|\psi_t, x_t) f(x_t|\psi_t, x_{t-1})}{\gamma(\phi_t, V_t)\, {}^{I}f(y_t|\psi_t, x_t)\, {}^{I}f(x_t|\psi_t, x_{t-1})}\right) dy_t dx_t.$$

Proof: See (Kárný and Guy, 2004) □

The following proposition describes the key result of this paper.

Proposition 3 (Solution of FPD for $\mathring{t} \to \infty$) *For a given randomized admissible strategy*

$$\{f(u_t|d(t-1)\}_{t=1}^{\mathring{t}}, \ \mathring{t} < \infty,$$

the KL divergence is expected value of an additive loss function.

Let there is such a controller for which the state estimate reaches its stationary form ${}^{s}f(x_t|V_t)$ *and expectation of the partial loss forming the Kullback-Leibler divergence is bounded even for* $\mathring{t} \to \infty$ *by a finite constant* K.

Then, for the horizon $\mathring{t} \to \infty$, *the optimal admissible control strategy in FPD sense is stationary randomized one given by the pdfs* ${}^{o}f(u_t|\phi_{t-1}, V_{t-1})$

$$= {}^{I}f(u_t|\phi_{t-1})\frac{\exp[-\omega(\psi_t, V_{t-1})]}{\gamma(\phi_{t-1}, V_{t-1})}, \quad t \in t^*, \quad (7)$$

$$\gamma(\phi_{t-1}, V_{t-1}) \equiv \int {}^{I}f(u_t|\phi_{t-1})\exp[-\omega(\psi_t, V_{t-1})]\, du_t.$$

The function $\omega(\psi_t, V_{t-1})$ *fulfills the following equation with* ${}^{s}f(x_t|V_t)$ *updated as in the Proposition 1*

$$\omega(\psi_t, V_{t-1}) \equiv \int f(y_t|\psi_t, x_t) f(x_t|\psi_t, x_{t-1}) \quad (8)$$
$$\times \ln\left(\frac{f(y_t|\psi_t, x_t)}{\int {}^{I}f(u_{t+1}|\phi_t)\exp[-\omega(\psi_{t+1}, V_t)]\, du_{t+1}} \times \frac{f(x_t|\psi_t, x_{t-1})}{{}^{I}f(x_t|\psi_t, x_{t-1})\, {}^{I}f(y_t|\psi_t, x_t)}\right)$$
$$\times\, {}^{s}f(x_{t-1}|V_{t-1})\, dy_t\, dx_t dx_{t-1}.$$

Proof: The KL divergence is an expectation of the logarithm containing ratio of products. Thus, it represents an expected value of the additive loss function. According to the assumptions, there is a strategy that makes expectations of the partial losses bounded. The loss function, which equals to the KL divergence $-\mathring{t}\ln(K)$ for any constant $K > 0$ is minimised by the same control law as the original KL divergence. At the same time, there exists K such that the shifted loss is bounded from the above for any $\mathring{t}$ and thus its limit superior exists. The minimising strategy depends on the reached minima γ (whose constant shifts do not change the minimising strategy), which converges, too. This implies convergence of ω and, finally, stationarity of the strategy obtained for growing horizon. The function ω, determining it, meets stationary version of non-stationary equations in Proposition 2. By excluding the intermediate functions γ, Ω, the claimed final version can be obtained. □

Example 2 (FPD for normal state-space model) *Let us assume controlled system described by the normal state-space model as in the Example 1. Let us consider the regulation problem, which implies that we try to push all dynamics to zero while leaving the uncontrollable innovations to their fate. Therefore, the ideal pdf is* ${}^{I}f(x_t|\psi_t, x_{t-1}) = \mathcal{N}_{x_t}(0, R)$, ${}^{I}f(y_t|\psi_t, x_t) = \mathcal{N}_{y_t}(0, r)$, *while requiring* ${}^{I}f(u_t|d(t-1)) = \mathcal{N}_{u_t}(0, q)$.

In this case, the optimal stationary control law is

$${}^{o}f(u_t|d(t-1)) = \mathcal{N}(L\hat{x}_{t-1}, {}^{o}q)$$

with

$$\begin{aligned} {}^{o}q &= (B'Q^{-1}B + D'r^{-1}D + q^{-1})^{-1} \\ L &= {}^{o}q^{-1}(B'Q^{-1}A + D'r^{-1}C) \text{ and} \\ Q^{-1} &= A'Q^{-1}A + C'r^{-1}C - L'\,{}^{o}qL + R^{-1}. \end{aligned}$$

Note that the non-standard equation for stationary Riccati matrix is caused by non-standard presence of the term Du_t *in the observation model and by the non-standard attempt to optimise jointly output and the state. Without this, the mean value of the optimal controller is usual stationary control law obtained in linear quadratic design with the state penalisation* R^{-1} *and input penalisation* q^{-1}.

Interpretation of ^oq and inversion of Riccati matrix Q is non-standard: they represent stationary covariance matrices of the optimal inputs and states, respectively.

4 DISCUSSION

Design of the optimal strategy reduces to the solution of the stationary version of the integral filtering equation (5) and of the solution of the integral equation (8). For the normal state-space model and the normal ideal pdf, it reduces to the stationary version of Kalman filter and design minimising quadratic criterion (Meditch, 1969). Even in this case, the FPD interpretation brings practical advantages, as it interprets the penalisation matrices as inversions of the ideal covariance matrices and thus guides on their choice. Moreover, when they are recursively (approximately) estimated, the weights adapts to the varying noise level, which generally spares the input effort as the control of uncontrollable innovations is given up.

Generally, the closed-form solutions of discussed equations exist rarely but the explicit form of control laws simplifies numerical approximations substantially. The stationary form of the solution prepares such approximations even better as ””only” the stationary functions $f(x_t|V_t)$ and $\omega(\psi_t, V_{t-1})$ have to be approximated (not sequences of such functions).

Non-linear character of the filtering and the design equations together with a generic high dimensionality of their domain restrict supply of available approximation techniques. Essentially, a global approximation suitable for higher dimensions has to be used. The neural networks (Haykin, 1994), ideally interpreted as finite probabilistic mixtures (Titterington et al., 1985) and general ANOVA-like approximations (Rabitz and Alis, 1999) seem to be prime candidates. Especially, the mixture versions look promising as there are approximate techniques for FPD with them (Murray-Smith and Johansen, 1997; Kárný et al., 2003).

ACKNOWLEDGEMENTS

This research has been partially supported by the grants GAČR 102/03/0049, GAČR 102/03/P010 and by the grant 1M6798555601 of the Czech Ministry of Education, Youth and Sports.

REFERENCES

Berger, J. (1985). *Statistical Decision Theory and Bayesian Analysis*. Springer-Verlag, New York.

Haykin, S. (1994). *Neural networks: A comprehensive foundation*. Macmillan College Publishing Company, New York.

Kárný, M. (1996). Towards fully probabilistic control design. *Automatica*, 32(12):1719–1722.

Kárný, M., Böhm, J., Guy, T., Jirsa, L., Nagy, I., Nedoma, P., and Tesař, L. (2005). *Optimized Bayesian Dynamic Advising: Theory and Algorithms*. Springer, London. to appear.

Kárný, M., Böhm, J., Guy, T. V., and Nedoma, P. (2003). Mixture-based adaptive probabilistic control. *International Journal of Adaptive Control and Signal Processing*, 17(2):119–132.

Kárný, M. and Guy, T. (2004). Fully probabilistic control design. *Systems & Control Letters*. submitted.

Kullback, S. and Leibler, R. (1951). On information and sufficiency. *Annals of Mathematical Statistics*, 22:79–87.

Kushner, H. (1971). *Introduction to stochastic control*. Holt, Rinehart and Winston, New York, San Francisco, London.

Lee, J.M. and Lee, J.H. (2004). Approximate dynamic procgramming strategies and their applica bility for process control: A review and future directions. *International Journal of Control*, 2(3):263–278.

Meditch, J. (1969). *Stochastic Optimal Linear Estimation and Control*. Mc. Graw Hill.

Murray-Smith, R. and Johansen, T. (1997). *Multiple Model Approaches to Modelling and Control*. Taylor & Francis, London.

Peterka, V. (1981). Bayesian system identification. In Eykhoff, P., editor, *Trends and Progress in System Identification*, pages 239–304. Pergamon Press, Oxford.

Rabitz, H. and Alis, O. (1999). General foundations of high-dimensional model representations. *Journal of Mathematical Chemistry*, 25:197–233.

Titterington, D., Smith, A., and Makov, U. (1985). *Statistical Analysis of Finite Mixtures*. John Wiley & Sons, Chichester, New York, Brisbane, Toronto, Singapore. ISBN 0 471 90763 4.