Received: 30 January 2021 | Revised: 9 June 2022 | Accepted: 1 October 2022

DOI: 10.1002/asjc.2971

RESEARCH ARTICLE

WILEY

Fully probabilistic control for uncertain nonlinear stochastic systems

Ana Zafar[1] | **Randa Herzallah**[2]

[1]Systems Analytics Research Institute, Aston University, Birmingham, UK

[2]Warwick Mathematics Institute, University of Warwick, Coventry, UK

Correspondence
Randa Herzallah, Warwick Mathematics Institute, University of Warwick, Coventry, UK.
Email: Randa.Herzallah@warwick.ac.uk

Funding information
Leverhulme Trust, Grant/Award Number: RPG-2017-337

Abstract

This paper develops a novel probabilistic framework for stochastic nonlinear and uncertain control problems. The proposed framework exploits the Kullback–Leibler divergence to measure the divergence between the distribution of the closed-loop behavior of a dynamical system and a predefined ideal distribution. To facilitate the derivation of the analytic solution of the randomized controllers for nonlinear systems, transformation methods are applied such that the dynamics of the controlled system becomes affine in the state and control input. Additionally, knowledge of uncertainty is taken into consideration in the derivation of the randomized controller. The derived analytic solution of the randomized controller is shown to be obtained from a generalized state-dependent Riccati solution that takes into consideration the state- and control-dependent functional uncertainty of the controlled system. The proposed framework is demonstrated on an inverted pendulum on a cart problem, and the results are obtained.

KEYWORDS

functional uncertainty, optimal control, probabilistic control, stochastic systems

1 | INTRODUCTION

In the contemporary field of control engineering, stochastic control is of much interest due to the stochastic and uncertain nature of current and future real-world systems [1–4]. Moreover, real-world systems are characterized by nonlinearity and high dimensionality. Therefore, the design of robust cautious controllers that take knowledge of noises and uncertainties into consideration is of paramount importance. Thus, a number of methods that consider the stochastic nature of the controlled system have been proposed and developed in the control literature. For example, the control problem that considers state-dependent noise using the H_2/H_∞ control methodology is studied in Chen and Zhang [5]. The work in Gajic and Losada [6] solved the algebraic Ricatti equation of a linear quadratic optimal control problem with state-dependent noise by the means of algebraic Lyapunov iterations. In addition, the infinite horizon linear quadratic control problem with state- and control-dependent noise is studied in Mukaidani et al. [7]. Moreover, the regulation problem of stochastic dynamical systems is developed following a fully probabilistic design (FPD) approach [8]. This approach considers the full distribution of the stochastic system dynamics for the derivation of randomized controllers. This approach has then been further developed to consider various aspects of stochastic and uncertain systems [3, 9–11]. In FPD, the Kullback–Leibler divergence (KLD), defined in (1), is implemented to measure the distance between the actual and ideal joint probability density functions.

$$\mathcal{D}(f\|f^I) = \int f(D)\ln\left(\frac{f(D)}{f^I(D)}\right)dD, \tag{1}$$

where $D = \{x_1, \ldots, x_H, u_1, \ldots, u_H\}$, with x and u being the state and control input vector, respectively, and H is the control horizon. The FPD minimizes the KLD given in (1) by designing a randomized controller that brings the actual joint pdf of the closed-loop system closer to the ideal joint pdf of the closed-loop system. This method was further extended by Herzallah [3] to derive a generalized FPD controller that considers stochastic systems which are affected by functional uncertainty. The existence of functional uncertainty is unavoidable and arises from the poor modeling of complex systems. Thus, an intelligent control technique that incorporates functional uncertainty in the optimization of the randomized controller was proposed by Herzallah [3]. This method was shown to be effective in improving the performance of the stochastic controlled system which also assumed to have unknown dynamics and are thus estimated online. The consideration of functional uncertainty in the design of the randomized controller yielded less transient overshoot due to the development of a cautious controller. However, the derived generalized FPD in Herzallah [3] is limited to linear stochastic systems.

In this paper, a generalized FPD is derived for even more complex stochastic systems by extending it to systems that are governed by nonlinear dynamics. Since the physical description of most real-world systems is unknown, our proposed adaptive control method will utilize the neural network methods to estimate the dynamics of the controlled system. In particular, multilayer perceptron (MLP) neural networks will be implemented in our work to estimate the system dynamics of the controlled system as well as estimate the functional uncertainty of the optimization process. The stochasticity, nonlinearity, and uncertainty of the considered class of dynamical systems emphasize the importance of the consideration of the knowledge of functional uncertainties in the derivation of the optimal control laws. On the other hand, the derivation of an analytic control solution for nonlinear systems cannot be obtained in a FPD framework. This is due to the nonlinearity of the parameters of the probability density functions that characterize the system state and control input. Consequently, this paper proposes a novel approach to the derivation of analytic solutions to the randomized controllers for nonlinear systems that at the same time takes knowledge of uncertainty into consideration in the optimization process of the controllers. The derivation of analytic control solutions is facilitated by the means of transformation methods. The introduced novelty allows the FPD to be extended to more realistic control problems which are nonlinear and uncertain and does not require linearization of the systems. It should be noted that in comparison with the conventional Riccati equation obtained from the conventional FPD, our derived Riccati equations is state-dependent Riccati equation (SDRE) due to the dependency of the nonlinear parameters of the state and control distributions on previous state values. Furthermore, the proposed methodology can be considered to fall under model-based control approaches where instead of modeling the dynamics of the system by a deterministic equation, a probabilistic description is obtained through the utilization of neural network models. For more information on model-based versus model-free control methods and the challenges they encounter, the readers are referred to Hou and Wang [12] and the references therein.

2 | PRELIMINARIES

2.1 | FPD objective

The FPD is an optimal control method which is based on the minimization of a predefined performance index. This performance index is derived from the KLD measure (1) which is the foundation of the FPD controller. The objective of the FPD is to control the distribution of the system state to a predefined desired state distribution. This objective can be achieved by minimizing the discrepancy between the joint distribution of the system state and control input, $f(x_k, u_k|x_{k-1})$, and a predefined ideal joint distribution, $f^I(x_k, u_k|x_{k-1})$. Using the chain rule, $f(x_k, u_k|x_{k-1})$ can be factorized as follows,

$$f(x_k, u_k|x_{k-1}) = s(x_k|u_k, x_{k-1})c(u_k|x_{k-1}), \tag{2}$$

where $s(x_k|u_k, x_{k-1})$ is the conditional distribution of the dynamics of the system state and $c(u_k|x_{k-1})$ is the conditional distribution of the controller. Similarly, the ideal joint probability distribution of the closed-loop system can be factorized as follows,

$$f^I(x_k, u_k|x_{k-1}) = s^I(x_k|u_k, x_{k-1})c^I(u_k|x_{k-1}), \tag{3}$$

where $s^I(x_k|u_k, x_{k-1})$ is the ideal distribution of the system state and $c^I(u_k|x_{k-1})$ is the ideal distribution of the controller. The minimization of the KLD is attained by finding a probabilistic control law $c(u_k|x_{k-1})$ which regulates the closed-loop system and brings it as close as possible to the ideal joint probability distribution $f^I(x_k, u_k|x_{k-1})$ of the closed-loop system.

To re-emphasize, the objective in this paper is to derive a randomized controller $c(u_k|x_{k-1})$ that minimizes the KLD given in (1) subject to the actual joint distribution of the system dynamics given in (2) and an ideal joint distribution given in (3). Using the concept of dynamic programming, the minimum cost-to-go function resulting from the minimization of (1) with respect to admissible control sequence is shown [9] to be given by the following recurrence equation,

$$-\ln(\gamma(x_{k-1})) = \min_{\{c(u_k|x_{k-1})\}} \int s(x_k|u_k, x_{k-1})c(u_k|x_{k-1}) \times \left[\ln\left(\frac{s(x_k|u_k, x_{k-1})c(u_k|x_{k-1})}{s^I(x_k|u_k, x_{k-1})c^I(u_k|x_{k-1})}\right) - \ln(\gamma(x_k))\right] \times d(x_k, u_k), \quad (4)$$

where $-\ln(\gamma(x_k))$ is the expected minimum cost-to-go function. The optimal control law, $c^*(u_k|x_{k-1})$, as obtained from (4) is then shown to be given by Herzallah and Kárny [10],

$$c^*(u_k|x_{k-1}) = \frac{c^I(u_k|x_{k-1})\exp[-\beta_1(u_k, x_{k-1}) - \beta_2(u_k, x_{k-1})]}{\gamma(x_{k-1})}, \quad (5)$$

where,

$$\beta_1(u_k, x_{k-1}) = \int s(x_k|u_k, x_{k-1})\left(\ln\frac{s(x_k|u_k, x_{k-1})}{s^I(x_k|u_k, x_{k-1})}\right)dx_k, \quad (6)$$

$$\beta_2(u_k, x_{k-1}) = -\int s(x_k|u_k, x_{k-1})\ln(\gamma(x_k))dx_k, \quad (7)$$

$$\gamma(x_{k-1}) = \int c^I(u_k|x_{k-1})\exp[-\beta_1(u_k, x_{k-1}) - \beta_2(u_k, x_{k-1})]du_k. \quad (8)$$

The details of the complete proof of the performance index $-\ln(\gamma(x_{k-1}))$ defined in (4) and the optimal control law $c^*(u_k|x_{k-1})$ given in (5) can be found in Herzallah and Kárny [10].

2.2 | Problem formulation

In real world, many systems are governed by nonlinear dynamics. As such, this paper considers a class of nonlinear discrete-time dynamical stochastic systems that are described by the following state space model,

$$x_k = \tilde{h}(x_{k-1}) + \hat{g}(x_{k-1})u_k + \epsilon_k, \quad (9)$$

where $k = 1, \dots, H$ denotes the discrete-time step, $x_k \in \mathbb{R}^n$ describes the state, and $u_k \in \mathbb{R}^m$ is the control input of the system. Also, $\epsilon_k \in \mathbb{R}^n$ is a Gaussian noise with zero mean and fixed arbitrary covariance, $\bar{\Sigma}$. The nonlinear state vector and control matrix are represented by $\tilde{h}(x_{k-1}) \in \mathbb{R}^n$ and $\hat{g}(x_{k-1}) \in \mathbb{R}^{n\times m}$, respectively. Note that although the nonlinear system model in (9) has zero delay between the control input and state values, it is causal. This is a requirement in the original form of the FPD method [8]. The extension of the FPD method for a delay greater than one can be found in Herzallah [13] and is beyond the scope of the current paper.

For the nonlinear system (9), the presence of the noise ϵ_k means that the previous state and current control input specify the probability distribution of the present state, $s(x_k|u_k, x_{k-1})$, rather than their actual values. To clarify, given the assumption that ϵ_k is a Gaussian noise, the distribution of the present state of the nonlinear system (9) can be characterized by a Gaussian distribution with mean given by $\tilde{h}(x_{k-1}) + \hat{g}(x_{k-1})u_k$ and a global covariance matrix given by $\bar{\Sigma}$. However, this nonlinearity of the parameters of the distribution that characterizes the dynamics of the nonlinear system (9) means that the optimal solution of the randomized controller given in (5) can only be obtained using numerical methods, and an analytic or closed-form solution of the randomized controller cannot be obtained. Therefore, as will become clear from further development, the derivation of the analytic solution of the randomized controller will be facilitated by first transforming Equation (9) such that it becomes nonlinear affine in the system state as follows,

$$x_k = \hat{h}(x_{k-1})x_{k-1} + \hat{g}(x_{k-1})u_k + \epsilon_k, \quad (10)$$

where $\tilde{h}(x_{k-1}) = \hat{h}(x_{k-1})x_{k-1}$ and where,

$$\hat{h}(x_{k-1}) = \begin{bmatrix} \hat{h}_{11}(x_{k-1}) & \dots & \hat{h}_{1n}(x_{k-1}) \\ \vdots & \ddots & \vdots \\ \hat{h}_{n1}(x_{k-1}) & \dots & \hat{h}_{nn}(x_{k-1}) \end{bmatrix}. \quad (11)$$

The functions $\hat{h}(x_{k-1})$ and $\hat{g}(x_{k-1})$ are unknown and are hence required to be estimated. The stochastic evolution of the system state defined in Equation (10) can be captured during the control process by estimating its generative distribution, $s(x_k|x_{k-1}, u_k)$, from the observed data as will be explained in the next section.

3 | PDF ESTIMATION OF THE DYNAMICS OF THE SYSTEM

For the estimation of the probability distribution of the system state variables, two neural networks (both are optimized online) are implemented in this paper to provide predictions for the conditional expectation of the system state and the covariance of the estimation error. As can be seen from Figure 1, the output of the first neural network (an MLP here) provides an estimation for the conditional expectation of the actual state of the system defined in (10) and is given by the following equation,

$$\tilde{x}_k = \text{mlp}(x_{k-1}) = h(x_{k-1})x_{k-1} + g(x_{k-1})u_k, \quad (12)$$

where $h(x_{k-1})$ and $g(x_{k-1})$ are the estimates of the actual state, $\hat{h}(x_{k-1})$, and control, $\hat{g}(x_{k-1})$, matrices, respectively. The parameters of the MLP model (12) are optimized online at each instant of time by computing the sum of squares error between the actual state values x_k as obtained from (10) and the estimated $\tilde{x}_k$ as obtained from (12).

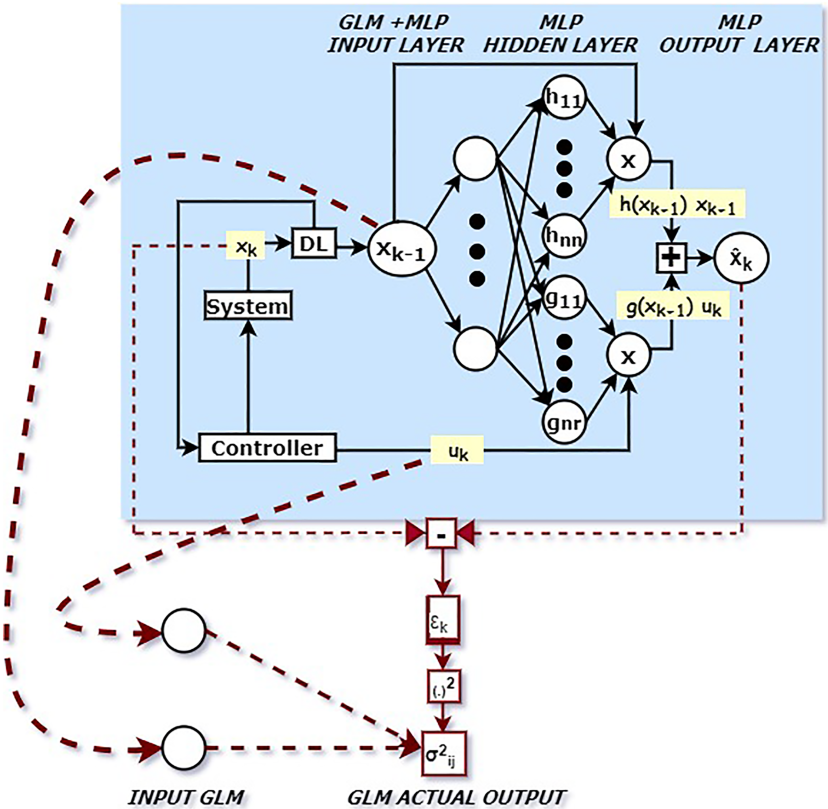

FIGURE 1 The shaded area in blue represents the process of the multilayer perceptron (MLP) for the estimation of the system dynamics. The outer part and dashed lines correspond to the generalized linear neural network model (GLM) process. For both, the MLP and GLM, x_k is the input layer to the neural network. DL is the delay line [Color figure can be viewed at wileyonlinelibrary.com]

Once the estimation has been completed, the following stochastic model can be established,

$$x_k = \tilde{x}_k + e(x_{k-1}, u_k), \tag{13}$$

where the estimation error $e(x_{k-1}, u_k)$ represents the functional uncertainty of the estimated model at time k which can be shown to be close to Gaussian noise [14] with zero mean and an input-dependent covariance matrix given by $\tilde{\Sigma}_k = E((x_k - \tilde{x}_k)^T(x_k - \tilde{x}_k)$. This input-dependent covariance matrix $\tilde{\Sigma}_k$ can then be estimated using a second generalized linear neural network model (GLM) (which is a single layer neural network that transforms the input vector using predefined nonlinear functions). To clarify, in this work, $\tilde{\Sigma}$ is estimated using a GLM that takes the state variables and control signal as inputs and that uses the identity function as the transformation function,

$$\Sigma_k = Dx_{k-1} + Gu_k, \text{ where}$$
$$\Sigma_k = \begin{pmatrix} \sigma_{11;k} & \sigma_{12;k} & \dots & \sigma_{1n;k} \\ \sigma_{21;k} & \sigma_{22;k} & \dots & \sigma_{2n;k} \\ \vdots & \vdots & \vdots & \vdots \\ \sigma_{n1;k} & \sigma_{n2;k} & \dots & \sigma_{nn;k} \end{pmatrix}, D = \begin{pmatrix} D_{11} & D_{12} & \dots & D_{1n} \\ D_{21} & D_{22} & \dots & D_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ D_{n1} & D_{n2} & \dots & D_{nn} \end{pmatrix}$$
$$\text{and } G = \begin{pmatrix} G_{11} & G_{12} & \dots & G_{1n} \\ G_{21} & G_{22} & \dots & G_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ G_{n1} & G_{n2} & \dots & G_{nn} \end{pmatrix} \tag{14}$$

are partitioned matrices and are updated online at each instant of time with the aim to minimize the error between the actual covariance matrix and the estimated one. Here the submatrices $D_{ij} \in \mathbb{R}^n$ and $G_{ij} \in \mathbb{R}^m$. To re-emphasize, D and G are partitioned matrices that are obtained from the parameters of the GLM to reconstruct the covariance matrix in the correct way. Furthermore, we introduced checks to make sure that the estimated covariances are always positive. To elaborate, whenever the covariance value goes negative, it is replaced by a small positive number. Therefore, the conditional distribution of the system state at time k can be represented by a Gaussian distribution,

$$s(x_k|u_k, x_{k-1}) = \mathcal{N}_{x_k}(\tilde{x}_k, \Sigma_k), \tag{15}$$

where the mean, $\tilde{x}_k$, is defined in (12) and Σ_k is the covariance matrix given in (14). To re-emphasize, the parameters of the estimated pdf of the system state are state and control input dependent. The estimated covariance matrix given in (14) characterizes the functional uncertainty that results due to the discrepancy between the actual and estimated behavior of the system dynamics. This estimate of the input-dependent covariance accounts for any uncertainty in the system dynamics including uncertain information in the nonlinear functions $\hat{h}$ and $\hat{g}$ as well as uncertainty due to the stochastic element, ϵ, affecting the system dynamics. No other information would be required. The only data available to us are the measurable state x_k and the input to the system defined as u_k. Furthermore, a prior assumption about the dynamics of the system is that it is governed by affine nonlinearities.

4 | FPD ALGORITHM FOR NONLINEAR SYSTEMS WITH FUNCTIONAL UNCERTAINTY

This section demonstrates the methodology of obtaining the feedback control law that takes functional uncertainty into consideration for the nonlinear class of systems described in the previous section. It verifies the optimal performance index for the GFPD, derives the solution to the GFPD that yields a generalized SDRE and explains the derivation of the optimal control law.

As discussed earlier, the FPD method requires the determination of ideal distribution for the system state and control input. Since the developed solution to the problem in this paper consists of a regulation problem, the predefined ideal distribution of the dynamics of the state is expressed as,

$$s^I(x_k|u_k, x_{k-1}) = \mathcal{N}_{x_k}(0, \Sigma_2), \tag{16}$$

where the zero mean reflects the regulation around zero objective. The ideal covariance matrix Σ_2 is smaller than the actual covariance Σ_k. This is permissible as the covariance matrix Σ_k is state and control-dependent and can, thus, be driven to the ideal covariance matrix. Generally, at initial stages of control, the covariance of the actual distribution of the system state will be bigger than the covariance of the ideal distribution. However, since the objective of the considered regulation problem is to minimize the discrepancy between the system distribution and a predefined ideal one, it is expected that at steady state, the covariances of the two distributions will become close to each other. This yields to the following assumption.

Assumption 1. At steady state, that is, $|\Sigma_k\Sigma_2^{-1} - I| < 1$. This is a valid assumption which can be satisfied by a proper selection of the parameters (in particular the covariance matrix) of the ideal distribution [15].

Finally, the ideal distribution of the controller is specified to be Gaussian and given by,

$$c^I(u_k|x_{k-1}) = \mathcal{N}_{u_k}(0, \Gamma), \tag{17}$$

where Γ identifies the permissible range of optimal control inputs.

Given the pdfs of the system state (15) and ideal distributions of the system state and control input given in (16) and (17), respectively, the optimal cost-to-go function is given by the following theorem.

Theorem 1. *Based on the estimated distribution of the system state given in (15) and Assumption 1, it can be shown that the minimum cost-to-go function, $-\ln(\gamma(x_k))$, is given by,*

$$-\ln(\gamma(x_k)) = 0.5x_k^T M_k x_k + 0.5T_k x_k + 0.5\omega_k, \tag{18}$$

where M_k is the discrete-time SDRE (19), T_k is the linear equation (20), and ω_k is the constant term (21). Equation T_k is a key adjustment to the conventional form of the FPD. This term arises from the consideration of input-dependent noise leading to a randomized controller that provides cautiousness, that is, controller that takes the uncertainty of the estimates into consideration when calculating the control law [3, 16].

$$M_{k-1} = h^T(x_{k-1})\left[J_k - J_k g(x_{k-1})\left[\Gamma^{-1} + g^T(x_{k-1})J_k g(x_{k-1})\right]^{-1} \times g^T(x_{k-1})J_k\right] h(x_{k-1}), \tag{19}$$

$$T_{k-1} = \left\{ T_k h(x_{k-1}) + tr(DM_k) - 2\left(\frac{1}{2}T_k g(x_{k-1}) + \frac{1}{2}tr(GM_k)\right) \times\left[\Gamma^{-1} + g^T(x_{k-1})J_k g(x_{k-1})\right]^{-1} g^T(x_{k-1})J_k h(x_{k-1}) \right\}, \tag{20}$$

$$\omega_{k-1} = \omega_k - \frac{1}{2}\left(T_k g(x_{k-1}) + tr(GM_k)\right)\left[\Gamma^{-1} + g^T(x_{k-1})J_k \times g(x_{k-1})\right]^{-1}\frac{1}{2}\left(g^T(x_{k-1})T_k + tr(GM_k)\right) + \ln|\Gamma| + \ln|\Gamma^{-1} + g^T(x_{k-1})J_k g(x_{k-1})| + tr(\mathcal{O}((\Sigma_k\Sigma_2{}^{-1} - I)^2)), \tag{21}$$

where $J_k = (\Sigma_2^{-1} + M_k)$ and $\mathcal{O}$ is the big-O notation.

The solutions of Equations (19)–(21) can be obtained by reversing the direction of time as discussed in Herzallah [17].

Proof. The specified forms of elements (18), (19), (20), and (21) can be obtained by evaluating $\gamma(x_{k-1})$ from (8) for which the evaluation of $\beta_1(u_k, x_{k-1})$ (6) and $\beta_2(u_k, x_{k-1})$ (7) are required.

$$\begin{aligned}\beta_1(u_k, x_{k-1}) &= \int \mathcal{N}_{x_k}(\tilde{x}_k, \Sigma_k)\left\{\frac{1}{2}\left[(x_k - \tilde{x}_k)^T(\Sigma_2^{-1} - \Sigma_k^{-1}) \times (x_k - \tilde{x}_k) + 2x_k^T(\Sigma_2^{-1} - \Sigma_k^{-1})\tilde{x}_k - \tilde{x}_k^T(\Sigma_2^{-1} - \Sigma_k^{-1})\tilde{x}_k\right]\right\} dx_k \\ &\quad + \frac{1}{2}\tilde{x}_k^T\Sigma_k^{-1}\tilde{x}_k - \frac{1}{2}\ln\left(\frac{|\Sigma_k|}{|\Sigma_2|}\right), \\ &= \frac{1}{2}\mathrm{tr}(\Sigma_2^{-1}\Sigma_k) - \frac{n}{2} + \frac{1}{2}\tilde{x}_k^T\Sigma_2^{-1}\tilde{x}_k - \frac{1}{2}\ln\left(\frac{|\Sigma_k|}{|\Sigma_2|}\right),\end{aligned} \tag{22}$$

where n is the dimension of state variable x_k.

Equation (22) can be further evaluated by simplifying $\ln\left(\frac{|\Sigma_k|}{|\Sigma_2|}\right)$. This can be achieved by using the useful identity from Dunford and Schwarz [18], namely,

$$\ln(\det(W)) = \mathrm{tr}(\ln(W)),$$

which is applicable if matrix W is positive definite. As the covariance matrices are positive definite, this rule can be applied directly to give,

$$\ln(|\Sigma_k|\,|\Sigma_2|^{-1}) = \ln(|\Sigma_k \Sigma_2{}^{-1}|) = \mathrm{tr}(\ln(\Sigma_k \Sigma_2{}^{-1})).$$

Considering the Maclaurin series expansion for logarithms, the $\ln(W)$ can be expressed as,

$$\sum_{j=1}^{\infty} (-1)^{j+1} \frac{(W-I)^j}{j} = (W-I) - \frac{(W-I)^2}{2} + \frac{(W-I)^3}{3} \ldots$$

If $||W-I|| < 1$ the higher order terms will become significantly small[19], thus the above Maclaurin series can be approximated using the big-O notation, namely $\mathcal{O}$ as follows,

$$\sum_{j=1}^{\infty} (-1)^{j+1} \frac{(W-I)^j}{j} = (W-I) - \mathcal{O}((W-I)^2).$$

Therefore, $\mathrm{tr}(\ln(\Sigma_k \Sigma_2{}^{-1}))$ can be computed as,

$$\begin{aligned} \mathrm{tr}(\ln(\Sigma_k \Sigma_2{}^{-1})) &= \mathrm{tr}(\Sigma_k \Sigma_2{}^{-1} - I + \mathcal{O}((\Sigma_k \Sigma_2{}^{-1} - I)^2)) \\ &= \mathrm{tr}(\Sigma_k \Sigma_2{}^{-1}) - n + \mathrm{tr}(\mathcal{O}((\Sigma_k \Sigma_2{}^{-1} - I)^2)). \end{aligned} \tag{23}$$

By substituting (23) in (22), and then using (12), we obtain,

$$\begin{aligned} \beta_1(u_k, x_{k-1}) = \frac{1}{2}(h(x_{k-1})x_k + g(x_{k-1})u_k))^T \Sigma_2^{-1} \\ \times (h(x_{k-1})x_k + g(x_{k-1})u_k)) + \mathrm{tr}(\mathcal{O}((\Sigma_k \Sigma_2{}^{-1} - I)^2)). \end{aligned} \tag{24}$$

Note that the trace of the higher order term given by $\mathrm{tr}(\mathcal{O}((\Sigma_k \Sigma_2{}^{-1} - I)^2))$ represents the trace of error term as a result of the approximation of the series (23). Being small, this error term will be absorbed in the constant term ω_{k-1} as an error term, thus will not contribute to the subsequent steps of the evaluation the control signal.

To solve $\gamma(x_{k-1})$, $\beta_2(u_k, x_{k-1})$ needs to be computed which can be evaluated by substituting (18) in (7) yielding,

$$\begin{aligned} \beta_2(u_k, x_{k-1}) &= -\int s(x_k\,|u_k, x_{k-1})\ln(\gamma(x_k))dx_k, \\ &= \int \mathcal{N}(\hat{x}_k, \Sigma_k)\left(\frac{1}{2}\left[x_k^T M_k x_k + T_k x_k + \omega_k\right]\right)dx_k, \\ &= \frac{1}{2}\left[T_k(h(x_{k-1})x_{k-1} + g(x_{k-1})u_k) + (h(x_{k-1})x_{k-1}\right. \\ &\quad + g(x_{k-1})u_k)^T M_k(h(x_{k-1})x_{k-1} + g(x_{k-1})u_k) + \mathrm{tr}(M_k\Sigma_k) \\ &\quad \left. + \omega_k + \mathrm{tr}(\mathcal{O}((\Sigma_k \Sigma_2{}^{-1} - I)^2))\right]. \end{aligned} \tag{25}$$

The integrals in (22) and (25) are special cases of the general multiple integral given in Theorem (10.5.1) in [20]. The results in (24) and (25) can now be substituted into (8) which gives,

$$\begin{aligned} \gamma(x_{k-1}) &= \exp\left\{-\frac{1}{2}\left[x_{k-1}^T h^T(x_{k-1})J_k h(x_{k-1})x_{k-1} + T_k h(x_{k-1})\right.\right. \\ &\quad \left.\left. \times x_{k-1} + \mathrm{tr}(DM_k)x_{k-1} + \omega_k + \mathrm{tr}(\mathcal{O}((\Sigma_k \Sigma_2{}^{-1} - I)^2))\right]\right\} \\ &\quad \times \int \exp\left\{-\frac{1}{2}\left\{u_k^T\left[\Gamma^{-1} + g^T(x_{k-1})J_k g(x_{k-1})\right]u_k + 2u_k^T\right.\right. \\ &\quad \left.\left. \times\left[g^T(x_{k-1})J_k h(x_{k-1})x_{k-1} + \frac{1}{2}(g^T(x_{k-1})T_k + \mathrm{tr}(GM_k))\right]\right\}\right\} du_k. \end{aligned} \tag{26}$$

For notational convenience, let us define $R_k = [\Gamma^{-1} + g^T(x_{k-1})J_k g(x_{k-1})]^{-1}$. The integral in (26) can be solved by completing the square over u_k and further simplification yields,

$$\begin{aligned} \gamma(x_{k-1}) &= \exp\left\{-\frac{1}{2}x_{k-1}^T\left\{h^T(x_{k-1})\left(J_k - J_k g(x_{k-1})R_k g^T(x_{k-1})\right.\right.\right. \\ &\quad \left.\left. \times J_k\right)h(x_{k-1})\right\}x_{k-1} - \frac{1}{2}\left\{T_k h(x_{k-1}) + \mathrm{tr}(DM_k) - (T_k g(x_{k-1})\right. \\ &\quad \left. + \mathrm{tr}(GM_k))R_k g^T(x_{k-1})J_k h(x_{k-1})\right\}x_{k-1} - \frac{1}{2}\left\{\omega_k\right. \\ &\quad + \mathrm{tr}(\mathcal{O}((\Sigma_k \Sigma_2{}^{-1} - I)^2)) - (T_k g(x_{k-1}) + \mathrm{tr}(GM_k))\frac{R_k}{4} \\ &\quad \left.\left. \times\left(g^T(x_{k-1})T_k^T + \mathrm{tr}(GM_k)\right) + \ln|\Gamma| + \ln|R_k^{-1}|\right\}\right\}. \end{aligned} \tag{27}$$

From (27), it can be seen that the assumed form of the performance index given in (18) holds and is, thus, verified for M_{k-1} (19), T_{k-1} (20), and ω_{k-1} (21). □

Finally, the solution of the SDRE Equation (19) of the probabilistic uncertain controller can be obtained using standard methods proposed in Ogata [21], which guarantee the existence and uniqueness of the solution. Same methods can be implemented to obtain the solution of Equation (20).

Following the evaluation of the optimal cost-to-go function, it is straight forward to calculate the optimal control law. This can be done through the substitution of Equation (18) in Equation (5) which results to the optimal control law specified in the following theorem.

Theorem 2. *The optimal control law that minimizes the cost-to-go function (4) can be shown to be given by,*

$$c^*(u_k|x_{k-1}) = \mathcal{N}(u_k^*, R_k), \text{ where } u_k^* = -K_k x_{k-1} - P_k, \tag{28}$$

$$\text{and where } K_k = R_k\left[g^T(x_{k-1})J_k h(x_{k-1})\right], \tag{29}$$

$$P_k = R_k\left[\frac{1}{2}g^T(x_{k-1})T_k + \frac{1}{2}tr(GM_k)\right], \tag{30}$$

$$R_k = \left[\Gamma^{-1} + g^T(x_{k-1})J_k g(x_{k-1})\right]^{-1}. \tag{31}$$

Proof. The proof of the form of the above randomized controller can be obtained by evaluating (5). The numerator (num) of this equation equates to,

$$\begin{aligned} num = & \exp\left\{-\frac{1}{2}\left[x_{k-1}^T h^T(x_{k-1})J_k h(x_{k-1})x_{k-1} + T_k h(x_{k-1})x_{k-1} \right.\right.\\ & \left.\left. + \ \omega_k + \mathrm{tr}(\mathcal{O}((\Sigma_k\Sigma_2^{-1} - I)^2)) + \mathrm{tr}(DM_k)x_{k-1}\right]\right\} \exp\left\{-\frac{1}{2}\left[\left(u_k\right.\right.\right.\\ & \left. + R_k\left[g^T(x_{k-1})J_k h(x_{k-1})x_{k-1} + \frac{1}{2}g^T(x_{k-1})T_k + \frac{1}{2}\mathrm{tr}(GM_k)\right]\right)^T\\ & \times R_k^{-1}\left(u_k + R_k\left[g^T(x_{k-1})J_k h(x_{k-1})x_{k-1} + \frac{1}{2}g^T(x_{k-1})T_k\right.\right.\\ & \left.\left.\left.\left. + \ \frac{1}{2}\mathrm{tr}(GM_k)\right]\right)\right]\right\} \exp\left\{\frac{1}{2}\left(x_{k-1}^T h^T(x_{k-1})J_k g(x_{k-1})\right.\right.\\ & \left. + \frac{1}{2}T_k g(x_{k-1}) + \frac{1}{2}\mathrm{tr}(GM_k)\right) R_k\left(g^T(x_{k-1})J_k h(x_{k-1})x_{k-1}\right.\\ & \left.\left. + \frac{1}{2}g^T(x_{k-1})T_k^T + \frac{1}{2}\mathrm{tr}(GM_k)\right)\right\}. \end{aligned} \tag{32}$$

The denominator can be obtained from (27). Using the laws of exponentials, $\frac{\exp(a)}{\exp(b)} = \exp(a-b)$, the control law can be derived using (27) and (32) to give,

$$\begin{aligned} c^*(u_k|x_{k-1}) = & \exp\left\{-\frac{1}{2}\left[\left(u_k + R_k\left[g^T(x_{k-1})J_k h(x_{k-1})x_{k-1}\right.\right.\right.\right.\\ & \left.\left. + \frac{1}{2}g^T(x_{k-1})T_k^T + \frac{1}{2}\mathrm{tr}(GM_k)\right]\right)^T R_k^{-1}\left(u_k + R_k\left[g^T(x_{k-1})J_k\right.\right.\\ & \left.\left.\left.\left. \times h(x_{k-1})x_{k-1} + \frac{1}{2}g^T(x_{k-1})T_k^T + \frac{1}{2}\mathrm{tr}(GM_k)\right]\right) + \ln|R_k^{-1}|\right]\right\}, \end{aligned} \tag{33}$$

which yields the claimed form of the randomized controller. □

5 | SIMULATION

The efficiency of the proposed method is demonstrated using the inverted pendulum on a cart problem [22, 23]. The performance of the proposed method is evaluated against the generalized SDRE derived by Zafar and Herzallah [24] that does not account for functional uncertainty. The original inverted pendulum system equation that is used in Wang et al. [23] does not involve any noise which is not the case in real-world situations. Therefore, to simulate the system properly as it operates in real world, the equation obtained from Wang et al. [22, 23] is modified by the addition of noise to it. The discrete-time equation of the nonlinear inverted pendulum is thus given by,

$$x_k = h(x_{k-1})x_{k-1} + g(x_{k-1})u_k + \epsilon_k, \tag{34}$$

where $x_k = \begin{bmatrix} x_{1,k} & x_{2,k} & x_{3,k} & x_{4,k} \end{bmatrix}^T$,

$$h(x_{k-1}) = \begin{bmatrix} 1 & T & 0 & 0 \\ 0 & a_{22} & a_{23} & a_{24} \\ 0 & 0 & 1 & T \\ 0 & a_{42} & a_{43} & a_{44} \end{bmatrix}, g(x_{k-1}) = \begin{bmatrix} 0 \\ b_2 \\ 0 \\ b_4 \end{bmatrix}. \tag{35}$$

In the above equation, $x_{1,k}$ represents the cart position, $x_{2,k}$ represents the velocity of the cart, $x_{3,k}$ is the angle of the beam, and $x_{4,k}$ is regarded as the angular velocity. Furthermore, u_k is the control input representing the external force F to the pendulum at discrete-time k. The noise ϵ_k in (34) is Gaussian with zero mean and covariance $Q = 0.001 \times I$. The matrices $h(x_{k-1})$ and $g(x_{k-1})$ are the state and control matrices, respectively, of which the elements of the matrices are defined as follows,

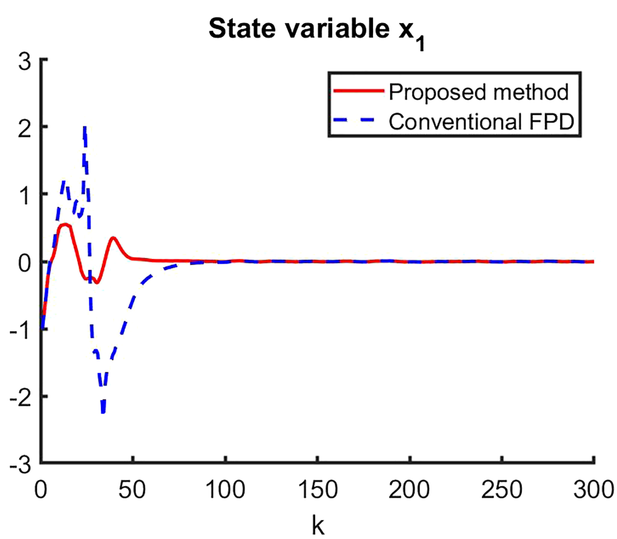

FIGURE 2 Comparison between the proposed method and generalized state-dependent Riccati equation (SDRE) [24] on state x_1 [Color figure can be viewed at wileyonlinelibrary.com]

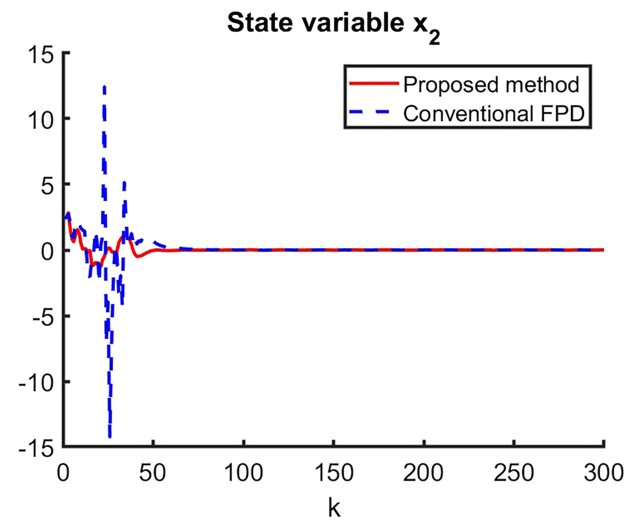

FIGURE 3 Comparison between the proposed method and generalized state-dependent Riccati equation (SDRE) [24] on state x_2 [Color figure can be viewed at wileyonlinelibrary.com]

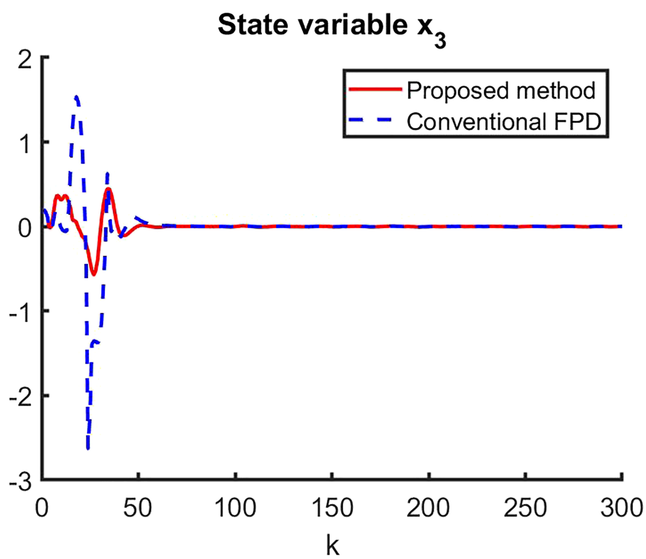

FIGURE 4 Comparison between the proposed method and generalized state-dependent Riccati equation (SDRE) [24] on state variable x_3 [Color figure can be viewed at wileyonlinelibrary.com]

$$a_{24} = a_{22} = 1 + T\frac{-b}{\Omega_2},\ a_{23} = T\frac{m^2L^2g\cos(x_{3,k})\sin(x_{3,k})}{\Omega_2(l + mL^2)},$$
$$a_{24} = T\frac{mL\sin(x_{3,k})}{\Omega_2}(x_{4,k}),\ a_{42} = T\frac{mLb\cos(x_{3,k})}{(M+m)\Omega_1},$$
$$a_{43} = -T\frac{mgL\sin(x_{3,k})}{\Omega_1(x_{3,k})},\ b_4 = -T\frac{mL\cos(x_{3,k})}{(M+m)\Omega_1},$$
$$b_2 = \frac{T}{\Omega_2},\ a_{44} = 1 - T\frac{m^2L^2\cos(x_{3,k})\sin(x_{3,k})(x_{4,k})}{(M+m)\Omega_1}, \tag{36}$$

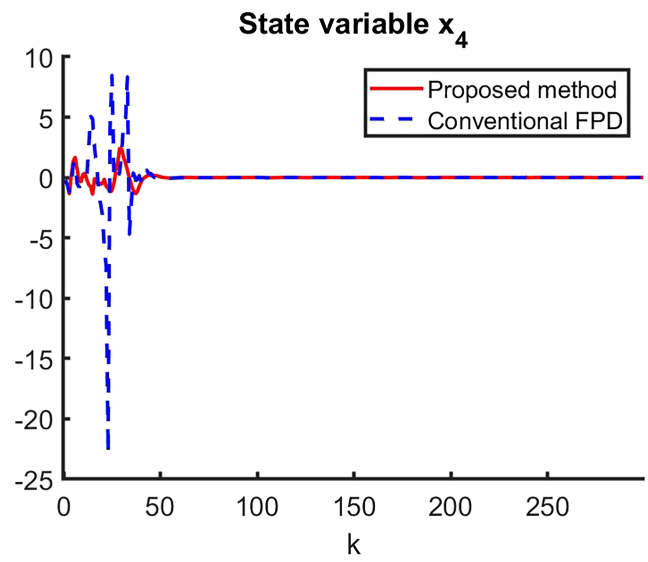

FIGURE 5 Comparison between the proposed method and generalized state-dependent Riccati equation (SDRE) [24] on state variable x_4 [Color figure can be viewed at wileyonlinelibrary.com]

where,

$$\Omega_1 = l + mL^2 - \frac{m^2L^2\cos^2(x_{3,k})}{M+m},$$
$$\Omega_2 = M + m - \frac{m^2L^2\cos^2(x_{3,k})}{l + ML^2}.$$

Additionally, the parameters of the pendulum are taken to be as follows,

$$M = 0.5\text{kg},\ m = 0.5\text{kg},$$
$$b = 0.1\text{N}.\frac{\text{sec}}{\text{m}},\ L = 0.3\text{m, and } l = 0.06\text{kg.m}^2.$$

The description of the parameters can be found in Wang et al. [22]. It should be noted that the dynamics of the system given by (34) are estimated using the process outlined in Section 3. The variance Γ of the controller is chosen to be 100 to give the controller more freedom and achieve a faster convergence rate. The consideration of functional uncertainty in this paper means that the covariance matrix is dependent on the state and control input which was taken into account in the derivation of the proposed method. On the contrary, the method implemented in this paper for comparison, namely, the generalized SDRE [24], does not consider the functional uncertainty of the system dynamics. There, the ideal covariance matrix Σ_2 is assumed to be the same as the covariance matrix of the actual noise. Furthermore, the method in Zafar and Herzallah [24] requires the evaluation of a SDRE which is the same as Equation (19). The equation of cautiousness given in (20) as well as the additional linear term (30) in the control law do not exist in this method. These additional equations emerge in the proposed method in this paper only. This is due to the consideration of the system functional uncertainty as explained earlier in the paper.

To clarify, two sets of experiments were conducted. In the first experiment, the conventional generalized SDRE [24] is used to derive the optimal randomized controller. In this experiment, the covariance matrix of the ideal distribution of the system state is taken to be equal to its actual covariance matrix, that is, $\Sigma_2 = 0.001I$, where I is the identity matrix. In the second experiment, the proposed method in this paper is used to derive the optimal randomized controller. Here, the covariance matrix, Σ_2, of the ideal distribution of the system state is taken to be as follows,

$$\Sigma_2 = 10^{-2} \times \begin{bmatrix} 0.0001 & 0 & 0 & 0 \\ 0 & 0.0004 & 0 & 0 \\ 0 & 0 & 0.0009 & 0 \\ 0 & 0 & 0 & 0.0013 \end{bmatrix}. \tag{37}$$

For both, the conventional FPD and the proposed method, the initial state of the pendulum is taken to be, $x_0 = \begin{bmatrix} -1 & 2.4 & 0.2 & -0.2 \end{bmatrix}$, and the control objective is to bring the four states of the pendulum from their initial values to zero. The results of both experiments are shown in Figures 2–5 from which it can be seen that the states of the pendulum (34) have converged to zero for both generalized SDRE [24] and the proposed method. However, compared with the generalized SDRE [24], the states converge faster and with less oscillations using the proposed method in this paper which accounts for functional uncertainty and input-dependent noises. Hence, the transient response is reached quicker for the proposed GFPD. Thus, it can be concluded that the converging speed of the proposed design is better. Also, having a controller that takes uncertainties into consideration ensures that the system does not overshoot which can also be clearly seen from Figures 2–5.

6 | CONCLUSION

This paper derived the solution of the FPD controller for nonlinear uncertain stochastic systems. The developed solution takes the system functional uncertainty into consideration in the derivation of the randomized controller. The developed method in this paper is innovative and extends the results of the conventional FPD methods to nonlinear systems, thus allowing the implementation of this method to real engineering systems. The derived solution to the FPD problem leads to a SDRE due to the nonlinearities found in the mean of the probability distribution of the system dynamics. As the covariance matrix is state and control input dependent, it is possible to drive the actual covariance matrix to a smaller ideal covariance matrix, thus allowing a reduction in randomness in the data. The discussed approach allows the derivation of an analytic solution for control problems for nonlinear systems by a simple transformation of the nonlinear state function of the system dynamics to a nonlinear state function which is affine in the state. Incorporating these novelties into the FPD framework resulted in a generalized Riccati equation which consists of an additional term that represents the equation of cautiousness. The simulation demonstrated the reduction in overshoots when considering functional uncertainty in the derived optimal control law. It proved to perform better in terms of the transient response than the conventional SDRE which does not consider functional uncertainties.

ACKNOWLEDGMENT

This work was supported by the Leverhulme Trust under grant RPG-2017-337.

CONFLICT OF INTEREST

The authors declare no potential conflict of interests.

AUTHOR CONTRIBUTIONS

R.H.: conceptualization, methodology, validation, and proof outline. R.H and A.Z.: conceived and planned the experiments. R.H. and A.Z. verified the analytical methods. R.H.: encouraged A.Z. to investigate the proposed idea and supervised the findings of this work. A.Z.: performed the numerical simulations for the experiments suggested by R.H.

ORCID

Ana Zafar https://orcid.org/0000-0001-5104-1767
Randa Herzallah https://orcid.org/0000-0001-9128-6814

REFERENCES

1. J. Xu, *Stochastic maximum principle for delayed doubly stochastic control systems and their applications*, Int. J. Control **93** (2018), no. 6, 1–10.
2. X. Guo, A. Hernández-del-Valle, and O. Hernández-Lerma, *First passage problems for nonstationary discrete-time stochastic control systems*, European J. Control **18** (2012), no. 6, 528–538.
3. R. Herzallah, *Generalised probabilistic control design for uncertain stochastic control systems*, Asian J. Control **20** (2018), no. 6, 2065–2074.
4. R. Herzallah, *Fully probabilistic control for stochastic nonlinear control systems with input dependent noise*, Neural Netw. **63** (2015), 199–207.
5. B. S. Chen and W. Zhang, *Stochastic H_2/H_∞ control with state-dependent noise*, IEEE Trans. Autom. Control **49** (2004), 45–57.

6. Z. Gajic and R. Losada, *Solution of the state-dependent noise optimal control problem in terms of Lyapunov iterations*, Automatica **35** (1994), no. 5, 951–954.
7. H. Mukaidani, H. Xu, and V. Dragan, *Stochastic optimal control for weakly coupled large-scale systems via state and static output feedback*, IET Control Theory Appl. **4** (2010), no. 9, 1849–1858.
8. M. Kárny, *Towards fully probabilistic control design*, Automatica **32** (1996), no. 12, 1719–1722.
9. R. Herzallah, *Probabilistic DHP adaptive critic for nonlinear stochastic control systems*, Neural Netw. **42** (2013), 74–82.
10. R. Herzallah and M. Kárny, *Fully probabilistic control design in an adaptive critic framework*, Neural Netw. **24** (2011), no. 10, 1128–1135.
11. Y. Zhou, R. Herzallah, and A. Zafar, *Fully probabilistic design for stochastic discrete system with multiplicative noise*, 2019 IEEE 15th International Conference on Control and Automation (ICCA), 2019, pp. 940–945.
12. Z. S. Hou and Z. Wang, *From model-based control to data-driven control: Survey, classification and perspective*, Inform. Sci. **235** (2013), 3–35.
13. R. Herzallah, *A fully probabilistic design for stochastic systems with input delay*, Int. J. Control **94** (2021), no. 11, 2934–2944.
14. R. Herzallah, *Adaptive critic methods for stochastic systems with input-dependent noise*, Automatica **43** (2007), no. 8, 1355–1362.
15. R. Herzallah and Y. Zhou, *A tracking error-based fully probabilistic control for stochastic discrete-time systems with multiplicative noise*, J. Vibr. Control **26** (2020), no. 23-24, 2329–2339.
16. R. Herzallah, *A novel cautious controller to uncertain models arising in stochastic control*, Int. J. Modell. Identif. Control **6** (2009), no. 2, 110–118.
17. R. Herzallah, *Decentralised probabilistic consensus control for stochastic complex dynamical networks*, IEEE Control Syst. Lett. **5** (2021), no. 2, 277–582.
18. N. Dunford and J. T. Schwarz, *Linear operators, part II*, John Wiley and Sons, 1963.
19. B. Hall, *Lie groups, Lie algebras, and representations: An elementary introduction. John Wiley and Sons*, Vol. **222**, Springer, 2015.
20. F. A. Graybill, *Matrices with Applications in Statistics*, Duxbury Classic, Wadsworth, Belmont, CA, 2001.
21. K. Ogata, *Discrete time control systems*, Prentice Hall, Inc., Englewood Cliffs, NJ, 1987.
22. X. Wang, E. E. Yaz, and C. S. Jeong, *Robust nonlinear feedback control of discrete-time nonlinear systems with mixed performance criteria*, Proceedings of the American Control Conference, Vol. **2010**, Baltimore, MD, 2010, pp. 6357–6362.
23. X. Wang, E. E. Yaz, S. C. Schneider, and Y. I. Yaz, $H_2 - H\infty$ *control of discrete-time nonlinear systems using the state-dependent Riccati equation approach*, Syst. Sci. Control Eng. **5** (2017), no. 1, 215–223.
24. A. Zafar and R. Herzallah, *Generalised fully probabilistic controller design for nonlinear affine systems*, 2019 Tenth International Conference on Intelligent Control and Information Processing (ICICIP), 2019, pp. 137–144.

AUTHOR BIOGRAPHIES

Ana Zafar received her B.Sc. degree in Mathematics from Aston University, Birmingham, UK, in 2015, and her Ph.D. degree in Control Theory from Aston University, Birmingham, UK, in 2022 with a thesis titled “Consensus, control and message passing in complex control systems.” Her work primarily focuses on the modeling and control of complex dynamical systems.

Randa Herzallah received her B.Sc. degree in Industrial Engineering, her M.Sc. degree in Industrial Engineering/Manufacturing and Design from Jordan University, Jordan, and her Ph.D. degree in Control Theory from Aston University, Birmingham, UK. She is currently an Associate Professor in the Warwick Mathematics Institute at University of Warwick. She previously held the positions of Reader in the College of Engineering and Physical Sciences at Aston University, UK, Associate Professor at Al-AHliyya Amman University, and Professor at Al-Balqa Applied University, Jordan. Her general research focus is on the general topic of systems modeling, optimization, design, and control. Her current research interests include nonlinear control theory, probabilistic, and stochastic control; intelligent control; pinning control; adaptive control; distributed and decentralized control; optimal control; signal processing; time-delay systems; networked and complex systems control with applications to power grids; and manufacturing systems and biomedical signals.

How to cite this article: A. Zafar and R. Herzallah, *Fully Probabilistic Control for Uncertain Nonlinear Stochastic Systems*, Asian J. Control, (2022), 1-10. https://doi.org/10.1002/asjc.2971