Proceedings of the
44th IEEE Conference on Decision and Control, and
the European Control Conference 2005
Seville, Spain, December 12-15, 2005

MoA11.1

Well-posedness of Filtering Problems for Stochastic Linear DAE Models

Markus Gerdin, Torkel Glad, and Lennart Ljung
Division of Automatic Control, Department of Electrical Engineering
Linköping University, SE-581 83 Linköping, Sweden
(gerdin,torkel,ljung)@isy.liu.se

Abstract— **Modern modeling tools often give descriptor or DAE models, i.e., models consisting of a mixture of differential and algebraic relationships. The introduction of stochastic signals into such models in connection with filtering problems raises several questions of well-posedness. The main problem is that the system equations may contain hidden relationships affecting variables defined as white noise. The result might be that certain physical variables get infinite variance or contain formal differentiations of white noise. The paper gives conditions for well-posedness in terms of certain subspaces defined by the system matrices.**

I. Introduction

The increased use of modeling in many fields of engineering has lead to development of new techniques to make modeling of complex dynamic systems easier. An increased use of model libraries and so-called object-oriented modeling is one result of this trend. The output from these modeling tools is usually not a state-space model but a model containing a mixture of equations, some of which contain variables differentiated with respect to time. These models are usually called descriptor models or DAE (differential-algebraic equation) models. There is a vast recent literature about these models but we have not found a complete treatment of continuous-time filtering problems, in particular concerning the well-posedness of DAE models with white noise inputs.

The problem centers around the fact that careless introduction of noise in DAE models might lead to derivatives of the noise appearing. The question whether these derivatives are well-defined has to some extent been discussed by other authors. In [1], it is guaranteed that the noise is not differentiated by assuming that the system is *index 1* (see, e.g., [2]). The assumption that the system is index 1 is more restrictive than is necessary, and rules out some applications such as many mechanics systems. This assumption will not be made in this paper, instead an exact condition for when the model is well-defined is derived. [1] also notes that some internal variables actually may be *generalized stochastic processes*, that is equal to a white noise process. [3] makes the same assumption as [1], but also treats a class of nonlinear DAE.

In [4] index 1 is assumed and a Kalman filter is constructed. However, in the estimation procedure the authors seem to overlook the fact that some variables may have

infinite variance. In [5], the original system specification may actually specify derivatives of white noise, but a controller is designed that removes any derivatives. In [6] the restrictive assumption that $\mathcal{R}\left(\begin{bmatrix} F & G \end{bmatrix}\right) \subseteq \mathcal{R}(E)$ guarantees that no derivatives appear (see the next section for notation), although this is not stated explicitly. Finally, in [7] nonlinear semi-explicit DAE (see, e.g., [2]) are discussed. Here well-posedness is guaranteed by only adding noise to the state-space part of the system. This is a sufficient, but not a necessary condition to avoid differentiated noise.

In [8] a transformation to a standard form is used to study when the filter problem is well-defined. Since the transformation matrices are involved it is fairly difficult to see the conditions in terms of the original matrices. We will therefore formulate our results in terms of certain subspaces instead.

II. Problem Formulation and Basic Structure

Consider a physical system described by a DAE model

$$E\dot{z} = Fz + Gu + Jw \tag{1a}$$
$$y = Hz + e. \tag{1b}$$

Here z is an n-vector of physical variables, y an n_y-vector of measurements, u an n_u-vector of inputs, w an n_w-vector of process noise variables and e a vector of measurement noise variables. E and F are square matrices while G, H and J are rectangular matrices of appropriate dimensions. With w and e modeled as stochastic processes, we call (1) a *stochastic differential-algebraic equation*, SDAE.

It is a feature of the modeling techniques mentioned in the introduction that they often introduce a number of variables that only play a role in intermediate calculations and are of no interest in themselves. Therefore we introduce the variable $\bar{z}$ where all (linear combinations of) components of z that are of interest are collected,

$$\bar{z} = Mz$$

for some rectangular matrix M.

The following assumption is standard in the theory of descriptor systems and guarantees the existence of a unique solution in the deterministic case (provided appropriate initial conditions are specified).

Assumption 1: There exists a scalar λ such that $\lambda E + F$ is nonsingular.

Now choose a λ such that $\lambda E + F$ is nonsingular and define the matrix $\bar{E}$ according to

$$\bar{E} = (\lambda E + F)^{-1} E \tag{2}$$

Let $\mathcal{V}$ and $\mathcal{N}$ denote the range and null spaces of a matrix. Then

$$\mathcal{V}(\bar{E}) \supset \mathcal{V}(\bar{E}^2) \supset \cdots \tag{3}$$
$$\mathcal{N}(\bar{E}) \subset \mathcal{N}(\bar{E}^2) \subset \cdots \tag{4}$$

Since all subspaces are finite-dimensional the inclusions must become equalities above some power of $\bar{E}$. Since $\dim(\mathcal{N}(\bar{E}^j)) + \dim(\mathcal{V}(\bar{E}^j)) = n$ for every j, this occurs for the same power in both sequences. We can then write

$$\mathcal{V}(\bar{E}) \supset \mathcal{V}(\bar{E}^2) \supset \cdots \supset \mathcal{V}(\bar{E}^k) = \mathcal{V}(\bar{E}^{k+1}) = \cdots \tag{5}$$
$$\mathcal{N}(\bar{E}) \subset \mathcal{N}(\bar{E}^2) \subset \cdots \subset \mathcal{N}(\bar{E}^k) = \mathcal{N}(\bar{E}^{k+1}) = \cdots \tag{6}$$

for some integer k. We assume that k is taken as the smallest integer for which (5), (6) holds. It is then called the *index* of the system (1). We introduce

$$X_{ss} = \mathcal{V}(\bar{E}^k), \quad X_{av} = \mathcal{N}(\bar{E}^k) \tag{7}$$

and define

$$n_s = \dim(X_{ss}), \quad n_a = \dim(X_{av}) \tag{8}$$

The following properties are easily shown.

Proposition 1: The spaces X_{ss} and X_{av} have the properties

 (i) X_{ss}, X_{av} are invariant under $\bar{E}$
 (ii) $X_{ss} \cap X_{av} = \{0\}$
 (iii) X_{ss}, X_{av} together span $\mathbb{R}^n$
 (iv) $\bar{E}$ is nonsingular when restricted to X_{ss}
 (v) $\bar{E}$ is nilpotent when restricted to X_{av}
 (vi) X_{ss}, X_{av} are independent of the choice of λ in (2) (as long as the inverse is defined)

Proof: Properties (i)–(v) follow directly from the relations

$$X_{ss} = \mathcal{V}(\bar{E}^k) = \mathcal{V}(\bar{E}^{k+1}), \quad X_{av} = \mathcal{N}(\bar{E}^k) = \mathcal{N}(\bar{E}^{k+1})$$

Property (vi) can be seen by defining

$$\bar{E}_\sigma = ((\sigma + \lambda)E + F)^{-1} E.$$

It follows that

$$\bar{E}_\sigma = ((\lambda E + F)(\sigma \bar{E} + I))^{-1} E = (\sigma \bar{E} + I)^{-1} \bar{E}.$$

It is easy to see that

$$(\sigma \bar{E} + I)^{-1} \bar{E} = \bar{E}(\sigma \bar{E} + I)^{-1}$$

so that for any positive integer j

$$\bar{E}_\sigma^j = (\sigma \bar{E} + I)^{-j} \bar{E}^j = \bar{E}^j (\sigma \bar{E} + I)^{-j}.$$

Since the range space of a matrix is not affected by multiplication with a nonsingular matrix from the right, X_{ss} remains the same if it is computed from $\bar{E}_\sigma^k$ or from $\bar{E}^k$. Since the null space remains the same when multiplying from the left

by a nonsingular matrix a similar statement holds for X_{av}. ∎

It follows that we can think of X_{ss} as being the state-space of the model and of X_{av} as being a space of auxiliary "non-dynamic" variables.

To make the division into X_{ss} and X_{av} explicit it is convenient to make a state transformation.

Proposition 2: Let T_1 and T_2 be matrices whose columns span X_{ss} and X_{av} respectively (and consequently together span $\mathbb{R}^n$). Then the transformation

$$z = \underbrace{(T_1 \quad T_2)}_{T} \underbrace{\begin{pmatrix} x_s \\ x_a \end{pmatrix}}_{x} \tag{9}$$

gives a system description of the form

$$E_s \dot{x}_s = (I - \lambda E_s) x_s + G_s u + J_s w \tag{10a}$$
$$N \dot{x}_a = (I - \lambda N) x_a + G_a u + J_a w \tag{10b}$$

where

$$\begin{pmatrix} E_s & 0 \\ 0 & N \end{pmatrix} = T^{-1} \bar{E} T \tag{11a}$$

$$\begin{pmatrix} J_s \\ J_a \end{pmatrix} = T^{-1} (\lambda E + F)^{-1} J \tag{11b}$$

$$\begin{pmatrix} G_s \\ G_a \end{pmatrix} = T^{-1} (\lambda E + F)^{-1} G \tag{11c}$$

The matrix E_s is invertible while N is nilpotent, with k as its index of nilpotency.

Proof: Adding $\lambda E z$ to each side of the first equation in (1) and then multiplying from the left with $(\lambda E + F)^{-1}$ gives

$$\bar{E}(\dot{z} + \lambda z) = z + (\lambda E + F)^{-1}(Gu + Jw)$$

Substituting $z = Tx$ and multiplying from the left with T^{-1} gives

$$T^{-1} \bar{E} T (\dot{x} + \lambda x) = x + T^{-1}(\lambda E + F)^{-1}(Gu + Jw)$$

From Proposition 1 it follows that $T^{-1} \bar{E} T$ has the block structure

$$T^{-1} \bar{E} T = \begin{pmatrix} E_s & 0 \\ 0 & N \end{pmatrix}$$

while the nonsingularity of E_s and nilpotency of N follow from Proposition 1 (iv) and (v) respectively. ∎

Proposition 3: The auxiliary variables x_a can be solved from (10b) to give

$$x_a = -\left(I + \left(\frac{d}{dt} + \lambda \right) N + \cdots \right.$$
$$\left. \cdots + \left(\frac{d}{dt} + \lambda \right)^{k-1} N^{k-1} \right) \left(G_a u + J_a w \right) \tag{12}$$

Proof: Writing (10b) as

$$x_a = N \left(\frac{d}{dt} + \lambda \right) x_a - \left(G_a u + J_a w \right) \tag{13}$$

and differentiating successively gives

$$N\Big(\frac{d}{dt}+\lambda\Big)x_a = N^2\Big(\frac{d}{dt}+\lambda\Big)^2 x_a - N\Big(\frac{d}{dt}+\lambda\Big)\Big(G_a u + J_a w\Big)$$

$$\vdots$$

$$N^{k-1}\Big(\frac{d}{dt}+\lambda\Big)x_a = -N^{k-1}\Big(\frac{d}{dt}+\lambda\Big)^{k-1}\Big(G_a u + J_a w\Big)$$

where we have used $N^k = 0$ in the last equation. A successive substitution from these equations into (13) then gives (12). ∎

III. Well-posedness of the filtering problem

We now consider the question of well-posedness of the solution z of (1) when w is interpreted as white noise. To formulate the result, the concept of an oblique projection will be defined.

Definition 1: Let B and C be spaces with $B \cap C = \{0\}$ that together span $\mathbb{R}^n$. Let the matrices $\bar{B}$ and $\bar{C}$ be bases for B and C respectively. The *oblique projection* of a matrix A along B on C is defined as

$$A/_B C \triangleq \begin{pmatrix} 0 & \bar{C} \end{pmatrix} \begin{pmatrix} \bar{B} & \bar{C} \end{pmatrix}^{-1} A \tag{14}$$

Note that the projection is independent of the choice of bases for B and C.

This definition basically follows the definition in [9, Section 1.4.2]. However, we here consider projections along column spaces instead of row spaces. Also, the conditions on the spaces B and C give a simpler definition. The more general version in [9] is not necessary here.

Note the important properties

$$\bar{C}/_B C = \bar{C} \tag{15a}$$
$$\bar{B}/_B C = 0 \tag{15b}$$

of the oblique projection. They will be utilized later. The main result can now be formulated.

Theorem 1: Consider the DAE model (1) with w regarded as white noise. There exists a unique well defined solution z, in the sense of linear stochastic differential equations, with all components of z having finite variance, provided the condition

$$(\lambda E + F)^{-1} J \in X_{ss} \tag{16}$$

is satisfied. If it is only required that $\bar{z}$ is well defined with finite variance, then the condition can be relaxed to

$$\big[\bar{E}^j(\lambda E + F)^{-1}J\big]/_{X_{ss}} X_{av} \in \mathcal{N}(M), \quad j = 0, \ldots, k-1. \tag{17}$$

Proof: Transforming the system into the form (10), we see that the equation for x_s can be interpreted as a stochastic differential equation

$$dx_s = (E_s^{-1} - \lambda I)x_s dt + E_s^{-1} G_s u dt + E_s^{-1} J_s dw \tag{18}$$

so that x_s gets a well-defined solution. From Proposition 2, equation (11b) we see that the condition (16) means that $J_a = 0$ in (12). It follows that x_a is a deterministic variable and that

$$z = T_1 x_s + T_2 x_a$$

is well-defined with finite variance. If it is only required that $\bar{z} = Mz$ is well-defined, then it is sufficient that $MT_2 x_a$ is well-defined. Equation (12) shows that this is equivalent to

$$MT_2 N^j J_a = 0, \quad j = 0, 1, \ldots, k-1.$$

This condition can be rewritten according to

$$
\begin{aligned}
0 &= MT_2 N^j J_a \\
&= M\,[T_1 E_s^j J_s + T_2 N^j J_a]/_{\mathcal{V}(T_1)}\,\mathcal{V}(T_2) \\
&= M\left[\begin{pmatrix} T_1 & T_2 \end{pmatrix}\begin{pmatrix} E_s^j & 0 \\ 0 & N^j \end{pmatrix}\begin{pmatrix} J_s \\ J_a \end{pmatrix}\right]\Big/_{\mathcal{V}(T_1)}\,\mathcal{V}(T_2) \\
&= M\,[\bar{E}^j(\lambda E + F)^{-1}J]/_{\mathcal{V}(T_1)}\,\mathcal{V}(T_2) \\
&= M\,[\bar{E}^j(\lambda E + F)^{-1}J]/_{X_{ss}}\,X_{av} \tag{19}
\end{aligned}
$$

where the relations $\mathcal{V}(T_1) = X_{ss}$ and $\mathcal{V}(T_2) = X_{av}$ were used in the last step. This gives the condition (17). ∎

We can now consider the filtering problem for (1).

Theorem 2: Consider the DAE model (1) with w and e regarded as white noises. A filter estimate of $\bar{z} = Mz$, with z given by the DAE model (1) can be computed using standard Kalman filter techniques, provided the conditions

$$\big[\bar{E}^j(\lambda E + F)^{-1}J\big]/_{X_{ss}} X_{av} \in \mathcal{N}(M), \quad j = 0, \ldots, k-1 \tag{20a}$$

$$\big[\bar{E}^j(\lambda E + F)^{-1}J\big]/_{X_{ss}} X_{av} \in \mathcal{N}(H), \quad j = 1, \ldots, k-1. \tag{20b}$$

are satisfied.

Proof: In the transformed variables the measurement equation can be written

$$y = Hz + e = HT_1 x_s + HT_2 x_a + e$$

Replacing x_a using (12) gives

$$y = HT_1 x_s - HT_2\Big(I + \Big(\frac{d}{dt}+\lambda\Big)N + \cdots$$
$$\cdots + \Big(\frac{d}{dt}+\lambda\Big)^{k-1}N^{k-1}\Big)\Big(G_a u + J_a w\Big) + e.$$

In order to avoid derivatives of the white noise w in the measurement equation we have to require that

$$HT_2 N^j J_a = 0, \quad j = 1, \ldots, k-1.$$

Repeating the calculation (19) gives that this is equivalent to

$$H\big[\bar{E}^j(\lambda E + F)^{-1}J\big]/_{X_{ss}} X_{av} = 0, \quad j = 1, \ldots, k-1$$

which gives the condition (20b). If this condition is satisfied, the measurement equation becomes

$$y = HT_1 x_s -$$
$$HT_2\Big(I + \cdots + \Big(\frac{d}{dt}+\lambda\Big)^{k-1}N^{k-1}\Big)G_a u -$$
$$HT_2 J_a w + e. \tag{21}$$

Together with the state-space description for x_s (18),

$$dx_s = (E_s^{-1} - \lambda I)x_s dt + E_s^{-1} G_s u dt + E_s^{-1} J_s dw, \tag{22}$$

(21) gives a well-defined filtering problem. The continuous-time Kalman filter can thus be used to compute an estimate of x_s from (21) and (22). As in Theorem 1, the condition (20a) ensures that $MT_2 x_a$ is a function of the deterministic input u,

$$MT_2 x_a = -MT_2 \left(I + \cdots + \left(\frac{d}{dt} + \lambda \right)^{k-1} N^{k-1} \right) G_a u.$$

The estimate of $\bar{z}$ is thus given as a linear combination of x_s and the deterministic input u,

$$\bar{z} = Mz = MT_1 x_s + MT_2 x_a.$$

∎

It can be noted that the initial value $z(t_0)$ cannot be chosen freely, since x_a is determined by the deterministic input u. Only the part of $z(t_0)$ determined by x_s, that is

$$[z(t_0)]/_{X_{av}} X_{ss}, \tag{23}$$

can be given an arbitrary value. Any conflicting values of $z(t_0)$ will be ignored and have no consequence for the estimation of $z(t)$, $t > t_0$.

IV. AN EXAMPLE

Consider two bodies, each with unit mass, moving in one dimension with velocities v_1 and v_2 and subject to external forces w_1 and w_2 respectively. If the two bodies are joined together the situation is described by the following set of equations

$$\begin{aligned} \dot{v}_1 &= f + w_1 \\ \dot{v}_2 &= -f + w_2 \\ 0 &= v_1 - v_2 \end{aligned} \tag{24}$$

where f is the force acting between the bodies. It is typical of the models obtained when joining components from model libraries that too many variables are included. (In this simple case it is of course obvious to the human modeler that this model can be simplified to that of a body with mass 2 accelerated by $w_1 + w_2$.) In the notation of (1) we have, with $z = [v_1 \ v_2 \ f]^T$,

$$E = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad F = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & -1 \\ 1 & -1 & 0 \end{pmatrix} \quad J = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}.$$

With $\lambda = 1$ we get

$$\bar{E} = \frac{1}{2} \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & -1 & 0 \end{pmatrix}$$

which gives $k = 2$ and

$$X_{ss} = sp \left\{ \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \right\}, \quad X_{av} = sp \left\{ \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\}$$

With

$$x_s = x_1 = \frac{v_1 + v_2}{2}, \quad x_a = \begin{pmatrix} x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} (v_1 - v_2)/2 \\ f \end{pmatrix}$$

the transformed system (10a), (10b) becomes

$$\begin{aligned} \dot{x}_1 &= \frac{1}{2}(w_1 + w_2) \\ 0 &= x_2 \\ \dot{x}_2 &= -x_2 + x_3 + \frac{1}{2}(w_1 - w_2) \end{aligned} \tag{25}$$

Since

$$(\lambda E + F)^{-1} J = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & 1 \\ 1 & -1 \end{pmatrix} \notin X_{ss}$$

it follows that not all physical variables are well defined with finite variance. This is in agreement with the equations above that show x_3 with a white noise component. Using the condition (17) we get that

$$\left[\bar{E}^j (\lambda E + F)^{-1} J \right]/_{X_{ss}} X_{av} = \begin{cases} \frac{1}{2} \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 1 & -1 \end{pmatrix} & j = 0 \\ 0 & j > 0. \end{cases}$$

The condition is thus satisfied as soon as the last column of M is zero, showing that all linear combinations of v_1 and v_2 are well-defined with finite variance. Furthermore, condition (20b) gives that we can have any measurement matrix H and still get a well-defined Kalman filter problem. This is consistent with (25), since x_3 contains white noise that is allowed in the Kalman filter measurement but not any derivative of white noise (which would not be allowed).

V. FORMULATIONS WITHOUT CONTINUOUS-TIME WHITE NOISE

The stochastic state-space formulations naturally contain continuous-time "white noises", like w and e in (1). These are not mathematically solid objects, but the interpretations of the equations can be made rigorous, in terms of stochastic differential equations, like (18). See, e.g., [10]. Yet, it is natural to be ambivalent about these signals, since they are not physically realizable. We shall in this section give a formulation of a SDAE-filtering problem that only explicitly employs stochastic variables with finite variance. We shall then investigate if it corresponds to a mathematically well-posed problem.

A. Problem Formulation

Consider the SDAE

$$E\dot{z}(t) = Fz(t) + Gu(t) + \sum_{r=1}^{n_w} J_r w_r(t) \tag{26a}$$

$$z(t_0) = z_0 \tag{26b}$$

$$\dim z(t) = n \tag{26c}$$

where $w_r(t)$ is a Gaussian second order stationary process with spectrum

$$\Phi_{w_r}(\omega) \tag{26d}$$

which is rational in ω with pole excess $2p_r$. This means that

$$0 < \lim_{\omega \to \infty} \omega^{2p_r} \Phi_{w_r}(\omega) < \infty \tag{26e}$$

The input $u(t)$ is known for all $t \in [0, T]$. It will also be assumed that it is differentiable a sufficient number of times. The condition that the input is known for every t typically means that it is given at a finite number of sampling instants, and its intersample behavior between these is known. (like piecewise constant, piecewise linear, or band-limited).

An output vector is measured at sampling instants $y(t_k)$:

$$y(t_k) = Hz(t_k) + e(t_k), \quad k = 1, \ldots, N \quad (26\text{f})$$

where $e(t_k)$ is a Gaussian random vector with covariance matrix R_k, such that $e(t_k)$ and $e(t_s)$ are independent for $k \neq s$ and also independent of all the processes $w_r(t)$.

Definition 2: The problem to estimate $\bar{z}(t) = Mz(t)$ from $y(t_k), t_k < t; u(s), s \leq t$ is *well posed* if $Hz(t)$ and $Mz(t)$ have finite variances.

We shall find that a well posed filtering problem can be solved by the regular Kalman filter. As discussed previously, the initial value z_0 may not be chosen freely. The possibly conflicting values in z_0 will be ignored, and actually have no consequence for the computation of $z(t)$ for $t > t_0$.

Of course, one approach to dealing with this problem would be to realize the stochastic processes $w_r(t)$ as white noises through linear filters in state-space form, and then apply the results of Section III. Instead we shall here give a result that can be directly applied to (26).

B. Main Result

Theorem 3: Consider (26). Let λ be a scalar such that $(\lambda E + F)$ is invertible. Let

$$\bar{E} = (\lambda E + F)^{-1} E. \quad (27)$$

Suppose $\begin{bmatrix} H^T & M^T \end{bmatrix}$ has full rank. Then the estimation problem (26) is well posed if and only if

$$\bar{E}^{p_r}(\lambda E + F)^{-1} J_r \in X_{ss} \ \forall r. \quad (28\text{a})$$

For general H, M, the problem is well posed if and only if

$$\left[\bar{E}^j(\lambda E + F)^{-1} J_r \right]\big/_{X_{ss}} X_{av} \in \mathcal{N}\left(\begin{bmatrix} M \\ H \end{bmatrix} \right)$$
$$j = p_r, p_{r+1}, \ldots, k-1, \forall r. \quad (28\text{b})$$

Here k is the index of the SDAE, defined as in (5) and p_r is defined as in (26e).

Proof: The proof uses calculations similar to those of Theorems 1 and 2. It is possible to group H and M together since the condition is that neither $Hz(t)$ nor $Mz(t)$ contains any white noise component. (This differs from Theorem 2 where white noise is allowed in the measurement equation, following continuous-time Kalman filter theory.) The conditions in (28) need only apply to $\bar{E}^{p_r}$ and higher powers of $\bar{E}$, since $w_r(t)$ can be differentiated $p_r - 1$ times. ∎

Now consider the problem to estimate $\bar{z}(t)$ using the Kalman filter. First note that since the disturbances $w_r(t)$

have rational spectra, they can be written as outputs from linear filters driven by white noise,

$$\dot{z}_w(t) = A_w z_w(t) + B_w v(t) \quad (29\text{a})$$
$$w(t) = C_w z_w(t) + D_w v(t) \quad (29\text{b})$$

where

$$w(t) = \begin{pmatrix} w_1(t) & \cdots & w_{n_w}(t) \end{pmatrix}^T \quad (30)$$

and $v(t)$ is white noise. With

$$J = \begin{pmatrix} J_1 & \cdots & J_{n_w} \end{pmatrix}, \quad (31)$$

(26a) and (29) can be combined to give

$$\begin{pmatrix} E & 0 \\ 0 & I \end{pmatrix} \begin{pmatrix} \dot{z}(t) \\ \dot{z}_w(t) \end{pmatrix} = \begin{pmatrix} F & JC_w \\ 0 & A_w \end{pmatrix} \begin{pmatrix} z(t) \\ z_w(t) \end{pmatrix} + \begin{pmatrix} G \\ 0 \end{pmatrix} u(t) + \begin{pmatrix} JD_w \\ B_w \end{pmatrix} v(t) \quad (32)$$

Using the methods described in Section II, this SDAE can be transformed into

$$\begin{pmatrix} z(t) \\ z_w(t) \end{pmatrix} = \begin{pmatrix} T_1 & T_2 \end{pmatrix} \begin{pmatrix} x_s(t) \\ x_a(t) \end{pmatrix} \quad (33\text{a})$$

$$\dot{x}_s(t) = (E_s^{-1} - \lambda I)x_s(t) + E_s^{-1}G_s u(t) + E_s^{-1} J_s v(t) \quad (33\text{b})$$

$$\begin{pmatrix} M & 0 \end{pmatrix} T_2 x_a(t) = - \begin{pmatrix} M & 0 \end{pmatrix} T_2 \times \left(I + \cdots + \left(\frac{d}{dt} + \lambda \right)^{k-1} N^{k-1} \right) G_a u(t) \quad (33\text{c})$$

$$\begin{pmatrix} H & 0 \end{pmatrix} T_2 x_a(t) = - \begin{pmatrix} H & 0 \end{pmatrix} T_2 \times \left(I + \cdots + \left(\frac{d}{dt} + \lambda \right)^{k-1} N^{k-1} \right) G_a u(t) \quad (33\text{d})$$

provided that the estimation problem is well posed so that $Hz(t)$ and $Mz(t)$ do not contain white noise components. Together with the measurement equation

$$y(t_k) = Hz(t_k) + e(t_k) = \begin{pmatrix} H & 0 \end{pmatrix} \begin{pmatrix} T_1 & T_2 \end{pmatrix} \begin{pmatrix} x_s(t_k) \\ x_a(t_k) \end{pmatrix} + e(t_k) \quad (34)$$

this finally gives the state-space description

$$\dot{x}_s(t) = (E_s^{-1} - \lambda I)x_s(t) + E_s^{-1}G_s u(t) + E_s^{-1} J_s v(t) \quad (35\text{a})$$

$$y(t_k) = \begin{pmatrix} H & 0 \end{pmatrix} T_1 x_s(t_k) + \begin{pmatrix} H & 0 \end{pmatrix} T_2 \times \left(I + \cdots + \left(\frac{d}{dt} + \lambda \right)^{k-1} N^{k-1} \right) G_a u(t_k) + e(t_k). \quad (35\text{b})$$

This state-space description gives a filtering problem with continuous-time dynamics and discrete-time measurements. The Kalman filter for this setting provided, e.g., in [11] can be used to estimate x_s. The estimate of $\bar{z}$ is then computed from the estimate of x_s and the deterministic input using (33).

C. Integrating sampling

It may happen that a selected output z_k has infinite instantaneous variance. This happens when condition (28) is violated. This could be handled by integrating sampling. This is best illustrated by an example: Let the DAE be

$$\dot{z}_1(t) = -2z_1(t) + v_1(t) \tag{36a}$$

$$0 = -z_2(t) + v_1(t) \tag{36b}$$

$$y(t) = z_1(t) + z_2(t) + v_2(t) \tag{36c}$$

where $v_k(t)$ are continuous-time white noises. This is not a well posed problem since z_2 has infinite variance. A common way to handle this would be to apply integrated sampling to y (see e.g. [12]). This would give finite variances to the sampled values $y(t_k)$, and it would also introduce correlation between the "process noise" v_1 and the measurement noise.

A more convenient way of dealing with this in a modeling situation, would be to explicitly introduce a presampling, low pass filter, to make the measured variable

$$z_3(t) = \frac{1}{0.01p + 1}\big(z_1(t) + z_2(t)\big)$$

Including this new variable in the DAE gives

$$\dot{z}_1(t) = -2z_1(t) + v_1(t)$$

$$\dot{z}_3(t) = -100z_3(t) + 100z_1(t) + 100v_1(t)$$

$$0 = -z_2(t) + v_1(t)$$

with the sampled measurements

$$y(t_k) = z_3(t_k) + e(t_k)$$

This is a well posed problem with no correlation between $v_1(t)$ and $e(t_k)$.

VI. Conclusions

We have stated conditions for stochastic differential-algebraic equations and related filtering problems to be well defined. The conditions are formulated in terms of certain subspaces generated by the system matrices. Since these calculations involve rank tests, they are not straightforward on numeric data involving roundoff errors. The task of calculating the subspaces accurately using numerical algebra software thus contains several topics for future research.

References

[1] 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*, vol. 100, no. 1, pp. 77–92, Nov. 1998.

[2] K. E. Brenan, S. L. Campbell, and L. R. Petzold, *Numerical Solution of Initial-Value Problems in Differential-Algebraic Equations*, ser. Classics In Applied Mathematics.   Philadelphia: SIAM, 1996.

[3] R. Winkler, "Stochastic differential algebraic equations of index 1 and applications in circuit simulation," *Journal of Computational and Applied Mathematics*, vol. 163, no. 2, pp. 435–463, Feb. 2004.

[4] M. Darouach, M. Boutayeb, and M. Zasadzinski, "Kalman filtering for continuous descriptor systems," in *Proceedings of the American Control Conference*.   Albuquerque, New Mexico: AACC, June 1997, pp. 2108–2112.

[5] V. Kučera, "Stationary LQG control of singular systems," *IEEE Transactions on Automatic Control*, vol. AC-31, no. 1, pp. 31–39, Jan. 1986.

[6] A. Germani, C. Manes, and P. Palumbo, "Kalman-Bucy filtering for singular stochastic differential systems," in *Proceedings of the 15th IFAC World Congress*, Barcelona, Spain, July 2002.

[7] V. M. Becerra, P. D. Roberts, and G. W. Griffiths, "Applying the extended Kalman filter to systems described by nonlinear differential-algebraic equations," *Control Engineering Practice*, vol. 9, pp. 267–281, 2001.

[8] T. Schön, M. Gerdin, T. Glad, and F. Gustafsson, "A modeling and filtering framework for linear differential-algebraic equations," in *Proceedings of the 42nd Conference on Decision and Control*, Dec 2003, pp. 892–897.

[9] P. Van Overschee and B. De Moor, *Subspace Identification for Linear Systems*.   Boston, London, Dordrecht: Kluwer Academic Publishers, 1996.

[10] I. I. Gihman and A. V. Skorohod, *Controlled Stochastic Processes*. Berlin: Springer-Verlag, 1979.

[11] A. H. Jazwinski, *Stochastic Processes and Filtering Theory*.   Academic Press, 1970.

[12] K. J. Åström, *Introduction to Stochastic Control Theory*.   New York: Academic Press, 1970.