A Modeling and Filtering Framework for Linear Differential-Algebraic Equations

Thomas Schön, Markus Gerdin, Torkel Glad, and Fredrik Gustafsson
Control & Communication
Department of Electrical Engineering
Linköping University
SE–581 83 Linköping, Sweden
(schon, gerdin, torkel, fredrik)@isy.liu.se

***Abstract*— General approaches to modeling, for instance using object-oriented software, lead to differential-algebraic equations (DAE). As the name reveals, it is a combination of differential and algebraic equations. For state estimation using observed system inputs and outputs in a stochastic framework similar to Kalman filtering, we need to augment the DAE with stochastic disturbances ("process noise"), whose covariance matrix becomes the tuning parameter. We will determine the subspace of possible causal disturbances based on the linear DAE model. This subspace determines all degrees of freedom in the filter design, and a Kalman filter algorithm is given. We illustrate the design on a system with two interconnected rotating masses.**

Keywords: Differential-algebraic equations, Implicit systems, Descriptor systems, Singular systems, White noise, Noise, Discretization, Kalman filters

I. Introduction

In recent years so-called object-oriented modeling software has increased in popularity. Examples of such software are Omola, Dymola, the SimMechanics toolbox for MATLAB, and Modelica [14], [20]. Such modeling software makes it possible to model physical systems by connecting submodels in a way which parallels the physical construction and without having to manually manipulate any equations. The available software usually gives the user the possibility to simulate the system, and perhaps also to extract a structured model in an automatic way. This model generally becomes a differential algebraic equation (DAE), which in the linear case can be written

$$E\dot{x}(t) + Fx(t) = B_u u(t), \tag{1a}$$

where $x(t)$ is the internal variable vector, $u(t)$ is the system input vector and E, F, B_u are matrices of appropriate dimensions. We assume that E is singular, otherwise we get an ordinary differential equation (ODE) by simply multiplying with E^{-1} from the left, and the standard Kalman filtering theory applies. Hence, when E is singular we obtain a differential-algebraic equation and the reason for the singularity is often that purely algebraic equations are present. Other common names for the model structure (1a) are e.g., implicit systems, descriptor systems, semi-state systems, generalized systems, and differential equations on a manifold [3].

We have the possibility to place sensors in the system to get a measurement equation

$$y(t) = Cx(t) + e(t), \tag{1b}$$

where $y(t)$ is the measurement and $e(t)$ the sensor noise. An important special case we will discuss separately is for computer controlled systems, where the measurements $y[k]$ are available at the sampling times $t = kT_s$,

$$E\dot{x}(t) + Fx(t) = B_u u(t), \tag{2a}$$
$$y[k] = Cx(kT_s) + e[k]. \tag{2b}$$

The estimation problem is to estimate $x(t)$ from $y[k]$. There are two reasons why we have to introduce process noise to (2a):

- There are unmodeled dynamics and disturbances acting on the system, that can only be included in the model as an unknown stochastic term.
- There is a practical need for tuning the filter in order to make a trade-off between tracking ability and sensor noise attenuation. This is in the Kalman filter done by keeping the sensor noise covariance matrix constant and tuning the process noise covariance matrix, or the other way around. Often, it is easier to describe the sensor noise in a stochastic setting, and then it is more natural to tune the process noise.

With process noise, the model (1) becomes

$$E\dot{x}(t) + Fx(t) = B_u u(t) + B_w w(t), \tag{3a}$$
$$y(t) = Cx(t) + e(t). \tag{3b}$$

The problem is to determine where in the system disturbances can occur. To fit the optimal filtering and Kalman filtering framework, $w(t)$ should be white noise. As will be demonstrated, adding white noise to all equations can lead to derivatives of white noise affecting internal variables of the system directly. This will be referred to as a non-causal system, with a physical interpretation of infinite forces, currents etc. Therefore, we will derive a basis for the subspace of all possible causal disturbances. This basis is taken as B_w in (3), and the process noise covariance matrix $Q = \mathrm{Cov}(w(t))$ is used as the design variable to rotate and scale this basis. This is a new way of defining the process noise as far as we know. The problem itself, however, is addressed in [3], where it is suggested to use band limited noise to avoid these problems. The idea is that the derivative of such noise exists, but the drawback is that the Kalman filter will become sub-optimal.

A system with the same structure as (3) but in discrete time will be referred to as a discrete time descriptor system. Such systems may also be non-causal, but are easier to handle since the non-causality here means dependence on future values of the noise or the input. An application for such systems is discrete time state-space systems with constraints. For an example see [19]. In the discrete time case much work has already been done, for example on Kalman filtering [4], [8], [15], [16], [7], [5]. In the continuous time case much less work has been done on statistical methods. However, some attempts to introduce white noise in the continuous case has been done as well, see e.g., [18], [22].

II. Derivation of the process noise subspace

We will omit the deterministic input in this derivation for notational convenience, so the continuous time linear invariant differential-algebraic equations considered has the form (4). The reader is referred to [9] for details on how the non-causality with respect to the input signal, $u(t)$, can be handled.

$$E\dot{x}(t) + Fx(t) = Bw(t) \tag{4a}$$

$$y(t) = Cx(t) + e(t) \tag{4b}$$

The E, F, and C matrices in (4) are constant matrices. For the purpose of this discussion we will assume that w and e are continuous time white noises. (See [1] for a thorough treatment of continuous time white noise). If $\det(Es+F)$ is not identically zero as a function of $s \in \mathbf{R}$, (4) can always be transformed into the *standard form* (6) (see [2]). Note that if $\det(Es + F)$ is identically zero, then $x(t)$ is not uniquely determined by $w(t)$ and the initial value $x(0)$. This can be realized by Laplace transforming (4). Therefore it is a reasonable assumption that $\det(Es + F)$ is not identically zero.

A. Time-domain derivation

First, a transformation to the standard form is needed. This is done by finding a suitable change of variables $x = Qz$ and a matrix P to multiply (4a) from the left. Both P and Q are non-singular matrices. By doing this we get

$$PEQ\dot{z}(t) + PFQz(t) = PBw(t), \tag{5}$$

which for suitably chosen P- and Q-matrices can be written in the following standard form:

$$\begin{bmatrix} I & 0 \\ 0 & N \end{bmatrix} \begin{bmatrix} \dot{z}_1(t) \\ \dot{z}_2(t) \end{bmatrix} + \begin{bmatrix} -A & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} z_1(t) \\ z_2(t) \end{bmatrix} = \begin{bmatrix} G_1 \\ G_2 \end{bmatrix} w(t), \tag{6}$$

where the N-matrix is *nilpotent*, i.e., $N^k = 0$ for some k. The matrices P and Q can be calculated using, e.g., ideas from [21] involving the generalized real Schur form and the generalized Sylvester equation. We can also write (6) on the form (7), see e.g., [6] or [13].

$$\dot{z}_1(t) = Az_1(t) + G_1 w(t), \tag{7a}$$

$$z_2(t) = \sum_{i=0}^{k-1} (-N)^i G_2 \frac{d^i w(t)}{dt^i}. \tag{7b}$$

From a theoretical point of view G_1 can be chosen arbitrarily, since it describes how white noise should enter an ordinary differential equation. However, constraints on G_1 can of course be imposed by the physics of the system that is modeled. When it comes to G_2, the situation is different, here we have to find a suitable parameterization. The problem is now that white noise cannot be differentiated, so we proceed to find a condition on the B-matrix in (4a) under which there does not occur any derivatives in (7b), i.e., $N^i G_2 = 0$ for all $i \geq 1$. This is equivalent to that $NG_2 = 0$. The result is given in the following theorem.

Theorem 2.1: The condition to avoid to differentiate white noise is equivalent to requiring that

$$B \in \mathcal{R}(M), \tag{8}$$

where M is a matrix derived from the standard form (6) (see the proof for details on how M is derived).

The expression $B \in \mathcal{R}(M)$ means that B is in the *range* of M, that is the columns of B are linear combinations of the columns of M.

Proof: Let the $n \times n$ matrix N in (6) have the singular value decomposition (SVD)

$$N = UDV^T. \tag{9}$$

Since it is nilpotent it is also singular, so m diagonal elements in D are zero. Partition $V = [V_1, V_2]$, where V_2 contains the last m columns of V having zero singular values. Then $NV_2 = 0$, and we can write $G_2 = V_2 T$, where T is an arbitrary $m \times m$ matrix, which parameterizes all matrices G_2 that satisfies $NG_2 = 0$.

According to (5) and (6) we have

$$B = P^{-1} \begin{bmatrix} G_1 \\ G_2 \end{bmatrix}. \tag{10}$$

If we now let $P^{-1} = \begin{bmatrix} R_1 & R_2 \end{bmatrix}$, we can write (10) as

$$\begin{aligned} B = P^{-1} \begin{bmatrix} G_1 \\ G_2 \end{bmatrix} &= \begin{bmatrix} R_1 & R_2 \end{bmatrix} \begin{bmatrix} G_1 \\ V_2 T \end{bmatrix} \\ &= \underbrace{[R_1 \ R_2 V_2]}_{M} \begin{bmatrix} G_1 \\ T \end{bmatrix} \end{aligned} \tag{11}$$

where both G_1 and T can be chosen arbitrarily. This calculation gives that

$$B \in \mathcal{R}(M) \tag{12}$$

is a condition for avoiding differentiation of the white noise signal $w(t)$. ■

The B-matrices satisfying (12) will thus allow us to incorporate white noise without having a problem with differentiation of white noise. The design parameters to be specified are G_1 and T defined in the proof above. Also note that the requirement that the white noise should not be differentiated is related to the concept of *impulse controllability* discussed in [6].

B. Frequency-domain derivation

The same condition on the noise can be derived in the frequency domain, as shown in this section. Throughout the section, we need some concepts from the theory of matrix fraction descriptions (MFD). We start by defining the *row degree* of a polynomial matrix and the concept of a *row reduced* MFD according to [17].

Definition 2.1: The i^{th} *row degree* of a polynomial matrix $P(s)$, written as $r_i[P]$, is the degree of the highest degree polynomial in the i^{th} row of $P(s)$.

Definition 2.2: If the polynomial matrix $P(s)$ is square and nonsingular, then it is called *row reduced* if

$$\deg[\det P(s)] = r_1[P] + \cdots + r_n[P]. \tag{13}$$

We will use the following theorem from [12]:

Theorem 2.2: If $D(s)$ is row reduced, then $D^{-1}(s)N(s)$ is proper if and only if each row of $N(s)$ has degree less than or equal the degree of the corresponding row of $D(s)$, i.e., $r_i[N] \leq r_i[D], i = 1, \ldots, n$.

To utilize the results we need to write (4a) as a matrix fraction description. A MFD of (4a) is

$$X(s) = (Es + F)^{-1}BW(s). \tag{14}$$

According to [17] a MFD can be converted to row reduced form by pre-multiplication of a unimodular[1] matrix $U(s)$. That is $D(s)$ is row reduced in the MFD

$$X(s) = D^{-1}(s)N(s)W(s) \tag{15}$$

where $D(s) = U(s)(Es + F)$ and $N(s) = U(s)B$ for a certain unimodular matrix $U(s)$. Now, Theorem 2.2 shows that the transfer function of the system is proper if the highest degree of the polynomials in each row in $N(s)$ is lower than or equal to the highest degree of the polynomials in the corresponding row of $D(s)$. This gives a condition on B in the following way:

Writing $U(s)$ as

$$U(s) = \sum_{i=0}^{m} U_i s^i \tag{16}$$

and writing the j^{th} row of U_i as U_{ij}, shows that the condition

$$U_{ij}B = 0 \quad i > r_j[D],\ j = 1 \ldots n \tag{17}$$

guarantees that the transfer function of the system is proper.

Conversely, assume that (17) does not hold. Then some row degree of N(s) is higher than the corresponding row degree of D(s), so the transfer function is then according to Theorem 2.2 not proper.

This discussion proves the following theorem.

Theorem 2.3: The transfer function of the system (4) is proper if and only if

$$U_{ij}B = 0 \quad i > r_j[D],\ j = 1 \ldots n. \tag{18}$$

Note that the criterion discussed in this section requires that the MFD is transformed to row reduced form, and an algorithm for finding this transformation is provided in [17].

We have now proved two theorems, one using time domain methods and one using frequency domain methods, that gives conditions which are equivalent to that no white noise is differentiated in (4). This means that these two conditions are equivalent as well. The frequency domain method is good in the sense that we do not have to compute the standard form (6). However if we want to discretize the equations it is worthwhile to compute the standard form. Once this is done the celebrated Kalman filter can be used to estimate the internal variables, $x(t)$. In the subsequent section we will discuss the discretization and the estimation problems.

III. Filtering

A. Discretization

If the noise enters the system according to a B-matrix satisfying Theorem 2.1 or 2.3 the original system (4) can be written as

$$\dot{z}_1(t) = Az_1(t) + G_1 w(t), \tag{19a}$$
$$z_2(t) = G_2 w(t), \tag{19b}$$
$$y(t) = CQz(t) + e(t). \tag{19c}$$

where $x = Qz$. Furthermore $w(t)$ and $e(t)$ are both assumed to be Gaussian white noise signals with covariances R_1 and R_2 respectively, and zero cross-covariance (the case of nonzero cross-covariance can be handled as well, the only difference is that the expressions are more involved).

The system (19) can be discretized using standard techniques from linear systems theory, see e.g., [17]. If we assume that $w(t)$ remains constant during one sample interval[2], we have (here it is assumed that sampling interval is one to simplify the notation)

$$w(t) = w[k], \qquad k \leq t < (k+1) \tag{20}$$

we obtain

$$z_1[k+1] = \tilde{A}z_1[k] + \tilde{G}_1 w[k], \tag{21a}$$
$$z_2[k] = G_2 w[k], \tag{21b}$$
$$y[k] = CQz[k] + e[k] \tag{21c}$$

[1] A polynomial matrix is called unimodular if its determinant is a nonzero real number [12].

[2] See e.g., [11] for a discussion on other possible assumptions on the stochastic process $w(t)$ when it comes to discretization.

where

$$\tilde{A} = e^{A} \qquad \tilde{G}_1 = \int_0^1 e^{A\tau} d\tau G_1. \tag{22}$$

Hence, Equation (21) and (22) constitutes a discrete time model of (4).

B. Kalman filter

In order to apply the Kalman filter to the discrete model (21) we start out by rewriting (21c) as

$$\begin{aligned} y[k] &= CQz[k] + e[k] = [\tilde{C}_1 \tilde{C}_2] \begin{bmatrix} z_1[k] \\ z_2[k] \end{bmatrix} + e[k] \\ &= \tilde{C}_1 z_1[k] + \tilde{C}_2 z_2[k] + e[k] \\ &= \tilde{C}_1 z_1[k] + \underbrace{\tilde{C}_2 G_2 w[k] + e[k]}_{\tilde{e}[k]} \end{aligned} \tag{23}$$

Combining (21a) and (23) we obtain

$$z_1[k+1] = \tilde{A} z_1[k] + \tilde{G}_1 w[k] \tag{24a}$$

$$y[k] = \tilde{C}_1 z_1[k] + \tilde{e}[k] \tag{24b}$$

Note that the measurement noise, $\tilde{e}[k]$, and the process noise, $w[k]$, are correlated. Now, the Kalman filter can be applied to (24) in order to estimate the internal variables $z_1[k]$ and the process noise $w[k]$. Finally an estimate of the internal variables $z_2[k]$ can be found using the estimated process noise, since $z_2[k] = G_2 w[k]$, according to (21b). Finally the internal variables, $x[k]$, are obtained by $x[k] = Q^{-1} z[k]$. For details on the Kalman filter see [10].

IV. EXAMPLE

In this example we will treat a system composed of two rotating masses as shown in Figure 1. The two rotating parts

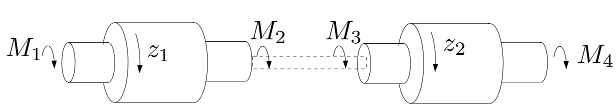

Fig. 1. Two interconnected rotating masses.

are described by the torques M_1, M_2, M_3 and M_4 and the angular velocities z_1 and z_2. The equations describing this system are

$$J_1 \dot{z}_1 = M_1 + M_2 \tag{25a}$$

$$J_2 \dot{z}_2 = M_3 + M_4 \tag{25b}$$

$$M_2 = -M_3 \tag{25c}$$

$$z_1 = z_2. \tag{25d}$$

Written on the form (4) these equations are

$$\begin{bmatrix} J_1 & 0 & 0 & 0 \\ 0 & J_2 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \dot{x} + \begin{bmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 1 \\ 1 & -1 & 0 & 0 \end{bmatrix} x = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} M_1 \\ M_4 \end{bmatrix}, \tag{26}$$

where $x = \begin{bmatrix} z_1 & z_2 & M_2 & M_3 \end{bmatrix}^T$. Note that the matrix in front of $\dot{x}$ is singular, hence (26) is a differential-algebraic equation. Using the following transformation matrices P and Q

$$P = \begin{bmatrix} 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ \frac{J_2}{J_1+J_2} & -\frac{J_1}{J_1+J_2} & \frac{J_2}{J_1+J_2} & 0 \end{bmatrix}, \tag{27}$$

$$Q = \begin{bmatrix} \frac{1}{J_1+J_2} & \frac{J_2}{J_1+J_2} & 0 & 0 \\ \frac{1}{J_1+J_2} & -\frac{J_1}{J_1+J_2} & 0 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \tag{28}$$

the equations can be written in the standard form (6):

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & \frac{J_1 J_2}{J_1+J_2} & 0 & 0 \end{bmatrix} \dot{z} + \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} z = \begin{bmatrix} 1 & 1 \\ 0 & 0 \\ 0 & 0 \\ \frac{J_2}{J_1+J_2} & -\frac{J_1}{J_1+J_2} \end{bmatrix} \begin{bmatrix} M_1 \\ M_2 \end{bmatrix} \tag{29}$$

Now to the important part, if we want to incorporate noise into the differential-algebraic equation (26), by adding Bw to (26), which B-matrices are allowed?

To answer this question Theorem 2.1 can be consulted. We begin by calculating the matrices R_1, R_2 and V_2 from (27) and (29). We have that

$$N = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ \frac{J_1 J_2}{J_1+J_2} & 0 & 0 \end{bmatrix} \quad \Rightarrow \quad V_2 = \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \tag{30}$$

and

$$P^{-1} = \begin{bmatrix} \frac{J_1}{J_1+J_2} & 0 & -1 & 1 \\ \frac{J_2}{J_1+J_2} & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \Rightarrow \tag{31}$$

$$R_1 = \begin{bmatrix} \frac{J_1}{J_1+J_2} \\ \frac{J_2}{J_1+J_2} \\ 0 \\ 0 \end{bmatrix}, \quad R_2 = \begin{bmatrix} 0 & -1 & 1 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \tag{32}$$

We can now calculate the M matrix:

$$M = \begin{bmatrix} R_1 & R_2V_2 \end{bmatrix} = \begin{bmatrix} \frac{J_1}{J_1+J_2} & -1 & 1 \\ \frac{J_2}{J_1+J_2} & 0 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad (33)$$

As the requirement was that $B \in \mathcal{R}(M)$ this simply means that we cannot directly add white noise to equation (25d) (if $J_1 > 0$ and $J_2 > 0$). This result makes physical sense, as a step change in the angular velocity would require an infinite torque.

The same condition on B can also be calculated in the frequency domain using Theorem 2.3. Transforming the system to row reduced form gives that

$$U(s) = \begin{bmatrix} -\frac{1}{J_1} & -\frac{1}{J_2} & 0 & s \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (34)$$

$$= \underbrace{\begin{bmatrix} -\frac{1}{J_1} & -\frac{1}{J_2} & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}}_{U_0} + \underbrace{\begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}}_{U_1} s \quad (35)$$

and that

$$D(s) = \begin{bmatrix} 0 & 0 & \frac{1}{J_1} & -\frac{1}{J_2} \\ 0 & J_2 s & 0 & -1 \\ 0 & 0 & 1 & 1 \\ 1 & -1 & 0 & 0 \end{bmatrix} \quad (36)$$

with notation from section II-B.

This gives that the row degrees of $D(s)$ are $r_1[D] = 0$, $r_2[D] = 1$, $r_3[D] = 0$, and $r_4[D] = 0$. Theorem 2.3 thus gives that the transfer function is proper if and only if

$$\begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} B = 0. \quad (37)$$

What equation (37) says is that the last row of B must be zero, which is the same conclusion as was reached using the time domain method, Theorem 2.1.

V. Discrete time linear descriptor systems

The discrete linear time invariant descriptor system is an equation on the form

$$Ex[k+1] + Fx[k] = Bw[k], \quad (38a)$$

$$y[k] = Cx[k] + e[k], \quad (38b)$$

where E, F, and C are constant matrices and $w[k]$ and $e[k]$ are white noise sequences, i.e., sequences of independent (identically distributed) random variables. For this case it is possible to make the same transformation as for a continuous differential-algebraic equation if $\det(Ez + F)$ is not identically zero as a function of $z \in \mathbf{R}$ (Section II) since the structure is similar. Similarly to the continuous time case, $x[k]$ will not be uniquely determined by $w(k)$ if $\det(Ez+F)$ is identically zero. A certain transformation

$$PEQx[k+1] + PFQx[k] = PBw[k] \quad (39)$$

with non-singular matrices P and Q will thus give us the form

$$\begin{bmatrix} I & 0 \\ 0 & N \end{bmatrix} \begin{bmatrix} z_1[k+1] \\ z_2[k+1] \end{bmatrix} + \begin{bmatrix} -A & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} z_1[k] \\ z_2[k] \end{bmatrix} = \begin{bmatrix} G_1 \\ G_2 \end{bmatrix} w[k]. \quad (40)$$

As in the continuous time case, we can write (40) in the form

$$z_1[k+1] = Az_1[k] + G_1 w[k] \quad (41a)$$

$$z_2[k] = \sum_{i=0}^{n-1} (-N)^i G_2 w[k+i]. \quad (41b)$$

The system (38) is thus well defined for all B-matrices, since no derivatives occur in this case. However, $z_2[k]$ will depend on future values of the noise. To avoid this, the noise sequence can be time shifted. If we let $\tilde{w}[k] = w[k+n-1]$ Equation (41) can be written

$$z_1[k+1] = Az_1[k] + G_1 \tilde{w}[k-n+1] \quad (42a)$$

$$z_2[k] = \sum_{i=-n+1}^{0} (-N)^i G_2 \tilde{w}[k+i] \quad (42b)$$

which can be transformed to a normal state-space description. This state-space description can then be used to implement a Kalman filter, which is discussed in [4]. Other approaches to Kalman filtering of discrete time linear descriptor systems are discussed in, e.g., [8], [15], [16], [7], [5].

The sequences $w[k]$ and $\tilde{w}[k]$ will have the same statistical properties since they both are white noise sequences.

It can be also be noted that the same requirement that was put on B in the continuous time case may also be used in the discrete time case. This would then guarantee that the system would not depend on future noise values and the noise sequence would not have to be time shifted.

A. Frequency domain

The ideas of time shifting the noise might become clearer if they are treated in the frequency domain. If we transform (38) to the frequency domain we get

$$X(z) = \underbrace{(Ez+F)^{-1}B}_{H(z)} W(z). \quad (43)$$

The only difference from a transfer function for a state-space system is that here $H(z)$ is non-causal in the general case.

If we rewrite (43) as

$$X(z) = \underbrace{H(z)z^{-T}}_{\tilde{H}(z)}\underbrace{z^{T}W(z)}_{\tilde{W}(z)}, \tag{44}$$

then $\tilde{W}(z)$ will be a time shifted version of $W(z)$ and $\tilde{H}(z)$ will be a causal transfer function if T is large enough.

VI. Conclusions

We have in this article proposed a framework for modeling and filtering of systems composed of linear differential-algebraic equations. The main reason for studying these systems is that they occur as the natural description delivered from object-oriented modeling software. At the core of this problem we find the question of how to incorporate stochastics into linear differential-algebraic equations. This has been solved in this paper in the case where white noise is used. The result was presented as two equivalent theorems, one in the time domain and one in the frequency domain. The resulting model fits into the optimal filtering framework and standard methods such as the Kalman filter applies. An example was also given, which showed that the conditions derived for how the noise can enter the system gives requirements which are physically motivated.

VII. Acknowledgments

The authors would like to thank the Swedish Research Council and the Foundation for Strategic Research for supporting this work.

VIII. References

[1] K.J. Åström. *Introduction to Stochastic Control Theory*. Mathematics in Science and Engineering. Academic Press, New York and London, 1970.

[2] K.E. Brenan, S.L. Campbell, and L.R. Petzold. *Numerical Solution of Initial-Value Problems in Differential-Algebraic Equations*. SIAM's Classics in Applied Methematics. SIAM, New York, 1996.

[3] S.L. Campbell. Descriptor systems in the 90's. In *Proceedings of the 29th Conference on Decision and control*, pages 442–447, Honolulu, Hawaii, USA, December 1990. IEEE.

[4] L. Dai. State estimation schemes for singular systems. In *Preprints of the 10th IFAC World Congress, Munich, Germany*, volume 9, pages 211–215, 1987.

[5] L. Dai. Filtering and LQG problems for discrete-time stochastic singular systems. *IEEE Transactions on Automatic Control*, 34(10):1105–1108, October 1989.

[6] L. Dai. *Singular Control Systems*. Lecture Notes in Control and Information Sciences. Springer-Verlag, Berlin, New York, 1989.

[7] M. Darouach, M. Zasadzinski, and D. Mehdi. State estimation of stochastic singular linear systems. *International Journal of Systems Sciences*, 24(2):345–354, 1993.

[8] Z.L. Deng and Y.M. Liu. Descriptor kalman filtering. *International Journal of Systems Science*, 30(11):1205–1212, 1999.

[9] M. Gerdin, T. Glad, and L. Ljung. Parameter estimation in linear differential-algebraic equations. In *Proceedings of the 13th IFAC symposium on system identification*, Rotterdam, the Netherlands, 2003. Accepted for publication.

[10] T. Glad and L. Ljung. *Control Theory, Multivariable and Nonlinear Methods*. Taylor and Francis, 2000.

[11] F. Gustafsson. *Adaptive Filtering and Change Detection*. John Wiley & Sons, 2000.

[12] T. Kailath. *Linear systems*. Prentice Hall, Englewood Cliffs, N.J., 1980.

[13] L. Ljung and T. Glad. *Modellbygge och simulering*. Studentlitteratur, 2003. To appear. In Swedish.

[14] S.E. Mattsson, H. Elmqvist, and M. Otter. Physical system modeling with Modelica. *Control Engineering Practice*, 6:501–510, 1998.

[15] R. Nikoukhah, S.L. Campbell, and F. Delebecque. Kalaman filtering for general discrete-time linear systems. In *Proceedings of the 37th Conference on Decision and Control*, pages 2886–2891, Tampa, Florida, USA, Dec 1998.

[16] R. Nikoukhah, S.L. Campbell, and F. Delebecque. Kalaman filtering for general discrete-time linear systems. *IEEE Transactions on Automatic Control*, 44(10):1829–1839, Oct 1999.

[17] W.J. Rugh. *Linear System Theory*. Information and system sciences series. Prentice Hall, Upper Saddle River, New Jersey, 2 edition, 1996.

[18] O. Schein and G. Denk. Numerical solution of stochastic differential-algebraic equations with applications to transient noise simulation of microelectronic circuits. *Journal of Computational and Applied Mathematics*, 100:77–92, 1998.

[19] T. Schön, F. Gustafsson, and A. Hansson. A note on state estimation as a convex optimization problem. In *IEEE International Conference on Acoustics, Speech, and Signal Processing*, volume 6, pages 61–64, Hong Kong, April 2003.

[20] M. Tiller. *Introduction to Physical Modeling with Modelica*. Kluwer, Boston, Mass., 2001.

[21] A. Varga. Numerical algorithms and software tools for analysis and modelling of descriptor systems. In *Prepr. of 2nd IFAC Workshop on System Structure and Control, Prague, Czechoslovakia*, pages 392–395, 1992.

[22] R. Winkler. Stochastic Differential Algebraic Equations of Index 1 and Applications in Circuit Simulation. Humboldt-University, Berlin, Institut fur Mathematik, 2001. To appear in J.Comp.Math.Appl.