

Copyright © 2017 by John Wiley & Sons. This is the peer reviewed version of the following article: Generalised Probabilistic Control Design for Uncertain Stochastic Control Systems Herzallah, R. 22 Sep 2017 In : Asian Journal of Control. p. 1 11 p., which has been published in final form at <http://dx.doi.org/10.1002/asjc.1717>. This article may be used for non-commercial purposes in accordance with Wiley Terms and Conditions for Self-Archiving.

Generalised Probabilistic Control Design for Uncertain Stochastic Control Systems

Randa Herzallah

NCRG

Aston University, UK

Email: r.herzallah@aston.ac.uk

Abstract

In this paper a novel generalised fully probabilistic controller design for the minimisation of the Kullback-Leibler divergence between the actual joint probability density function (pdf) of the closed loop control system, and an ideal joint pdf is presented for a linear Gaussian uncertain class of stochastic systems. A single layer neural network is used to approximate the probability density function of the system dynamics. The generalised probabilistic control law is obtained by solving the recurrence equation of dynamic programming to the fully probabilistic design control problem while taking into consideration the dependency of the parameters of the estimated probability density function of the system dynamics on the input values. It is shown to be of the class of cautious type controllers which accurately minimises the value of the Kullback-Leibler divergence without disregarding the variance of the model prediction as an element to be minimised. Comparison of theoretical and numerical results obtained from the F-16 fighter aircraft application with existing state-of-the-art demonstrates the effectiveness of the proposed method.

Keywords

Generalised fully probabilistic controller design, cautious controller, functional uncertainty, generalised Riccati equation.

I. INTRODUCTION

This paper considers the currently unsolved design problem of randomised controllers for linear dynamic discrete time stochastic systems characterised by functional uncertainties and nonstationary and input dependent noise. Many engineering problems are fraught with several sources of uncertainties including the aforementioned. Other examples include model based control design methods where further uncertainty is introduced through model discrepancy as a result of the lack of knowledge of the underlying physics that describes how a real system behaves [11]. Many practical applications, also, are stochastic with the statistics of the external noises affecting their dynamical behavior being unknown exactly, hence introducing models' uncertainty [9]. Another example is where the wind disturbance which affects the way an aircraft lands in an aircraft autolander system is a function of the altitude of the aircraft [20]. For these systems, better control algorithms can be devised by taking the uncertainties of the stochastic system into consideration when calculating optimal control inputs. Consequently, the development of such control algorithms has become an attractive research area in the past few

years. In the following we mention a few approaches. Papers which provide information on the theoretical development can be found in [3], [19], [25], and the references therein. A stochastic H_2/H_∞ control design is developed in [6] for systems with state dependent noise. The H_2/H_∞ is also considered in [29] for Itô-type nonlinear stochastic differential equation with state dependent noise and cross-coupled Hamilton-Jacobi equations are obtained. The research in [21] studied partially observed linear quadratic Gaussian (LQG) models where the stochastic disturbances depend on both the states and the controls, and the measurements are bilinear in the noise and the states/controls. The infinite horizon discrete time LQ control with state and control-dependent white noise is considered in [15] and a discussion of the properties of the obtained generalised algebraic Riccati equation is also provided.

The design and development of control laws that consider systems' uncertainty and input dependent noise are the main objectives of this paper. However, in contrast to earlier discussed approaches the design process in this article will be based on the minimisation of the Kullback–Leibler divergence distance between the joint distribution of the system dynamics and a pre-defined ideal joint distribution. This method of control is known as Fully Probabilistic Design (FPD) [16]. The objective of FPD is to determine the pdf of a randomised optimal control law, $c(\mathbf{u}_t | \mathbf{x}_{t-1})$ that minimises the following explicit performance function derived from the Kullback–Leibler divergence,

$$\mathcal{D}(f || \mathcal{I}f) \equiv \int f(\mathbf{x}(H), \mathbf{u}(H)) \ln \left(\frac{f(\mathbf{x}(H), \mathbf{u}(H))}{\mathcal{I}f(\mathbf{x}(H), \mathbf{u}(H))} \right) d\mathbf{x}(H)d\mathbf{u}(H), \quad (1)$$

where $\mathbf{x}(H) = \{x_1, \dots, x_H\}$ is the sequence of measured states of the system, $\mathbf{u}(H) = \{u_1, \dots, u_H\}$ is the sequence of measured inputs, $H \leq \infty$ is a given control horizon, and where $f(\mathbf{x}(H), \mathbf{u}(H))$ and $\mathcal{I}f(\mathbf{x}(H), \mathbf{u}(H))$ represent the joint and ideal distributions of the system dynamics respectively.

Remark 1: The randomised controllers obtained by FPD method are optimal controllers that shape the joint pdf describing the closed loop behaviour of the system dynamics. Randomisation makes the controllers explorative where ideally actions should be sampled from their pdfs [17].

Remark 2: The ideal joint pdf of the closed loop system behaviour, $\mathcal{I}f(\mathbf{x}(H), \mathbf{u}(H))$ specifies the preferred form for the joint distribution of the system behaviour, $f(\mathbf{x}(H), \mathbf{u}(H))$. It assigns a high probability-mass to highly desirable realisations of the closed loop signals. For an axiomatic justification of the FPD, and the role of the ideal distribution the reader is referred to [17].

The main advantage of the FPD is that it provides an explicit form of the randomised optimal controller. However, since the evaluation of the randomised optimal controller involves multivariate integration steps, as can be seen from (1), which need to be computed by backward recursion the problem renders to be nontrivial and computationally very intensive. To overcome the difficulties arising in the FPD, a probabilistic Dual Heuristic Programming (DHP) adaptive critic method was proposed in [12], [14]. The DHP adaptive critic method uses a critic network to circumvent the need for explicitly evaluating the optimal value function, therefore, dramatically reducing computational requirements.

At present, most of the proposed FPD control and DHP adaptive critic control laws are obtained based on the assumptions that all required pdfs of the stochastic system are measurable and can be approximated from process data. However, these methods have only considered the noise to be global and independent of the input variables. On the other hand, as discussed earlier many engineering systems and systems with unknown dynamics are known to have functional uncertainties and are affected by input dependent noise. This means that input dependent noise should be considered and contribute to the derivation of the optimal control law.

The methodology of this paper is to introduce the more general solution of fully probabilistic control for stochastic linear Gaussian systems where the dynamics of the system is unknown and where the uncertainty introduced by the model discrepancy is estimated as a function of the system inputs. It will be seen that the proposed solution yields a cautious type controller which takes into consideration model uncertainty when calculating the optimal control law. Cautious type controllers apply caution but lack probing [8]. The approach followed in this paper is pragmatic and fully probabilistic. It is radically different to the state of the art control design methods [1], [4], [7], [13], [22], [26], [28] which are concerned with the minimisation of objective functions that are confined to be either the mean value or variance of the stochastic output.

Analogous to the standard solution of the fully probabilistic control problem, it will be shown that the optimal control law for uncertain systems with input dependent noise is again linear in the state of the system, x but is now rather critically dependent on the parameters of the estimated noise. This is in accordance with the generalised Riccati solutions that take into account the effect of uncertainties in the control design of linear uncertain systems with quadratic cost functions [2], [5], [18], [27].

To summarise, this paper provides three novel contributions: Firstly, we use the Kulback-Leibler distance between the closed loop system and an ideal pdf as the cost function within the previously discussed context where noise and models' uncertainty are dependent on the input values. This will lead to cautious control design that applies caution under highly uncertain situations, hence yielding better control characteristics; Secondly, we use advanced neural network methods to estimate the corresponding pdfs that are required in the FPD method such that their parameters are dependent on the input values. This allows us to characterise and estimate model discrepancy and functional uncertainty in the system dynamics; Thirdly, having obtained these pdfs, it allows us to obtain a generalised solution to the FPD control problem such that model uncertainty is explicitly taken into consideration in the derivation of the optimal control law. This generalised solution to the FPD control problem is referred to as Generalised Fully Probabilistic Control Design (GFPCD) and leads to higher order terms in the analysis and more accurate predictions in the numerical example shown later.

To emphasise, the GFPCD method proposed in this article elicits optimal control laws that are dependent on the parameters of the estimated system uncertainty. This characterisation and consideration of systems' uncertainty in the derivation of the optimal control law, allow elicitation of cautious controllers that apply caution under highly uncertain situations. This desirable caution feature of the elicited control inputs improves the performance of the controlled system and minimises regulation errors. This is particularly significant in the transient period of the control process where the discrepancy between the estimated model and actual system is large. The consideration of system uncertainty in the derivation of the optimal control laws represents the main novelty of the new generalised solution to the FPD control problem proposed in this paper and it leads to more robust, accurate estimates of control signals than current approaches.

II. PRELIMINARIES

A. Model Description

Discrete linear time invariant stochastic dynamical systems play an essential role in modern control theory. In this paper linear discrete time stochastic dynamical systems described by the following state space model are considered:

$$\mathbf{x}_t = \tilde{\mathbf{A}}\mathbf{x}_{t-1} + \tilde{\mathbf{B}}\mathbf{u}_t + \tilde{\mathbf{e}}_t, \quad (2)$$

where $\mathbf{x} \in \mathcal{R}^n$ is the measured state vector, $\mathbf{u} \in \mathcal{R}^r$ is the control input vector, $\tilde{\mathbf{A}} \in \mathcal{R}^{n \times n}$ and $\tilde{\mathbf{B}} \in \mathcal{R}^{n \times r}$ are fixed state and control matrices respectively. Section II-C provides a discussion on how these matrices can be determined. Also $\tilde{\epsilon}_t \in \mathcal{R}^n$ is a Gaussian vector noise process of zero mean and fixed arbitrary covariance. Because of the noise input $\tilde{\epsilon}_t$ the previous state and present control do not completely specify the present state, but instead determine only the probability distribution of these states, $s(\mathbf{x}_t | \mathbf{u}_t, \mathbf{x}_{t-1})$. In this paper, it is assumed that the noise input $\tilde{\epsilon}_t$ is unknown and hence the probability distribution of the states is unknown and needs to be estimated.

The estimation method of the unknown probability density functions will be discussed in Section II-C, but first we introduce the probabilistic control objective.

B. Probabilistic Control Objective

The control problem confronted here is to design a control strategy for the system in (2) to control the state of the system to a predefined desired state value. However, since only the probability distribution of the states can be determined, this control objective should be re-defined in terms of the probabilistic control theory. Therefore, to achieve this control objective we consider designing a probabilistic controller $c(\mathbf{u}_t | \mathbf{x}_{t-1})$ that shapes the joint pdf of the closed loop system, $f(\mathbf{x}_t, \mathbf{u}_t)$ and makes it as close as possible to a predefined desired pdf, $l_f(\mathbf{x}_t, \mathbf{u}_t)$. This design method was originally presented in [16] where the probabilistic controller is obtained such that it minimises the Kullback–Leibler divergence distance defined in (1). The minimum cost function resulting from minimisation of (1) with respect to admissible control sequence \mathbf{u}_t , $t \in \{1, \dots, H\}$, with H being the control horizon, is then shown to be given by the following recurrence equation [14],

$$\begin{aligned}
 -\ln(\gamma(\mathbf{x}_{t-1})) &= \min_{c(\mathbf{u}_t|\mathbf{x}_{t-1})} \int s(\mathbf{x}_t|\mathbf{u}_t, \mathbf{x}_{t-1}) c(\mathbf{u}_t|\mathbf{x}_{t-1}) \\
 &\times \left[\underbrace{\ln \left(\frac{s(\mathbf{x}_t|\mathbf{u}_t, \mathbf{x}_{t-1}) c(\mathbf{u}_t|\mathbf{x}_{t-1})}{l_s(\mathbf{x}_t|\mathbf{u}_t, \mathbf{x}_{t-1}) l_c(\mathbf{u}_t|\mathbf{x}_{t-1})} \right)}_{\text{partial cost} \Rightarrow \mathbf{U}(\mathbf{x}_t, \mathbf{u}_t)} - \underbrace{\ln(\gamma(\mathbf{x}_t))}_{\text{optimal cost-to-go}} \right] d(\mathbf{x}_t, \mathbf{u}_t), \quad (3)
 \end{aligned}$$

where $-\ln(\gamma(\mathbf{x}_{t-1}))$ is the expected minimum cost-to-go function and where

$$f(\mathbf{x}_t, \mathbf{u}_t) = s(\mathbf{x}_t|\mathbf{u}_t, \mathbf{x}_{t-1}) c(\mathbf{u}_t|\mathbf{x}_{t-1}), \quad (4)$$

is the decomposition of the actual joint pdf by the chain rule [24], which represents the most complete probabilistic description of the closed loop system. Here the pdf $s(x_t|u_t, x_{t-1})$ describes the dynamics of the observed state vector x_t . Similarly

$${}^I f(x_t, u_t) = {}^I s(x_t|u_t, x_{t-1}) {}^I c(u_t|x_{t-1}), \quad (5)$$

is the decomposition of the ideal joint pdf of the closed loop system where, ${}^I s(x_t|u_t, x_{t-1})$ and ${}^I c(u_t|x_{t-1})$ represent the pdfs of the desired dynamics of the observed state vector and ideal controller respectively. Full derivation of (3) can be found in [14]. It constitutes the recurrence equation of the dynamic programming solution to the FPD control problem.

The minimisation of the cost-to-go function (3) with respect to control law, $c(u_t | x_{t-1})$ is shown in [12], [14] to be given by

$$\begin{aligned} c^*(u_t|x_{t-1}) &= \frac{{}^I c(u_t|x_{t-1}) \exp[-\beta_1(u_t, x_{t-1}) - \beta_2(u_t, x_{t-1})]}{\gamma(x_{t-1})}, \\ \gamma(x_{t-1}) &= \int {}^I c(u_t|x_{t-1}) \exp[-\beta_1(u_t, x_{t-1}) - \beta_2(u_t, x_{t-1})] du_t, \\ \beta_1(u_t, x_{t-1}) &= \int s(x_t|u_t, x_{t-1}) \left[\ln \frac{s(x_t|u_t, x_{t-1})}{{}^I s(x_t|u_t, x_{t-1})} \right] dx_t, \\ \beta_2(u_t, x_{t-1}) &= - \int s(x_t|u_t, x_{t-1}) \ln(\gamma(x_t)) dx_t. \end{aligned} \quad (6)$$

The solution of this probabilistic control problem to linear stochastic uncertain systems which have probability density functions characterised by input dependent parameters will be introduced soon. However, we first discuss the estimation of the probabilistic distributions of the systems dynamics and hence model uncertainty which will be required for the derivation of the probabilistic control solution.

C. pdf of the system dynamics

To estimate the probabilistic model of the linear stochastic system (2) we adopt the method proposed in [11], where neural network models are used to provide a prediction for the conditional expectation of the system state values and calculate the input dependent variance of its residual error. For the class of linear systems (2) considered in this paper, a single layer neural network known as a Generalised Linear Model (GLM) can be used to estimate the expected values of the system state such that the inequality,

$$|x_t - N_f(u_t, x_{t-1})| \leq \delta, \quad (7)$$

holds, where $\delta > 0$ is a known small number and $N_f(\mathbf{u}_t, \mathbf{x}_{t-1}) = \mu_t$ is a GLM approximation of the state \mathbf{x}_t . Considering the linear transformation of previous state values and present control inputs, $N_f(\mathbf{u}_t, \mathbf{x}_{t-1}) = A\mathbf{x}_{t-1} + B\mathbf{u}_t$ in which A and B are estimates of the state \tilde{A} and control \tilde{B} matrices respectively, the stochastic system (2) can be re-expressed as,

$$\mathbf{x}_t = A\mathbf{x}_{t-1} + B\mathbf{u}_t + e(\mathbf{x}_{t-1}, \mathbf{u}_t). \quad (8)$$

Here, $e(\mathbf{x}_{t-1}, \mathbf{u}_t)$ represents the approximation error satisfying $|e(\mathbf{x}_{t-1}, \mathbf{u}_t)| \leq \delta$. This means that the resulting conditional distribution of the system dynamics $s(\mathbf{x}_t | \mathbf{u}_t, \mathbf{x}_{t-1})$ is Gaussian distribution function with conditional expectation of the distribution being given by the GLM approximation and an input dependent covariance given by [11],

$$\tilde{\Sigma}_t = E((\mathbf{x}_t - \mu_t)(\mathbf{x}_t - \mu_t)'), \quad (9)$$

where $E(\cdot)$ denotes the expected value, and $'$ means transpose. We emphasise here that $\tilde{\Sigma}_t$ is input dependent and can be calculated for each input pattern. Following the procedure in [11] another generalised linear model which has the same structure and same inputs as that of the state model is then used to predict the covariance matrix, Σ_t of the error function $e(\mathbf{x}_{t-1}, \mathbf{u}_t)$,

$$\Sigma_t = D\mathbf{x}_{t-1} + G\mathbf{u}_t, \quad (10)$$

where D and G are partitioned matrices and are updated such that the error between the actual covariance matrix and the estimated one is minimised. The architecture of this probabilistic model estimation method is shown in figure 1.

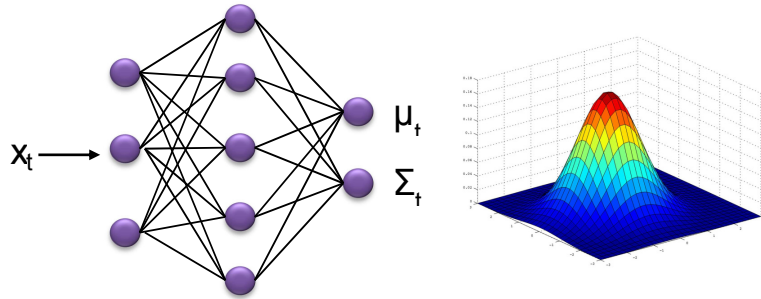


Fig. 1. The architecture of the probabilistic model estimation method

III. SOLUTION TO THE PROBABILISTIC UNCERTAIN CONTROL PROBLEM

A. Basic Elements

In this section we discuss the basic elements of the linear stochastic system defined in (2) that are required for the derivation of the generalised fully probabilistic control solution. For presentational clarity and simplicity, the solution to this problem will be developed for a regulation problem where the objective is to reach a zero state with a spread determined by a specified covariance matrix. Generalisation to a state value that is different than zero can be obtained in a straight forward manner by taking the mean of the the ideal probability density function of the forward model of the system dynamics equal to the mean value of the desired state trajectory.

As discussed in Section II-C, the conditional distribution of the linear system (2) is estimated as a Gaussian distribution described by,

$$\begin{aligned} \mathbf{x}_t &= \mathbf{A}\mathbf{x}_{t-1} + \mathbf{B}\mathbf{u}_t + \mathbf{e}(\mathbf{x}_{t-1}, \mathbf{u}_t), \\ s(\mathbf{x}_t | \mathbf{u}_t, \mathbf{x}_{t-1}) &\rightsquigarrow \mathcal{N}_{\mathbf{x}_t}(\boldsymbol{\mu}_t, \boldsymbol{\Sigma}_t), \end{aligned} \quad (11)$$

where

$$\boldsymbol{\mu}_t = \mathbf{A}\mathbf{x}_{t-1} + \mathbf{B}\mathbf{u}_t, \quad (12)$$

$$\boldsymbol{\Sigma}_t = \mathbf{D}\mathbf{x}_{t-1} + \mathbf{G}\mathbf{u}_t. \quad (13)$$

For the considered regulation problem, the system is initially in state \mathbf{x}_{t-1} and the aim is to return the system state to the origin. Hence, the distribution of the ideal state of the system is taken to be,

$${}^I s(\mathbf{x}_t | \mathbf{u}_t, \mathbf{x}_{t-1}) = \mathcal{N}_{\mathbf{x}_t}(0, \boldsymbol{\Sigma}_t), \quad (14)$$

where here the desired mean value of the state is taken to be zero and where $\boldsymbol{\Sigma}_t$ specifies the covariance of the innovation of the state values.

The distribution of the ideal controller is also assumed to be Gaussian

$${}^I c(\mathbf{u}_t | \mathbf{x}_{t-1}) = \mathcal{N}_{\mathbf{u}_t}(0, \boldsymbol{\Gamma}), \quad (15)$$

where $\boldsymbol{\Gamma}$ determines the allowed range of optimal control inputs.

Remark 3: To re-emphasise, regulation control problems with the objective to reach zero states that have a spread determined by a specified covariance matrix is considered. This objective is

reflected in our selection for $I_s(x_t|u_t, x_{t-1})$ and $I_c(u_t|x_{t-1})$ which are defined to have zero mean and covariance matrices Σ_t and Γ respectively as can be seen from Equations (14) and (15). Generalisation to servo control problems can be obtained in a straight forward manner by making the mean of the ideal distribution of the system state, $I_s(x_t|u_t, x_{t-1})$ equal to the desired state value. To be more specific, the desired output needs to be selected as usual in a model reference control problem and then set to be equal to the mean of the desired output of the ideal distribution of the system state. Other generalisations where for example alternative ideal joint pdfs with different covariance matrices than that of the actual joint pdf are possible as well.

B. Generalised Probabilistic Control Law

In this section we derive the generalised fully probabilistic control solution of the regulation control problem of stochastic linear systems with functional uncertainty that is discussed in Section III-A. We show here that the obtained optimal control law and optimal performance index from the proposed GFPCD are different to current theoretical approaches to this problem. In particular we show that the optimal control law is a state feedback law which depends on the parameters of the estimated input dependent uncertainty and that the optimal performance index is quadratic in the state x_t which is also dependent on the estimated uncertainty. The derivation of optimal control laws depends on the evaluation of the optimal performance index, $\ln(-\gamma(x_t))$. This in turn requires the calculation of $\beta_1(u_t, x_{t-1})$, $\beta_2(u_t, x_{t-1})$ and $\gamma(x_t)$ as defined by equation (6) and yields the optimal performance index specified in the following theorem.

Theorem 1: Using the elements defined in Equations (11), (12), (13), (14), and (15) in Equation (6) yields the following optimal performance index

$$-\ln(\gamma(x_t)) = 0.5x_t' M_t x_t + 0.5V_t x_t + 0.5\omega_t, \quad (16)$$

where

$$\begin{aligned} M_{t-1} &= A' \left\{ (\Sigma_t^{-1} + M_t) - (\Sigma_t^{-1} + M_t)B[\Gamma^{-1} + B'(\Sigma_t^{-1} + M_t)B]^{-1} \right. \\ &\quad \left. \times B'(\Sigma_t^{-1} + M_t) \right\} A, \end{aligned} \quad (17)$$

$$\begin{aligned} V_{t-1} &= \left\{ V_t A + \text{tr}(DM_t) - 2A'(\Sigma_t^{-1} + M_t)B \right. \\ &\quad \left. \times [\Gamma^{-1} + B'(\Sigma_t^{-1} + M_t)B]^{-1} (0.5B'V_t + 0.5\text{tr}(GM_t)) \right\}, \end{aligned} \quad (18)$$

$$\begin{aligned} \omega_{t-1} &= \omega_t - \{0.5V_t B + 0.5\text{tr}(GM_t)\} [\Gamma^{-1} + B'(\Sigma_t^{-1} + M_t)B]^{-1} \{0.5B'V_t + 0.5\text{tr}(GM_t)\} \\ &\quad + \ln | I + (B\Gamma^{0.5})'(\Sigma_t^{-1} + M_t)(B\Gamma^{0.5}) |. \end{aligned} \quad (19)$$

Remark 4: Equation (17) is called the discrete time Riccati equation. It is similar to that obtained for probabilistic controller with input independent noise. The derivation of Equation (18) is a key modification to the original standard FPD. The manifestation of this equation is due to the consideration of input dependent noise. It provides cautiousness to the optimal probabilistic quadratic controller, therefore will be referred to as the equation of cautiousness. It can also be seen that the solution of Equation (18) depends on the solution of the Riccati equation (17).

Proof: The claimed form of $-\ln(\gamma(x_t))$ can be verified subsequently by backward induction. We start the proof by evaluating $\beta_1(u_t, x_{t-1})$ as defined by Equation (6). It can be easily evaluated to give

$$\begin{aligned} \beta_1(u_t, x_{t-1}) &= \int s(x_t | u_t, x_{t-1}) \ln \frac{s(x_t | u_t, x_{t-1})}{I_s(x_t | u_t, x_{t-1})} dx_t, \\ &= \int \mathcal{N}_{x_t}(\mu_t, \Sigma_t) [\mu_t' \Sigma_t^{-1} x_t - \frac{1}{2} \mu_t' \Sigma_t^{-1} \mu_t] dx_t, \\ &= \frac{1}{2} (Ax_{t-1} + Bu_t)' \Sigma_t^{-1} (Ax_{t-1} + Bu_t), \end{aligned} \quad (20)$$

where $s(x_t | u_t, x_{t-1}) \rightsquigarrow \mathcal{N}_{x_t}(\mu_t, \Sigma_t)$ as specified by Equation (11), $I_s(x_t | u_t, x_{t-1}) = \mathcal{N}_{x_t}(0, \Sigma_t)$ as specified by Equation (14), and where $\ln \frac{s(x_t | u_t, x_{t-1})}{I_s(x_t | u_t, x_{t-1})} = [\mu_t' \Sigma_t^{-1} x_t - \frac{1}{2} \mu_t' \Sigma_t^{-1} \mu_t]$. The evaluation of $\beta_2(u_t, x_{t-1})$ is based on the assumed form for $-\ln(\gamma(x_t))$ as specified in Equation (16). It

can be evaluated as follows

$$\begin{aligned}
\beta_2(\mathbf{u}_t, \mathbf{x}_{t-1}) &= - \int s(\mathbf{x}_t | \mathbf{u}_t, \mathbf{x}_{t-1}) \ln(\gamma(\mathbf{x}_t)) d\mathbf{x}_t, \\
&= \int \mathcal{N}_{\mathbf{x}_t}(\boldsymbol{\mu}_t, \boldsymbol{\Sigma}_t) \left[0.5 \left(\mathbf{x}_t' \mathbf{M}_t \mathbf{x}_t + \mathbf{V}_t \mathbf{x}_t + \boldsymbol{\omega}_t \right) \right], \\
&= 0.5 \left\{ (\mathbf{A} \mathbf{x}_{t-1} + \mathbf{B} \mathbf{u}_t)' \mathbf{M}_t (\mathbf{A} \mathbf{x}_{t-1} + \mathbf{B} \mathbf{u}_t) + \text{tr}(\mathbf{M}_t \boldsymbol{\Sigma}_t) + \boldsymbol{\omega}_t \right. \\
&\quad \left. + \mathbf{V}_t (\mathbf{A} \mathbf{x}_{t-1} + \mathbf{B} \mathbf{u}_t) \right\}. \tag{21}
\end{aligned}$$

The integrals in Equations (20) and (21) are special cases of the general multiple integral given in Theorem (10.5.1) in [10]. Hence for more details on their evaluation the interested reader is referred to [10]. Now according to Equation (6) the function $\gamma(\mathbf{x}_{t-1})$ has to satisfy the following identity,

$$\begin{aligned}
\gamma(\mathbf{x}_{t-1}) &= \int \mathcal{I}_{\mathbf{c}}(\mathbf{u}_t | \mathbf{x}_{t-1}) \exp[-\beta_1(\mathbf{u}_t, \mathbf{x}_{t-1}) - \beta_2(\mathbf{u}_t, \mathbf{x}_{t-1})] d\mathbf{u}_t, \\
&= \int \mathcal{N}_{\mathbf{u}_t}(0, \boldsymbol{\Gamma}_t) \exp \left[-0.5 \left\{ (\mathbf{A} \mathbf{x}_{t-1} + \mathbf{B} \mathbf{u}_t)' (\boldsymbol{\Sigma}_t^{-1} + \mathbf{M}_t) (\mathbf{A} \mathbf{x}_{t-1} + \mathbf{B} \mathbf{u}_t) \right. \right. \\
&\quad \left. \left. + \mathbf{V}_t (\mathbf{A} \mathbf{x}_{t-1} + \mathbf{B} \mathbf{u}_t) + \text{tr}(\mathbf{M}_t \boldsymbol{\Sigma}_t) + \boldsymbol{\omega}_t \right\} \right] d\mathbf{u}_t, \\
&= \exp \left[-0.5 \left(\mathbf{x}_{t-1}' \mathbf{A}' (\boldsymbol{\Sigma}_t^{-1} + \mathbf{M}_t) \mathbf{A} \mathbf{x}_{t-1} + \mathbf{V}_t \mathbf{A} \mathbf{x}_{t-1} + \text{tr}(\mathbf{D} \mathbf{M}_t) \mathbf{x}_{t-1} + \boldsymbol{\omega}_t \right) \right] \\
&\quad \times \int \exp \left[-0.5 \left\{ \mathbf{u}_t' [\boldsymbol{\Gamma}^{-1} + \mathbf{B}' (\boldsymbol{\Sigma}_t^{-1} + \mathbf{M}_t) \mathbf{B}] \mathbf{u}_t + 2 \mathbf{u}_t' \left(\mathbf{B}' (\boldsymbol{\Sigma}_t^{-1} + \mathbf{M}_t) \mathbf{A} \mathbf{x}_{t-1} \right. \right. \right. \\
&\quad \left. \left. \left. + 0.5 \mathbf{B}' \mathbf{V}_t + 0.5 \text{tr}(\mathbf{G} \mathbf{M}_t) \right) \right\} \right] d\mathbf{u}_t.
\end{aligned}$$

Completing the square over \mathbf{u}_t in the above equation gives,

$$\begin{aligned}
\gamma(\mathbf{x}_{t-1}) &= \exp \left(-0.5 \mathbf{x}_{t-1}' \mathbf{A}' \left\{ (\boldsymbol{\Sigma}_t^{-1} + \mathbf{M}_t) - (\boldsymbol{\Sigma}_t^{-1} + \mathbf{M}_t) \mathbf{B} [\boldsymbol{\Gamma}^{-1} + \mathbf{B}' (\boldsymbol{\Sigma}_t^{-1} + \mathbf{M}_t) \mathbf{B}]^{-1} \right. \right. \\
&\quad \times \left. \left. \mathbf{B}' (\boldsymbol{\Sigma}_t^{-1} + \mathbf{M}_t) \right\} \mathbf{A} \mathbf{x}_{t-1} - 0.5 \left\{ \mathbf{V}_t \mathbf{A} + \text{tr}(\mathbf{D} \mathbf{M}_t) - 2(0.5 \mathbf{V}_t \mathbf{B} + 0.5 \text{tr}(\mathbf{G} \mathbf{M}_t)) \right. \right. \\
&\quad \times \left. \left. [\boldsymbol{\Gamma}^{-1} + \mathbf{B}' (\boldsymbol{\Sigma}_t^{-1} + \mathbf{M}_t) \mathbf{B}]^{-1} \mathbf{B}' (\boldsymbol{\Sigma}_t^{-1} + \mathbf{M}_t) \mathbf{A} \right\} \mathbf{x}_{t-1} - 0.5 \left\{ \boldsymbol{\omega}_t \right. \right. \\
&\quad \left. \left. - \{0.5 \mathbf{V}_t \mathbf{B} + 0.5 \text{tr}(\mathbf{G} \mathbf{M}_t)\} [\boldsymbol{\Gamma}^{-1} + \mathbf{B}' (\boldsymbol{\Sigma}_t^{-1} + \mathbf{M}_t) \mathbf{B}]^{-1} \{0.5 \mathbf{B}' \mathbf{V}_t + 0.5 \text{tr}(\mathbf{G} \mathbf{M}_t)\} \right\} \right. \\
&\quad \left. + \ln | \mathbf{I} + (\mathbf{B} \boldsymbol{\Gamma}^{0.5})' (\boldsymbol{\Sigma}_t^{-1} + \mathbf{M}_t) (\mathbf{B} \boldsymbol{\Gamma}^{0.5}) | \right). \tag{22}
\end{aligned}$$

Noting that $-\ln(\gamma(x_{t-1})) = 0.5x'_{t-1}M_{t-1}x_{t-1} + 0.5V_{t-1}x_{t-1} + 0.5\omega_{t-1}$, it can be seen that the identity is satisfied for M_{t-1} , V_{t-1} and ω_{t-1} as defined in Equations (17), (18), and (19) respectively. Hence the claimed quadratic nature of the performance function is proved.

Now it is straight forward to calculate the optimal control law. This can be done through the exploitation of Equation (16) in Equation (6) which leads to the optimal control law specified in the following theorem.

Theorem 2: The optimal control law minimising the Kulback-Leibler distance defined in Equation (3) is given by

$$\mathbf{u}_t^* = -\mathbf{K}_t \mathbf{x}_{t-1} - \mathbf{Z}_t, \quad (23)$$

where

$$\mathbf{K}_t = \Gamma_t [\mathbf{B}'(\Sigma_t^{-1} + \mathbf{M}_t)\mathbf{A}], \quad (24)$$

$$\mathbf{Z}_t = \Gamma_t \frac{1}{2} [\text{tr}(\mathbf{G}\mathbf{M}_t) + \mathbf{V}_t \mathbf{B}], \quad (25)$$

and where

$$\Gamma_t = [\mathbf{B}'(\Sigma_t^{-1} + \mathbf{M}_t)\mathbf{B} + \Gamma^{-1}]^{-1}. \quad (26)$$

Proof: The optimal control law of the fully probabilistic control problem is defined in Equation (6), repeated here

$$\mathbf{c}^*(\mathbf{u}_t | \mathbf{x}_{t-1}) = \frac{{}^I \mathbf{c}(\mathbf{u}_t | \mathbf{x}_{t-1}) \exp[-\beta_1(\mathbf{u}_t, \mathbf{x}_{t-1}) - \beta_2(\mathbf{u}_t, \mathbf{x}_{t-1})]}{\gamma(\mathbf{x}_{t-1})} \quad (27)$$

For the system defined in Equation (11), and the elements in Equations (12), (13), (14), and (15), the numerator of Equation (27) evaluates to,

$$\begin{aligned} & {}^I \mathbf{c}(\mathbf{u}_t | \mathbf{x}_{t-1}) \exp[-\beta_1(\mathbf{u}_t, \mathbf{x}_{t-1}) - \beta_2(\mathbf{u}_t, \mathbf{x}_{t-1})] = \exp \left\{ -0.5 \left[\mathbf{u}'_t \Gamma^{-1} \mathbf{u}_t + (\mathbf{A}\mathbf{x}_{t-1} + \mathbf{B}\mathbf{u}_t)'(\Sigma_t^{-1} + \mathbf{M}_t) \right. \right. \\ & \left. \left. \times (\mathbf{A}\mathbf{x}_{t-1} + \mathbf{B}\mathbf{u}_t) + \text{tr}(\mathbf{D}\mathbf{M}_t)\mathbf{x}_{t-1} + \text{tr}(\mathbf{G}\mathbf{M}_t)\mathbf{u}_t + \omega_t + \mathbf{V}_t(\mathbf{A}\mathbf{x}_{t-1} + \mathbf{B}\mathbf{u}_t) \right] \right\} \\ & = \exp \left\{ -0.5 \left[\mathbf{x}'_{t-1} \mathbf{A}'(\Sigma_t^{-1} + \mathbf{M}_t)\mathbf{A}\mathbf{x}_{t-1} + \text{tr}(\mathbf{D}\mathbf{M}_t)\mathbf{x}_{t-1} + \omega_t + \mathbf{V}_t \mathbf{A}\mathbf{x}_{t-1} \right] \right\} \\ & \times \exp \left\{ -0.5 \left[\mathbf{u}'_t (\Gamma^{-1} + \mathbf{B}'(\Sigma_t^{-1} + \mathbf{M}_t)\mathbf{B})\mathbf{u}_t + 2\mathbf{u}'_t \left\{ \mathbf{B}'(\Sigma_t^{-1} + \mathbf{M}_t)\mathbf{A}\mathbf{x}_{t-1} + 0.5\mathbf{B}'\mathbf{V}_t \right. \right. \right. \\ & \left. \left. \left. + 0.5\text{tr}(\mathbf{G}\mathbf{M}_t) \right\} \right] \right\}. \end{aligned} \quad (28)$$

The denominator of Equation (27) is evaluated in Equation (22). It can be rewritten as follows,

$$\begin{aligned}
\gamma(x_{t-1}) &= \exp \left[-0.5 \left(x'_{t-1} A' (\Sigma_t^{-1} + M_t) A x_{t-1} + V_t A x_{t-1} + \text{tr}(D M_t) x_{t-1} + \omega_t \right) \right. \\
&\times \exp \left\{ 0.5 \left[x'_{t-1} A' (\Sigma_t^{-1} + M_t) B [\Gamma^{-1} + B' (\Sigma_t^{-1} + M_t) B]^{-1} B' (\Sigma_t^{-1} + M_t) A x_{t-1} \right. \right. \\
&+ 2(0.5 V_t B + 0.5 \text{tr}(G M_t)) [\Gamma^{-1} + B' (\Sigma_t^{-1} + M_t) B]^{-1} B' (\Sigma_t^{-1} + M_t) A x_{t-1} \\
&+ \left. \left. \{0.5 V_t B + 0.5 \text{tr}(G M_t)\} [\Gamma^{-1} + B' (\Sigma_t^{-1} + M_t) B]^{-1} \{0.5 B' V_t + 0.5 \text{tr}(G M_t)\} \right] \right\} \\
&+ \left. \ln | I + (B \Gamma^{0.5})' (\Sigma_t^{-1} + M_t) (B \Gamma^{0.5}) \right\}. \tag{29}
\end{aligned}$$

Using Equations (28) and (29) in Equation (27) yields

$$\begin{aligned}
c^*(u_t | x_{t-1}) &= \exp \left\{ -0.5 \left[u'_t (\Gamma^{-1} + B' (\Sigma_t^{-1} + M_t) B) u_t + 2 u'_t \left(B' (\Sigma_t^{-1} + M_t) A x_{t-1} + 0.5 B' V_t \right. \right. \right. \\
&+ \left. \left. \left. 0.5 \text{tr}(G M_t) \right) + x'_{t-1} A' (\Sigma_t^{-1} + M_t) B [\Gamma^{-1} + B' (\Sigma_t^{-1} + M_t) B]^{-1} B' (\Sigma_t^{-1} + M_t) A x_{t-1} \right. \right. \\
&+ 2(0.5 V_t B + 0.5 \text{tr}(G M_t)) [\Gamma^{-1} + B' (\Sigma_t^{-1} + M_t) B]^{-1} B' (\Sigma_t^{-1} + M_t) A x_{t-1} \\
&+ \left. \left. \left. \{0.5 V_t B + 0.5 \text{tr}(G M_t)\} [\Gamma^{-1} + B' (\Sigma_t^{-1} + M_t) B]^{-1} \{0.5 B' V_t + 0.5 \text{tr}(G M_t)\} \right] \right\} \\
&+ \left. \ln | I + (B \Gamma^{0.5})' (\Sigma_t^{-1} + M_t) (B \Gamma^{0.5}) \right\}. \tag{30}
\end{aligned}$$

Completing the square in Equation (30) for u_t gives

$$c^*(u_t | x_{t-1}) = \mathcal{N}_{u_t}(-K_t x_{t-1} - Z_t, \Gamma_t), \tag{31}$$

with K_t and Z_t as defined in Equations (24) and (25) respectively. The covariance matrix is given by

$$\Gamma_t = [\Gamma^{-1} + B' (\Sigma_t^{-1} + M_t) B]^{-1}. \tag{32}$$

Hence the generalised probabilistic optimal control law (23) is proved.

Remark 5: Compared to the standard randomised controller in FPD [16], the mean of the derived randomised controller of the GFPCD method is shifted by Z_t as can be seen from Equation (31). The introduction of this shift is the consequence of the consideration of input dependent noise. As can be seen from (25), the introduced shift Z_t is dependent on the parameters of the noise model (13) and the equation of cautiousness (18), adding the desirable caution feature to obtained control inputs. Although the derived probabilistic controller for the assumed probability distributions maintains the standard form of linear quadratic controllers, it is more

exploratory due to its probabilistic nature. Ideally control inputs should be sampled from the obtained pdf of the randomised controller. This however results in slightly worse control quality, but randomisation makes the controller more explorative.

The steady state solution of the derived generalised probabilistic optimal control law can be obtained using standard methods discussed in [23]. Here $K_t, Z_t, M_t,$ and V_t become constant and defined as follows,

$$K = [B^T M B + B^T \Sigma^{-1} B + \Gamma^{-1}]^{-1} [B^T \Sigma^{-1} A + B^T M A], \quad (33)$$

$$Z = [B^T M B + B^T \Sigma^{-1} B + \Gamma^{-1}]^{-1} \frac{1}{2} [\text{tr}[GM] + VB], \quad (34)$$

$$\begin{aligned} M &= A^T \Sigma^{-1} A + A^T M A \\ &- [B^T \Sigma^{-1} A + B^T M A]^T [B^T M B + B^T \Sigma^{-1} B + \Gamma^{-1}]^{-1} [B^T \Sigma^{-1} A + B^T M A], \end{aligned} \quad (35)$$

$$\begin{aligned} V &= \text{tr}[DM] + VA \\ &- [\text{tr}[GM] + VB] [B^T M B + B^T \Sigma^{-1} B + \Gamma^{-1}]^{-1} [B^T \Sigma^{-1} A + B^T M A]. \end{aligned} \quad (36)$$

The probabilistic uncertain optimal controller derived in this paper, is based on the solution of the Riccati equation as well as the solution of the equation of cautiousness, leading to the extra terms in the GFPCD compared to conventional FPD. To obtain the solution of the Riccati equation of the probabilistic uncertain controller (35), standard methods proposed in [23] can be implemented. The steady state solution of the cautiousness equation (36) can be obtained from the non-steady state equation of cautiousness (18), repeated here

$$V_{t-1} = \text{tr}[DM_t] + V_t A - [\text{tr}[GM_t] + V_t B] \Gamma_t [B^T \Sigma_t^{-1} A + B^T M_t A]. \quad (37)$$

By substituting the steady state matrix M and reversing the direction of time, Equation (37) can be modified to read,

$$\begin{aligned} V_t &= \text{tr}[DM] + V_{t-1} A \\ &- [\text{tr}[GM] + V_{t-1} B] [B^T M B + B^T \Sigma^{-1} B + \Gamma^{-1}]^{-1} [B^T \Sigma^{-1} A + B^T M A]. \end{aligned} \quad (38)$$

Then by starting the solution with $V_0 = 0$, Equation (38) is iterated until a stationary solution is obtained.

IV. CAUTIOUS PROBABILISTIC CONTROL ALGORITHM

The generalised fully probabilistic control law derived in Section III-B minimises the Kullback–Leibler divergence between the actual joint probability density function of the closed loop system and an ideal joint probability density function and at the same time minimises the variance of the estimated probabilistic model of the system dynamics. This allows considering model uncertainty in the derivation of the optimal control law. The description below is appropriate for direct application to uncertain linear stochastic control problems of the form stated in Section III-A.

- 1) Estimate the pdf of the stochastic model described by (11) as discussed in Section II-C.
- 2) Specify the ideal density functions of both the state and control inputs.
- 3) Obtain the steady state solution of the Riccati equation (35).
- 4) Obtain the steady state solution of the equation of cautiousness (36).
- 5) Use Equations (33) and (34) in Equation (23) to calculate the optimal control law.

V. AN APPLICATION ON A LINEARISED F-16 FIGHTER AIRCRAFT

In this section a discrete-time reduced linear model of F-16 fighter aircraft without actuator dynamics [30] will be used to demonstrate the proposed GFPCD method. This model simulates the linearised dynamics of a real F-16 aircraft based on the following description,

$$\begin{bmatrix} \alpha_{t+1} \\ q_{t+1} \end{bmatrix} = \begin{bmatrix} 0.9935 & 0.0093 \\ -0.0156 & 0.9912 \end{bmatrix} \begin{bmatrix} \alpha_t \\ q_t \end{bmatrix} + \begin{bmatrix} -1.8861e^{-5} \\ -0.0011 \end{bmatrix} u_t + \tilde{\epsilon}_t, \quad (39)$$

where

$$E[\tilde{\epsilon}_t \tilde{\epsilon}_t'] = \begin{bmatrix} 0.000001 & 0 \\ 0 & 0.000001 \end{bmatrix}, \quad (40)$$

and where α the angle of attack and q the pitch rate are the two longitudinal states of the aircraft. In this example only the elevator deflection angle is used as a control input, u to stabilise the aircraft and maintain the wings-level flight condition. The study in [30] has not considered external disturbance, i.e $\tilde{\epsilon}_t = 0$, however, we have explicitly added and considered the external noise in our simulation.

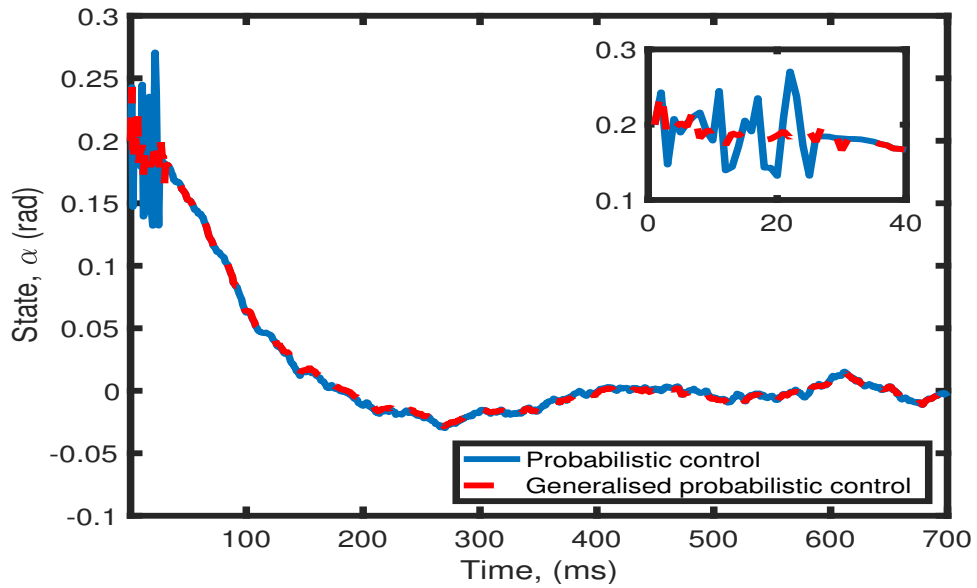
For purposes of comparison, the solution to this control problem is obtained using both the classical probabilistic control method [16] and the generalised control method proposed in Section III-B and summarised in Section IV. For both methods, two GLMs were used to estimate the Gaussian pdf of the stochastic model described by Equation (39). For the generalised probabilistic control method the input dependent covariance of the pdf of the system states, Σ_t is obtained using another GLM as discussed in Section II-C. The aircraft is initially in states, $\alpha(0) = 0.2$ and $q(0) = 0.1$ and the objective is to regulate the aircraft to achieve zero state values. Hence, the ideal pdf of the system state is assumed to be normal with zero mean and Σ_t covariance as estimated by the generalised linear model. The distribution of the ideal controller is also taken to be normal with zero mean and a variance of 0.01 which specifies the allowed range of optimal control inputs. The optimisation problem of the two methods starts with initialisation of the parameters of the generalised linear models. The generalised linear models were then estimated online and never subjected to an initial off line training phase. Both experiments start from the same initial state values and same noise sequence as defined in Equation (40).

Figure 2 shows the evolution of the states of the aircraft as obtained from both methods. It can be seen that the consideration of the input dependent noise and model uncertainty in

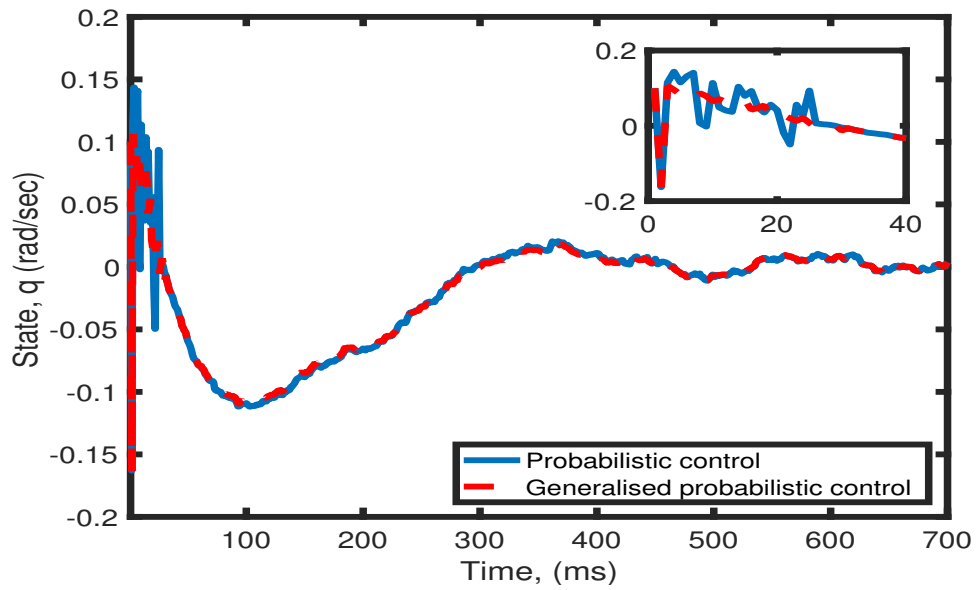
the proposed GFPCD method yields to better control results in the transient period where the parameters of the estimated models are still not converged. In particular, the proposed GFPCD controller can be seen from Figure 2 to apply conservative control which acts to dampen the initial large transients, compared to the existing FPD approach where the transients persist with much larger amplitude. This can be clearly seen from the magnified figures in figure 2.

VI. CONCLUSION

In this paper, an optimal probabilistic control framework is considered for linear stochastic uncertain systems with Gaussian random inputs that are dependent on the input and state values. A Kullback–Leibler divergence between the actual joint probability density function of the system dynamics and an ideal probability density function is used as a performance measure rather than the mean variance. Recursive optimal control laws are developed using dynamic programming such that the conditional output pdfs of the closed–loop systems can be made to follow the desired one. The uncertainty of estimated models is assumed to be input dependent and taken into consideration in the derivation of the optimal control law. Compared with previous works, the main results in this paper have the following distinct features: our generalised framework which incorporates input-dependent noise leads to additional terms in the control equations which depend on the parameters of the estimated uncertainty in the model dynamics; the existence of these extra terms collectively provide a more conservative control model damping out transients due to initial conditions and to model uncertainty, which is also borne out by the numerical simulation based on a real control problem in Section V. As a result, the control algorithm developed in this paper has a potential application for practical stochastic and deterministic systems with unknown dynamics. The F-16 fighter aircraft application demonstrated the effectiveness and feasibility of the proposed generalised probabilistic control, and showed the difference the extra control terms make on reducing transient fluctuations of the conventional approach.



(a)



(b)

Fig. 2. Controlled linearised F-16 fighter aircraft: (a) Angle of attack, α from the standard and generalised control methods. (b) Pitch rate, q from the standard and generalised control methods. In comparison to the standard FPD control model where the transients persist with much larger amplitude, the magnified figures show that the proposed GFPCD provides a more conservative control model damping out transients due to initial conditions and to the uncertainty in the parameters of the estimated state models.

PLACE
PHOTO
HERE

Randa Herzallah received her B.Sc. in Industrial Engineering and M.Sc. in Industrial Engineering/Manufacturing and Design at Jordan University, Jordan, and later her Ph.D. from Aston University, Birmingham, UK. She previously held the positions of Associate Professor at Al-AHliyya Amman University and Professor at Al-Balqa Applied University. She is currently a Senior Lecturer in the Faculty of Engineering and Applied Science at Aston University, Birmingham, UK. Her general research focus is on the general topic of Systems Modelling, Design and Control. Her current research interests are nonlinear control theory, probabilistic and stochastic control, intelligent control, pinning control, adaptive control, distributed control, and optimal control.

REFERENCES

- [1] Anderson B. D. O. and J. B. Moore. *Linear Optimal Control*, Prentice Hall, Englewood Cliffs, NJ, (1971).
- [2] Bernstein D. S. and W. M. Haddad. "Robust stability and performance via fixed order dynamic compensation with guaranteed cost bounds," *Mathematics of Control, Signals and Systems*, Vol. 3, pp. 139–163 (1990).
- [3] Tao Bian, Yu Jiang, and Zhong–Ping Jiang. "Adaptive dynamic programming for stochastic systems with state and control dependent noise," *IEEE Transactions on Automatic Control*, Vol. 61, pp. 4170–4175 (2016).
- [4] Blackmore L., M. Ono, A. Bektassov, and B. C. Williams. "A probabilistic particle-control approximation of chance-constrained stochastic predictive control," *IEEE Transactions on Robotics*, Vol. 26, No. 3, pp. 502–517 (2010).
- [5] Chang S. S. L. and T. K. C. Peng. "Adaptive guaranteed cost control of systems with uncertain parameters," *IEEE Transactions on Automatic Control*, Vol. 17, pp. 474–483 (1972).
- [6] Chen B. S. and Weihai Zhang. "Stochastic H_2/H_∞ control with state dependent noise," *IEEE Transactions on Automatic Control*, Vol. 49, pp. 45–57 (2004).
- [7] Everdij M. H. C. and H. A. P. Blom. "Embedding adaptive JLQG into LQ martingale control with a completely observable stochastic control matrix," *IEEE Transactions on Automatic Control*, Vol. 41, No. 3, pp. 424–430 (1996).
- [8] Fabri S. G. and V. Kadiramanathan. *Functional Adaptive Control: An Intelligent Systems Approach*, Springer-Verlag, February (2001).
- [9] Fu M., C. E. De Souza, and L. Xie. " H_∞ estimation for uncertain systems," *International Journal of Robust Nonlinear Control*, Vol. 2, pp. 87–105 (1992).
- [10] Franklin A. Graybill. *Matrices with Applications in Statistics*, Duxbury Classic. Duxbury Press (2001).
- [11] Randa Herzallah and D. Lowe. "A Bayesian perspective on stochastic neuro control," *IEEE Transactions on Neural Networks*, Vol. 19, No. 5, pp. 914–924, May (2008).
- [12] Randa Herzallah. "Probabilistic dhp adaptive critic for nonlinear stochastic control systems," *Neural Networks*, Vol. 42, pp. 74–82 (2013).
- [13] Randa Herzallah. "Fully probabilistic control systems for stochastic nonlinear control systems with input dependent noise," *Neural Networks*, Vol. 63, pp. 199–207 (2015).
- [14] Randa Herzallah and Miroslav. Kárný. "Fully probabilistic control design in an adaptive critic framework," *Neural Networks*, Vol. 24, No. 11, pp. 1128–1135 (2011).
- [15] Yulin Huang, Weihai Zhang, and Huanshui Zhang. "Infinite horizon linear quadratic optimal control for discrete time stochastic systems," *Asian Journal of Control*, Vol. 10, pp. 608–615 (2008).
- [16] Miroslav Kárný. "Towards fully probabilistic control design," *Automatica*, Vol. 32, No. 12, pp. 1719–1722 (1996).
- [17] Miroslav Kárný and Tomáš Kroupa. "Axiomatisation of fully probabilistic design," *Information Sciences*, Vol. 186, No. 1, pp. 105–113 (2012).

- [18] Kosmidou O. I. “Generalized riccati equations associated with guaranteed cost control: An overview of solutions and features,” *Applied Mathematics and Computation*, Vol. 191, pp. 511–520 (2007).
- [19] McLane P. “Optimal stochastic control of linear systems with state- and control-dependent disturbances,” *IEEE Transactions on Automatic Control*, Vol. 16, pp. 793–798 (1971).
- [20] Miller W. T., R. S. Sutton, and P. J. Werbos, editors. *Neural Networks for Control*, MIT Press, Cambridge, Massachusetts, London, England (1990).
- [21] John B. Moore, Xun Yu Zhou, and Andrew E. B. Lim. “Discrete time LQG controls with control dependent noise,” *Systems and Control Letters*, Vol. 36, pp. 199–206 (1999).
- [22] Yugang Niu, D. W.C. Ho, and Xingyu Wang. “Robust H_∞ control for nonlinear stochastic systems: A sliding-mode approach,” *IEEE Transactions on Automatic Control*, Vol. 53, No. 7, pp. 1695–1701 (2008).
- [23] Ogata K. *Discrete Time Control Systems*, Prentice Hall, Inc., EnglewoodCliffs, New Jersey (1987).
- [24] Peterka V. Bayesian system identification, In P. Eykhoff, editor, *Trends and Progress in System Identification*, pages 239–304. Pergamon Press, Oxford, 1981.
- [25] Zhang Qixia. H_2/H_∞ “control for stochastic systems with delay,” *International Journal of Control*, Vol. 88, pp. 2554–2562 (2015).
- [26] Solo V. “Stochastic adaptive control and martingale limit theory,” *IEEE Transactions on Automatic Control*, Vol. 35, No. 1, pp. 66–71 (1990).
- [27] Niladri Sekhar Tripathy, Indra N. Kar, and Kolin Paul. “Model based robust control law for linear event-triggered system,” *Asian Journal of Control*, Vol. 18, No. 5, pp. 1765–1780 (2016).
- [28] Werbos P. J. “Approximate dynamic programming for real-time control and neural modeling,” In D. A. White and D. A. Sofge, editors, *Handbook of Intelligent Control*, chapter 13, pages 493–526. Multiscience Press, Inc, New York, N.Y. (1992).
- [29] Weihai Zhang, Juné Feng, Bor-Sen Chen, and Zhaolin Cheng. “Nonlinear stochastic H_2/H_∞ control with state dependent noise,” In *2005 American Control Conference*, pp. 380–385, Portlan, OR , USA (2005).
- [30] Zhou Y., E. J. Van Kampen, and Q. P. Chu. “Incremental approximate dynamic programming for nonlinear flight control design,” In *Proceedings of the 3rd CEAS EuroGNC: Specialist Conference on Guidance, Navigation and Control*, pp. 1–18, Toulouse, France (2015).