A stochastic PID controller for a class of MIMO systems

Samer S. Saab

To cite this article: Samer S. Saab (2016): A stochastic PID controller for a class of MIMO systems, International Journal of Control, DOI: 10.1080/00207179.2016.1183176

To link to this article: http://dx.doi.org/10.1080/00207179.2016.1183176

Accepted author version posted online: 02 May 2016.
Published online: 13 Jul 2016.

Submit your article to this journal

Article views: 64

View related articles

View Crossmark data
A stochastic PID controller for a class of MIMO systems

Samer S. Saab
Department of Electrical and Computer Engineering, Lebanese American University, Byblos, Lebanon

ABSTRACT
This paper deals with the design of a controller possessing tracking capability of any realisable reference trajectory while rejecting measurement noise. We consider discrete-time-varying multi-input multi-output stable linear systems and a proportional-integral-derivative (PID) controller. A novel recursive algorithm estimating the time-varying PID gains is proposed. The development of the proposed algorithm is based on minimising a stochastic performance index. The implementation of the proposed algorithm is described and boundedness of trajectories and convergence characteristics are presented for a discretised continuous-time model. Simulation results are included to illustrate the performance capabilities of the proposed algorithm.

1. Introduction
Proportional-integral-derivative (PID) control is considered to be one of the earlier control strategies and remains to be the most common control scheme used in industry. The acceptance of PID controllers can be credited to its effective performance and its simple structure and implementation. In addition, the contribution produced by each of their three terms is well comprehended by plant operators and adjusting their gains becomes a relatively straightforward task. Usually, operators tune the three gains manually relying on a trial-and-error approach. Such operation can take considerable time if done diligently (Gawthrop & Nomikos, 1990). Several classical methods have been developed that help in systematically tuning the gains. The classical methods, which are summarised in Gawthrop and Nomikos (1990) and a survey of different adaptive techniques presented in Åström, Hägglund, Hang, and Ho (1993), are applicable to single-input single-output asymptotically stable systems. Automatic tuning or adequate selection of PID gains becomes more challenging when considering multi-input multi-output (MIMO) systems. Substantial amount of work pertaining to analytically determining or tuning PID gains for various MIMO systems has been proposed. One approach is based on optimal control (see, e.g., Cao & Chen, 2014a; Cao & Chen, 2014b; Choi & Chung, 2005; Reynoso-Meza, Garcia-Nieto, Sanchis, & Blasco, 2013; Wang, Davison, & Davison, 2013). A flexible multi-objective optimisation tuning design for a class of nonlinear systems has been recently proposed by Reynoso-Meza et al. (2013). Wang et al. (2013) successfully implemented their design to an industrial plant. Their approach does not require any knowledge of the plant model including the order of the system. However, the system under consideration is assumed to be asymptotically stable linear time-invariant subject to constant reference and disturbance signals. Other techniques based on Lyapunov stability analysis are proposed for PID tuning (see, e.g., Chuan-ke, Jiang, Wu, Yong, & Min, 2013; Mendoza, Zavala-Rio, Santibañez, & Reyes, 2015; Orrante-Sakanassi & Hernández-Guzmán, 2014). A relationship between delay margins and control gains is studied for delay-dependent load frequency control (Chuan-ke et al., 2013). Another recent technique for selecting appropriate PID gains for a class of MIMO systems is achieved by stabilising an augmented system where a solution is obtained by treating the augmented system as a static output feedback (SOF) problem (Saab & Toukharian, 2015). Another category for tuning PID gains is based on employing various intelligent control schemes. Such methods mostly consider nonlinear systems, among those methods; techniques that employ simulated annealing or genetic algorithms (see, e.g., Hametner, Mayr, Kozek, & Jakubek, 2013; Ming-Hao Hung, Li-Sun, Shinn-Jang, Shioi-Fen, & Shinn-Ying, 2008; Neath, Swain, Madawala, & Thrimawithana, 2014), neural network (see, e.g., Cong & Liang, 2009; Jing & Cheng, 2013; Wen & Rosen, 2013) and several techniques that are based on fuzzy strategies (see, e.g., Harinath & Mann, 2008; Mannani & Talebi, 2007; Mannani & Talebi, 2011; Tao, Taur, Chang, & Chang, 2010). All the proposed
methods above and many others come with a number of unique features.

A problem with PID controller is in its derivative term. This term amplifies higher frequency measurement noise that can cause significant degradation in the tracking performance. Many proposed PID controllers embed different robust designs limiting the effects of higher frequency noise. Even if the approach was based on a holistic multi-objective optimisation design, difficulties in assigning weights that relate different performance criteria and robustness would arise (Romero, Hägglund, & Åström, 2014). During the past few years, considerable efforts have been orientated towards the remedy of measurement noise (see, e.g., Chang & Jung, 2009; Isaksson & Graebe, 2002; Kristiansson & Lennartson, 2006; Larsson & Hägglund, 2011; Lee, Shin, & Chung, 2012; Romero et al., 2014; Sekara & Matausek, 2009; Skogestad, 2006; Tan et al., 2012). In order to reduce the effects of noise, a method for tuning the gains is proposed (Skogestad, 2006). Another approach is the implementation of a filter with some specific time constant. The selections of PID gains and filter time constant are optimally determined based on a performance index (Isaksson & Graebe, 2002; Kristiansson & Lennartson, 2006; Romero et al., 2014; Sekara & Matausek, 2009). A significant reduction in noise is achieved when employing a second order Butterworth filter at a cost of a moderate decrease of performance (Romero et al., 2014).

However, most of the proposed published works do not explicitly deal with rejecting measurement noise as they are mostly based or derived on a deterministic approach including minimising a deterministic performance index such as the integral square error, the integral absolute error and the integral time absolute error.

This paper presents a novel recursive algorithm for producing the PID gains. Unlike the traditional performance index employed in the PID literature, we use a stochastic performance index. In particular, the development of the algorithm is based on minimising the covariance of the state error and the input error. In this work, we consider MIMO discrete-time-varying asymptotically stable linear systems with zero-mean white process noise and measurement noise. A sufficient condition is presented for the boundedness of all trajectories and convergence requiring the product of the output-coupling matrix and the input-coupling matrix to be full-column rank. The key contribution of this paper is in showing that a standalone PID controller can reject measurement noise without the use of any filter and, yet, providing boundedness of all trajectories and tracking of realisable reference trajectories in the absence of process noise. The advantages of the proposed recursive algorithm are as follows:

- Discrete-time implementation and automatic PID gains selection.
- Estimation of the covariance of the state error and the covariance of the input error at every time samples. That is, the plant operator can estimate the overall performance of a controlled plant at all times by simply running the proposed recursive algorithm.
- Ability of rejecting measurement noise and errors due to erroneous initial conditions.
- Capability of achieving arbitrary small errors, while considering a discretised continuous-time model, for sufficiently small sampling period, in the presence of measurement noise and erroneous initial conditions.

It is important to note that the ability of rejecting measurement noise along with the above characteristics are also demonstrated for Stochastic Iterative Learning Control (SILC) in Saab (2001) and Saab (2005), and for other SILC algorithms the reader is referred to Shen and Wang (2014) and relevant references therein. However, the major drawbacks of Iterative Learning Control (ILC) algorithms are (1) implementation requires repetitions, (2) does not work well under non-repetitive dynamics and (3) different implementation is needed for different reference trajectories. The latter drawbacks are irrelevant to our proposed approach.

The effectiveness of the proposed method is confirmed numerically for a linear time-varying system with two inputs and two outputs, and performance is compared with a method recently proposed in Saab and Toukharian (2015). In addition, we compare the proposed controller to a model reference adaptive control (MRAC) (Guo, Liu, & Tao, 2009) while considering a lateral motion dynamic model of a large transport airplane.

The rest of the paper is organised as follows. Section 2 defines the system under consideration and the associated control problem formulation. The proposed recursive algorithm and its characteristics are given in Section 3, followed by numerical examples in Section 4. The paper ends with conclusions. The proofs of all results are provided in the Appendix.

2. System descriptions and problem formulation

The system under consideration is a MIMO discrete-time-varying system described by the following state-space equation:

\[ x(k + 1) = A(k)x(k) + B(k)u(k) + w(k) \]
\[ y(k) = C(k)x(k) + v(k) \quad (1) \]
where the argument $k \in \mathbb{Z}_+$ is the discrete-time index, $x(k) \in \mathbb{R}^n$ is the system state, $u(k) \in \mathbb{R}^p$ is the process noise or random state disturbance, $y(k) \in \mathbb{R}^q$ is the measured output and $C(k)x(k)$ is the output, $u(k) \in \mathbb{R}^p$ is the system input and the measurement error $v(k) \in \mathbb{R}^q$. It is worthwhile noting that the values of the parameters and variables in Equation (1) depend on the magnitude of the sampling period whenever the system in Equation (1) is the result of a discretised continuous system. It is assumed that either the plant is inherently discrete-time, or that model (1) is described by an underlying linear continuous-time-varying model, which is sampled with $T_s > 0$ sampling period and with a zero-order hold. The values of the parameters and variables in Equation (1) depend on the magnitude of the sampling period whenever the system in Equation (1) is the result of a discretised continuous-time system.

The control objective is to design an output feedback control vector $u(k)$ to make all signals in the closed-loop system bounded under initialisation errors, measurement noise and random state disturbance. In addition, it is desired to have the output vector $y(k)$ asymptotically track a given reference vector $y_d(k)$ in the presence of measurement noise and initialisation errors. The problem becomes similar to a SOF problem. Although many approaches and techniques exist to approach different versions of the problem, no efficient non-iterative algorithmic solutions are available and even the simplest SOF case is rather involved. On the other hand, PID controllers do not normally require full-state feedback and, due to their integral action, can achieve asymptotic tracking (see, e.g., Saab & Toukhitanian, 2015). However, the desired tracking should be achieved in the presence of measurement noise. Consequently, the selection of the PID gains could be based on minimising a stochastic performance index, which may require some model-based designs. Such characteristics inspire the consideration of a stochastic-based PID controller for the problem at hand.

**Notations:** We denote by $E[.]$ to be the expectation operator, $I_m \in \mathbb{R}^{m \times m}$ the identity matrix, $\lambda(M)$ the eigenvalues of $M$ and $\text{tr}(.)$ the trace operator.

**Assumptions:**

(A1) The system matrices $[A(k) B(k) C(k)]$ are assumed to be bounded for all $k$ and as $k \rightarrow \infty$.

(A2) For any realisable output trajectory or reference signal, $y_d(k) = [y_{1,d}(k) \ldots y_{q,d}(k)]^T$, and an appropriate initial condition $x_d(0)$, there exists a control input $u_d(k) \in \mathbb{R}^p$ generating the desired output trajectory, $y_d(k)$, for the nominal plant. That is, the following difference equation is satisfied:

\[
x_d(k + 1) = A(k)x_d(k) + B(k)u_d(k) \\
y_d(k) = C(k)x_d(k)
\]

where $x_d(k)$ is the state response due to $u_d(k)$ with a given $x_d(0)$ such that $y_d(0) = C(0)x_d(0)$.

(A3) $v(.)$ and $w(.)$ are zero-mean white noise processes mutually uncorrelated with each other and with $x(0)$. $E[v(k)w(k)^T] = Q(k)$ and $E[v(k)v(k)^T] = R(k)$ where $Q(k) = Q(k)^T$ positive semi-definite matrix and $R(k) = R(k)^T$ positive-definite matrix.

In addition, the matrix $E[(x_d(0) - x(0))(x_d(0) - x(0))^T]$ is positive semi-definite.

(A1)–(A3) are assumed to hold throughout this paper.

The control law under consideration is given by

\[
u(k + 1) = u(k) + \sum_{i=1}^{3} K_i(k) e(k + 2 - i)
\]

where the matrices $K_i(k)$, $i \in \{1, 2, 3\}$, represent the $(p \times q)$ learning gains, $e(k) \equiv y_d(k) - y(k)$ is the output measurement error due to the control action $u(k)$ and $u(k) = 0$ for $k \leq 1$. Consequently, we set $K_i(0) \equiv 0$, $i \in \{1, 2, 3\}$.

Define the state and the input errors as $\delta x(k) \equiv x_d(k) - x(k)$ and $\delta u(k) \equiv u_d(k) - u(k)$, respectively.

For compactness of presentation, we denote

\[
A \equiv A(k), B \equiv B(k), C \equiv C(k), C^+ \equiv C(k + 1), \\
A^- \equiv A(k - 1), B^- \equiv B(k - 1), C^- \equiv C(k - 1), \\
K_i \equiv K_i(k), i \in \{1, 2, 3\}.
\]

The error model corresponding to the update law can be derived from Equation (3) as follows:

\[
\delta u(k + 1) = \delta u(k) - K_1 e(k + 1) - K_2 e(k) - K_3 e(k - 1) + \Delta u_d(k)
\]

where $\Delta u_d(k) \equiv u_d(k + 1) - u_d(k)$. Equations (1) and (2) yield

\[
\delta x(k + 1) = A\delta x(k) + B\delta u(k) - w(k)
\]

which further lead to

\[
e(k + 1) = C^+ A^- \delta x(k - 1) + B^- \delta u(k - 1) - w(k - 1) + C^+ B\delta u(k) - C^+ w(k) - v(k + 1)
\]

Similarly, we can obtain

\[
e(k) = C^- A\delta x(k - 1) + C B^- \delta u(k - 1) - C w(k - 1) - v(k)
\]

\[
e(k - 1) = C^- \delta x(k - 1) - v(k - 1)
\]
Inserting Equations (6)–(8) in Equation (4) and rearranging terms, we arrive to

\[ \delta u(k + 1) = (I - K_1C^+B)\delta u(k) - (K_1C^+AB^- + K_2CB^-)\delta u(k - 1) - (K_1C^+A^- + K_2CA^- + K_3C^-)\delta x(k - 1) + K_1\nu(k) + K_2\nu(k) + K_3\nu(k - 1) + \Delta u_d(k) \]  

(9)

In order to arrange Equations (10) and (14) into an augmented system, we define

\[
X = \begin{bmatrix}
\delta u(k) \\
\delta u(k - 1) \\
\delta x(k - 1) \\
\nu(k) \\
\nu(k - 1) \\
w(k) \\
w(k - 1)
\end{bmatrix},
X^+ = \begin{bmatrix}
\delta u(k + 1) \\
\delta u(k) \\
\delta x(k) \\
\nu(k + 1) \\
\nu(k) \\
w(k) \\
w(k - 1)
\end{bmatrix},
\Omega = \begin{bmatrix}
\Delta u_d(k) \\
0 \\
0
\end{bmatrix}.
\]

We combine Equations (5) and (9) to obtain

\[ X^+ = \Phi X + \Gamma Z + \Omega \]

(10)

where

\[
\Phi = \begin{bmatrix}
\Psi_1 & \Psi_2 & \Psi_3 \\
1 & 0 & 0 \\
0 & B^- & A^- \\
K_1 & K_2 & K_1C^+ \\
0 & 0 & 0 \\
0 & 0 & 1
\end{bmatrix},
\Gamma = \begin{bmatrix}
K_1C^+A^- + K_2CA^- + K_3C^- \\
0 \\
0
\end{bmatrix},
\Omega = \begin{bmatrix}
\Delta u_d(k) \\
0 \\
0
\end{bmatrix}.
\]

Problem 2.1. Statement: Consider the system presented in Equation (1) under Assumptions (A1)–(A3). Find the gains \(K_i(k), \ i \in \{1, 2, 3\}\), of the PID controller (3), that minimise the mean-square errors corresponding to \(u(k + 1)\) and \(x(k + 1)\).

3. Main results

This section presents the proposed recursive algorithm and its characteristics pertaining to the boundedness and convergence of trajectories. By boundedness, it is meant that all system trajectories including the state, input and output are bounded for all \(k \geq 0\).

3.1. PID gains

In this section, we develop the proposed recursive algorithm.

Notations: Let us denote, for convenience,

\[
P^+ \equiv E[X^+X^{+T}], \ P \equiv E[XX^T], \ P_u \equiv E[\delta u(k)\delta u(k)^T], \ P_{au} \equiv E[\delta u(k)\delta u(k+1)^T], \ P_{au^-} \equiv E[\delta u(k)\delta u(k-1)^T], \ P_x \equiv E[\delta x(k)\delta x(k)^T], \ P_x^- \equiv E[\delta x(k-1)\delta x(k-1)^T], \ P_u^- \equiv E[\delta u(k-1)\delta u(k-1)^T]
\]

\[
(P_{+} X^+)_{\text{symmetric and positive definite}}.
\]

Minimisation of the mean-square errors of \(u(k + 1), u(k)\) and \(x(k)\) is equivalent to minimisation of the mean-square errors of \(X^+, \) which is also equivalent to minimisation of \(\text{tr}(P^+).\) Note that \(\text{tr}(P^+) = E(\delta u^2(k + 1)) + E(\delta x^2(k)) + E(\delta x^2(k))\) is a sum of positive terms. Consequently, in order to minimise \(\text{tr}(P^+).\) with respect to \(K_i, \ i \in \{1, 2, 3\}, \) at each \(k\)th instant in time, we set

\[
\frac{\partial P^+}{\partial K_i} = 0, \ i \in \{1, 2, 3\}
\]

(11)

where

\[
P^+ = E[(\Phi X - \Gamma Z + \Omega)(\Phi X - \Gamma Z + \Omega)^T]
\]

(12)

The proposed recursive algorithm requires that a matrix, \(S(k),\) to be non-singular. The following proposition defines this matrix and guarantees its positive definiteness.

Proposition 3.1: For all \(k \geq 0,\) the \(3q \times 3q\) matrix, \(S(k),\) with its block elements \(S_{ij}\) defined in the following, is symmetric and positive definite.

\[
S_{11} \equiv C^+BP_u(C^+B)^T + C^+AA^-P_x^-(C^+AA^-)^T + C^+AB^-P_x^-(C^+AB^-)^T + C^+BP_{au^-}-(C^+AB^-)^T + C^+QC^T
\]

\[
S_{12} \equiv C^+AA^-P_x^-(CA^-) + (C^+AB^-)P_x^-(CB^-)^T + C^+AQ^-C^T + C^+BP_{au^-}-(CB^-)
\]

(13)
\[ S_{13} \equiv C^T A A^{-1} P_x (C^-)^T, \]
\[ S_{22} \equiv C A^T P_x (C^-)^T + (CB^-) P_x (CB^-)^T + R + CQ C^T, \]
\[ S_{23} \equiv C A^- P_x (C^-)^T, \quad S_{33} \equiv C^- P_x (C^-)^T + R^- , \]
\[ S_{21} \equiv S_{12}, \quad S_{31} \equiv S_{13} \text{ and } S_{32} \equiv S_{23}, \]
\[ \Psi_1 \]

Proposed recursive algorithm: The following four equations reveal the proposed recursive algorithm for updating the gains \( K_i(k), \ i \in \{1, 2, 3\} \), and updating the estimates of the state error and input error covariance matrices, for all \( k > 0 \):

\[
\begin{align*}
K(k) &= \left[ K_1(k) \ \ K_2(k) \ \ K_3(k) \right] \\
&= \left[ \begin{array}{c}
(P_u(k)(C(k + 1)B(k))^T + P_{uw}(k)(C(k + 1)A(k)B(k - 1))^T \\
	imes (P_{uw}(k) \times \ \ \\
(C(k)B(k - 1))^T \end{array} \right] S(k)^{-1} \\
P_x(k) &= A(k - 1)P_x(k - 1)A(k - 1)^T \\
&+ B(k - 1)P_u(k - 1)B(k - 1)^T + Q(k - 1) \\
P_{uw}(k + 1) &= \Psi_1(k)P_u(k) + \Psi_2(k)P_{uw}(k) \\
P_u(k + 1) &= \Psi_1(k)P_u(k) + \Psi_2(k)P_{uw}(k) + \Pi_d(k)
\end{align*}
\]

where \( \Psi_1(k) \equiv I - K_1(k)C(k + 1)B(k), \quad \Psi_2(k) \equiv -K_1(k)C(k + 1)A(k)B(k - 1) - K_2(k)(C)(k)B(k - 1) \), the matrix \( S(k) \) is as defined in Proposition 3.1, and \( P_u(1) \) and \( P_{uw}(1) \) are initially set such that they are symmetric positive-definite matrices.

The detailed development of the recursive algorithm is provided in the Appendix. Although the above recursive algorithm is developed based on minimizing the trace of the error covariance matrix, the proposed algorithm cannot be considered optimal since its development neglects the correlation between the input error and state error.

### 3.2. Boundedness and convergence characteristics

This section presents basic characteristics of the proposed algorithm. The presented characteristics basically assume that \( C(k + 1)B(k) \) is full-column rank, \( \forall \ k \geq 0 \) and system (1) is asymptotically stable. Proposition 3.2 states that \( P_u(k) \) and \( P_{uw}(k) \) are always symmetric, the first part of Lemma 3.1 shows that \( \hat{P}_u(k) \) is always symmetric and positive definite, and provides a contraction condition relevant to the input error. Proposition 3.3 presents a relation between the couple \( [P_u(k + 1), P_{uw}(k + 1)] \) and \( [P_u(k), P_{uw}(k)] \), whereas, Corollary 3.1 provides a contraction relationship among the pairs. The latter characteristics do not assume stability of system (1). Theorems 3.1 and 3.2 disclose boundedness of all trajectories and some convergence characteristics assuming system stability. The proofs of all results are provided in the Appendix.

Proposition 3.2: If, for \( k = 1 \), \( P_u(k) \) and \( P_{uw}(k) \) are symmetric matrices, then they are symmetric for all \( k \geq 1 \).

The subsequent results employ a \( 2q \times 2q \) matrix, \( \hat{S}_{2q}^{-1} \). The development of this matrix is first elaborated. Let \( \hat{S} \equiv S - \bar{N}E_1\bar{N}^T \). Equation (A1), of the Appendix, shows that \( S = \bar{N}E_1\bar{N}^T + \) a symmetric positive-definite matrix. Therefore, \( \hat{S} \) is a symmetric positive-definite matrix. Consequently, \( \hat{S}^{-1} \) is also a positive-definite matrix including all its leading principal minors (Sylvester’s criterion). Let \( \hat{S}_{2q}^{-1} \) be the 2qth leading principal minor of \( \hat{S}^{-1} \) (first \( 2q \times 2q \) block of matrix \( \hat{S}^{-1} \)). It is worthwhile noting that when \( C(k + 1)B(k) \) is full-column rank \( \forall k \), then \( \hat{N} \) becomes full-column rank.

Proposition 3.3: If \( \hat{P}_u \) is symmetric positive definite, then

\[
\begin{bmatrix}
P_u^+ \\
P_{uw}^+
\end{bmatrix} = \begin{bmatrix}
I_p & 0 \\
I_p & 0
\end{bmatrix} \left( I_{2p} + \hat{P}_u \bar{N} \bar{N}^T \hat{S}_{2q}^{-1} \hat{N} \right)^{-1} \begin{bmatrix}
P_u \\
P_{uw}
\end{bmatrix} + \begin{bmatrix}
\Pi_d \\
0
\end{bmatrix}
\]

(17)

Lemma 3.1: If \( C(k + 1)B(k) \) is full-column rank, \( \forall \ k \geq 0 \), \( P_u(1) \) and \( P_{uw}(1) \) are symmetric positive-definite matrices, then the update law given in Equation (3) and the recursive algorithm presented by Equations (13)–(16) guarantee that \( P_u(k), P_{uw}(k) \) and \( \hat{P}_u(k) \) are symmetric positive-definite matrices \( \forall \ k > 0 \). Moreover, the eigenvalues of \( (I_{2p} + \hat{P}_u \bar{N} \bar{N}^T \hat{S}_{2q}^{-1} \hat{N})^{-1} \) are positive and less than one, that is, \( \forall \ k > 0 \)

\[
0 < \lambda \left( (I_{2p} + \hat{P}_u \bar{N} \bar{N}^T \hat{S}_{2q}^{-1} \hat{N})^{-1} \right) < 1
\]

(18)

Corollary 3.1: There exists a consistent norm \( \| \cdot \| \) such that \( \forall \ k > 0 \):

\[
\begin{bmatrix}
P_u(k + 1) \\
P_{uw}(k + 1)
\end{bmatrix} \leq \begin{bmatrix}
P_u(k) \\
P_{uw}(k)
\end{bmatrix} + \begin{bmatrix}
\Pi_d \\
0
\end{bmatrix}
\]

(19)

Theorem 1: Assume that \( C(k + 1)B(k) \) is full-column rank, \( \forall \ k \geq 0 \), \( P_u(1) \) and \( P_{uw}(1) \) are symmetric positive-definite matrices, and the system in Equation (1) is asymptotically stable. If there exists a \( k_1 > 0 \) such that \( \mu_d(k) = \mu_d(k + 1) \quad \forall \ k \geq k_1 \), then the update law given in Equation (3) and the recursive algorithm presented by Equations (13)–(16) guarantee that all trajectories are
bounded. In addition,

\[ \lim_{k \to \infty} P_r(k) = 0 \]  

(20)

Furthermore, if \( \forall k, w(k) = 0 \),

\[ \lim_{k \to \infty} P_r(k) = 0 \]  

(21)

**Remark 3.1:** The assumption that \( u_d(k) = u_d(k+1) \), \( \forall k \geq k_1 \) is rather restrictive and especially for time-varying systems. However, Theorem 3.1 indicates the capability of the controller in rejecting random measurement errors, random initial conditions and random state disturbances. However, zero-error output convergence (21) requires absence of random state disturbances. The capability of rejecting random measurement errors is further justified with a numerical example in Section 4 whenever \( u_d(k) \neq u_d(k+1) \), \( \forall k \).

**Remark 3.2:** \( \lim_{k \to \infty} P_r(k) = 0 \Rightarrow \lim_{k \to \infty} P_{au-1}(k) = 0. \) Consequently, Equation (13) implies that \( \lim_{k \to \infty} K(k) = 0. \)

In what follows, we consider a discretised plant.

**Definition 3.1:** A trajectory \( u_d(k) \) is said to be a smooth function provided that for any given sampling period \( T_s \) and any consistent norm \( \| \cdot \| \), \( \exists c_u > 0 \) such that \( \forall k \geq 0, \| u_d(k+1) - u_d(k) \| \leq c_u T_s. \)

**Theorem 3.2:** Assume that \( C(k+1)B(k) \) is full-column rank, \( \forall k \geq 0, P_a(1) \) and \( P_{au-1}(1) \) are symmetric positive-definite matrices and the system in Equation (1) is asymptotically stable. The update law given in Equation (3) and the recursive algorithm presented by Equations (13)–(16) guarantee that all trajectories are bounded. Furthermore, if the trajectory of \( u_d(k) \) is a smooth function, then there exists a positive constant \( c_u \) such that

\[ \lim_{k \to \infty} \| P_u(k) \| \leq c_u T_s \]  

(22)

and if \( \forall k, w(k) = 0 \), then there exists a positive constant \( c_x \) such that

\[ \lim_{k \to \infty} \| P_x(k) \| \leq c_x T_s \]  

(23)

**Remark 3.3:** Equation (23) implies that the steady-state output error decreases as the sampling period decreases in the presence of measurement noise and in the absence of process noise. That is, arbitrary small errors can be achieved at steady state for sufficiently large sampling rate. Theorem 3.2 does not indicate what sampling rate to choose in order to achieve a desired error bound. However, while proving \( \lim_{k \to \infty} \| P_x(k) \| \leq c_x T_s, c_x \) is shown to be proportional to \( c_u \) and also \( c_x \) is proportional to \( c_u \), where \( c_u \) is defined in Definition 3.1. That is, smaller \( c_u \) leads to smaller \( c_x \) and smaller bound on the steady-state error. In addition, \( c_u \), which depends on both the plant and desired trajectory bandwidth, should decrease as the bandwidth of the desired output trajectory decreases. Therefore, for a specific tolerance of the steady-state error, the choice of the sampling rate should depend on the bandwidth of the desired trajectory. That is, larger sampling rate should be employed for desired trajectories with higher bandwidth.

**Remark 3.4:** Based on numerical results, it is observed that whenever the size of process noise and/or measurement noise is large, better performance is obtained when \( T_s \) is made smaller right after the transient period. As shown in the proof of Theorem 3.2, the error due to erroneous initial condition is reflected in \( m_o(k) = \prod_{j=0}^{k-1} M(j)P(0) \) where \( 0 < M(j) < 1, \forall j. \) However, as \( T_s \) is made smaller, \( M(j) \) gets closer to 1. This phenomenon can lead to ineffectively rejecting the errors due to erroneous initial condition. Consequently, whenever the overall performance is not satisfactory while using a fixed sampling rate, it is suggested considering implementation of an adaptive sampling rate strategy where the sampling rate is increased after the transient period.

**Remark 3.5:** Assuming that \( C(k+1)B(k) \) is full-column rank requires that the number of outputs is greater than or equal to the number of inputs. The latter is a consequence for contracting the covariance of the input and state errors associated while considering the proposed recursive algorithm that generates the PID gains. However, in the case where the number of inputs is greater than the number of outputs, another approach may be needed where the development of the PID gains can based on minimising the trace of the output error covariance.

**Remark 3.6:** In this work, the plant relative degree \( \mu = 1 \) is assumed. However, for \( \mu > 1 \), it is worthwhile considering a modified PID controller, as suggested in Saab (2007), while finding appropriate gains that result in closed-loop stability. In particular, the modified controller has the form: \( u(k+1) = u(k - \mu + 1) + \sum_{i=1}^{\mu-1} K_i(k)e(k + 2 - i) \). The \( \mu - 1 \) delays in \( u(k - \mu + 1) \) should compensate for observing the associated output error for plants with arbitrary relative degree.

**Remark 3.7:** In this work, it is assumed that the plant is asymptotically stable. For a class of unstable plants, as illustrated in Saab and Toukhtarian (2015), a higher-order controller may be needed, that is, \( u(k+1) = u(k) + \sum_{i=1}^{l} K_i(k)e(k + 2 - i) \) with \( l > 3 \). The latter can be motivated while considering root-locus analysis for single-input single-output linear time-invariant system. In this case, the controller (with \( l > 3 \)) would result
in adding more zeros inside the unit circle attracting more unstable modes.

4. Numerical example

In this section, we illustrate the implementation and performance of the proposed recursive algorithm while considering two examples with different models. All numerical simulations are carried out using MATLAB. All random errors are assumed to be zero-mean white Gaussian noise. The norm, ||.||, employed throughout this example is the 2-norm.

Example 4.1: We consider a two-input two-output system with plant nominal parameters \( A_n(k) B_n(k) C \) to update the output of the plant. The role of the nominal model is basically to depict the actual plant, which may not be available to the controller. An erroneous model of the plant is considered as a reference model in order to only update the recursive formulas, \( A(k) B(k) C \) used to update the recursive formulas (13)–(16). However, since the algorithm implicitly assumes knowledge of the nominal plant, in few scenarios, we compare the performance whenever Equations (13)–(16) employ the nominal plant as an 'optimal' case versus the erroneous plant as a non-optimal case. In the following, we present the nominal and reference parameters employed in this study.

\[
A_n(k) = \begin{bmatrix}
-2.85 & -1.9 & 1.05 \\
0.95 & -4.75 + \cos(a_n k T_s) & -1 \\
1.9 & -1.05 & -6.3 + \sin(b_n k T_s)
\end{bmatrix}
\]

\[
I_3 + T_s \begin{bmatrix}
-3 & -2 & 1 \\
1 & -5 + \cos(a_i k T_s) & -1 \\
2 & -1 & -6 + \sin(b_i k T_s)
\end{bmatrix}
\]

\[
A(k) = \begin{bmatrix}
0 & 0 \\
4.75 & \sin(a_n k T_s) \\
\cos(b_n k T_s) & 5.25
\end{bmatrix}
\]

\[
B_n(k) = T_s \begin{bmatrix}
0 & 0 \\
\cos(b_n k T_s) & 5.25
\end{bmatrix}
\]

\[
B(k) = T_s \begin{bmatrix}
0 & 0 \\
\cos(b_i k T_s) & 5
\end{bmatrix}
\]

\( T_s \) is the sampling period, \( a_n = 1.05 \), \( a_i = 1 \), \( b_n = 1.9 \) and \( b_i = 2 \). The output-coupling matrix is given by \( C = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \). That is, \( y_1(k) = x_2(k) \) and \( y_2(y) = x_5(k) \). The following two sets of desired output trajectories are considered, \( \forall k \geq 0 \):

- Trajectory A: \[ y_{d_1}(k) = -15 \]
  \[ y_{d_2}(k) = 10 \]
- Trajectory B: \[ y_{d_1}(k) = -50 \sin(k T_s) \]
  \[ y_{d_2}(k) = 50 \cos(2k T_s) \]

4.1. Sensitivity of modelling errors and time variation

In order to check for the sensitivity of modelling errors, we compute the input corresponding to Trajectory A using the inverse dynamics of the reference model and then apply the resulting open-loop control to the nominal model. We consider zero error at the initial state; that is, we set \( x(0) \equiv x_d(0) \). Figure 1 depicts the resulting response. By examining Figure 1, it can be observed that the output diverges away from desired trajectory as time increases. Figure 2 shows the nominal desired trajectories of \( u_{d_1}(k) \) and \( u_{d_2}(k) \) needed to generate Trajectory A in the absence of process noise. Although the desired outputs are constants, the controller possesses significant time variation.

4.2. Robustness and rejection of measurement noise

We consider the following performance measures for the errors corresponding to \( \xi \in \{ x_1, x_2, x_3, u_1, u_2 \} : |\delta \xi(k)| \) is the absolute error and \( \text{sqrt}(P_\xi(k)) \) is the square root of the diagonal element corresponding to \( \xi \) extracted from the recursive Equations (13) and (16). However, Equations (13)–(16) employ the reference model and not the nominal one. The latter measure is an estimate of the error standard deviation estimated by the proposed recursive algorithm. These two measures are mostly used in various figures presented in this example. avg \( x \in [a, b] |\delta \xi(k)| \),
Figure 2. $u_1(k)$ and $u_2(k)$ for Trajectory A.

Figure 3. $|\delta \xi(k)|$ and $\sqrt{P_\xi(k)}$, $\xi \in [x_2, x_3]$ with $Q_k = 4I_3$, $R_k = 2500I_2$: Trajectory A with $T_s = 0.1$.

Figure 4. $|\delta \xi(k)|$ and $\sqrt{P_\xi(k)}$, $\xi \in [u_1, u_2]$ with $Q_k = 4I_3$, $R_k = 2500I_2$: Trajectory A with $T_s = 0.1$.

Table 1. Robustness and rejection of measurement errors: Trajectory A with $T_i = 0.01$ and $kT_i \in [0, 100]$ sec.

| $\sqrt{Q_k}$ | $\sqrt{R_k}$ | $\text{avg}_{\xi} |\delta x_1(k)|$ | $\text{avg}_{\xi} |\delta x_2(k)|$ | $\text{avg}_{\xi} |\delta x_3(k)|$ |
|--------------|--------------|-------------------------------|-------------------------------|-------------------------------|
| 100$I_2$    | 2$I_3$      | 7.74                          | 5.83                          | 5.57                          |
| 50$I_2$     | 2$I_3$      | 7.75                          | 5.73                          | 5.45                          |
| 50$I_2$     | $I_3/2$     | 4.22                          | 3.46                          | 3.45                          |
| 50$I_2$     | $0.5I_3$    | 2.69                          | 2.66                          | 2.81                          |

$\text{std}_{t \in [t_1, t_2]}(\delta \xi(k))$ and $\text{avg}_{t \in [t_1, t_2]} \sqrt{P_\xi(k)}$ are the average of absolute errors, standard deviation average of errors and average of $\sqrt{P_\xi(k)}$ over $kT_i \in [t_1, t_2]$, respectively. These three measures are included in the tables. Since the average errors are relatively zero when employing the proposed control law, they are not presented.

The updated law (3) is applied to the nominal model using the reference model for updating the proposed gains presented in Equations (13)–(16). The (erroneous) initial values for the state and input are always set to zero, that is, $x(0) \equiv 0$ and $u(0) \equiv 0$. Consequently, Trajectory A is used to analyse the step response characteristics of the proposed updated law. Trajectory B examines the tracking capability of sinusoidal trajectories with two embedded different frequencies and phases along with an initial step response (the case of $y_{d_2}(0) = 50$).

We first illustrate the boundedness of trajectories in the presence of very large random errors, in particular, we use $Q_k = 4I_3$, $R_k = 2500I_2$, $T_i = 0.1$ sec, we consider Trajectory A, and we run for 1000 sec. Figures 3 and 4 illustrate boundedness of trajectories. Tables 1 and 2 list the performance for different values of $Q_k$ and $R_k$, and run time of 100 sec with $T_s = 0.01$.

Examining Figures 1 and 2 and Tables 1 and 2, the following assessments can be concluded from this example:

- Robustness: Regardless of how large the random measurement errors and random state disturbance are, all trajectories remain bounded – as indicated in Theorem 3.2.

Table 2. Actual standard deviations versus model-reference-based predicted standard deviations: Trajectory A with $T_i = 0.01$ and $kT_i \in [0, 100]$ sec.

<table>
<thead>
<tr>
<th>$\sqrt{Q_k}$</th>
<th>$\sqrt{R_k}$</th>
<th>100$I_2$</th>
<th>50$I_2$</th>
<th>50$I_2$</th>
<th>50$I_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\text{std}<em>{t \in [t_1, t_2]}(\delta \xi(k))^{\text{avg}</em>{t \in [t_1, t_2]} \sqrt{P_\xi(k)}}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\xi = x_1$</td>
<td>9.7 (8.9)</td>
<td>9.7 (8.9)</td>
<td>5.3 (4.5)</td>
<td>3.4 (2.3)</td>
<td></td>
</tr>
<tr>
<td>$\xi = x_2$</td>
<td>7.4 (6.6)</td>
<td>7.3 (6.6)</td>
<td>4.5 (3.4)</td>
<td>3.5 (1.9)</td>
<td></td>
</tr>
<tr>
<td>$\xi = x_3$</td>
<td>6.9 (6.8)</td>
<td>6.7 (6.7)</td>
<td>4.2 (3.5)</td>
<td>3.4 (2.0)</td>
<td></td>
</tr>
<tr>
<td>$\xi = u_1$</td>
<td>3.8 (8.2)</td>
<td>3.2 (6.0)</td>
<td>3.1 (5.9)</td>
<td>3.2 (6.0)</td>
<td></td>
</tr>
<tr>
<td>$\xi = u_2$</td>
<td>3.4 (9.0)</td>
<td>3.6 (6.6)</td>
<td>3.4 (6.6)</td>
<td>3.3 (6.6)</td>
<td></td>
</tr>
</tbody>
</table>
• Estimated errors by the recursive algorithm: The recursive algorithm employs the erroneous reference model. However, by comparing the actual values to their corresponding ones, in Table 2, we find that the estimated values can be indicative to a certain extent. Such estimates can vary to the better or to the worse for different sampling periods or different scenarios but no general conclusion can be drawn as far as accuracy of the estimates extracted from the proposed recursive algorithm is concerned. In addition, there exists some sort of periodic component in the actual errors (e.g., Figures 5 and 6) that are not reflected in $P_k(k)$. The latter issue is unclear to the author and is worthwhile investigating in future work.

4.3. Transient and steady-state performance

This section examines the transient and steady-state characteristics for various sampling periods. Trajectory A is considered for $kT_s \in [0, 5]$ sec with $Q_k = 0$, $R_k = I_2$. In this section, we introduce the following common three measures: $t_r$, OS$_{\%}$ and $e_{\%}$ are the rise time (the time required for the response to rise from 0% to 100% of its final value), maximum per cent overshoot and steady-state error computed by taking the average of the absolute errors during the last two seconds of the run, respectively. Figure 5 shows the response, Figure 6 depicts $|\delta \xi(k)|$ and $\sqrt{P_x(k)}$, $\xi \in [x_1, x_2]$ employing $T_s = 10^{-4}$ and Figure 7 shows how the norm of each PID gain varies with

**Figure 5.** $\xi(k)$ (solid) and $\xi_d(k)$ (dashed), $\xi \in [x_1, x_2, x_3]$ with $Q_k = 0, R_k = I_2$: Trajectory A with $T_s = 10^{-4}$.

**Figure 6.** $|\delta \xi(k)|$ and $\sqrt{P_x(k)}$, $\xi \in [x_2, x_3]$ with $Q_k = 0, R_k = I_2$: Trajectory A with $T_s = 10^{-4}$.

**Figure 7.** $\|K_i(k)\|$, $i \in \{1, 2, 3\}$, $Q_k = 0, R_k = I_2$: Trajectory A with $T_s = 10^{-3}$.
Table 3. Transient and steady-state performance at different sampling rates: Trajectory A with $Q_k = 0$ and $R_k = I_2$.

<table>
<thead>
<tr>
<th>$T_s$</th>
<th>$10^{-1}$</th>
<th>$10^{-2}$</th>
<th>$10^{-3}$</th>
<th>$10^{-4}$</th>
<th>$10^{-5}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$t_s$ sec</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\delta x_2$</td>
<td>1.2</td>
<td>0.61</td>
<td>0.41</td>
<td>0.35</td>
<td>0.34</td>
</tr>
<tr>
<td>$\delta x_3$</td>
<td>1.7</td>
<td>0.64</td>
<td>0.45</td>
<td>0.38</td>
<td>0.37</td>
</tr>
<tr>
<td>OS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\delta x_2$</td>
<td>11%</td>
<td>6.2%</td>
<td>13%</td>
<td>25%</td>
<td>27%</td>
</tr>
<tr>
<td>$\delta x_3$</td>
<td>20%</td>
<td>17.6%</td>
<td>30%</td>
<td>35%</td>
<td>39%</td>
</tr>
<tr>
<td>$e_{ss}$</td>
<td>0.93</td>
<td>0.42</td>
<td>0.2</td>
<td>0.12</td>
<td>0.05</td>
</tr>
<tr>
<td>$\delta x_3$</td>
<td>1.12</td>
<td>0.47</td>
<td>0.2</td>
<td>0.14</td>
<td>0.10</td>
</tr>
</tbody>
</table>

**Table 4.** Steady-state performance at different sampling rates: Trajectory B with $Q_k = 0$ and $R_k = I_2$.

<table>
<thead>
<tr>
<th>$T_s$</th>
<th>$10^{-1}$</th>
<th>$10^{-2}$</th>
<th>$10^{-3}$</th>
<th>$10^{-4}$</th>
<th>$10^{-5}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$e_{ss}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\delta x_2$</td>
<td>2.74</td>
<td>0.79</td>
<td>0.51</td>
<td>0.17</td>
<td>0.07</td>
</tr>
<tr>
<td>$\delta x_3$</td>
<td>2.55</td>
<td>0.79</td>
<td>0.4</td>
<td>0.14</td>
<td>0.06</td>
</tr>
</tbody>
</table>

Figure 8. $\xi(k)$ (solid) and $\xi_d(k)$ (dashed), $\xi \in [x_1, x_2, x_3]$ with $Q_k = 0$, $R_k = I_2$: Trajectory B with $T_s = 10^{-4}$.

Figure 9. $|\delta \xi(k)|$ and $\sqrt{P_\xi(k)}$, $\xi \in [x_2, x_3]$ with $Q_k = 0$, $R_k = I_2$: Trajectory A with $T_s = 10^{-4}$.

- The steady-state error decreases as the sampling rate increases (as indicated in Theorem 3.2). In particular, as sampling rate is increased by an order of magnitude, the error is decreased by 50% for Trajectory A, and at a faster rate for Trajectory B (Table 4).

In addition, since the steady-state errors decrease at a geometric rate as the sampling rate is increased in the presence of measurement noise ($R_k = I_2$), it can be concluded that the updated law can completely reject measurement noise, and can achieve arbitrary small transient period and arbitrary small steady-state errors for sufficiently small sampling period at a cost of larger overshoot.

In order to show the advantage of the proposed stochastic algorithm over deterministic approach for the selection of PID gains, we compare the performance of our proposed method with the deterministic PID method that was recently published in Saab and Toukhtarian (2015). Figure 9 considers $|\delta \xi(k)|$ and $\sqrt{P_\xi(k)}$, $\xi \in [x_1, x_2]$ with $T_s = 10^{-4}$. As expected, the tracking errors pertaining to the method in Saab and Toukhtarian (2015) are tainted with measurement noise.

**Example 4.2:** The main purpose of this example is to compare the performance of proposed controller to a multivariable MRAC. In this example, we consider the linearised lateral motion dynamic model of a large transport airplane (Guo et al., 2009) described by

$$\dot{x} = A_x x + B_x u, \quad x = \begin{bmatrix} v_b & p_b & r_b & \varphi & \psi \end{bmatrix}^T, \quad u = \begin{bmatrix} d_r & d_a \end{bmatrix}^T$$

where $v_b$ is the lateral velocity, $p_b$ is the roll rate, $r_b$ is the yaw rate, $\varphi$ is the roll angle, $\psi$ is the yaw angle, $d_r$ is the rudder position and $d_a$ is the aileron position. The state
matrix, $A$ and the output-coupling matrix $B$, as in Guo et al. (2009), are given by

$$A_c = \begin{bmatrix}
-0.13858 & 14.326 & -219.04 & 323.167 & 0 \\
-0.02073 & -2.1692 & 0.91315 & 0.000256 & 0 \\
0.00289 & -0.16444 & -0.15768 & -0.00489 & 0 \\
0 & 1 & 0.00618 & 0 & 0 \\
0 & 0 & 1 & 0 & 0
\end{bmatrix}$$

$$B_c = \begin{bmatrix}
0.15935 & 0.00211 \\
0.01264 & 0.021326 \\
-0.012879 & 0.00171 \\
0 & 0 \\
0 & 0
\end{bmatrix}$$

It is important to note that one eigenvalue of the continuous-time state matrix corresponding to the yaw angle is zero. The problem is in having the lateral velocity and the yaw angle tracking the following reference trajectories

$$v_{\text{ref}}^b = 5 \sin(0.01t) \text{ ft/sec and } \psi_{\text{ref}} = 5 - 5 \cos(0.01t) \text{ deg.}$$

Consequently, as in Guo et al. (2009), we use only two output measurements, the lateral velocity and the yaw angle. Thus, the output-coupling matrix

$$C = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

It is important to note that the product of the output–input coupling matrices $CB_c$, associated with the continuous-time plant is not full rank. However, due to discretisation, the corresponding product, $CB$, becomes full rank. We discretise the plant with sampling period $T_s = 8 \times 10^{-2}$ sec.

Unlike the work in Guo et al. (2009), we corrupt the output measurements with additive zero-mean white Gaussian noise. In particular, the standard deviations corresponding to the lateral velocity and the yaw angle are $\sigma_1 = 5 \times 10^{-4}$ ft/sec and $\sigma_2 = 0.1$ deg, respectively. We also use an erroneous model for the measurement noise, in particular, $R_k = \begin{bmatrix} 2 \times 10^{-3} & 0 \\ 0 & 10^{-9} \end{bmatrix}$. We set the initial control input to zero, $P_0 = 0$, and simulate the system for 1000 sec. The relevant aircraft outputs, lateral velocity, yaw rate and yaw angle are shown along with their corresponding reference signals in Figure 10. The outputs uniformly track their reference trajectories. The corresponding control signals, rudder position and aileron position shown in Figure 11, can be considered practically acceptable.

5. Conclusion and future work

In this paper, the ability of a PID controller in rejecting measurement noise has been demonstrated. The advantages of the proposed recursive algorithm is in: (1) withstanding process noise and measurement noise of any
size, (2) rejecting measurement noise and errors due to random initial conditions, (3) automatically producing the PID gains in a discrete-time fashion, (4) providing indicative estimates of errors at every time sample and (5) skilfully tracking any realisable reference trajectories in the presence of measurement noise. Although the presented theory is based on full knowledge of the plant model, simulation results showed that an erroneous model could achieve all presented characteristics of the proposed recursive algorithm.

Note

1. A more common form of a PID control law is given by

\[ u(k) = u(k-1) + \sum_{i=1}^{3} K_i (k) e(k + 1 - i). \]

Acknowledgments

The author would like to thank the anonymous reviewers for their constructive suggestions in improving this work.

Disclosure statement

No potential conflict of interest was reported by the authors.

References


Development of the proposed recursive algorithm.

Processes are uncorrelated since $\delta x(0), \nu(\cdot)$ and $w(\cdot)$ are assumed to be zero-mean white noise. Consequently, $E[(\Phi X)(-\Gamma Z + \Omega)\Gamma^T] = 0$ and $E[ZZ^T] = \begin{bmatrix} \rho & 0 & 0 \\ 0 & \rho & 0 \\ 0 & 0 & \rho \end{bmatrix}$ Therefore,

$$P^+ = \Phi \Phi^T + \Gamma E\left[ZZ^T\right] \Gamma^T + \Gamma E\left[\Omega \Omega^T\right]$$ (A2)

Since $u(k)$ generates $x(k + 1)$ (but not $x(k)$), we assume that $E[\delta x(k)\delta u(k)\Gamma^T] \equiv 0$ and $E[\delta x(k + 1)\delta u(k)^T] \equiv 0$ (but $P_u = E[\delta x(k + 1)\delta u(k)^T] \not\equiv 0$). That is,

$$P = \begin{bmatrix} P_{uu} & P_{uz} \\ P_{zu} & P_z \end{bmatrix}$$

Consequently, expanding Equation (A2) and taking the trace, yield

$$\text{trace}\{P^+\} = \begin{bmatrix} (\Psi_1 P_{uu} + \Psi_2 P_{uz}^T) \Psi_{11}^T + (\Psi_1 P_{uu} + \Psi_2 P_{uz}^T) \Psi_{12}^T \\ + \Psi_2 P_{uz}^T K_{T2} P_{uu} + B^T w_s + A' P_{2} (A')^T \\ + K_{T1} R_{K2} K_{T1}^T + K_{T1} R_{K2} K_{T1}^T + K_{T1} C^T Q K_{T1} \end{bmatrix} + \begin{bmatrix} (K_{T1} C + A) Q (K_{T1} C + A) \end{bmatrix} Q + \Pi_d$$

Expanding and collecting terms and with some straightforward tedious work, we obtain

$$\text{trace}(P^+) = \begin{bmatrix} -K_{T1} C + A \Psi_{11}^T K_{T1}^T - K_{T1} C + A^T \Psi_{12}^T K_{T1}^T \\ -P_{uu} (C' AB - \omega_u)^T K_{T1}^T - K_{T1} C + A^T \Psi_{12}^T K_{T1}^T \\ + K_{S11} K_{T1}^T + K_{S12} K_{T1}^T + K_{S13} K_{T1}^T \\ + K_{S21} K_{T1}^T + K_{S22} K_{T1}^T + K_{S23} K_{T1}^T \\ + K_{S31} K_{T1}^T + K_{S32} K_{T1}^T \\ + 2P_u + B' w_s (B')^T + A' P_{2} (A')^T + Q + \Pi_d \\ \end{bmatrix}$$

where the terms $S_{ij}, i, j \in \{1, 2, 3\}$, are defined in Proposition 3.1. Equation (11) leads to

$$\frac{\partial \text{trace}(P^+)}{\partial K_1} = 2 \begin{bmatrix} K_{S11} + K_{S21} + K_{S31} \\ -P_{uu} (C' B)^T - P_{uu} (C' AB - \omega_u)^T \end{bmatrix} \equiv 0$$

$$\frac{\partial \text{trace}(P^+)}{\partial K_2} = 2 \begin{bmatrix} K_{S22} + K_{S12} + K_{S32} - P_{uu} (C'B)^T \end{bmatrix} \equiv 0$$

$$\frac{\partial \text{trace}(P^+)}{\partial K_3} = 2 \begin{bmatrix} K_{S33} + K_{S13} + K_{S23} \end{bmatrix} \equiv 0$$

Putting the above three equations in a block matrix, we obtain

$$\begin{bmatrix} K_1 & K_2 & K_3 \end{bmatrix} \begin{bmatrix} S \\ S \end{bmatrix} = \begin{bmatrix} P_u (C' B)^T + P_{uu} (C' AB - \omega_u)^T P_{uu} (C'B)^T \end{bmatrix} 0$$

whence Equation (13). Inserting Equation (5) in $P_x = E[\delta x(k)\delta x(k)^T]$, expanding and cancelling the uncorrelated terms yield Equation (14). Equation (15) can be easily extracted from Equation (A2). Again, after expanding and arranging terms of the first $p \times p$ block of Equation
and making use of $S$, we obtain

\[
P_{un}^+ = P_u - K_1C^+BP_u - P_u(C^+B)^T K_1^T - K_2C^+AB\tilde{D}_{un}^T
- P_{un}(C^+AB)^T K_1^T - K_2CB\tilde{D}_{un}^T
- P_{un}(CB)^T K_2^T + KS K^T + \Pi_d \quad (A3)
\]

Inserting the value of $K$ presented in Equation (13) in Equation (A3), and cancelling redundant terms, we obtain Equation (16).

**Proof of Proposition 3.2:** Equation (A3) implies that $P_u(k), \forall k$, is symmetric (and positive semi-definite) matrix. Equations (15) and (16) imply that

\[
P_{un}^-(k + 1) = P_u(k + 1) - \Pi_d \quad (A4)
\]

Since $\Pi_d$ is symmetric, $P_{un}^-(k)$ is also symmetric. It is worthwhile noting that

\[
I_p - KSK^T = \Psi_1 P_u + \Psi_2 P_{un}^- = P_{un}^+
\]

**Claim 1:**

\[
(I_{2q} + \hat{P}_u \hat{N}^T (\hat{N} \hat{P}_u \hat{N}^T + \tilde{S}_{2q})^{-1}\hat{N})^{-1}
\]

**Proof:** Let $\Xi \equiv (\hat{N}^T \tilde{S}_{2q}^{-1} \hat{N} + \hat{P}_u)^{-1}$ and $\Lambda \equiv \hat{N}^T \tilde{S}_{2q}^{-1} \hat{N}$

Using a well-known matrix inversion lemma

\[
\hat{P}_u \hat{N}^T (\hat{N} \hat{P}_u \hat{N}^T + \tilde{S}_{2q})^{-1}\hat{N} = \hat{P}_u \Lambda - \hat{P}_u \Lambda \Xi \Lambda
\]

\[
= I_{2p} + \hat{P}_u \Lambda - \hat{P}_u \Lambda \hat{P}_u \Lambda - \hat{P}_u \Lambda \Xi \Lambda - \hat{P}_u \Lambda \Xi \hat{P}_u \Lambda
\]

Cancelling the second and third terms, and rearranging the others while using $\Xi^{-1} \equiv \Lambda + \hat{P}_u^{-1}$

\[
(I_{2p} - \hat{P}_u \Lambda - \hat{P}_u \Lambda \Xi \Lambda) (I_{2p} + \hat{P}_u \Lambda)
\]

\[
= I_{2p} - \hat{P}_u \Lambda \hat{P}_u \Lambda - \hat{P}_u \Lambda \Xi \hat{P}_u \Lambda = I_{2p}
\]

**Proof of Proposition 3.3:** Equation (13) implies

\[
\hat{K} = [I_p 0]\hat{P}_u \hat{N}^T [I_{2q} 0]S^{-1} [I_{2q} 0] \quad (A5)
\]

where $\hat{K} \equiv [K_1 K_2]$. Based on the definitions of $\hat{S}, \hat{N}, \hat{P}_u$ and $\hat{N}, S$ is presented as

\[
S = \hat{S} + [\hat{N}^T 0] \hat{P}_u [\hat{N}^T 0] \quad (A6)
\]

It is worthwhile noting that $\hat{S}$ does not include any direct term related to $\hat{P}_u$. Using a well-known matrix inversion lemma, we have $S^{-1} = \hat{S}^{-1} - \hat{S}^{-1} [\hat{N}^T 0] (\hat{N}^T \hat{S}_{2q}^{-1} \hat{N} + \hat{P}_u)^{-1} \hat{N}^T \hat{S}_{2q}^{-1}$

Consequently,

\[
\hat{S}_{2q}^{-1} = [I_{2q} 0]S^{-1} [I_{2q} 0]
\]

\[
= \hat{S}_{2q}^{-1} - \hat{S}_{2q}^{-1} \hat{N} (\hat{N}^T \hat{S}_{2q}^{-1} \hat{N} + \hat{P}_u)^{-1} \hat{N}^T \hat{S}_{2q}^{-1} \quad (A7)
\]

Again, using the same inversion lemma, we have

\[
[I_{2q} 0]S^{-1} [I_{2q} 0] = (\hat{S}_{2q} + \hat{N} \hat{P}_u \hat{N}^T)^{-1} \quad (A8)
\]

Inserting Equation (A8) into Equation (A5), we obtain

\[
\hat{K} = [I_p 0] \hat{P}_u \hat{N}^T (\hat{N} \hat{P}_u \hat{N}^T + \tilde{S}_{2q})^{-1} \quad (A9)
\]

Next, we evaluate

\[
P_u^+ = ([I_p 0] - \hat{K} \hat{N}) \hat{P}_u [I_p 0] + \Pi_d \quad (A10)
\]

Inserting Equation (A9) in Equation (A10), we get

\[
P_u^+ = [I_p 0] (I_{2p} - \hat{P}_u \hat{N}^T (\hat{N} \hat{P}_u \hat{N}^T + \tilde{S}_{2q})^{-1} \hat{N}) \hat{P}_u [I_p 0] + \Pi_d, \text{ where } \hat{P}_u [I_p 0] = [P_{un}^-]. \text{ Claim 1 implies}
\]

\[
P_u^+ = [I_p 0] (I_{2p} + \hat{P}_u \hat{N}^T \tilde{S}_{2q}^{-1} \hat{N})^{-1} [P_u 0] + \Pi_d \quad (A11)
\]

Equation (A4) implies that

\[
P_{un}^+ = [I_p 0] (I_{2p} + \hat{P}_u \hat{N}^T \tilde{S}_{2q}^{-1} \hat{N})^{-1} [P_u 0] \quad (A12)
\]

Combining Equations (A11) and (A12), we obtain Equation (17).

**Proof of Lemma 3.1:** The proof is preceded by an induction argument with respect to index $k$. Since $C(k + 1)B(k)$ is full-column rank $\forall k \geq 0$, $\hat{N} = \begin{bmatrix} C^B & C^A & AB \\ 0 & 0 & C^B \\ 0 & 0 & CB \end{bmatrix}$ is also of column rank $\forall k \geq 0$. $\hat{P}_u(1) = E([\delta_{u0}^T [\delta_{u0}^T])^T$ is by structure symmetric, semi-positive and diagonally dominant matrix. Since $P_u(1)$ and $P_{un^-}(1)$ are symmetric positive-definite matrices, $\hat{P}_u(1)$ is symmetric positive-definite matrix. Since $\tilde{S}_{2q}^{-1}$ is symmetric and positive definite and $\hat{N}$ is full-column rank, $\hat{N}^T \tilde{S}_{2q}^{-1} \hat{N}$ is symmetric positive definite. Since for, $k = 1$ $P_u(1) > 0$, then $\lambda(\hat{P}_u \hat{N}^T \tilde{S}_{2q}^{-1} \hat{N}) > 0$; hence, $\lambda(\hat{I}_{2p} + \hat{P}_u \hat{N}^T \tilde{S}_{2q}^{-1} \hat{N}) > 1$ and

\[
0 < \lambda((I_{2p} + \hat{P}_u \hat{N}^T \tilde{S}_{2q}^{-1} \hat{N})^{-1}) < 1. \text{ Since } [I_p 0] \text{ is of full}
\]
row rank, $[P_{uu}]$ is of full-column rank and $\Pi_d$ is symmetric semi-definite, Equations (A11) and (A12) imply that both $P_u(2)$ and $P_{uu}(2)$ are symmetric positive-definite matrices. The rest of the induction argument follows the same steps.

**Proof of Corollary 3.1:** Since $0 < \lambda((I_{2p} + \tilde{P}_u \tilde{N}^T \tilde{S}_{2q}^{-1} N)^{-1}) < 1$, then there exists a consistent norm such that $\|(I_{2p} + \tilde{P}_u \tilde{N}^T \tilde{S}_{2q}^{-1} N)^{-1}\| < 1$ and $\|I_p 0\| = 1$. Consequently, Equation (17) implies the desired results.

**Proof of Theorem 3.1:** There exists a consistent norm such that (Corollary 3.1)

$$
\begin{align*}
\left\| \begin{bmatrix} P_u(k + 1) \\ P_{uu}(k + 1) \end{bmatrix} \right\| &\leq \left\| \begin{bmatrix} I_{2p} + \bar{P}_u \tilde{N}^T \tilde{S}_{2q}^{-1} \tilde{N} \end{bmatrix}^{-1} \right\| \left\| \begin{bmatrix} P_u(k) \\ P_{uu}(k) \end{bmatrix} \right\| \\
&+ \left\| \begin{bmatrix} \Pi_d \\ 0 \end{bmatrix} \right\|
\end{align*}
$$

(A13)

Since $u_d(k) = u_d(k + 1) \forall k \geq k_1, \Pi_d = 0, \forall k \geq k_1$. Define $\bar{P}_u(k) \equiv [P_u(k) P_{uu}(k)]$. Now we have $\forall k > k_1$

$$
\|\bar{P}_u(k + 1)\| \leq \left\| (I_{2p} + \mathbf{\bar{P}}_u \bar{N}^T \mathbf{S}_{2q}^{-1} \bar{N})^{-1} \right\| \|\bar{P}_u(k)\|
$$

We show that $\lim_{k \to \infty} \bar{P}_u(k) = 0$. This part of the proof is preceded by a contradiction argument. Since $\|(I_{2p} + \tilde{P}_u \tilde{N}^T \tilde{S}_{2q}^{-1} \tilde{N})^{-1}\| < 1$, Equation (A13) implies that $\|\bar{P}_u(k)\|$ is strictly decreasing sequence and bounded from below by zero, $\lim_{k \to \infty} \|\bar{P}_u(k)\|$ exists so do $\lim_{k \to \infty} \|\mathbf{\bar{P}}_u(k)\|$ and $\lim_{k \to \infty} \|P_u(k)\|$. Assume that $\lim_{k \to \infty} \|\mathbf{\bar{P}}_u(k)\| \neq 0$. Consider the following ratio test for the sequence $\|\bar{P}_u(k)\|:

$$
\frac{|\bar{P}_u(k + 1)\|}{|P_u(k)\|} \leq \left\| (I_{2p} + \mathbf{\bar{P}}_u \bar{N}^T \mathbf{S}_{2q}^{-1} \bar{N})^{-1} \right\| < 1.
$$

Since $\forall k, w(k)$ and $P_u(k)$ are bounded, so is $u(k)$. Since the system is assumed to be asymptotically stable, $x(k) = \sup_{\forall k} \|x(k)\|$ and $\lim_{k \to \infty} \|u(k)\|$ is bounded. Consequently, $\lim_{k \to \infty} S(k)$ is bounded; hence, $\lim_{k \to \infty} \mathbf{S}_{2q}^{-1}(k)$ remains symmetric and positive definite as well as $\lim_{k \to \infty} \mathbf{\hat{N}}^T \mathbf{S}_{2q}^{-1}(k) \mathbf{\hat{N}}$. Since $\|\mathbf{\bar{P}}_u(k)\|$ is also bounded, $\lim_{k \to \infty} \lambda(I_{2p} + \mathbf{\bar{P}}_u \mathbf{\hat{N}}^T \mathbf{\hat{S}}_{2q}^{-1} \mathbf{\hat{N}})^{-1} < 1$ and, for this specific consistent norm, $\lim_{k \to \infty} \|\mathbf{\bar{P}}_u(k)\| < 1$. Therefore, $\lim_{k \to \infty} \|\mathbf{\bar{P}}_u(k)\| < 1$. Since $\|\mathbf{\bar{P}}_u(k)\| > 0$, the ratio test implies that $\lim_{k \to \infty} \|P_u(k)\| = 0$, which is equivalent to $\lim_{k \to \infty} P_u(k) = 0$. Next, having $w(k) = 0, \forall k \geq k_1$ then $Q(k) = 0, \forall k \geq k_1$. Since $\lim_{k \to \infty} P_u(k) = 0$ and system is assumed to be asymptotically stable, Equation (14) implies that $\lim_{k \to \infty} P_u(k) = 0$.

**Proof of Theorem 3.2:** $\Pi_d = E[\Delta u_d(k) \Delta u_d(k)^T]$, where $\Delta u_d(k) = u_d(k + 1) - u_d(k)$. Since $u_d(k)$ and its limit are bounded, $\|\Pi_d\|$ and its limit are bounded.

Define $M(k) \equiv \|(I_{2p} + \mathbf{\bar{P}}_u \bar{N}^T \mathbf{\hat{S}}_{2q}^{-1} \mathbf{\hat{N}})^{-1}\|, \quad P(k) \equiv \left\| \begin{bmatrix} P_u(k) \\ P_{uu}(k) \end{bmatrix} \right\|$ and $D(k, T_i) \equiv \left\| \begin{bmatrix} \Pi_d \\ 0 \end{bmatrix} \right\|$. It follows that $P(k + 1) \leq M(k)P(k) + D(k, T_i), \forall k, \Pi_d$ and $D(k, T_i)$ are bounded, $\lim_{k \to \infty} D(k, T_i)$ is bounded and

$$
P(k + 1) \leq M(k)P(k) + c_D(T_i)
$$

(A14)

where $c_D(T_i) = \sup_{\Pi_d} D(k, T_i)$. All variables become dependent on $T_i$ and especially whenever a continuous system is discretised with a sampling period $T_i$. However, for compactness, we drop the argument $T_i$ and use it when relevant. $\forall k, \lambda(I_{2p} + \mathbf{\bar{P}}_u \bar{N}^T \mathbf{\hat{S}}_{2q}^{-1} \mathbf{\hat{N}})^{-1} < 1, M(k) < 1$ and $P_u$ are bounded. Next, we show that $\lim_{k \to \infty} M(k) < 1$. If $\lim_{k \to \infty} \lambda(P_u \tilde{N}^T \tilde{S}_{2q}^{-1} \tilde{N}) > 0$, then $\lambda(I_{2p} + \mathbf{\bar{P}}_u \mathbf{\hat{N}}^T \mathbf{\hat{S}}_{2q}^{-1} \mathbf{\hat{N}})^{-1} < 1$. For all $k$, $\mathbf{\bar{P}}_u$ and $\tilde{N}^T \tilde{S}_{2q}^{-1}$ are positive definite, and $\lim_{k \to \infty} \mathbf{\hat{N}}$ is bounded. Equation (A1) implies that $\tilde{S} = H_2 E_2 H_3^T + H_3 E_2 H_2^T + T$, which indicates that $\tilde{S}$ is independent of $\mathbf{\bar{P}}_u$(refer to the proof of Proposition 3.1). In addition, $\forall k, H_2 E_2 H_3^T \geq 0, \lim_{k \to \infty} (H_2 E_2 H_3^T + T) > 0$ is positive definite with bounded limit, and $\lim_{k \to \infty} (H_2)$ is bounded. Since system is assumed to be asymptotically stable, $\lim_{k \to \infty} H_2 E_2 H_3^T \geq 0$ and bounded. Therefore, $\lim_{k \to \infty} \mathbf{\bar{P}}_u \bar{N}^T \mathbf{\hat{S}}_{2q}^{-1} \mathbf{\hat{N}} = \tilde{S}_{2q}^{-1}$ is positive definite, and $\lim_{k \to \infty} \mathbf{\hat{S}}_{2q}^{-1} \mathbf{\hat{N}}$ is bounded. Equation (A14) implies $P(k + 1) \leq \prod_{j=0}^{k-1} M(j)P(0) + c_D(T_i)$. In addition, there exists a $c > 0$ such that $\sup_{T_i} \lim_{k \to \infty} \sum_{j=0}^{k-1} \prod_{i=0}^{j-1} M(j + 1 - i) = c$. Equation (A14) implies $P(k + 1) \leq \prod_{j=0}^{k-1} M(j)P(0) + c_D(T_i)$. In addition, $\lim_{k \to \infty} \prod_{j=0}^{k-1} M(j + 1 - i) = 0$ or $P(k + 1) \leq \prod_{j=0}^{k-1} M(j)P(0) + c_D(T_i)$. In addition, when $u_d(k)$ is assumed to be a smooth trajectory, then $\exists \varepsilon > 0$ such that $c_D(T_i) \leq c_\varepsilon T_i$. We can deduce then that $\lim_{k \to \infty} P(k) \leq c_\varepsilon T_i$ and $\lim_{k \to \infty} \|P_u(k)\| \leq c_\varepsilon T_i$, where $c_\varepsilon \equiv c_\varepsilon$.

Since $\|P_u(k)\| \leq P(k)$, then the above further leads to

$$
\|P_u(k)\| \leq m_\varepsilon(k) + c_\varepsilon T_i
$$

(A15)
where \( m_o(k) \equiv \prod_{j=0}^{k-1} M(j)P(0) \), with \( \lim_{k \to \infty} m_o(k) = 0 \). Next, since system is assumed to be asymptotically stable, then, for discrete-time linear systems, we have \( \lambda(A(k)) < 1 \) and, consequently, there exists a consistent norm such that \( \|A(k)\| < 1 \) and \( \lim_{k \to \infty} \|A(k)\| < 1 \). Having \( u(k) = 0, \forall k \) then \( Q(k) = 0, \forall k \), then we can deduce from Equation (14) that 
\[
\|P_x(k)\| \leq \|A(k-1)\|^2\|P_x(k-1)\| + \|B(k-1)\|^2\|P_o(k-1)\|
\]
This, together with Equation (A15), yields
\[
\|P_x(k)\| \leq c_A^2 \|P_x(k-1)\| + c_B^2 (m_o(k) + ccT_s)
\]
(A16)

where \( 0 < c_A \equiv \sup_k \|A(k)\| < 1 \) and \( 0 < c_B \equiv \sup_k \|B(k)\| < \infty \) and \( \lim_{k \to \infty} m_o(k) = 0 \). We can write the inequality (A16) as 
\[
\|P_x(k)\| \leq (c_A^2)^k\|P_x(0)\| + c_B^2 (m_o(k) + ccT_s) \sum_{j=0}^{k-1} (c_A^2)^j.
\]
Consequently, \( \lim_{k \to \infty} \|P_x(k)\| \leq c_xT_s \), where \( c_x \equiv \frac{c_B^2}{1-c_A^2} > 0 \).