Achieving State Estimation Equivalence for Misassigned Disturbances in Offset-Free Model Predictive Control

Murali R. Rajamani
BP Research and Technology, Naperville, IL 60563

James B. Rawlings
Dept. of Chemical and Biological Engineering, University of Wisconsin, Madison, WI 53706

S. Joe Qin
The Mork Family Dept. of Chemical Engineering and Materials Science, Ming Hsieh Dept. of Electrical Engineering, Daniel J. Epstein Dept. of Industrial and Systems Engineering, University of Southern California, Los Angeles, CA 90089

DOI 10.1002/aic.11673
Published online January 6, 2009 in Wiley InterScience (www.interscience.wiley.com).

Integrated white noise disturbance models are included in advanced control strategies, such as Model Predictive Control, to remove offset when there are unmodeled disturbances or plant/model mismatch. These integrating disturbances are usually modeled to enter either through the plant inputs or the plant outputs or partially through both. There is currently a lack of consensus in the literature on the best choice for the structure of this disturbance model to obtain good feedback control. We show that the choice of the disturbance model does not affect the closed-loop performance if appropriate covariances are used in specifying the state estimator. We also present a data-based autocovariance technique to estimate the appropriate covariances regardless of the plant’s true unknown disturbance source. The covariances estimated using the autocovariance technique and the resulting estimator gain are shown to compensate for an incorrect choice of the source of the disturbance in the disturbance model.

Keywords: state estimation, disturbance models, model predictive control, model equivalence

Introduction

Model Predictive Control (MPC) has emerged as one of the most important advanced control methods in the chemical industry. MPC casts the control problem in the form of an optimization, which makes it convenient to handle constraints and nonlinear models explicitly. Most current formulations of MPC have a state estimator along with the dynamic constrained regulator and a target calculator. Nominal asymptotic stability for a deterministic linear MPC was shown in Ref. 2 under the assumption that the model is known perfectly and there is full state information. With random disturbances, however, the state is no longer known exactly and one requires dynamic feedback from the measurements to update the current estimate of the state. The function of the...
state estimator is to reconcile the past and current measurements to estimate the current state of the system. Nominal stability of the state estimator combined with the regulator was shown in Ref. 3. The regulator and the target calculator require estimates of the current state and the disturbance to make a prediction of optimal control moves over a future horizon.

Any correlated disturbance that enters the plant and is not included in the model results in biased predicted states. It is customary to assume that the disturbance enters through the inputs or the outputs. The bias can be removed by using the predicted and measured outputs to estimate the disturbance. In the target calculator formulation, the state is augmented with integrating disturbances. The target calculator then ensures offset free control in presence of unmodeled disturbances\(^4\) by shifting the steady state target for the regulator depending on the estimate of the bias (under the assumption that the bias remains constant in the future). The integrating disturbance can be added either to the input,\(^5,6\) the output, \(^7,8\) or a combination of both.\(^9\)–\(^11\) Most industrial MPC implementations add the bias term to the output.\(^1,12\) In Refs. 10 and 11, rank conditions that should be satisfied for the disturbance models to ensure offset-free control were independently derived, and the lack of consensus in the literature on disturbance models to ensure offset-free control were independently derived, and the lack of consensus in the literature on the choice of the disturbance model was also pointed out. The examples in these references show significant difference in the closed-loop behavior of the controller depending on the choice of the disturbance model. Recently,\(^1,13\) a method for a combined design of a disturbance model and its observer by solving an appropriate \(H_\infty\) control problem was presented. The results presented in Ref. 13 differ considerably from the autocorrelation-based technique to be discussed in this article. The autocorrelation-based technique uses steady state operating data to specify the estimator gain as opposed to designing the disturbance models for a specified disturbance.

The main contribution of this article is to show that for linear models, the choice of the disturbance model does not affect the performance of the closed-loop when the estimator gain is found from appropriately identified noise covariances. The equivalence in the estimator gains is shown by appealing to realization theory.\(^14\) Subsequently, the performance of any regulator that is a feedback on the current estimate of the state remains unaffected by the choice of the disturbance model. The covariances can be estimated from data using autocovariances at different time lags as shown in Ref. 15. The state estimator gain thus determined from the estimated covariances gives optimal closed-loop performance.

The rest of this article is organized as follows: In the next section the equations for estimation, regulation, and target calculation in MPC are given. We motivate the rest of the article by examining the consequences of an incorrect disturbance model in the section “Motivation.” In the section “Model Equivalence,” the equivalence transformation between arbitrary choices of the disturbance models is formulated. The section “Using Correlations to Estimate Filter Gain” presents the autocovariance least-squares technique for estimating the covariances for incorrectly chosen disturbance models, followed by examples in the section “Simulation Examples.” Finally, conclusions are presented in the last section.

### Background

We start with the following discrete time state-space model of the plant.

#### Plant

\[
x_{k+1} = Ax_k + Bu_k + Gw_k \\
y_k = Cx_k + v_k
\]  

(1)

in which \(x_k \in \mathbb{R}^n\) is the state of the plant at time \(t_k\) with inputs \(u_k \in \mathbb{R}^m\), \(y_k \in \mathbb{R}^p\) is the vector containing the measurements. Both the states and the measurements are assumed to have Gaussian zero mean white noises \(w_k\), \(v_k\) corrupting them. Let the covariances of \(w_k\), \(v_k\) be \(Q_w\), \(R_v\) respectively. We also assume that \(w_k\), \(v_k\) are not cross correlated. Dimensions of the matrices are \(A \in \mathbb{R}^{n \times n}\), \(B \in \mathbb{R}^{n \times m}\), \(G \in \mathbb{R}^{p \times m}\), and \(C \in \mathbb{R}^{p \times n}\). The pair \((A, C)\) is assumed to be observable and the pair \((A, B)\) assumed controllable. Let \(y_k^\text{f}\) be the set of controlled outputs that are a subset of the measurements \(y_k\) given by \(y_k^\text{f} = Hy_k\). In the rest of this article, we assume without loss of generality that the set of controlled outputs are the same as the measured outputs with \(H = I\). We also use the term “measurements” and “outputs” interchangeably to mean \(y_k\). The analysis in this article remains the same when \(y_k^\text{f}\) is only a subset of \(y_k\).

We assume that the deterministic parameters \((A, B, C)\) of the plant in Eq. 1 are known Modeled from first principles or estimated using identification techniques such as subspace ID\(^6,13,17\) or prediction error methods,\(^18\) and the stochastic parameters \(G, Q_w, R_v\) are unknown. Even though the deterministic parameters may be known perfectly, unmodeled disturbances can enter through the inputs or the outputs and result in an offset in the controller. A physical example of such a disturbance is a leak in the feed pipe causing decreased input flow rate or cooling liquid contamination decreasing the effective cooling rate. Such a disturbance can often be adequately described as a integrating white noise with some small unknown covariance \(Q_\xi\) for \(\xi_k\).

#### Unmodeled disturbance in the plant

\[
d_{k+1} = d_k + \xi_k
\]  

(2)

To ensure offset-free control in the presence of unmodeled disturbances MPC models include an integrating disturbance entering at either the inputs or the outputs of the system.\(^19,20\) Let \(d_k \in \mathbb{R}^m\) be the integrating disturbance entering the state through the matrix \(B_d \in \mathbb{R}^{n \times m}\) and the measurement through \(C_d \in \mathbb{R}^{p \times n}\). The state \(x\) in the model is augmented with the integrating disturbance.

#### Augmented model

\[
\begin{bmatrix} x \\ d \end{bmatrix}_{k+1} = \begin{bmatrix} A & B_d \\ 0 & I \end{bmatrix} \begin{bmatrix} x \\ d \end{bmatrix}_k + \begin{bmatrix} B \\ 0 \end{bmatrix} u_k + \begin{bmatrix} G \\ 0 \end{bmatrix}\begin{bmatrix} w \\ \xi_k \end{bmatrix}_k \\
y_k = \begin{bmatrix} C & C_d \end{bmatrix} \begin{bmatrix} x \\ d \end{bmatrix}_k + v_k
\]  

(3)
The structure of the disturbance $B_d$, $C_d$ is usually unknown and a general unstructured disturbance model can have disturbances entering both the states and the outputs. The conditions on the models $B_d$, $C_d$ for detectability and to ensure offset free control were proved in Refs. 10 and 11.

In addition to the covariances $GQwG^T$, $R$, the covariance $Q$ of the disturbance $\zeta_k$ is also required to specify the state estimator gain. Usual choices for the disturbance model are as follows:

(1) $B_d = 0$, $C_d = I$ called the output disturbance model, or
(2) $B_d = B$, $C_d = 0$ called the input disturbance model.

The number of disturbances $n_d$ has to be chosen as $n_d = p$ to ensure zero offset in all the outputs. Note that the input and output disturbance models have identical minimal realizations of a $n_y$-model and both add an uncontrollable, integrating mode (disturbance). Minimal realizations of both input/output models are therefore identical.

With Eq. 1 representing the plant, Eq. 2 for the unmodeled disturbance in the plant, and the augmented model in Eq. 3, the state estimation and regulation in MPC are described in the following subsections. The basic block diagram for the linear MPC scheme is shown in Figure 1.

**State estimation**

For linear time-invariant (LTI) models like the ones considered in this article, the Kalman filter is the optimal state estimator and consists of the set of simple recursive calculations shown below. Let $\hat{x}_{ijk}$ be the filtered estimate of the state $x_k$ given data up to $t_k$ and $\hat{\xi}_{ijk}$ be the predicted state given data up to $t_{k-1}$. The prediction equations are

$$\hat{x}_{k+1|k} = A\hat{x}_{ik} + Bu_t + B_d\hat{\xi}_{ijk}$$

and the filtering equations are

$$\hat{x}_{ijk} = \hat{x}_{ijk-1} + L_x(y_k - C\hat{x}_{ijk-1} - C_d\hat{\xi}_{ijk-1})$$
$$\hat{\xi}_{ijk} = \hat{\xi}_{ijk-1} + L_d(y_k - C\hat{x}_{ijk-1} - C_d\hat{\xi}_{ijk-1})$$

where $L_x$, $L_d$ are the steady state Kalman filter gains calculated using covariances $GQwG^T$, $Q$, $R$ and using the Riccati equation for the augmented system defined in Eq. 3. For a time invariant model the steady state Kalman gains $L_x$, $L_d$ are also time invariant.

$$P_a = A_aP_aA_a^T - A_aP_aC_a^T(C_aP_aC_a^T + R) - 1 C_aP_aA_a^T$$
$$+ [GQwG^T \ 0]$$

and

$$L_a = P_aC_a(C_aP_aC_a^T + R)^{-1}$$

in which, $L_a = [L_x \ L_d]^T$.

Other advanced state estimators, e.g. the Moving Horizon Estimator, are required when the model is nonlinear or the system has constraints.

**Target calculation**

To reject the unmodeled disturbances, the target calculator is used to calculate a steady state target $x_s$, $u_s$ by solving an optimization based on the current estimate $\hat{d}_{ijk}$. If $x_{sp}$ and $u_{sp}$ are the set points for the output and the input and $R_s$ is a positive definite quadratic penalty, the modified steady state targets $x_s$, $u_s$, calculated at each sampling time as a solution to the following convex quadratic programming problem as shown in Ref. 4:

$$\min_{x_s,u_s} \{u_s - u_{sp}\}^T R_s \{u_s - u_{sp}\}$$  (7a)

subject to,

$$\begin{bmatrix} I - A & -B \\ C & 0 \end{bmatrix} \begin{bmatrix} x_s \\ u_i \end{bmatrix} = \begin{bmatrix} B_a\hat{d}_{ijk} \\ -C_d\hat{\xi}_{ijk} + y_{sp} \end{bmatrix}$$  (7b)

and inequality constraints on $x_s$, $u_s$.

The above optimization has a feasible solution whenever the inequality constraints in Eq. 7b are not active. Both the objective function Eq. 7a and the inequality constraints (to include slack variables) are modified when there is no feasible solution.

**Regulation**

The receding horizon formulation for the regulator in MPC is given as the quadratic objective function with the model constraints and penalties $Q \geq 0, R \geq 0$. The state and the input are penalized for deviations from the steady state targets calculated from Eq. 7a, 4

$$\min_{u_t, u_{t-1}, X_{ik}} \sum_{j=k}^{k+N_c} (\hat{x}_{ijk} - x_s)^T Q (\hat{x}_{ijk} - x_s) + (u_j - u_{sp})^T R (u_j - u_{sp})$$

subject to, $\hat{x}_{ijk} = A\hat{x}_{ijk} + Bu_i + B_d\hat{\xi}_{ijk}$, $j \geq k$.

The state and input constraints are not shown in the above formulation for ease of notation. Inclusion of the constraints does not change any of the analysis that follows. Here, the control horizon and the prediction horizon are taken to be the same: $N_c$.

**Remark 1.** Note that the state estimation in the subsection “State Estimation” requires knowledge of the covariances.
GQₜₜ, Qₜ, Rₜ, and the target calculation in the section “Target Calculation” requires the structure of the unmodeled disturbances given by the matrices Bₜₜ, Cₜₜ. In general, none of these are known from first principles modeling. The ideal choice of Bₜₜ, Cₜₜ is one that accurately represents the unmodeled disturbance in the plant. The choice of the disturbance model has been the subject of research22–24 (also see Ref. 25, chapter 5). These studies assume that the covariances GQₜₜGᵀₜₜ, Qₜ, Rₜ are known and the models Bₜₜ, Cₜₜ are unknown. However, in the sections “Model Equivalence” and “Using Correlations to Estimate Filter Gain,” we show that if an incorrect choice for Bₜₜ, Cₜₜ is made, then the covariances GQₜₜGᵀₜₜ, Qₜ, Rₜ estimated from data are within an equivalent transformation such that the particular choice of Bₜₜ, Cₜₜ has no effect on the closed-loop performance.

Motivation

The rules for choosing the penalties Q, R for the regulator are well researched and based on intuitive reasoning (see, for example, Ref. 26, chapter 6). The estimator on the other hand requires knowledge of the disturbance covariances GQₜₜGᵀₜₜ, Qₜ, Rₜ and the disturbance models Bₜₜ, Cₜₜ. Consider an unconstrained single input single output plant with the following values for the parameters:

A = 0.5, B = 0.5, C = 1, yₜₜ = 0.8

Let a deterministic unmodeled disturbance of magnitude 1 enter the input at sampling time 3 and is measured at sampling time 4. Let there be no other noises affecting the plant: GQₜₜGᵀₜₜ = 0, Rₜ = 0. An incorrect choice of the MPC disturbance model would be Bₜₜ = 0, Cₜₜ = 1 modeling the disturbance as entering at the output. The actual plant model on the other hand is Bₜₜ = B = 0.5, Cₜₜ = 0 (the input disturbance model).

We now compare the closed-loop performance for the incorrect disturbance model against the plant disturbance model. The regulator is chosen to be the LQR feedback: uk = (xₜₜ − xₜₜ). The rejection of the deterministic disturbance in the two cases is shown in Figure 2. With the input disturbance model, the disturbance in the input is estimated and rejected much more quickly than when using the incorrect model although the same covariances GQₜₜGᵀₜₜ, Rₜ, Qₜ are used in the calculation by the Kalman filter Eq. 6. The incorrect Kalman gains for the input disturbance model is

L₁ = \[
\begin{bmatrix}
1 & 2 \\
0 & 1
\end{bmatrix}
\]

and for the incorrect output disturbance model is

L₂ = \[
\begin{bmatrix}
1 & 0 \\
0 & 1
\end{bmatrix}
\].

The control law is the same both cases and given by the LQR feedback: uk = (xₜₜ − xₜₜ). The rejection of the deterministic disturbance in the two cases is shown in Figure 2. With the input disturbance model, the disturbance in the input is estimated and rejected much more quickly than when using the incorrect model although the same covariances GQₜₜGᵀₜₜ, Rₜ, Qₜ are used in the calculation by the Kalman filter Eq. 6. The input disturbance model is estimated and rejected much more quickly than when using the incorrect model although the same covariances GQₜₜGᵀₜₜ, Rₜ, Qₜ are used in the calculation by the Kalman filter Eq. 6. The sluggish behavior of the output disturbance model and the decrease in closed-loop performance is due to the choice of the disturbance model and the chosen estimator gain. A similar difference in the performance with the incorrect choice of the disturbance model was noted by many others.11,22,25,27

Figure 2. Disturbance rejection using two disturbance models.

Since we usually lack knowledge about the structure of the unmodeled disturbances (matrices Bₜₜ, Cₜₜ), the only other handle to change the input-output performance is the estimator gain L. Once the choice of the disturbance model is fixed, we can use a data based technique that estimates L such that the closed-loop performance is not affected by the choice of the disturbance model (Section “Using Correlations to Estimate Filter Gain”). For example, if in the above problem, the choice of the disturbance model is already fixed incorrectly as the output disturbance model, then a transformation of the covariances GQₜₜGᵀₜₜ, Qₜ, Rₜ can be estimated from data to give the same closed-loop performance as achieved using the input disturbance model and the original covariances.

Model Equivalence

The choice of the disturbance model for offset-free control in MPC is often ambiguous with no real guidelines available. In Ref. 28, a min-max problem is solved to find the best performing disturbance model. Most industrial MPC implementations such as IDCOM,7 DMC,8 and QDMC29 use an output disturbance model to remove offset. The choice of the output disturbance model has been criticized in the literature as being sluggish in rejecting disturbances [Refs. 22 and 25 (chapter 5)]. In this section we show that for LTI systems, disturbance models are equivalent as long as the estimator is determined with the appropriate covariances.

Remark 2. It is interesting to examine some of the early history of disturbance modeling. Box and Jenkins clearly pointed out that one can identify equivalent disturbance models that represent different disturbance sources by what they call “transfer of disturbance origins” (see Ref. 30, p. 469). Note that interested readers must obtain the first edition, because in the third edition11 this terminology and entire discussion had been deleted. They also pointed out that identifying a disturbance model with data collected in a fashion that does not include all relevant disturbances that are encountered in closed-loop operation does not produce an optimal control because the disturbance model consequently has significant error. They recommend using closed-loop identification
(Ref. 30, p. 471) to avoid the accidental neglect of disturbance sources.

In this work, we are emphasizing again Box and Jenkins’s conclusion that one is free to choose the source of the disturbance in the disturbance model, and one naturally requires data that include all relevant disturbances. It is irrelevant, however, if these data are collected under closed-loop or open-loop control. Also, we are not interested in identifying the disturbance model, which was the primary focus of Box and Jenkins’s earlier work. We fully expect the practitioner to employ an integrated white noise disturbance model to represent essentially all of the plant disturbances. The motivation is to remove steady offset in the closed-loop system and show that such rough disturbance models are often completely adequate to handle disturbances entering from other sources with different dynamics, if one estimates the statistics of the driving noise from data.

Consider two choices of the disturbance models \((B_{d1}, C_{d1})\) and \((B_{d2}, C_{d2})\) satisfying detectability conditions. If \(n_d\) is the dimension of the disturbance vector, the assumption of detectability implies:

\[
\text{rank} \begin{bmatrix} A - I & B_{d1} \\ C & C_{d1} \end{bmatrix} = n + n_d, \quad \text{rank} \begin{bmatrix} A - I & B_{d2} \\ C & C_{d2} \end{bmatrix} = n + n_d
\]

(9)

Note that if the disturbance in the actual plant is not detectable, the state estimator is defined for a transformed, lower dimensional model that is detectable.

The systems augmented with the disturbance model are simple non-minimal realizations of the same unaugmented model where \(d_k\) is an uncontrollable state. The augmented systems with the two general disturbance models can be written as follows:

**Model 1:**

\[
\begin{aligned}
\begin{bmatrix} x^1 \\ d^1 \end{bmatrix}_{t+1} &= \begin{bmatrix} A & B_{d1} \\ 0 & I \end{bmatrix} \begin{bmatrix} x^1 \\ d_1 \end{bmatrix}_t + \begin{bmatrix} 0 \\ B \end{bmatrix} u_k + G_1 \begin{bmatrix} w \\ \xi \end{bmatrix}_k \\
y_k &= \begin{bmatrix} C & C_{d1} \end{bmatrix} \begin{bmatrix} x^1 \\ d_1 \end{bmatrix}_t + v_k
\end{aligned}
\]

(10)

**Model 2:**

\[
\begin{aligned}
\begin{bmatrix} x^2 \\ d^2 \end{bmatrix}_{t+1} &= \begin{bmatrix} A & B_{d2} \\ 0 & I \end{bmatrix} \begin{bmatrix} x^2 \\ d_2 \end{bmatrix}_t + \begin{bmatrix} 0 \\ B \end{bmatrix} u_k + G_2 \begin{bmatrix} w \end{bmatrix}_k \\
y_k &= \begin{bmatrix} C & C_{d2} \end{bmatrix} \begin{bmatrix} x^2 \\ d_2 \end{bmatrix}_t + v_k
\end{aligned}
\]

(11)

If the two models have the same input to output relationship, there exists a similarity transformation matrix \(T\), such that

\[
TA_1T^{-1} = A_2, \quad C_1T^{-1} = C_2, \quad TB_d = B_d
\]

(12)

Also for Models 1 and 2 the unknown stochastic parameters \(G_1\) and \(G_2\) must satisfy:

\[
TG_1 = G_2
\]

(13)

**Lemma 1.** If the augmented models in Eqs. 10 and 11 are assumed to be detectable, the transformation relations in Eq. 12 can be expressed equivalently as

\[
T_{11} = I, \quad \begin{bmatrix} A - I & B_{d2} \\ C & C_{d2} \end{bmatrix} \begin{bmatrix} T_{13} \\ T_{22} \end{bmatrix} = \begin{bmatrix} B_{d1} \\ C_{d1} \end{bmatrix}, \quad T_{21} = 0
\]

where, \(T\) is partitioned as (with appropriate dimensions):

\[
T = \begin{bmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{bmatrix}
\]

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

**Theorem 1.** For the two augmented models in Eqs. 10 and 11 and satisfying the detectability condition in Eq. 9, a transformation matrix \(T\) satisfying the relations in Eq. 12 exists and is unique if and only if the following condition holds:

\[
\text{rank} \begin{bmatrix} A - I & B_{d1} & B_{d2} \\ C & C_{d1} & C_{d2} \end{bmatrix} = n + n_d
\]

(14)

**Proof.** The proof follows from the relations in Eq. 12 expressed equivalently from Lemma 1 and the detectability conditions in Eq. 9. □

Note that for zero offset in all the outputs \(y_k\), we need \(n_d = p\). The condition in Theorem 1 is then automatically satisfied for all detectable disturbance models.

**Remark 3.** Notice that we are considering a known and fixed minimal input–output realization of the triple \((A, B, C)\). The LQR cost in Eq. 8 then remains unchanged for the two non-minimal augmented models in Eqs. 10 and 11 related by the conditions given in Lemma 1.

**Special case of equivalence**

For an input disturbance model we have \(B_d = B\) and \(C_d = 0\) and for an output disturbance model we have \(B_d = 0\) and \(C_d = I\). For these particular choices of the disturbance models with \(n_d = m = p\), we get the following transformation matrix:

**Model 1:** \(B_{d1} = B, C_{d1} = 0\); **Model 2:** \(B_{d2} = 0, C_{d2} = I\)

\[
T = \begin{bmatrix} I & (A - I)^{-1}B \\ 0 & -C(A - I)^{-1}B \end{bmatrix}
\]

(15)

**Remark 4.** Note that when \(A\) has an integrator, the inverses above are not defined. In this case the output disturbance model is not detectable and fails Theorem 1. This has been long known in the literature as a potential problem with industrial MPC implementations such as IDCOM, DMC, QDRC, which use a constant output step disturbance model. Various other fixes have also been suggested in the literature."
Equivalent closed-loop performance

Given Models 1 and 2 related by the equivalent transformation matrix $T$ satisfying the relations in Eq. 12, from realization theory, equivalent closed-loop performance for the two models implies that given a sequence of control actions $\{u_1, \ldots, u_{k-1}\}$ and noise sequences $\{w_1, \ldots, w_{k-1}\}$, $\{\xi_1, \ldots, \xi_{k-1}\}$ and $\{v_1, \ldots, v_k\}$, then the output sequence from the two models: $\{y_1, \ldots, y_k\}$ are the same. When these models are used in MPC, the noises $w_k, \xi_k, v_k$ are unknown.

If $L_1$ and $L_2$ are the state estimator gains (not necessarily the Kalman filter gains) for Models 1 and 2, then equivalent closed-loop performance implies that given the inputs $\{u_1, \ldots, u_{k-1}\}$ and outputs $\{y_1, \ldots, y_k\}$, the innovations sequence $\{y_1, \ldots, y_k\}$ are the same, where $y_k = (y_k - \hat{y}_{i|k-1})$. The innovations for the augmented system in Eq. 3 are given by Eqs. 4 and 5:

$$\begin{align*}
\dot{x}_{j|k} &= A_x \dot{x}_{j|k-1} + B_x u_{k} + A_w L_w (y_k - \hat{y}_{i|k-1}) \\
\gamma_{j|k} &= y_k - C_x \dot{x}_{j|k-1}
\end{align*}$$  \hspace{1cm} (16)

**Lemma 2.** Given a set of measurements $\{y_1, \ldots, y_k\}$ and inputs $\{u_1, \ldots, u_{k-1}\}$ and the two disturbance models in Eqs. 10 and 11 related to each other through the unique transformation matrix $T$ satisfying Eq. 16, the innovations sequence defined as $\gamma_k = (y_k - \hat{y}_{i|k-1})$ are the same for the two models if and only if the state estimator gains (can be suboptimal gains) satisfy the relation:

$$L_2 = TL_1$$  \hspace{1cm} (17)

**Proof.** The proof follows using simple algebraic manipulations using Eqs. 10 and 11 in Eq. 16 and the transformations in Eq. 12. \qed

**Remark 5.** Since Lemma 2 makes no assumptions on the optimality of the estimator gains $L_1$ and $L_2$, the result also holds when the optimal Kalman filter gains are used instead.

For the Kalman filter, the calculation of the estimator gains follow from the Riccati Eq. 6.

For Model 1:

$$\begin{align*}
P_1 &= A_1 P_1 A_1^T - A_1 P_1 C_{1T} (C_1 P_1 C_{1T} + R_1) \quad (C_1 P_1 C_{1T} + R_1)^{-1} C_1 P_1 A_1^T + G_1 Q_1 G_1^T \\
L_1 &= P_1 C_{1T} (C_1 P_1 C_{1T} + R_1)^{-1} + G_1 Q_1 G_1^T \\
L_2 &= P_2 C_{2T} (C_2 P_2 C_{2T} + R_2)^{-1} (C_2 P_2 C_{2T} + R_2)^{-1}
\end{align*}$$

and for Model 2:

$$\begin{align*}
P_2 &= A_2 P_2 A_2^T - A_2 P_2 C_{2T} (C_2 P_2 C_{2T} + R_2) \quad (C_2 P_2 C_{2T} + R_2)^{-1} C_2 P_2 A_2^T + G_2 Q_2 G_2^T \\
L_2 &= P_2 C_{2T} (C_2 P_2 C_{2T} + R_2)^{-1} (C_2 P_2 C_{2T} + R_2)^{-1}
\end{align*}$$

in which,

$$Q_* = \begin{bmatrix} Q_w & 0 \\ 0 & Q_\xi \end{bmatrix}$$

To obtain optimal closed-loop performance, we need to know the transformation matrix $T$ relating the plant disturbance model to the controller disturbance model in addition to the original covariances $Q_{an}, R_v$. Since the disturbance in the plant is unmodeled and the source unknown, the transformation matrix $T$ is also unknown. Hence, the transformation in the state estimator gains as shown in Eq. 12 cannot be used. However, using Eqs. 18 and 19 we see that if covariances $G_1 Q_1 G_1^T, R_v$ for Model 1 and $G_2 Q_2 G_2^T, R_v$ for Model 2 are used to calculate the Kalman filter, where $T G = G$, then Eq. 12 is satisfied.

Using Correlations to Estimate Filter Gain

Again consider the plant given by a LTI state-space model as shown in Eq. 1:

$$\begin{align*}
x_{k+1} &= A x_k + B u_k + G w_k \\
y_k &= C x_k + v_k
\end{align*}$$  \hspace{1cm} (20)

Let $d_k$ be an unmodeled integrating white noise disturbance entering the plant through the disturbance model structures $B_{dm}, C_{dm}$. The augmented plant model is then given by:

$$\begin{align*}
\begin{bmatrix} \dot{x}^p \\ \dot{d}^p \end{bmatrix}_{k+1} &= \begin{bmatrix} A & B_{dm} \\ 0 & I \end{bmatrix} \begin{bmatrix} \dot{x}^p \\ \dot{d}^p \end{bmatrix}_k + \begin{bmatrix} B \\ 0 \end{bmatrix} u_k + \begin{bmatrix} G & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} w \\ \zeta_k \end{bmatrix}_k \\
y_k &= \begin{bmatrix} C & C_{dm} \end{bmatrix} \begin{bmatrix} \dot{x}^p \\ \dot{d}^p \end{bmatrix}_k + v_k
\end{align*}$$

Since knowledge of the disturbance structure is unknown, assume the disturbance model in MPC to remove offset is incorrectly chosen to have the structure $B_{dm}, C_{dm}$:

$$\begin{align*}
\begin{bmatrix} \dot{x}^m \\ \dot{d}^m \end{bmatrix}_{k+1} &= \begin{bmatrix} A & B_{dm} \\ 0 & I \end{bmatrix} \begin{bmatrix} \dot{x}^m \\ \dot{d}^m \end{bmatrix}_k + \begin{bmatrix} B \\ 0 \end{bmatrix} u_k + \begin{bmatrix} G_m & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} w \\ \zeta_k \end{bmatrix}_k \\
y_k &= \begin{bmatrix} C & C_{dm} \end{bmatrix} \begin{bmatrix} \dot{x}^m \\ \dot{d}^m \end{bmatrix}_k + v_k
\end{align*}$$

The superscripts $p$ and $m$ are used to denote states for the actual plant and the model respectively. Assume that the conditions in Theorem 1 are satisfied and there exists an unknown transformation matrix $T$ that connects models in Eqs. 21 and 22 through the relations in Eq. 12. Then the following holds from realization theory:

$$T \begin{bmatrix} \dot{x}^p \\ \dot{d}^p \end{bmatrix}_k = \begin{bmatrix} \dot{x}^m \\ \dot{d}^m \end{bmatrix}_k, \quad TA_p T^{-1} = A_m, \quad C_p T^{-1} = C_m$$

We assume $(A, B, C)$ to be known from some system identification scheme, the stochastic parameters $G, Q_{an}, Q_\xi, R_v$ are unknowns and $B_{dm}, C_{dm}$ are incorrectly chosen.

Given some arbitrary (stable, perhaps suboptimal) initial estimator $L$, the state estimates given the incorrect model are given by:
By substituting from Eq. 23, we get:

\[
\begin{bmatrix}
y_{k} \\
\hat{y}_{k|k-1}
\end{bmatrix} = C_p \begin{bmatrix} x_p^m \\ d_p^m \end{bmatrix}_{k|k-1} + v_k
\]

Let the innovation sequence be given by \( \gamma_k = y_k - \hat{y}_{k|k-1} \), using Eqs. 21 and 24, we get:

\[
\gamma_k = C_m \begin{bmatrix} x_m^m \\ d_m^m \end{bmatrix}_{k|k-1} + v_k
\]

We can then write the evolution of the state estimate error as:

\[
\begin{bmatrix}
\hat{e}_{k+1} \\
\hat{y}_{k}
\end{bmatrix} = \begin{bmatrix} A_m - A_mL \end{bmatrix} \begin{bmatrix} \hat{e}_k \\ \hat{y}_k \end{bmatrix} + \begin{bmatrix} \hat{G} \\ 0 \end{bmatrix}
\]

We define the autocovariance as

\[
E[\gamma_k \gamma_j^T] = C_mPC_m^T + R_v
\]

\[
E[\gamma_k \gamma_{k+j}^T] = C_mA^jPC_m^T - C_mA^jA_mL R_v, \quad j \geq 1
\]

which are independent of \( k \) when we assume that data are collected when operating at steady state. Again using Eq. 25 we note that \( \hat{P} = E[\hat{e}_k \hat{e}_k^T] \) the steady-state state estimate error covariance satisfies the Lyapunov equation:

\[
P = \hat{A}P\hat{A}^T + [TG_p - A_mL] \begin{bmatrix} Q_0 & 0 \\ 0 & R_v \end{bmatrix} G^T,
\]

where \( Q_0 = \begin{bmatrix} Q_w & 0 \\ 0 & Q_v \end{bmatrix} \).

A computation of the first column block of the Autocovariance Matrix of innovations process over some user defined window \( N \) then gives:

\[
R_v(N) = \begin{bmatrix} C_0 \\ C_1 \\ \vdots \\ C_{N-1}
\end{bmatrix}
\]

\[
= \begin{bmatrix} C_m \\ C_mA^1 \\ \vdots \\ C_mA^{N-1} \end{bmatrix} P C_m^T R_v
\]

Define the following covariance product:

\[
\Psi \triangleq TG_pQ_aG_p^T T^T
\]

Notice that in the above Eq. 29 with \( P \) given by Eq. 28, the only unknowns are the transformed covariance \( Q \) and \( R_v \). If \( R_v(N) \) is an estimate of \( R_v \), from data, then using properties of Kronecker products, 39,40 we can express Eq. 29 as a weighted least-squares problem in a vector of unknowns \( \Psi_s \). Here, the subscript “s” denotes the column-wise stacking of the elements of the matrix in a vector. We can further add semidefinite constraints to ensure the covariances to be positive semidefinite. We then get the following convex optimization problem:

\[
\Phi = \min_{\Psi, R_v} \left\| A \begin{bmatrix} (Q)_s \\ (R_v)_s \end{bmatrix} - b \right\|_W^2
\]

subject to, \( \Psi \geq 0, \ R_v \geq 0, \ Q = Q^T, \ R_v = R_v^T \)

where \( A = \begin{bmatrix} A_1 & A_2 \end{bmatrix} \)

\[
A_1 = \left[(C_m \otimes \hat{O})(I_{(n+\nu)_s} - \hat{A} \otimes \hat{A})^{-1}\right]
\]

\[
A_2 = \left[(C_m \otimes \hat{O})(I_{(n+\nu)_s} - \hat{A} \otimes \hat{A})^{-1}(A_mL \otimes A_mL) + (I_p \otimes \hat{T})\right]
\]

The \( A \) matrix is found using the system matrices and the initial estimator gain \( L \). The matrix \( R_v(N) \) contains the estimated autocovariances from data and the vector \( b \) is defined as \( b = (R_v)_s \). The estimation method in Eq. 30 will be referred to as the autocovariance least-squares (ALS) technique in the rest of this article. For more details about the ALS technique, interested readers are referred to Ref. 38.

**Theorem 2.** If \( A \) has full column rank then in the limit of large data: \( N_d \to \infty \), where \( N_d \) is the length of the innovations sequence used to estimate \( R_v(N) \), the ALS technique gives unbiased estimates for the covariances: \( \lim_{N_d \to \infty} Q \to TG_pQ_aG_p^T T^T \) and \( \lim_{N_d \to \infty} R_v \to R_v \).

**Proof.** See Ref. 15 for proof, showing that the estimates are unbiased.

**Lemma 3.** If covariances \( TG_pQ_aG_p^T T^T \) and \( R_v \) are used to calculate the Kalman filter for the incorrect disturbance
model in Eq. 22, then the resulting innovations sequence is the same as when using the optimal Kalman filter with the actual plant disturbance model in Eq. 21.

Proof. Substituting $TG_pQ_cG_p^T$ and $R_c$ in the Kalman filter Eq. 6 for the incorrect model in Eq. 22 and using the transformations in Eq. 23, we get:

$$L_m = TL_p$$

Equivalence and optimality of the innovations sequence then follows from Lemma 2 and noting that $L_m$ is the optimal Kalman gain for the disturbance model in Eq. 21.

Thus, even when the disturbance model is chosen different from the actual plant disturbance location, the ALS technique can be used to estimate transformed covariances such that the closed-loop performance is equivalent to that achieved using the optimal filter and the actual plant disturbance model.

Necessary and sufficient conditions for uniqueness of the ALS problem in Eq. 30 are proved in Ref. 38. The advantages of using the ALS technique to estimate the covariances as opposed to the maximum likelihood technique are detailed in Ref. 41 (chapter 6). The modification of the technique when the cross-covariances is nonzero is addressed in Refs. 41 (chapter 4) and 42.

Table 1. Equivalence in the Estimator Gains Identified from Data

<table>
<thead>
<tr>
<th>Deadbeat Estimator</th>
<th>ALS Estimator Gain</th>
<th>Transformation of the Estimator Gain</th>
</tr>
</thead>
<tbody>
<tr>
<td>Input DM</td>
<td>$L_{opt}$</td>
<td>$L_{inp}$</td>
</tr>
<tr>
<td></td>
<td>$[-1 1.24]$</td>
<td>$[-0.891 1.11]$</td>
</tr>
<tr>
<td></td>
<td>$[0 -2.12]$</td>
<td>$[-0.08 -1.95]$</td>
</tr>
<tr>
<td></td>
<td>$[0 -0.88]$</td>
<td>$[-0.05 -0.93]$</td>
</tr>
<tr>
<td></td>
<td>$[1.58 -1.96]$</td>
<td>$[1.44 -1.73]$</td>
</tr>
<tr>
<td></td>
<td>$[0 2.18]$</td>
<td>$[0.1 2.00]$</td>
</tr>
<tr>
<td>Output DM</td>
<td>$T_{opt}$</td>
<td>$T_{out}$</td>
</tr>
<tr>
<td></td>
<td>$[0.58 -0.72]$</td>
<td>$[0.59 -0.75]$</td>
</tr>
<tr>
<td></td>
<td>$[0 2.51]$</td>
<td>$[0.16 2.30]$</td>
</tr>
<tr>
<td></td>
<td>$[0 7.15]$</td>
<td>$[0.34 6.88]$</td>
</tr>
<tr>
<td></td>
<td>$[1.58 2.41]$</td>
<td>$[1.68 2.31]$</td>
</tr>
<tr>
<td></td>
<td>$[0 2.18]$</td>
<td>$[0.1 2.04]$</td>
</tr>
</tbody>
</table>

Figure 4. Closed-loop performance using estimator gains from the ALS technique (D.M. is disturbance model).
The input disturbance model is given as follows:

\[ B_d \text{ Riccati Eq. 6. The deadbeat estimator for the} \]

The pure output disturbance model (model)

The controller penalties used in Eq. 8 were \( Q = I, R = I \).

The initial data were simulated with the output disturbance model (incorrect) with the above initial suboptimal estimator gain. The data were then processed using the ALS technique from the previous section to estimate (transformed) covariances \( G Q G^T, R_c \) and then used to calculate the estimator gain. The computation times were less than 5 s in both cases running Octave on a standard AMD-64 desktop. The estimator gains from the different cases are presented in Table 1. The optimum gain is the deadbeat estimator with the input disturbance model. As seen in Table 1, using the ALS technique, the estimator gains are close to the optimum values, even when the incorrect disturbance model is used. A snapshot of the performance of the closed-loop with the different models and estimator gains are shown in Figure 4.

As seen in the Figure 4, once the estimator gain is calculated from the covariances estimated using the ALS technique, the inputs and the outputs become equivalent and thus indistinguishable using either of the disturbance models.

The closed-loop cost benefits for the deterministic disturbances shown in Figure 3 can be easily calculated and is presented in Table 2. (The closed-loop performance values presented in Table 2 are not identical for the two models because the ALS technique is data based and depends on the length of the data set and the tolerances set in the semidefinite programming algorithm. As shown in Theorem 2, the estimates are unbiased in the limit \( N_d \to \infty \).)

### Stochastic disturbances

Instead of the deterministic disturbance, in this section we consider slow-moving integrated white noise disturbances entering each input as shown in Figure 5. The actual plant

\[ L_1 = \begin{bmatrix} -1 & 1.24 \\ 0 & -2.12 \\ 0 & -0.88 \\ 1.58 & -1.96 \\ 0 & 2.18 \end{bmatrix} \]

On the other hand, if the disturbance is incorrectly modeled as entering entirely in the output (the output disturbance model), the same covariances \( G = I, Q_w = 0, R_{\varepsilon} \approx 0, Q_{\zeta} \neq 0 \) in the estimator Riccati Eq. 6. The deadbeat estimator for the input disturbance model is given as follows:

\[ L_2 = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \]

![Figure 5. Disturbances entering the inputs.](image)

### Simulation Examples

The following examples illustrate the ideas developed in the previous sections. Closed-loop performance of a pure input disturbance model (\( B_d = B, C_d = 0 \)) is compared with the pure output disturbance model (\( B_d = 0, C_d = I_p \) for the following MIMO system (\( n_d = p \) in both models):

\[
\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} \frac{1}{\tau_T} & \frac{2}{\tau_T^2} \\ 0 & \frac{1}{(\lambda-1)(\lambda+2)} \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}
\]

A minimal state-space realization (discrete time with a sampling time of 1 sec) gives:

\[
\begin{bmatrix} A & B \\ C & D \end{bmatrix} = \begin{bmatrix} 0.368 & 0 & 0 \\ 0 & 0.455 & 0.050 \\ 0 & 0.453 & 0.629 \end{bmatrix}, \quad \begin{bmatrix} -0.632 & 0 \\ 0 & -0.971 \\ 0 & -0.401 \end{bmatrix},
\]

\[
\begin{bmatrix} -1 & -0.471 & -0.272 \\ 0 & -0.471 & 0 \end{bmatrix}
\]

### Deterministic disturbances

In the plant, deterministic unmodeled disturbances are added to both the inputs as shown in Figure 3. The disturbances entering are of different magnitude and are uncorrelated. Other stochastic disturbances are absent in the simulation: \( Q_w, R_{\varepsilon} = 0 \). An input disturbance model is an accurate description of the unmodeled disturbance in the plant. The optimal state estimator is the deadbeat estimator with the gain calculated by using \( G = I, Q_w = 0, R_{\varepsilon} \approx 0, Q_{\zeta} \neq 0 \) in the estimator Riccati Eq. 6. The deadbeat estimator for the input disturbance model is given as follows:

\[
L_1 = \begin{bmatrix} -1 & 1.24 \\ 0 & -2.12 \\ 0 & -0.88 \\ 1.58 & -1.96 \\ 0 & 2.18 \end{bmatrix}
\]

\[
L_2 = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}
\]

<table>
<thead>
<tr>
<th>Input Disturbance Model</th>
<th>( \sum \sigma_v^2 Q_x )</th>
<th>( \sum (u - u_i)^2 R(u - u_i) )</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial</td>
<td>N/A</td>
<td>1.82</td>
<td>17.58</td>
</tr>
<tr>
<td>ALS</td>
<td>N/A</td>
<td>1.69</td>
<td>17.30</td>
</tr>
<tr>
<td>Benefits</td>
<td>N/A</td>
<td>15.76</td>
<td>N/A</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Output Disturbance Model</th>
<th>( \sum \sigma_v^2 Q_x )</th>
<th>( \sum (u - u_i)^2 R(u - u_i) )</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial</td>
<td>34.48</td>
<td>15.60</td>
<td>-49%</td>
</tr>
<tr>
<td>ALS</td>
<td>1.69</td>
<td>17.30</td>
<td></td>
</tr>
</tbody>
</table>

The closed-loop cost benefits for the deterministic disturbances shown in Figure 3 can be easily calculated and is presented in Table 2. (The closed-loop performance values presented in Table 2 are not identical for the two models because the ALS technique is data based and depends on the length of the data set and the tolerances set in the semidefinite programming algorithm. As shown in Theorem 2, the estimates are unbiased in the limit \( N_d \to \infty \).)
The disturbance model in this case is again the input disturbance model and the output disturbance model performs poorly without the covariance estimates using the ALS technique. The covariance of the integrating disturbances is $Q = 10^{-3}I$. In addition, the inputs and the outputs are corrupted by zero-mean Gaussian white noise having covariances $Q_w = 10^{-2}I, R_v = 10^{-3}I$.

The estimator gains for the input and the output disturbance models were calculated separately using the ALS technique (Eq. 30) from the same data set and a snapshot of the performance of the two disturbance models is shown in Figure 6. Again the inputs and outputs perform in a similar manner irrespective of the disturbance model chosen once the ALS technique is used to calculate the estimator gains.

The comparison between the expected closed-loop costs (see Appendix B for calculation of closed-loop cost) for the three cases is presented in Table 3. Note that the optimal closed-loop cost is $1.788 \times 10^{-4}$.

### Conclusions

Although the choice of the disturbance model for offset-free MPC has been the subject of active research, the choice of an appropriate estimator gain for the disturbance model has not been studied until recently. The focus of most of the research has been to design appropriate disturbance models to accurately represent the actual plant disturbance source while assuming that the estimator gain is known. In the section “Model Equivalence” it was shown that since different augmented disturbance models are nonminimal realizations of the same input–output process, there exists estimator gain matrices for each disturbance model that achieve the same closed-loop input–output performance.

In the section “Using Correlations to Estimate Filter Gain,” an ALS technique was presented to estimate the covariances of the disturbance model from steady state data (closed or open loop) irrespective of the true disturbance source. The estimation of these covariances and hence the estimator gain was shown to compensate automatically for the incorrect choice of the model’s disturbance location. Examples were presented to show the benefits of using the ALS technique to estimate the covariances and to illustrate the

<table>
<thead>
<tr>
<th>Table 3. Expectation of Closed-Loop Cost for Different Estimator Gains</th>
</tr>
</thead>
<tbody>
<tr>
<td>Expectation of the</td>
</tr>
<tr>
<td>Closed-Loop Cost</td>
</tr>
<tr>
<td>Input DM $E[x^T_k Q_{D1} + u^T_k R_{u0}]$</td>
</tr>
<tr>
<td>Output DM $E[x^T_k Q_{D1} + u^T_k R_{u0}]$</td>
</tr>
</tbody>
</table>

Optimal cost is $1.788 \times 10^{-4}$. See Appendix B.
equivalent closed-loop performance using different disturbance models.

A direct implication of the results presented in this article is that an industrial practitioner can model the disturbance as an integrator at any location in the plant but then use the autocovariance of the data to estimate the disturbance covariances to calculate the estimator gain. The control performance using this rough disturbance model and data-based estimator gain is often comparable to that obtained using a disturbance model that accurately represents the true plant disturbance. Modeling effort can then be directed to obtaining the estimator gain from data rather than trying to model accurately the dynamics of the unknown disturbance, which is a much more complex modeling task.

Acknowledgments

The authors thank Dr. T. A. Badgwell for helpful discussions. The authors acknowledge the financial support from NSF through Grant CNS-0540147 and from PFR through Grant 43321-AC9.

Literature Cited


Appendix A: Proof of Lemma 1

Since the pair (A, C) is assumed observable, we have rank \[
\begin{bmatrix}
A - I & C
\end{bmatrix}
\] = n. From the detectability conditions in Eq. 9 we then get:

\[
\begin{align*}
\text{rank} \begin{bmatrix}
B_{d1} \\
C_{d1}
\end{bmatrix} &= n_1, \\
\text{rank} \begin{bmatrix}
B_{d2} \\
C_{d2}
\end{bmatrix} &= n_d
\end{align*}
\] (A1)
Substituting $T$ partitioned as $T = \begin{bmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{bmatrix}$ into the equivalence relations in Eq. 12 we get:

1. Equivalence between the augmented input matrices:
   \[ TB_a = B_a \]
   \[ \Rightarrow T_{11}B = B, \quad T_{21}B = 0 \]  
   
   Since $T$ is an invertible transformation matrix, the diagonal block $T_{11}$ is also invertible. This implies:
   \[ T_{11} = I \]  

2. Equivalence between the augmented state matrices:
   \[ TA_1 = A_2T \]
   \[ \Rightarrow [A - I \quad B_{d2}] \begin{bmatrix} T_{12} \\ T_{22} \end{bmatrix} = B_{d1}, \quad B_{d2}T_{21} = 0 \]  

3. Equivalence between the augmented output matrices:
   \[ C_1 = C_2T \]
   \[ \Rightarrow [C C_{d2}] \begin{bmatrix} T_{12} \\ T_{22} \end{bmatrix} = C_{d1}, \quad C_{d2}T_{21} = 0 \]

Using Eqs. A3 and A4 we get:
\[ \begin{bmatrix} A - I \quad B_{d2} \\ C \quad C_{d2} \end{bmatrix} \begin{bmatrix} T_{12} \\ T_{22} \end{bmatrix} = \begin{bmatrix} B_{d1} \\ C_{d1} \end{bmatrix}, \quad B_{d2}T_{21} = 0 \]  

Using the rank condition in Eq. A1 in the above equation implies:
\[ T_{21} = 0 \]  

Putting Eqs. A2, A5, and A6 together, we get:
\[ T_{11} = I, \quad [A - I \quad B_{d2}] \begin{bmatrix} T_{12} \\ T_{22} \end{bmatrix} = \begin{bmatrix} B_{d1} \\ C_{d1} \end{bmatrix}, \quad T_{21} = 0 \]

\[ \check{x}_{k|k} = \hat{x}_{k|k-1} + L(y_k - C\hat{x}_{k|k-1}) \quad \text{(B4)} \]

where $\hat{x}_{k|k}$ is the filtered estimate given the measurements and inputs up to time $t_k$ and $\check{x}_{k|k-1}$ is the predicted estimate given measurements and inputs up to time $t_{k-1}$.

The deterministic LQR objective is given as follows:
\[ \Phi = \sum_{k=0}^{\infty} x_k^T Q x_k + u_k^T R u_k \]

giving the feedback law $u_k = K\hat{x}_{k|k}$. For the stochastic case, the objective can be defined as
\[ \Phi = \lim_{N \to \infty} \frac{1}{N} E \left[ \sum_{k=0}^{N} x_k^T Q x_k + u_k^T R u_k \right] \]

The closed-loop cost for the continuous time state-space model is derived in Ref. 26 (p. 220). Simple algebraic manipulations on Eqs. B1 and B3 give the following augmented system:
\[ \begin{bmatrix} x_{k+1} \\ \check{x}_{k+1|k+1} - x_{k+1} \end{bmatrix} = \begin{bmatrix} A + BK & BK \\ 0 & A - LCA \end{bmatrix} \begin{bmatrix} x_k \\ \check{x}_{k|k} - x_k \end{bmatrix} + \begin{bmatrix} G \\ (LC - I)G \end{bmatrix} \begin{bmatrix} w_k \\ v_k \end{bmatrix} \]

The covariance matrix for the augmented system:
\[ S = \begin{bmatrix} S_{11} & S_{12} \\ S_{12}^T & S_{22} \end{bmatrix} = E \left[ \begin{bmatrix} x_k \\ \check{x}_{k|k} - x_k \end{bmatrix} \begin{bmatrix} x_k^T & \check{x}_{k|k}^T - x_k^T \end{bmatrix} \right] \quad \text{(B7)} \]

$S$ is given by the following Lyapunov equation:
\[ S = \bar{A}S\bar{A}^T + \bar{G} \begin{bmatrix} Q_w & 0 \\ 0 & R_v \end{bmatrix} \bar{G}^T \quad \text{(B8)} \]

The closed-loop cost in Eq. B5 can now be written as follows (using rules for trace and expectations):
\[ \Phi = \lim_{N \to \infty} \frac{1}{N} E \left[ \sum_{k=0}^{N} x_k^T Q x_k + u_k^T R u_k \right] \]
\[ \quad = \lim_{N \to \infty} \frac{1}{N} \sum_{k=0}^{N} \text{tr} [Q E(x_k, x_k^T) + K^T R K (\check{x}_{k|k} - x_k)] \]
\[ \quad = \lim_{N \to \infty} \frac{1}{N} \sum_{k=0}^{N} \text{tr} [Q S_{11} + K^T R K (S_{11} + S_{12} + S_{12}^T + S_{22})] \]
\[ \quad = \text{tr} [Q S_{11} + K^T R K (S_{11} + S_{12} + S_{12}^T + S_{22})] \]

\[ \text{Manuscript received Aug. 9, 2007, revision received May 29, 2008, and final revision received Aug. 23, 2008.} \]