

# A Fully Probabilistic Design for Stochastic Systems with Input Delay

## Abstract

This paper is concerned with the regulation problem of discrete time stochastic systems involving input delays. This problem has attracted resurgent interests in recent years due to its relevance to networked control systems. The problem is formulated in a fully probabilistic framework and the control solution is obtained by minimising the Kullback-Leibler Divergence (KLD), as a performance function, between the actual and desired joint probability density functions of the system dynamics. A closed form solution for the randomised controller is obtained for stochastic systems that can be described by arbitrary probability density functions. Furthermore, the analytic solution for a class of linear Gaussian stochastic systems is obtained. It is shown that the optimal randomised controller for this class of linear Gaussian stochastic systems is state feedback controller which is modified by an extra linear term that is related to the lagged and future control inputs. The developed method is demonstrated on a simulation example and the results are compared with the standard fully probabilistic design control method.

## Keywords

Stochastic systems. Input delay. Fully probabilistic design. Kullback-Leibler Divergence.

## I. INTRODUCTION

Time delays are inherent to many real-world processes such as process control, and biological and communication systems. Time delay arises either as a result of inherent delays in the components of the system, online sensors data acquisition, or information processing time. It also arises naturally in networked control systems and distributed and decentralised control where fixed, varying, and random delays in the communication networks become inevitable. This phenomenon usually causes severe degradation in the system performance and may even yield instability in the controlled system. Therefore, the development of control methods and algorithms for systems with time delays have received paramount attention since the 1960s [Chyung, 1969], [Kojima and Shijima, 2003], [Pindyck, 1972], [Wang, 1975], [Richard, 2003], [Zhang and Skliar, 2007].

Depending on the type of delay, the nature of the system equations, and the performance criterion used for optimisation, some progress was made towards the development of the control

solution to these delay systems. For continuous time systems the time delay problems have mostly been dealt with by the infinite dimension system theory where the controller solution is given in terms of operator Riccati equations [Delfour, 1986], [Keulen, 1993]. For example, the optimal linear quadratic regulator solution for linear systems with multiple time delays in the control input was developed in [Basin and Gonzalez, 2006]. A solution to the  $H_\infty$  control for systems with single input delays has been developed in [Tadmor and Mirkin, 2005]. A complete solution to the  $H_\infty$  control problem has been developed in [Meinsma and Mirkin, 2005] by converting the delay problem into a nested sequence of elementary problems.

On the other hand, the rapid development in digital techniques and the emerge of new fields such as networked control and congestion control have attracted attention to discrete time systems with delays. One of the well-established approaches for discrete time systems with delays is the augmentation of the system state where a delay problem is transformed into a delay free problem [Hongguo Zhao et al,2008], [Tadmor and Mirkin, 2005], [Theodor and Shaked, 1994]. This approach however, results in higher state dimension and therefore increases the computational cost which becomes a real obstacle when the system under investigation involves multiple and large delays. Other more efficient methods for optimal estimation and control have consequently been sought and developed. Based on the maximum principle, the analytic solution for linear, deterministic, discrete time tracking control problem with multiple delays in the state and control input has been developed in [Wang, 2009]. The solution to the linear quadratic tracking problem is developed in [Pindyck, 1972] for discrete time systems with single delay in the control input. Moreover, linear quadratic control of systems with random network delays is studied in [Nilsson et al, 1998], [Hirano et al, 2005], [Zhang et al, 2005]. However, most of the previously mentioned approaches assume that the system dynamics are deterministic which is not true for most real-world processes. Real world processes are usually subjected to several sources of uncertainties including functional uncertainties, random noises and disturbances introduced by measurement devices and other surrounding environmental conditions thus, are far from being deterministic. For these systems where high levels of uncertainties and stochasticity and input and/or state delays occur simultaneously, the derivation of a control solution becomes even much more challenging. As such, a significant amount of research has been focusing on developing solutions to the stochastic control problems with delays in the input and/or the state [Huanshui Zhang et al, 2015], [Larssen, 2002], [Chang, 2011]. Despite this fact the stochastic problem with delays remains challenging and largely unsolved. In particular, a

general solution that considers the stochastic nature of the system dynamics and the delays in its input and/or state is still lacking.

This paper develops a general probabilistic framework for the derivation of an optimal controller for systems that are affected by stochastic noise and functional uncertainties and at the same time have time delays. The proposed framework is based on the utilisation of the Kullback-Leibler divergence between the joint probability density function (pdf) of the system dynamics and a predefined desired joint pdf as a cost function. This method is called the fully probabilistic design method (FPD). The exploitation of the Kullback-Leibler divergence yields to a randomised controller as opposed to single deterministic estimate of control laws. It also provides a systematic approach to the consideration of noise and uncertainty in the derivation of the optimal control law. The FPD method was originally proposed in [Karny, 1996], [Herzallah, 2011] to obtain a general solution for stochastic systems subject to random inputs and deterministic systems characterised by functional uncertainty with unknown pdfs. However, the FPD method insists on zero delay between the input and the system state, therefore, does not provide optimal solutions for systems with delays. As such, the objective of the current paper is to extend this FPD method and explain how a randomised controller can be derived when input delays exist. The extended method will be referred to by Time Delay FPD (TDFPD) method. We start by developing the general solution for stochastic systems with delays which can be described by arbitrary probabilistic models. The analytic solution for a class of linear stochastic discrete time systems with input delays will then be obtained and discussed in details. For this class of systems, the solution is obtained as a feedback control law, which is linear in the state and whose gain matrix satisfies a generalised Riccati equation which has its state and control input penalisation specified by the ideal distributions of the state and control input respectively. The linear state feedback control law is also modified by an additional linear term that is related to the lagged control input and future state values.

The proposed FPD framework is different and more general to what have already been established in the literature. To be more specific, although the method developed in [Huanshui Zhang et al, 2015] considered stochastic systems, the derived control law in [Huanshui Zhang et al, 2015] is deterministic as opposed to the randomised controller derived in our FPD approach. Our derived randomised controller provides complete information in the decision making process where not only a deterministic value of the control signal is estimated, but also the uncertainty around this estimation. Moreover, the derived solution in our paper is for general nonlinear

systems and is not restricted by the linearity of the system. Additionally, as discussed earlier in this section, the state augmentation method [Hongguo Zhao et al,2008], [Tadmor and Mirkin, 2005], [Theodor and Shaked, 1994] increases the computational cost when the considered system has multiple delays. When the considered system is stochastic and has delays, thus needs to be dealt with using the FPD approach as followed in this paper, the computational problem from the state augmentation approach becomes more stringent and even unsolvable. This is because the FPD requires full specification of the joint density of the trajectory of the state and the control over all times. As such, we follow a different approach that facilitate the derivation of the required optimal randomised controller for delay systems. However, as will be seen from further development, since in this approach the derived solution involves time-delay and future terms in the optimal control law the approximate solution of the TDFPD need to be sought by iteratively solving the corresponding difference equations involved in the calculation of the optimal randomised controller.

To re-emphasise, compared with the existing results on the topic, the proposed framework has three distinct features. Firstly, it is a general probabilistic framework that extends the FPD approach such that it takes systems delays into consideration when deriving optimal control laws. In this framework, the system dynamics are completely characterised by their pdfs which moreover are not assumed to be known apriori, therefore are estimated online. This is opposite to most of the existing literature, where the system dynamics are described by dynamical equations which do not provide a complete characterisation for stochastic systems that operate under high levels of uncertainty and noise. Secondly, both the stochastic nature and delays in the system dynamics are considered simultaneously, and a randomised controller is obtained. Contrary to existing deterministic control laws for systems with delays, the obtained randomised controller is more explorative and provides complete information in the decision making process, therefore is the natural solution to stochastic and uncertain systems. Thirdly, although an analytic solution of the proposed method is only feasible for linear Gaussian systems, the solution can be obtained in a closed form for any linear or nonlinear stochastic system with input delays that can be described by an arbitrary pdf.

The rest of this article is organised as follows. In Section II we state the objectives of the proposed TDFPD and demonstrates its general solution. Section III applies the TDFPD control theory to a class of linear, Gaussian, Stochastic dynamical systems with input delay and derive the corresponding randomised controller. In addition, it summarises and explains the

numerical solution of the derived randomised controller. Section IV contains some simulation results to show the effectiveness of the proposed TDFPD controller. The conclusion is provided in Section V.

## II. TDFPD CONTROL PROBLEM FOR STOCHASTIC SYSTEMS WITH INPUT DELAYS

This section will state the control objectives and demonstrate the general solution of the TDFPD control problem for stochastic systems with input delays that can be described by an arbitrary probabilistic dynamical model.

### A. Control Objectives of the TDFPD Control Problem

Assume that the stochastic system can be represented at each time instant,  $t$  by the following conditional pdf,

$$s(x_t|x_{t-1}, u_t, u_{t-h}), \quad (1)$$

where  $x_t \in \mathbb{R}^n$  represents the system state,  $u_t \in \mathbb{R}^m$  represents the control input, and  $h > 0$  is a positive time delay. The initial values  $x_0, u_i, i = -h, \dots, 0$ , are known. In general  $s(\cdot)$  needs not to be known and needs not to be constrained by the Gaussian assumption. Note that because of the stochasticity of the system, the probabilistic description of the system dynamics as given in (1) provides a complete specification of the present state as a function of the previous state and present and previous control. For these stochastic systems, the objective of the TDFPD control problem is to design a randomised controller,  $c(u_t|x_{t-1}, u_{t-h})$  that minimises the Kullback-Leibler divergence between the joint pdf of the closed loop description of the system dynamics,  $f(\mathcal{X}(t, T))$  and the ideal joint pdf  $I_f(\mathcal{X}(t, T))$ ,

$$\mathcal{D}(f||I_f) \equiv \int f(\mathcal{X}(t, T)) \ln \left( \frac{f(\mathcal{X}(t, T))}{I_f(\mathcal{X}(t, T))} \right) d\mathcal{X}(t, T), \quad (2)$$

where  $\mathcal{X}(t, T) = \{x_t, \dots, x_T, u_t, \dots, u_T\}$  is the closed loop observed data sequence, and  $T \leq \infty$  is a given control horizon. The joint pdf of the closed loop description of the system dynamics is the most complete probabilistic description of its behaviour. For the stochastic systems with input delays given in (1), it can be evaluated using the chain rule [Peterka, 1981] as follows,

$$f(\mathcal{X}(t, T)) = \prod_{t=1}^T s(x_t|x_{t-1}, u_t, u_{t-h})c(u_t|x_{t-1}, u_{t-h}), \quad (3)$$

where the pdf  $s(x_t|x_{t-1}, u_t, u_{t-h})$  describes the dynamics of the observed state vector  $x_t$ , and  $c(u_t|x_{t-1}, u_{t-h})$  represents the pdf of the required randomised controller as mentioned earlier.

Similarly, the ideal joint pdf of the closed loop data can be factorised as follows,

$${}^I f(\mathcal{X}(t, T)) = \prod_{t=1}^T {}^I s(x_t | x_{t-1}, u_t, u_{t-h}) {}^I c(u_t | x_{t-1}, u_{t-h}). \quad (4)$$

where the pdf  ${}^I s(x_t | x_{t-1}, u_t, u_{t-h})$  describes the ideal distribution of the system state vector  $x_t$ , and  ${}^I c(u_t | x_{t-1}, u_{t-h})$  represents the ideal pdf of the randomised controller. Given the definitions of the joint pdf of the closed loop system and the ideal pdf as specified by equations (3) and (4) respectively, minimisation of (2) can be obtained recursively by introducing the following definition,

$$\begin{aligned} -\ln(\gamma(x_{t-1})) &= \min_{c(u_\tau | x_{\tau-1}, u_{\tau-h})} \sum_{\tau=t}^T \int f(\mathcal{X}_t, \dots, \mathcal{X}_T | x_{t-1}) \\ &\times \ln \left( \frac{s(x_\tau | x_{\tau-1}, u_\tau, u_{\tau-h}) c(u_\tau | x_{\tau-1}, u_{\tau-h})}{{}^I s(x_\tau | x_{\tau-1}, u_\tau, u_{\tau-h}) {}^I c(u_\tau | x_{\tau-1}, u_{\tau-h})} \right) d(\mathcal{X}_t, \dots, \mathcal{X}_T), \end{aligned} \quad (5)$$

for arbitrary  $\tau \in \{1, \dots, T\}$ . Here  $\mathcal{X}_t = (x_t, u_t)$ , and  $-\ln(\gamma(x_{t-1}))$  is the optimal performance index. Justification of why the  $\ln$  function is required here, can be found in Appendix B.

The definition in Equation (5) leads to the recursive formula for the cost function specified in the following theorem. This recursive formula will be used later for the derivation of the optimal randomised controller.

**Theorem 1:** Using the definition given in (5), the minimisation of the Kullback-Leibler divergence (2) can be performed recursively to give the following recurrence functional equation

$$\begin{aligned} -\ln(\gamma(x_{t-1})) &= \min_{c(u_t | x_{t-1}, u_{t-h})} \int s(x_t | x_{t-1}, u_t, u_{t-h}) c(u_t | x_{t-1}, u_{t-h}) \\ &\left[ \underbrace{\ln \left( \frac{s(x_t | x_{t-1}, u_t, u_{t-h}) c(u_t | x_{t-1}, u_{t-h})}{{}^I s(x_t | x_{t-1}, u_t, u_{t-h}) {}^I c(u_t | x_{t-1}, u_{t-h})} \right)}_{\equiv \text{partial cost} \implies U(x_t, u_t)} - \underbrace{\ln(\gamma(x_t))} \right] \\ &- \underbrace{\int s(x_{t+h} | x_{t+h-1}, u_{t+h}, u_t) c(u_{t+h} | x_{t+h-1}, u_t) \ln(\gamma(x_{t+h})) d(x_{t+h}, u_{t+h}) \delta(t+h)}_{\text{optimal cost-to-go}} \Big] d(x_t, u_t), \end{aligned} \quad (6)$$

where,

$$\delta(t+h) = \begin{cases} 0, & t = T-h, T-h+1, \dots \\ 1, & t = 0, 1, 2, T-h-1. \end{cases} \quad (7)$$

**Proof:** The proof is given in Appendix A.

### B. Solution to the TDFPD Control Problem

The general solution of the optimisation of the TDFPD control problem as obtained from the minimisation of the cost-to-go function with respect to control law,  $c(\mathbf{u}_t|\mathbf{x}_{t-1}, \mathbf{u}_{t-h})$  can be shown to be given by the following theorem.

**Theorem 2:** The pdf of the optimal randomised controller minimising the cost-to-go function (6) subject to the conditional distribution of the stochastic system,  $s(\mathbf{x}_t|\mathbf{x}_{t-1}, \mathbf{u}_t, \mathbf{u}_{t-h})$  is given by,

$$c^*(\mathbf{u}_t|\mathbf{x}_{t-1}, \mathbf{u}_{t-h}) = \frac{\int c(\mathbf{u}_t|\mathbf{x}_{t-1}, \mathbf{u}_{t-h}) \exp \left[ -\beta_1(\cdot) - \beta_2(\cdot) - \int c(\mathbf{u}_{t+h}|\mathbf{x}_{t+h-1}, \mathbf{u}_t) \beta_3(\cdot) \delta(t+h) d\mathbf{u}_{t+h} \right]}{\gamma(\mathbf{x}_{t-1})}, \quad (8)$$

where,

$$\begin{aligned} \gamma(\mathbf{x}_{t-1}) &= \int \int c(\mathbf{u}_t|\mathbf{x}_{t-1}, \mathbf{u}_{t-h}) \exp \left[ -\beta_1(\mathbf{u}_t, \mathbf{x}_{t-1}, \mathbf{u}_{t-h}) - \beta_2(\mathbf{u}_t, \mathbf{x}_{t-1}, \mathbf{u}_{t-h}) \right. \\ &\quad \left. - \int c(\mathbf{u}_{t+h}|\mathbf{x}_{t+h-1}, \mathbf{u}_t) \beta_3(\mathbf{u}_{t+h}, \mathbf{x}_{t+h-1}, \mathbf{u}_t) \delta(t+h) d\mathbf{u}_{t+h} \right] d\mathbf{u}_t, \\ \beta_1(\mathbf{u}_t, \mathbf{x}_{t-1}, \mathbf{u}_{t-h}) &= \int s(\mathbf{x}_t|\mathbf{x}_{t-1}, \mathbf{u}_t, \mathbf{u}_{t-h}) \left[ \ln \frac{s(\mathbf{x}_t|\mathbf{x}_{t-1}, \mathbf{u}_t, \mathbf{u}_{t-h})}{\int s(\mathbf{x}_t|\mathbf{x}_{t-1}, \mathbf{u}_t, \mathbf{u}_{t-h})} \right] d\mathbf{x}_t, \\ \beta_2(\mathbf{u}_t, \mathbf{x}_{t-1}, \mathbf{u}_{t-h}) &= - \int s(\mathbf{x}_t|\mathbf{x}_{t-1}, \mathbf{u}_t, \mathbf{u}_{t-h}) \ln(\gamma(\mathbf{x}_t)) d\mathbf{x}_t, \\ \beta_3(\mathbf{u}_{t+h}, \mathbf{x}_{t+h-1}, \mathbf{u}_t) &= - \int s(\mathbf{x}_t|\mathbf{x}_{t-1}, \mathbf{u}_t, \mathbf{u}_{t-h}) s(\mathbf{x}_{t+h}|\mathbf{x}_{t+h-1}, \mathbf{u}_{t+h}, \mathbf{u}_t) \ln(\gamma(\mathbf{x}_{t+h})) d(\mathbf{x}_t, \mathbf{x}_{t+h}), \end{aligned} \quad (9)$$

**Proof:** This proposition can be proven by adapting the proof of Proposition 2 in [Karny, 2006].

**Remark 1:** To re-emphasise, the solution to the TDFPD as specified in Theorem 1 does not assume any specific form on the generative probabilistic model of the system dynamics. It provides a general solution for any stochastic system with input delays that can be described by an arbitrary pdf.

### III. LINEAR GAUSSIAN STOCHASTIC SYSTEMS WITH INPUT DELAYS

The theory developed in the previous section is applied here to a class of linear, Gaussian, Stochastic dynamical systems with input delay,

$$\mathbf{x}_t = \tilde{\mathbf{A}}\mathbf{x}_{t-1} + \tilde{\mathbf{B}}\mathbf{u}_t + \tilde{\mathbf{B}}_1\mathbf{u}_{t-h} + \tilde{\mathbf{e}}_t, \quad (10)$$

Where  $\tilde{\mathbf{A}}$  is the system state matrix,  $\tilde{\mathbf{B}}$  is the control input matrix,  $\tilde{\mathbf{B}}_1$  is the matrix of the delayed control input, and  $\tilde{\mathbf{e}}_t$  is a zero mean Gaussian noise with covariance  $\tilde{\Sigma}$ . The parameters

of the stochastic model (10) are not known in general, thus need to be estimated. However, since the current value of the system state is affected by noise, its value cannot be completely specified by the current and previous control and previous state values. Therefore, the generative probabilistic model of the stochastic model (10) can be estimated on-line using observed data from the stochastic system dynamics to describe the probabilistic evolution of the system state. The estimation process of the generative probabilistic model will be discussed in Section III-A. There, it will be shown that the generative probabilistic model of the class of linear stochastic systems given in (10) can be estimated by a Gaussian pdf,

$$\begin{aligned} s(x_t|x_{t-1}, \mathbf{u}_t, \mathbf{u}_{t-h}) &= \mathcal{N}_{x_t}(\boldsymbol{\mu}_t, \boldsymbol{\Sigma}), \\ \boldsymbol{\mu}_t &= \mathbf{A}x_{t-1} + \mathbf{B}\mathbf{u}_t + \mathbf{B}_1\mathbf{u}_{t-h}, \end{aligned} \quad (11)$$

where  $\mathcal{N}_{x_t}$  represents the Gaussian pdf of  $x_t$ . In addition,  $\mathbf{A}$ ,  $\mathbf{B}$ ,  $\mathbf{B}_1$  and  $\boldsymbol{\Sigma}$  are the estimates of the matrices  $\tilde{\mathbf{A}}$ ,  $\tilde{\mathbf{B}}$ ,  $\tilde{\mathbf{B}}_1$  and  $\tilde{\boldsymbol{\Sigma}}$  respectively.

The following sections will discuss the development of the analytic solution for the stochastic linear Gaussian system defined in (11) such that its state is regulated around zero, but first we discuss the estimation process of the pdf of the state of the system.

#### A. Conditional Distribution of the System State

As discussed in Section III, the generative probabilistic model of the system state,  $s(x_t|x_{t-1}, \mathbf{u}_t, \mathbf{u}_{t-h})$  is unknown in reality and hence needs to be estimated. The estimation process of this probabilistic generative model is based on the use of linear optimisation methods to estimate the parameters of this model as given in equation (11), represented in the  $\mathbf{A}$ ,  $\mathbf{B}$ ,  $\mathbf{B}_1$ , and  $\boldsymbol{\Sigma}$  matrices. The parameters can be estimated online where their estimates are updated at each time instant,  $t$  when a new measurement of the state value becomes available. In particular, given a new observation of the system state  $x_t$  and the prior knowledge of linear dynamics, the required model can be written as follows,

$$\begin{aligned} x_t &= \begin{bmatrix} \mathbf{A} & \mathbf{B} & \mathbf{B}_1 \end{bmatrix} \begin{bmatrix} x_{t-1} \\ \mathbf{u}_t \\ \mathbf{u}_{t-h} \end{bmatrix}, \\ &= \boldsymbol{\Theta}\Phi_t, \end{aligned} \quad (12)$$



where  $\Phi_t = [x_{t-1} \ u_t \ u_{t-h}]^T$ , and  $\Theta = [A \ B \ B_1]$ . Having formulated Equation (12), all observed data up to time  $t$  is used in the estimation process in order to guarantee accurate estimation of the parameters  $\Theta$ . To be more specific the observed values of the state  $x_t$  and input vector  $\Phi_t$  are retained up to time  $t$  (i.e  $x_t = [x_0, \dots, x_t]$  and  $\Phi_t = [\Phi_0, \dots, \Phi_t]$ ) and used in the optimisation process. Here,  $x_t$  will have dimension  $n \times N$ , and  $\Phi_t$  will have dimension  $(n + 2m) \times N$ , with  $N$  being the number of observations up to time  $t$ . Then to be able to invert the problem and find the estimated parameters,  $\Theta$  both sides of the above equation are multiplied by  $\Phi_t^T$ ,

$$x_t \Phi_t^T = \Theta \Phi_t \Phi_t^T. \quad (13)$$

The matrix  $\Phi_t \Phi_t^T$  in equation (13) is a square matrix of dimension  $(n + 2m) \times (n + 2m)$ . Provided it is nonsingular we may invert it to obtain the required estimates of the parameters,  $\Theta$ ,

$$x_t \Phi_t^\dagger = \Theta, \quad (14)$$

where  $\Phi_t^\dagger$  is an  $N \times (n + 2m)$  matrix known as the pseudo-inverse of  $\Phi_t$  and is given by,

$$\Phi_t^\dagger = \Phi_t^T (\Phi_t \Phi_t^T)^{-1}. \quad (15)$$

**Remark 2:** As can be seen from the above discussion  $\Phi_t$  is, in general, a non-square matrix. Therefore  $\Phi_t$  does not itself have a true inverse. Thus, the introduction of the pseudo inverse matrix in the optimisation process is necessary. On the other hand, as can be seen from Equation (15), the pseudo inverse matrix does have the property that  $\Phi_t \Phi_t^\dagger = I$  where  $I$  is the identity matrix. However, note that  $\Phi_t \Phi_t^\dagger \neq I$  in general. If the matrix  $\Phi_t \Phi_t^T$  is singular then Equation (13) does not have a unique solution. In this case, if the pseudo inverse is defined as,

$$\Phi_t^\dagger = \lim_{\epsilon \rightarrow 0} \Phi_t^T (\Phi_t \Phi_t^T + \epsilon I)^{-1}, \quad (16)$$

then, the limit can be shown to be always exist and that the limiting value guarantees the optimal solution of Equation (13).

Following the estimation of these parameters, the conditional distribution of the system state can be shown to be given by a Gaussian distribution as follows,

$$s(x_t | x_{t-1}, u_t, u_{t-h}) = \mathcal{N}(Ax_{t-1} + Bu_t + B_1u_{t-h}, \Sigma), \quad (17)$$

where  $Ax_{t-1} + Bu_t + B_1u_{t-h}$  is simply the mean of the state evaluated based on the estimated parameters,  $A$ ,  $B$  and  $B_1$  of their corresponding actual values  $\tilde{A}$ ,  $\tilde{B}$  and  $\tilde{B}_1$  respectively. In addition,  $\Sigma$  is the global covariance matrix of the residual error,  $\epsilon_t$  between the actual and estimated states,  $\Sigma = E\{(x_t - [Ax_{t-1} + Bu_t + B_1u_{t-h}])(x_t - [Ax_{t-1} + Bu_t + B_1u_{t-h}])^T\}$ .

## B. Analytic solution to the Linear Gaussian TDFPD

Within the FPD control framework, the control objective can be achieved through the specification of the appropriate parameters of the ideal distribution that will realise the desired objectives of the control problem. Therefore, the solution in this section will, without any loss of generality, be demonstrated on regulating the state of the stochastic system where the system is initially in state  $x_0$  and the aim of control is to return the system state to the origin. Thus, given this control objective, for the stochastic system defined in Equation (11), the ideal distribution of the system state is assumed to be given by,

$${}^1s(x_t|x_{t-1}, u_t, u_{t-h}) = \mathcal{N}_{x_t}(0, \Sigma). \quad (18)$$

The zero mean of the ideal distribution reflects the realistic objective of the regulation problem of reaching zero state values and the covariance determines the spread of the state values around zero.

**Remark 3:** This papers develops the control solution of the time delay control problem for a class of stochastic systems where the noise affecting the system does not depend on the state or control inputs. For this class of systems where the noise affecting the system is independent of the input and/or state, the spread of the state values around the origin will be specified by the covariance of innovations  $\Sigma$  and cannot be made smaller. Therefore, we assume that the desired state covariance is equal to the process noise covariance. For other classes of stochastic systems where the noise affecting the system is dependent on the state or input values, the covariance of the ideal distribution can be specified to be different than the covariance of the system innovations with the aim of designing controllers that reduce the effect of the noise on the dynamics of the system. This has been addressed in our recent work on the development of FPD control methods for stochastic systems with multiplicative noise [Zhou et al, 2019].

Similarly, the ideal distribution of the randomised controller is also assumed to be Gaussian,

$${}^1c(u_t|x_{t-1}, u_{t-h}) = \mathcal{N}_{u_t}(0, \Gamma), \quad (19)$$

where  $\Gamma$  determines the admissible range of optimal control inputs. To re-emphasise, the ideal distribution of the randomised controller  ${}^1c(u_t|x_{t-1}, u_{t-h})$  specifies the admissible range of optimal control inputs only. It is not used in the selection process of the optimal control input. Optimal control inputs are selected using the distribution of the optimised randomised controller,  $c(u_t|x_{t-1}, u_{t-h})$  which as will be seen from further development is a state feedback controller whose gain matrix satisfies a generalised Riccati equation.

Following the above definitions of the pdfs of the system state, its ideal distribution and the ideal distribution of the randomised controller, the derivation of optimal randomised control laws can then be obtained through the evaluation of the optimal performance index,  $-\ln(\gamma(x_{t-1}))$ . This in turn requires the calculation of  $\beta_1(\cdot)$ ,  $\beta_2(\cdot)$ ,  $\beta_3(\cdot)$  and  $\gamma(x_{t-1})$  as defined by Equation (9). The aforementioned, leads to the following theorem.

**Theorem 5:** The randomised optimal controller minimising the optimal cost-to-go function (5) subject to the pdf of the system dynamics (11) and ideal distributions given by (18) and (19) is

$$c(\mathbf{u}_t | x_{t-1}, \mathbf{u}_{t-h}) = \mathcal{N}_{\mathbf{u}_t}(\mathbf{v}_t, \Gamma_t), \quad (20)$$

where

$$\begin{aligned} \mathbf{v}_t &= -\mathbf{K}_t x_{t-1} + \mathbf{L}_t, \\ \mathbf{K}_t &= \Gamma_t^{-1} \mathbf{B}^\top (\Sigma^{-1} + \mathbf{M}_t) \mathbf{A}, \\ \mathbf{L}_t &= -\Gamma_t^{-1} \mathbf{B}^\top \left\{ (\Sigma^{-1} + \mathbf{M}_t) \mathbf{B}_1 \mathbf{u}_{t-h} + 0.5 \mathbf{g}_t^\top \right\} - \Gamma_t^{-1} \rho_t, \\ \Gamma_t &= [\Gamma^{-1} + \mathbf{B}^\top (\Sigma^{-1} + \mathbf{M}_t) \mathbf{B}]^{-1}, \\ \rho_t &= \mathbf{B}_1^\top (\mathbf{M}_{t+h} \bar{\boldsymbol{\mu}}_{t+h} + 0.5 \mathbf{g}_{t+h}^\top) \delta(t+h), \\ \bar{\boldsymbol{\mu}}_{t+h} &= \mathbf{A} x_{t+h-1} + \mathbf{B} \mathbf{v}_{t+h} + \mathbf{B}_1 \mathbf{u}_t. \end{aligned} \quad (21)$$

In addition,

$$-\ln(\gamma(x_t)) = 0.5(x_t^\top \mathbf{M}_t x_t + \mathbf{g}_t x_t + \omega_t), \quad (22)$$

with

$$\mathbf{M}_{t-1} = \mathbf{A}^\top \left\{ (\Sigma^{-1} + \mathbf{M}_t) - (\Sigma^{-1} + \mathbf{M}_t) \mathbf{B} \Gamma_t^{-1} \mathbf{B}^\top (\Sigma^{-1} + \mathbf{M}_t) \right\} \mathbf{A}, \quad (23)$$

$$\begin{aligned} \mathbf{g}_{t-1} &= \mathbf{A}^\top \left\{ 2(\Sigma^{-1} + \mathbf{M}_t) \mathbf{B}_1 \mathbf{u}_{t-h} + \mathbf{g}_t^\top - 2(\Sigma^{-1} + \mathbf{M}_t) \mathbf{B} \Gamma_t^{-1} \mathbf{B}^\top (\Sigma^{-1} + \mathbf{M}_t) \mathbf{B}_1 \mathbf{u}_{t-h} \right. \\ &\quad \left. - (\Sigma^{-1} + \mathbf{M}_t) \mathbf{B} \Gamma_t^{-1} \mathbf{B}^\top \mathbf{g}_t^\top - 2(\Sigma^{-1} + \mathbf{M}_t) \mathbf{B} \Gamma_t^{-1} \rho_t \right\}, \end{aligned} \quad (24)$$

$$\begin{aligned} \omega_{t-1} &= \mathbf{u}_{t-h}^\top \mathbf{B}_1^\top (\Sigma^{-1} + \mathbf{M}_t) \mathbf{B}_1 \mathbf{u}_{t-h} + \text{tr}(\mathbf{M}_t \Sigma) + \mathbf{u}_{t-h}^\top \mathbf{B}_1^\top \mathbf{g}_t^\top + \omega_t \\ &\quad - \mathbf{u}_{t-h}^\top \mathbf{B}_1^\top (\Sigma^{-1} + \mathbf{M}_t) \mathbf{B} \Gamma_t^{-1} \mathbf{B}^\top (\Sigma^{-1} + \mathbf{M}_t) \mathbf{B}_1 \mathbf{u}_{t-h} - 0.25 \mathbf{g}_t \mathbf{B} \Gamma_t^{-1} \mathbf{B}^\top \mathbf{g}_t^\top - \rho_t \Gamma_t^{-1} \rho_t \\ &\quad - 2 \mathbf{u}_{t-h}^\top \mathbf{B}_1^\top (\Sigma^{-1} + \mathbf{M}_t) \mathbf{B} \Gamma_t^{-1} (0.5 \mathbf{B}^\top \mathbf{g}_t + \rho_t) - \mathbf{g}_t^\top \mathbf{B} \Gamma_t^{-1} \rho_t + \{\mathbf{v}_{t+h}^\top \mathbf{B}^\top \mathbf{M}_{t+h} \bar{\boldsymbol{\mu}}_{t+h} \\ &\quad + x_{t+h-1}^\top \mathbf{A}^\top \mathbf{M}_{t+h} \bar{\boldsymbol{\mu}}_{t+h} + 0.5 \mathbf{v}_{t+h}^\top \mathbf{B}^\top \mathbf{g}_{t+h}^\top + 0.5 x_{t+h-1}^\top \mathbf{A}^\top \mathbf{g}_{t+h}^\top + \omega_{t+h} + \text{tr}(\mathbf{M}_{t+h} \Sigma)\} \delta(t+h), \end{aligned} \quad (25)$$

is the quadratic cost function.

**Proof:** The proof is given in Appendix B.

**Remark 5:** Equation (23) is called the generalised discrete time Riccati equation. It is similar to that obtained for the conventional FPD method. The derivation of Equation (24) is a key modification to the original conventional FPD. The manifestation of this equation is due to the existence of the lagged control input in the system equation. Its solution is dependent on the solution of the Riccati equation (23). Note also that the solution of this equation requires knowledge of future time terms  $\rho_t$  and time delay terms that are dependent on  $u_{t-h}$ .

**Remark 6:** Compared with the conventional randomised controller in FPD [Karny, 1996], the mean of the derived randomised controller of the TDFPD method is modified by a linear term,  $L_t$  as can be seen from Equation (21). The introduction of this linear term is the consequence of the presence of the lagged control input. This additional linear term in the mean of the optimised control input is also dependent on delayed and future control input values.

As stated in remark 5 and remark 6 the solution of the optimal cost-to-go function as well as the optimal randomised controller requires knowledge of delayed and future control input values, thus rendering the problem hard to solve analytically. The work in [Wang, 2009] introduced an approximation method to obtain the optimal solution of a two point boundary value problem based on LQ control for tracking control problems. The method is called successive approximation approach, SAA. It will be adopted in the current paper to find the numerical solution of the randomised optimal control inputs obtained from the TDFPD method. This numerical solution to the proposed TDFPD method is discussed in the next section.

### *C. Numerical Solution to the Linear Gaussian TDFPD*

The solution to the TDFPD control problem for linear stochastic systems with input delay (10) is given in Equation (21). Using this derived randomised controller, the optimally controlled conditional distribution of the system state conditioned on previous state and control input values can be obtained as follows,

$$s(x_t|x_{t-1}, u_{t-h}) = \int s(x_t|x_{t-1}, u_t, u_{t-h})c(u_t|x_{t-1}, u_{t-h})du_t \quad (26)$$

where  $s(x_t|x_{t-1}, u_t, u_{t-h})$  defined in Equation (11) is the Gaussian distribution of the system state, and  $c(u_t|x_{t-1}, u_{t-h})$  defined in Equation (20) is the optimised randomised controller.

Substituting Equations (11) and (20) in Equation (26) and integrating over  $\mathbf{u}_t$  yields,

$$\begin{aligned} s(\mathbf{x}_t|\mathbf{x}_{t-1}, \mathbf{u}_{t-h}) &= \mathcal{N}(\bar{\mathbf{x}}_t, \Sigma_t), \\ \bar{\mathbf{x}}_t &= [\mathbf{I} - \mathbf{B}\Gamma_t^{-1}\mathbf{B}^\top(\Sigma^{-1} + \mathbf{M}_t)](\mathbf{A}\mathbf{x}_{t-1} + \mathbf{B}_1\mathbf{u}_{t-h}) - 0.5\Gamma_t^{-1}\mathbf{B}^\top\mathbf{g}_t^\top - \Gamma_t^{-1}\rho_t, \\ \Sigma_t &= \mathbf{B}\Gamma_t\mathbf{B}^\top + \Sigma. \end{aligned} \tag{27}$$

The approximate solution of the TDFPD control problem can then be obtained by iteratively solving the corresponding difference equations involved in the calculation of the optimal randomised controller. This will be done by implementing a slight variation of the SAA proposed in [Wang, 2009]. This approach is based on iteratively solving a sequence of the corresponding nonhomogeneous linear equations in the LQ control problem where each sequence is solved as a standard numerical problem. Following this approach, the future terms in the optimal control law and corresponding optimality equations are obtained from previous iterations, thus overcoming the requirement of predicting these future values. For more details about the SAA the readers are referred to [Wang, 2009]. The following slight variation of the SAA can be readily applied to the TDFPD control problem for obtaining the approximate solution of the optimal randomised controller,

- 1) Let  $\mathbf{g}_{t-1}^{(0)} = 0$ , and  $l = 1$ .
- 2) Estimate the pdf of the stochastic model described by (10) as discussed in Section III-A,
- 3) Solve the Riccati equation,  $\mathbf{M}_{t-1}$  using Equation (23),
- 4) Obtain the  $l$ -th order vector  $\mathbf{g}_{t-1}^{(l)}$  using,

$$\begin{aligned} \mathbf{g}_{t-1}^{(l)} &= \mathbf{A}^\top \left\{ 2(\Sigma^{-1} + \mathbf{M}_t)\mathbf{B}_1\mathbf{u}_{t-h} + \mathbf{g}_t^{(l)\top} - 2(\Sigma^{-1} + \mathbf{M}_t)\mathbf{B}\Gamma_t^{-1}\mathbf{B}^\top(\Sigma^{-1} + \mathbf{M}_t)\mathbf{B}_1\mathbf{u}_{t-h} \right. \\ &\quad \left. - \mathbf{A}^\top(\Sigma^{-1} + \mathbf{M}_t)\mathbf{B}\Gamma_t^{-1}\mathbf{B}^\top\mathbf{g}_t^{(l)\top} - 2\mathbf{A}^\top(\Sigma^{-1} + \mathbf{M}_t)\mathbf{B}\Gamma_t^{-1}\rho_t^{(l-1)} \right\}, \end{aligned} \tag{28}$$

- 5) Calculate the mean of the optimal randomised controller using,

$$\begin{aligned} \mathbf{v}_t &= -\Gamma_t^{-1}\mathbf{B}^\top(\Sigma^{-1} + \mathbf{M}_t)\mathbf{A}\mathbf{x}_{t-1} - \Gamma_t^{-1}\mathbf{B}^\top \left\{ (\Sigma^{-1} + \mathbf{M}_t)\mathbf{B}_1\mathbf{u}_{t-h} + 0.5\mathbf{g}_t^{(l)\top} \right\} \\ &\quad - \Gamma_t^{-1}\rho_t^{(l-1)}, \end{aligned} \tag{29}$$

- 6) Apply the mean of the optimal randomised controller to the system represented by Equation (10) and evaluate its state value,

- 7) Calculate the value of the performance function using (2),
- 8) If  $\frac{\mathcal{D}^{(l)} - \mathcal{D}^{(l-1)}}{\mathcal{D}^{(l)}} < \delta$ , stop. Here  $\delta$  is a prespecified precession. Else,
- 9) Let  $l = l + 1$  and go to 2).

As can be seen from the above algorithm, the values of the future time terms are obtained from the previous iteration of the optimisation process, thus overcoming the requirement of the prediction of these future terms. This allows the numerical calculation of the randomised optimal control law and its implementation to stochastic systems with time delays.

#### IV. NUMERICAL EXAMPLE

Consider the stochastic system (11) with

$$A = \begin{bmatrix} 0.9 & 0.02 & 0 \\ 0 & 0.92 & 0.01 \\ -0.05 & -0.52 & 0.75 \end{bmatrix}, \quad B = \begin{bmatrix} 0 \\ 0 \\ 0.15 \end{bmatrix}, \quad B_1 = \begin{bmatrix} 0 \\ 0 \\ 0.2 \end{bmatrix}, \quad (30)$$

$$\Sigma = 0.0001, \quad x_0 = \begin{bmatrix} 0.2 \\ -0.5 \\ 1.0 \end{bmatrix}. \quad (31)$$

In order to rigorously test the efficiency of the proposed TDFPD control method, three different simulations are carried out. In the first two simulations, different time delays are considered, while in the third simulation the proposed TDFPD is tested when the algorithm assumes a different time delay,  $h'$  than the actual time delay,  $h$  of the stochastic system.

**Simulation 1:** In this simulation study, the time delay  $h$  is taken to be equal to 10. In addition, the initial values  $u_i, i = -10, \dots, 0$  are all taken to be equal to zero. For comparison purposes two sets of experiments were conducted. In the first set of experiments, the system is controlled using the TDFPD control approach for systems with delays as discussed in Section III. In the second experiment the system is controlled using the conventional FPD without consideration of time delays [Herzallah, 2011]. For fair comparison, both experiments assume that the system is initially in state  $x_0$  as given in (31) and the objective of their respective randomised controllers is to bring the state back to zero. As such the ideal distribution of the randomised controllers of both approaches is taken to be Gaussian with zero mean and  $\Gamma = 0.1$  covariance. In addition, the system parameters are assumed to be unknown in both experiments thus estimated online following the method discussed in Section III-A.

During testing, the optimal control signal,  $u_t$  is generated at each time instant,  $t$  from the derived randomised controllers of the conventional FPD and the TDFPD approaches. This optimal control signal is then forwarded to the plant which is specified by the parameters given in Equations (30) and (31) and the system output,  $x_t$  is measured. Figure 1, (a)-(c) and (d) shows histories of the TDFPD and conventional FPD states and control efforts respectively. As can be seen from the figure, both of the TDFPD and conventional FPD approaches managed to regularise the state of the system around zero as required. However, a better performance is achieved from the TDFPD method which shows minimal overshoot compared to the conventional FPD method. Moreover, the minimum cost-to-go function value is found to be considerably lower from the TDFPD method compared to that of the conventional FPD method. For the TDFPD controller,  $-\ln(\gamma(x_{t-1}))$  is found to be 101.772, while for the FPD approach it is equal to 503.4352. This is expected due to the lack of consideration of the lagged control input in the conventional FPD method. In the TDFPD method on the other hand the optimal randomised controller is derived such that the cost of the future control input due to the presence of its lagged value in the system equations is taken into consideration. This provides better anticipatory planning, thus consequently improves the performance of the derived optimal control law from the developed TDFPD method.

**Simulation Study 2:** In this simulation study, the time delay,  $h$  is increased to 30 with all previous input values,  $u_i, i = -30, \dots, 0$  taken to be equal to zero. This study shows that the proposed TDFPD controller performs better again with a considerably lower minimum cost-to-go function value,  $-\ln(\gamma(x_{t-1}))$ . For TDFPD,  $-\ln(\gamma(x_{t-1}))$  is equal to 101.78, while for the FPD approach it is equal to 518.891.

**Simulation Study 3:** This last simulation study tests the proposed TDFPD control method when it assumes a different time delay,  $h'$  than the actual time delay,  $h$  of the stochastic system. Here the actual time delay  $h$  of the stochastic system is taken to be equal to 10 with all previous input values  $u_i, i = -10, \dots, 0$  being uniformly generated from the interval  $(0, 1)$ . However, the proposed TDFPD control is implemented for different time delays to see the effect of the uncertainty of knowing the actual time delay of the stochastic system on its performance. The proposed TDFPD is implemented with time delays,  $h'$  equal to  $(1, 2, 3, 4, \dots, 20)$  and the results are obtained. The value of the minimum cost-to-go function,  $-\ln(\gamma(x_{t-1}))$  versus the assumed time delay  $h'$  is shown in Figure 2. The figure shows that when  $h' < h$ , the value of the minimum cost-to-go function,  $-\ln(\gamma(x_{t-1}))$  is comparable to that when  $h' = h$ . This is expected because

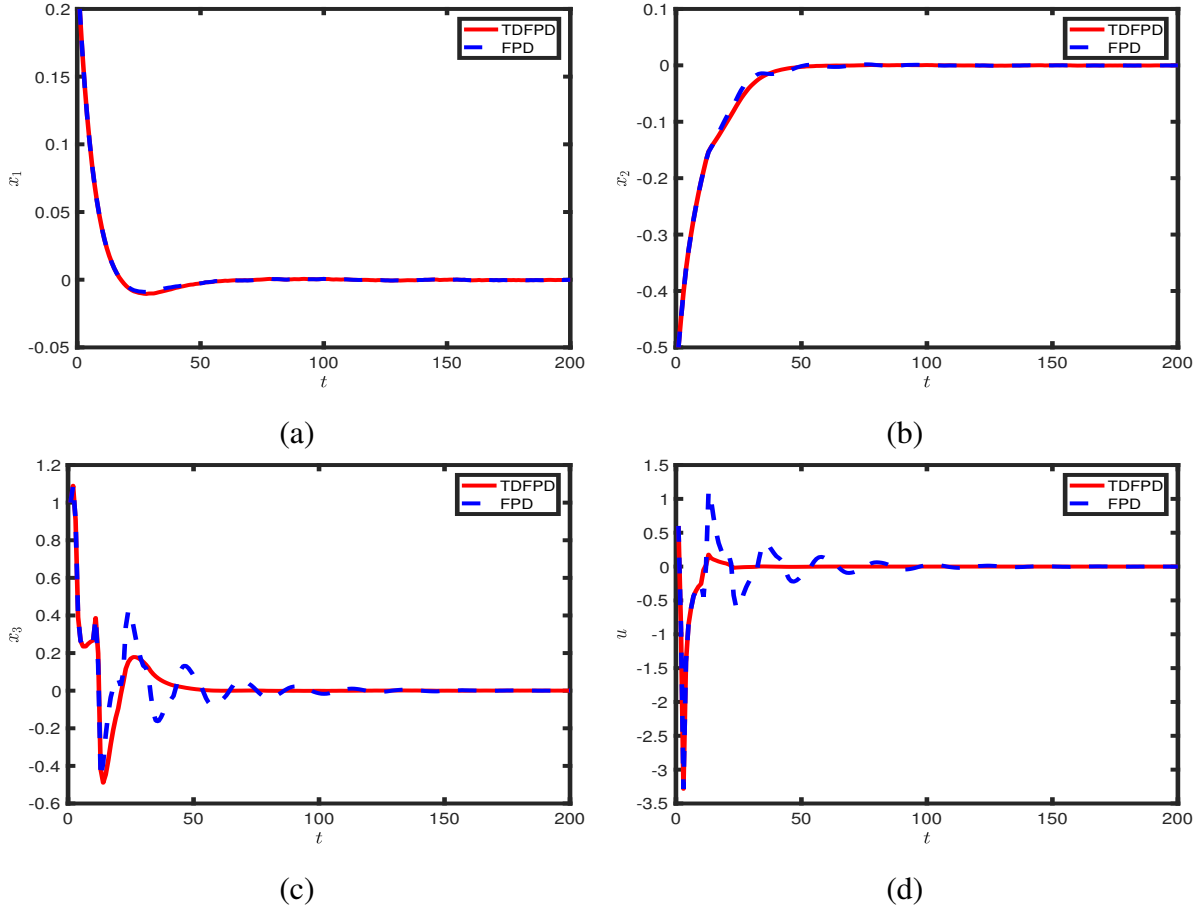


Fig. 1. The state and control inputs of the simulated linear stochastic system with input delay using the FPD and TDFPD control methods: (a) the first state of the stochastic system from the TDFPD shown in solid red and from the FPD shown in dashed blue. (b) the second state of the stochastic system from the TDFPD shown in solid red and from the FPD shown in dashed blue. (c) the third state of the stochastic system from the TDFPD shown in solid red and from the FPD shown in dashed blue. (d) the control effort from the TDFPD shown in solid red and from the FPD shown in dashed blue.

as can be seen from Equation (21) the controller will start adding corrections through the  $L_t$  term earlier assuming shorter delays in the input and therefore accounting for an earlier delayed input than the actual delay of the system. However, when  $h' > h$ , the corrections are added later than the actual delay of the system, ignoring its effect. Therefore, in this case the minimum cost-to-go function,  $-\ln(\gamma(x_{t-1}))$  shows larger values compared to that when  $h' = h$ . In addition, when  $h' = h \pm 3$  the minimum cost-to-go function value,  $-\ln(\gamma(x_{t-1}))$  does not show noticeable variation, with its value continuing to fluctuate around 398.2656 (the value when the time delay is equal to 10).

if the violation of the assumption is not too much, the result will be very close to the baseline.



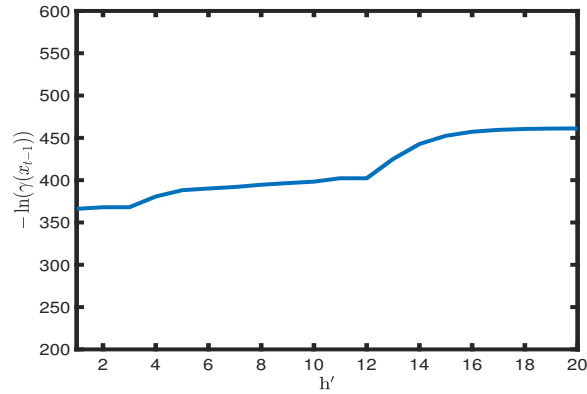


Fig. 2. The effect of the uncertainty in knowing the actual value of the time delay.

## V. CONCLUSION

This paper studied the design of a randomised controller and optimisation problems for stochastic discrete time systems with input delay. The general solution of stochastic systems that can be described by arbitrary distribution functions is derived and a closed form of the optimal randomised controller is obtained. The derived results are then demonstrated on linear Stochastic discrete time systems and the analytic solution of the linear optimal randomised controller is obtained. This obtained results show that the optimal solution for linear stochastic discrete time systems is different and much more complicated than that of the standard solution of the FPD control approach without input delay. The obtained solution still involves solving a generalised Riccati equation, but it also requires the solution of a new vector that takes delayed and future control values into consideration. In addition, the derived linear optimal randomised controller is a state feedback controller but with an additional linear term considering delayed and future inputs. Finally, although the current study considered stochastic systems with single input delay, the methodology presented in this paper provides a pedagogical methodology for the optimisation of control inputs for stochastic systems with multiple input and or state delays. Future work will discuss and demonstrate the extension of the proposed method when the considered control system has multiple input and or state delays. In addition, although this paper restricts its focus on systems with fixed time delays, the extension of the proposed solution to random and variable time delays will also be sought.

## APPENDIX

### A. Proof of Theorem 1

For the stochastic system with input delay as described by (1), the control input  $u_t$  regulates the system by applying the values of  $u_t$  at the current time  $t$  as well as a delayed time moment  $t - h$ . This assumes that the current system state depends not only on the current value of the control input  $u_t$ , but also on its values at certain time lag  $h$ . As stated earlier, the objective of the optimal control problem is to find a randomised optimal control input  $c(u_t|x_{t-1}, u_{t-h})$ ,  $t \in [t_0, T]$  that minimises the cost-to-go function specified in (5). To consider the effect of the delayed control input on future state values, the summation given in (5) can be expanded to be written as follows,

$$\begin{aligned}
 -\ln(\gamma(x_{t-1})) &= \min_{c(u_\tau|x_{\tau-1}, u_{\tau-h})} \left[ \sum_{\tau=t}^T \int f(\mathcal{X}_\tau|x_{t-1}) \right. \\
 &\times \ln \left( \frac{s(x_\tau|x_{\tau-1}, u_\tau, u_{\tau-h})c(u_\tau|x_{\tau-1}, u_{\tau-h})}{\int s(x_\tau|x_{\tau-1}, u_\tau, u_{\tau-h})c(u_\tau|x_{\tau-1}, u_{\tau-h})} \right) + \dots \\
 &+ f(\mathcal{X}_{t+h}|x_{t-1}) \sum_{\tau=t+h+1}^T \int f(\mathcal{X}_\tau|x_{t-1}) \\
 &\left. \times \ln \left( \frac{s(x_\tau|x_{\tau-1}, u_\tau, u_{\tau-h})c(u_\tau|x_{\tau-1}, u_{\tau-h})}{\int s(x_\tau|x_{\tau-1}, u_\tau, u_{\tau-h})c(u_\tau|x_{\tau-1}, u_{\tau-h})} \right) d(\mathcal{X}_\tau) \right] d(\mathcal{X}_t). \tag{32}
 \end{aligned}$$

The expanded form in (32) explicitly shows the future cost of the lagged control input as represented by the second summation starting at  $\tau = t + h + 1$ . Using the dynamic programming principle of optimality, the minimisation of the above equation can be done in two stages. The first stage considers the control input at time  $t$  and the second stage considers the control over

the remaining period. This can be represented mathematically as follows,

$$\begin{aligned}
-\ln(\gamma(x_{t-1})) &= \min_{c(u_t|x_{t-1}, u_{t-h})} \left\{ \min_{c(u_\tau|x_{\tau-1}, u_{\tau-h})} f(\mathcal{X}_t|x_{t-1}) \right. \\
&\times \left[ \ln \left( \frac{s(x_t|x_{t-1}, u_t, u_{t-h})c(u_t|x_{t-1}, u_{t-h})}{I_s(x_t|x_{t-1}, u_t, u_{t-h})I_c(u_t|x_{t-1}, u_{t-h})} \right) \right. \\
&+ \sum_{\tau=t+1}^T \int f(\mathcal{X}_\tau|x_{t-1}) \\
&\times \ln \left( \frac{s(x_\tau|x_{\tau-1}, u_\tau, u_{\tau-h})c(u_\tau|x_{\tau-1}, u_{\tau-h})}{I_s(x_\tau|x_{\tau-1}, u_\tau, u_{\tau-h})I_c(u_\tau|x_{\tau-1}, u_{\tau-h})} \right) d(\mathcal{X}_\tau) + \dots \\
&+ f(\mathcal{X}_{t+h}|x_{t-1}) \sum_{\tau=t+h+1}^T \int f(\mathcal{X}_\tau|x_{t-1}) \\
&\times \ln \left( \frac{s(x_\tau|x_{\tau-1}, u_\tau, u_{\tau-h})c(u_\tau|x_{\tau-1}, u_{\tau-h})}{I_s(x_\tau|x_{\tau-1}, u_\tau, u_{\tau-h})I_c(u_\tau|x_{\tau-1}, u_{\tau-h})} \right) d(\mathcal{X}_\tau) \left. \right] d(\mathcal{X}_t, \dots, \mathcal{X}_T) \left. \right\}. \quad (33)
\end{aligned}$$

Using the definition of the minimum cost to go function given in (5), observing that the last two terms in the above equation are independent of the optimised control, and noting that the last term in the above equation is zero when  $t = T - h, T - h + 1, \dots$ . Equation (33) simplifies to give Equation (6). This completes the proof.

### B. Proof of Theorem 2

The proof of Theorem 3 can be shown by calculating the various elements that were given in Theorem 2, Equation (9). The evaluation of  $\beta_1(u_t, x_{t-1}, u_{t-h})$  can be obtained as follows,

$$\begin{aligned}
\beta_1(u_t, x_{t-1}, u_{t-h}) &= \int \mathcal{N}_{x_t}(Ax_{t-1} + Bu_t + B_1u_{t-h}, \Sigma) \left\{ -0.5[x_t - (Ax_{t-1} + Bu_t + B_1u_{t-h})]^\top \Sigma^{-1} \right. \\
&[x_t - (Ax_{t-1} + Bu_t + B_1u_{t-h})] + 0.5x_t^\top \Sigma^{-1} x_t \left. \right\} dx_t, \\
&= 0.5(Ax_{t-1} + Bu_t + B_1u_{t-h})^\top \Sigma^{-1} (Ax_{t-1} + Bu_t + B_1u_{t-h}). \quad (34)
\end{aligned}$$

The evaluation of  $\beta_2(u_t, x_{t-1}, u_{t-h})$  is based on the assumed quadratic form of  $-\ln(\gamma(x_t))$  defined in Equation (22). This quadratic form of the optimal cost-to-go function will be verified by backward induction, where  $M_{t-1}$  and  $g_{t-1}$  will be obtained recursively as will be shown below. With the assumed quadratic form of  $-\ln(\gamma(x_t))$  we obtain,

$$\begin{aligned}
\beta_2(u_t, x_{t-1}, u_{t-h}) &= \int \mathcal{N}_{x_t}(Ax_{t-1} + Bu_t + B_1u_{t-h}, \Sigma) \left\{ 0.5x_t^\top M_t x_t + 0.5g_t x_t + 0.5\omega_t \right\} dx_t \\
&= 0.5(Ax_{t-1} + Bu_t + B_1u_{t-h})^\top M_t (Ax_{t-1} + Bu_t + B_1u_{t-h}) + 0.5\text{tr}(M_t \Sigma) \\
&+ 0.5\omega_t + 0.5g_t (Ax_{t-1} + Bu_t + B_1u_{t-h}). \quad (35)
\end{aligned}$$

The evaluation of  $\beta_3(\mathbf{u}_{t+h}, \mathbf{x}_{t+h-1}, \mathbf{u}_t)$  is also based on the assumed quadratic form of  $-\ln(\gamma(\mathbf{x}_t))$ .

It is evaluated to give,

$$\begin{aligned}
\beta_3(\mathbf{u}_{t+h}, \mathbf{x}_{t+h-1}, \mathbf{u}_t) &= \iint \mathcal{N}_{\mathbf{x}_t}(\mathbf{A}\mathbf{x}_{t-1} + \mathbf{B}\mathbf{u}_t + \mathbf{B}_1\mathbf{u}_{t-h}, \Sigma) \mathcal{N}_{\mathbf{x}_{t+h}}(\mathbf{A}\mathbf{x}_{t+h-1} + \mathbf{B}\mathbf{u}_{t+h} + \mathbf{B}_1\mathbf{u}_t, \Sigma) \\
&\quad \left\{ 0.5\mathbf{x}_{t+h}^\top \mathbf{M}_{t+h} \mathbf{x}_{t+h} + 0.5\mathbf{g}_{t+h}^\top \mathbf{x}_{t+h} + 0.5\omega_{t+h} \right\} d(\mathbf{x}_t, \mathbf{x}_{t+h}) \\
&= 0.5(\mathbf{A}\mathbf{x}_{t+h-1} + \mathbf{B}\mathbf{u}_{t+h} + \mathbf{B}_1\mathbf{u}_t)^\top \mathbf{M}_{t+h} (\mathbf{A}\mathbf{x}_{t+h-1} + \mathbf{B}\mathbf{u}_{t+h} + \mathbf{B}_1\mathbf{u}_t) \\
&\quad + 0.5\text{tr}(\mathbf{M}_{t+h}\Sigma) + 0.5\omega_{t+h} + 0.5\mathbf{g}_{t+h}^\top (\mathbf{A}\mathbf{x}_{t+h-1} + \mathbf{B}\mathbf{u}_{t+h} + \mathbf{B}_1\mathbf{u}_t), \\
&= \mathbf{u}_t^\top \mathbf{B}_1^\top \mathbf{M}_{t+h} \mathbf{u}_{t+h} + 0.5\mathbf{u}_t^\top \mathbf{B}_1^\top \mathbf{g}_{t+h}^\top + 0.5\text{tr}(\mathbf{M}_{t+h}\Sigma) + 0.5\omega_{t+h} + \mathbf{u}_{t+h}^\top \mathbf{B}^\top \mathbf{M}_{t+h} \mathbf{u}_{t+h} \\
&\quad + \mathbf{x}_{t+h-1}^\top \mathbf{A}^\top \mathbf{M}_{t+h} \mathbf{u}_{t+h} + 0.5\mathbf{u}_{t+h}^\top \mathbf{B}^\top \mathbf{g}_{t+h}^\top + 0.5\mathbf{x}_{t+h-1}^\top \mathbf{A}^\top \mathbf{g}_{t+h}^\top. \tag{36}
\end{aligned}$$

The function  $\gamma(\mathbf{x}_{t-1})$  can also be calculated as given in Theorem 2. It is given by,

$$\begin{aligned}
\gamma(\mathbf{x}_{t-1}) &= \int \mathcal{N}_{\mathbf{u}_t}(0, \Gamma) \exp \left\{ -0.5 \left( \left[ \mathbf{A}\mathbf{x}_{t-1} + \mathbf{B}\mathbf{u}_t + \mathbf{B}_1\mathbf{u}_{t-h} \right]^\top \Sigma^{-1} \left[ \mathbf{A}\mathbf{x}_{t-1} + \mathbf{B}\mathbf{u}_t + \mathbf{B}_1\mathbf{u}_{t-h} \right] \right. \right. \\
&\quad + \left. \left[ \mathbf{A}\mathbf{x}_{t-1} + \mathbf{B}\mathbf{u}_t + \mathbf{B}_1\mathbf{u}_{t-h} \right]^\top \mathbf{M}_t \left[ \mathbf{A}\mathbf{x}_{t-1} + \mathbf{B}\mathbf{u}_t + \mathbf{B}_1\mathbf{u}_{t-h} \right] + \mathbf{g}_t^\top (\mathbf{A}\mathbf{x}_{t-1} + \mathbf{B}\mathbf{u}_t + \mathbf{B}_1\mathbf{u}_{t-h}) \right. \\
&\quad + \text{tr}(\mathbf{M}_t\Sigma) + \omega_t + \left. \int c(\mathbf{u}_{t+h}|\mathbf{u}_t, \mathbf{x}_{t+h}) \left[ \text{tr}(\mathbf{M}_{t+h}\Sigma) + \omega_{t+h} + 2\mathbf{u}_t^\top \mathbf{B}_1^\top \mathbf{M}_{t+h} \mathbf{u}_{t+h} + \mathbf{u}_t^\top \mathbf{B}_1^\top \mathbf{g}_{t+h}^\top \right. \right. \\
&\quad + \left. \left. 2\mathbf{u}_{t+h}^\top \mathbf{B}^\top \mathbf{M}_{t+h} \mathbf{u}_{t+h} + 2\mathbf{x}_{t+h-1}^\top \mathbf{A}^\top \mathbf{M}_{t+h} \mathbf{u}_{t+h} + \mathbf{u}_{t+h}^\top \mathbf{B}^\top \mathbf{g}_{t+h}^\top + \mathbf{x}_{t+h-1}^\top \mathbf{A}^\top \mathbf{g}_{t+h}^\top \right] \delta(t+h) d\mathbf{u}_{t+h} \right\} d\mathbf{u}_t \\
&= \exp \left\{ -0.5 \left( \mathbf{x}_{t-1}^\top \mathbf{A}^\top (\Sigma^{-1} + \mathbf{M}_t) \mathbf{A} \mathbf{x}_{t-1} + \mathbf{u}_{t-h}^\top \mathbf{B}_1^\top (\Sigma^{-1} + \mathbf{M}_t) \mathbf{B}_1 \mathbf{u}_{t-h} \right. \right. \\
&\quad + \left. \left. 2\mathbf{x}_{t-1}^\top \mathbf{A}^\top (\Sigma^{-1} + \mathbf{M}_t) \mathbf{B}_1 \mathbf{u}_{t-h} + \text{tr}(\mathbf{M}_t\Sigma) + \mathbf{x}_{t-1}^\top \mathbf{A}^\top \mathbf{g}_t^\top + \mathbf{u}_{t-h}^\top \mathbf{B}_1^\top \mathbf{g}_t^\top + \omega_t + \{ \omega_{t+h} + \text{tr}(\mathbf{M}_{t+h}\Sigma) \right. \right. \\
&\quad + \left. \left. 2\mathbf{v}_{t+h}^\top \mathbf{B}^\top \mathbf{M}_{t+h} \bar{\mathbf{u}}_{t+h} + 2\mathbf{x}_{t+h-1}^\top \mathbf{A}^\top \mathbf{M}_{t+h} \bar{\mathbf{u}}_{t+h} + \mathbf{v}_{t+h}^\top \mathbf{B}^\top \mathbf{g}_{t+h}^\top + \mathbf{x}_{t+h-1}^\top \mathbf{A}^\top \mathbf{g}_{t+h}^\top \} \delta(t+h) \right) \right\} \\
&\quad + \int \exp \left\{ -0.5 \left( \mathbf{u}_t^\top [\Gamma^{-1} + \mathbf{B}^\top (\Sigma^{-1} + \mathbf{M}_t) \mathbf{B}] \mathbf{u}_t \right. \right. \\
&\quad + \left. \left. 2\mathbf{u}_t^\top \mathbf{B}^\top (\Sigma^{-1} + \mathbf{M}_t) (\mathbf{A}\mathbf{x}_{t-1} + \mathbf{B}_1\mathbf{u}_{t-h}) + \mathbf{u}_t^\top \mathbf{B}^\top \mathbf{g}_t^\top + 2\mathbf{u}_t^\top \rho_t \right) \right\} d\mathbf{u}_t. \tag{37}
\end{aligned}$$

Complete the square with respect to  $\mathbf{u}_t$  and integrate yields,

$$\begin{aligned}
\gamma(\mathbf{x}_{t-1}) = \exp \left\{ -0.5 \left( \mathbf{x}_{t-1}^T \mathbf{A}^T (\Sigma^{-1} + \mathbf{M}_t) \mathbf{A} \mathbf{x}_{t-1} + \mathbf{u}_{t-h}^T \mathbf{B}_1^T (\Sigma^{-1} + \mathbf{M}_t) \mathbf{B}_1 \mathbf{u}_{t-h} \right. \right. \\
+ 2\mathbf{x}_{t-1}^T \mathbf{A}^T (\Sigma^{-1} + \mathbf{M}_t) \mathbf{B}_1 \mathbf{u}_{t-h} + \text{tr}(\mathbf{M}_t \Sigma) + \mathbf{x}_{t-1}^T \mathbf{A}^T \mathbf{g}_t^T + \mathbf{u}_{t-h}^T \mathbf{B}_1^T \mathbf{g}_t^T + \omega_t + \{\omega_{t+h} \\
+ \text{tr}(\mathbf{M}_{t+h} \Sigma) + 2\mathbf{v}_{t+h}^T \mathbf{B}^T \mathbf{M}_{t+h} \bar{\mu}_{t+h} + 2\mathbf{x}_{t+h-1}^T \mathbf{A}^T \mathbf{M}_{t+h} \bar{\mu}_{t+h} + \mathbf{v}_{t+h}^T \mathbf{B}^T \mathbf{g}_{t+h}^T \\
+ \mathbf{x}_{t+h-1}^T \mathbf{A}^T \mathbf{g}_{t+h}^T \} \delta(t+h) - \left[ \mathbf{B}^T (\Sigma^{-1} + \mathbf{M}_t) (\mathbf{A} \mathbf{x}_{t-1} + \mathbf{B}_1 \mathbf{u}_{t-h}) + 0.5 \mathbf{B}^T \mathbf{g}_t^T + \rho_t \right]^T \\
\left. \left. \times \Gamma_t^{-1} \left[ \mathbf{B}^T (\Sigma^{-1} + \mathbf{M}_t) (\mathbf{A} \mathbf{x}_{t-1} + \mathbf{B}_1 \mathbf{u}_{t-h}) + 0.5 \mathbf{B}^T \mathbf{g}_t^T + \rho_t \right] \right) \right\}. \quad (38)
\end{aligned}$$

Taking the  $-\ln$  of both sides of the above equation and noting that  $-\ln(\gamma(\mathbf{x}_{t-1})) = 0.5(\mathbf{x}_{t-1}^T \mathbf{M}_{t-1} \mathbf{x}_{t-1} + \mathbf{g}_{t-1}^T \mathbf{x}_{t-1} + \omega_{t-1})$ , yield the results given in Equations (23), (24), and (25). To be more specific equating quadratic, linear and constant terms from the left and right hand sides of the above equation gives  $\mathbf{M}_{t-1}$ ,  $\mathbf{g}_{t-1}$  and  $\omega_{t-1}$  as defined by Equations (23), (24), and (25) respectively. This proves the claimed quadratic nature of the performance function.

Similarly, the optimal randomised control law can be obtained as specified in Theorem 2. Using Equations (34), (35), (36) and (38) in Equation (8) of Theorem 2 and simplifying, yields,

$$\begin{aligned}
c(\mathbf{u}_t | \mathbf{x}_{t-1}, \mathbf{u}_{t-h}) = \exp \left\{ -0.5 \left( \mathbf{u}_t^T [\Gamma^{-1} + \mathbf{B}^T (\Sigma^{-1} + \mathbf{M}_t) \mathbf{B}] \mathbf{u}_t \right. \right. \\
+ 2\mathbf{u}_t^T \mathbf{B}^T (\Sigma^{-1} + \mathbf{M}_t) (\mathbf{A} \mathbf{x}_{t-1} + \mathbf{B}_1 \mathbf{u}_{t-h}) + \mathbf{u}_t^T \mathbf{B}^T \mathbf{g}_t^T + 2\mathbf{u}_t^T \rho_t \\
+ \left[ \mathbf{B}^T (\Sigma^{-1} + \mathbf{M}_t) (\mathbf{A} \mathbf{x}_{t-1} + \mathbf{B}_1 \mathbf{u}_{t-h}) + 0.5 \mathbf{B}^T \mathbf{g}_t^T + \rho_t \right]^T \\
\left. \left. \times \Gamma_t^{-1} \left[ \mathbf{B}^T (\Sigma^{-1} + \mathbf{M}_t) (\mathbf{A} \mathbf{x}_{t-1} + \mathbf{B}_1 \mathbf{u}_{t-h}) + 0.5 \mathbf{B}^T \mathbf{g}_t^T + \rho_t \right] \right) \right\} \quad (39)
\end{aligned}$$

Completing the square in Equation (39) with respect to  $\mathbf{u}_t$  gives the randomised controller defined in Equation (20) with  $\mathbf{v}_t$ ,  $\Gamma_t$  and  $\rho_t$  as defined in Equation (21).

## REFERENCES

- [Chyung, 1969] Chyung, D., (1969), Discrete systems with delays in control. *IEEE Transactions on Automatic Control*, 14(2): 196–197.
- [Kojima and Shijima, 2003] Kojima, A. & Shijima, S., (2003), Formulas on preview and delayed  $h_\infty$  control. *Proceedings of the 42th IEEE CDC*, Hawaii. 6532–6538.
- [Pindyck, 1972] Pindyck, R. S., (1972), The discrete-time tracking problem with a time delay in the control. *IEEE Transactions on Automatic Control*, 17(3): 397–398.
- [Wang, 1975] Wang, P., (1975), Optimal control of discrete time systems with time lag control. *IEEE Transactions on Automatic Control*, 20(3): 425–426.

- [Richard, 2003] Richard, J., (2003), Time delay systems: an overview of some recent advances and open problems. *Automatica*, 39(10): 1667–1694.
- [Zhang and Skliar, 2007] Zhang, M. B. H. & Skliar, M., (2007), Itô-Volterra Optimal State Estimation with Continuous, Multirate, Randomly sampled, and Delayed Measurements. *IEEE Transactions on Automatic Control*, 52(3): 401–416.
- [Delfour, 1986] Delfour, M. C., (1986), The linear quadratic control problem with delays in space and control variables: A state space approach. *SIAM Journal of Control and Optimisation*, 24: 835–883.
- [Keulen, 1993] Keulen van B., (1993), *H<sub>∞</sub> Control for Distributed Parameter Systems: A State Space Approach*. Birkhäuser Basel.
- [Basin and Gonzalez, 2006] Basin, M., & Rodriguez-Gonzalez, J., (2006), Optimal control for linear systems with multiple time delays in control input. *IEEE Transactions on Automatic Control*, 51(1): 91–97.
- [Tadmor and Mirkin, 2005] Tadmor, G., & Mirkin, L., (2005), h<sub>∞</sub> control and estimation with preview part ii: Fixed size ARE solutions in discrete time. *IEEE Transactions on Automatic Control*, 50(1): 29–40.
- [Meinsma and Mirkin, 2005] Meinsma, G., & Mirkin, L., (2005), h<sub>∞</sub> control of systems with multiple i/o delays via decomposition to adobe problems. *IEEE Transactions on Automatic Control*, 50(2): 199–211.
- [Hongguo Zhao et al.2008] Hongguo Zhao, H. W., Huanshui Zhang, & Zhang, C., (2008), Linear quadratic regulation for discrete-time systems with input delay: Spectral factorization approach. *Journal of Systems Science and Complexity*, 21(1): 46–59.
- [Theodor and Shaked, 1994] Theodor, Y., & Shaked, U., (1994), Game theory approach to h<sub>∞</sub> optimal discrete-time fixed-point and fixed-lag smoothing. *IEEE Transactions on Automatic Control*, 39(9): 1944–1948.
- [Wang, 2009] Wang, H. H., & Tang, G. Y., (2009), Observer-based optimal output tracking for discrete-time systems with multiple state and input delays. *International Journal of Control, Automation, and Systems*, 7(1):57–66.
- [Huanshui Zhang et al, 2015] Huanshui Zhang, J. X., Lin Li, & Fu, M., (2015) Linear quadratic regulation and stabilization of discrete-time systems with delay and multiplicative noise. *IEEE Transactions on Automatic Control*, 60(10): 2599–2613.
- [Larssen, 2002] Larssen, B., (2002), Dynamic programming in stochastic control of systems with delay. *Stochastics and Stochastic Reports* 74(3-4): 651–673.
- [Chang, 2011] Chang, T. P. M., & Yang, Y., (2011), A stochastic portfolio optimization model with bound memory. *Mathematics of Operations Research*, 36(4): 604–619.
- [Herzallah, 2011] Herzallah, R., & Kárný, M., (2011), Fully probabilistic control design in an adaptive critic framework. *Neural Networks*, 24(4): 1128–1135.
- [Peterka, 1981] Peterka, V., (1981), Bayesian system identification. *Trends and Progress in System Identification*, P. Eykhoff, Ed. Oxford: Pergamon Press. 239–304.
- [Karny, 2006] Kárný, M., & Guy, T. V., (2006), Fully probabilistic control design. *Systems & Control Letters*, 55(4): 259–265.
- [Karny, 1996] Kárný, M., (1996), Towards fully probabilistic control design. *Automatica*, 32(12): 1719–1722.
- [Nilsson et al, 1998] Nilsson, J., Bernhardsson, B., & Wittenmark, B., (1998), Stochastic Analysis and Control of Real-time Systems with Random Time Delays. *Automatica*, 34(1): 57–64.
- [Hirano et al, 2005] Hirano, H., Mukai, M., Azuma, T., & Fujita, M., (2005), Optimal Control of Discrete-Time Linear Systems with Network-Induced Varying Delay. *American Control Conference*, June 8-10. Portland, OR, USA.
- [Zhang et al, 2005] Zhang, L., Shi, Y., Chen, T., & Huang, B., (2005), A New Method for Stabilization of Networked Control Systems With Random Delays. *IEEE Transactions on Automatic Control*, 50(8): 1177–1181.
- [Zhou et al, 2019] Zhou, Y., Herzallah, R., & Zafar, A., (2019), Fully Probabilistic Design for Stochastic Discrete System with Multiplicative Noise. . *The 15th IEEE International Conference on Control and Automation*, July 16-19. Edinburgh, Scotland, United Kingdom..