FINAL REPORT

Grant AFOSR-89-0241

January 15, 1989 - October 14, 1990

Estimation and Control of Nonlinear and Hybrid Systems
with Applications to Air-to-Air Guidance

by

A. H. Haddad

Department of EECS
Northwestern University
Evanston, IL 60208-3118

with subcontract to

E. I. Verriest

School of Electrical Engineering
Georgia Institute of Technology
Atlanta, GA 30332-0250

Prepared for:

Air Force Office of Scientific Research
Dr. Marc Jacobs/NM
Bolling AFB, Washington DC 20332
The research covers several aspects of the basic issues that are needed to develop and implement nonlinear filtering and control of maneuvering vehicles in uncertain environments and nonlinear geometry. The research is involved in modeling maneuvering nonlinear vehicles as switched linear Markov models. The research therefore leads in several directions investigating various aspects of such models which in general are called hybrid systems. Three different aspects are considered: The first involves realization and other generic properties of hybrid systems, including controllability and stability as well as simulation and averaging. The second involves estimation and detection systems for hybrid systems, including various related models and approximate filtering schemes. The third involves the application of switched Markov filtering schemes to the tracking of maneuvering vehicles.
SUMMARY

This is the final report of Grant AFOSR-89-0241 with the Air Force Office of Scientific Research, which is the continuation of Grant AFOSR-87-0308 to the Georgia Institute of Technology. The research was performed at Northwestern University with subcontract to Georgia Tech.

The research covers several aspects of the basic issues that are needed to develop and implement nonlinear filtering and control schemes for hybrid systems with applications to tracking, guidance, and control of maneuvering vehicles in uncertain environments and nonlinear geometry. The research is involved in modeling maneuvering nonlinear vehicles as switched linear Markov models. The research therefore leads in several directions investigating various aspects of such models which in general are called hybrid systems. Three different aspects are considered: The first involves realization and other generic properties of hybrid systems, including controllability and stability as well as simulation and averaging. The second involves estimation and detection systems for hybrid systems, including various related models and approximate filtering schemes. The third involves the application of switched Markov filtering schemes to the tracking of maneuvering vehicles.

The research culminated in the conclusion of two Ph.D. thesis by J. Ezzine and M. A. Ingram at Georgia Tech and one M.S. project by K. S. Lee at Northwestern University. It also supported the initial stages of three Ph.D. students, P. D. West, C. C. Tsai, and D. R. Shin who are in various stages of completing their Ph.D. dissertations.
SECTION I

INTRODUCTION

The objective of this research was to develop nonlinear filtering and tracking schemes for systems subject to complex geometries and uncertainties. These attributes characterize the air-to-air engagement scenario. The approach was based on the approximation of the original nonlinear stochastic model with a piecewise linear model. Then the resulting model was further approximated by a switched Markov linear model. The resulting model becomes then a typical representation of hybrid systems involving both continuous and discrete dynamics as shown:

\[ \dot{X}(t) = A[r(t)] X(t) + B[r(t)] U(t) \]  \hspace{1cm} (1a)

\[ Y(t) = C[r(t)] X(t) + V(t) \]  \hspace{1cm} (1b)

where the state vector is \( X(t) \), the observation vector is \( Y(t) \), \( U(t) \) can serve as either the control vector when considering a control problem, or as the process noise model representing the model uncertainties for the filtering problem, and \( V(t) \) is the observation noise vector. The noise processes are assumed to be white and Gaussian. The process \( r(t) \) is called either the form index or the macro-state process and is assumed to be a finite state Markov process taking the values in the set \( \{1, 2, \ldots, N\} \). The resulting system is known as either a switched Markov linear model or a hybrid system model since the state \( X(t) \) is continuous and the vector \( r(t) \) is discrete. The linear system switches among the finite number of realizations (\( A[i], B[i], C[i] \)) depending on the value of \( r(t) \), and the switching follows a Markov chain rule.

The research reported under the earlier grants covered both the analysis of the switched Markov approximation to the modeling of nonlinear systems as well as realization and characterization results on hybrid systems. These reports also discussed filtering schemes for such systems and similar models that involve the dependence of the Markov process parameters on the system state.

This report addresses continuation of these efforts and resulted in the conclusion of two Ph.D. theses at Georgia Tech. The two completed theses covered two different aspects of the mixed models that include both discrete and continuous variables. The first is by Jelel Ezzine (Reference 1) considered the properties of hybrid systems involving both discrete and continuous states which in our case reflects the switched linear Markov models used to represent the maneuvering vehicles to be tracked and/or controlled. The thesis studied the stability and controllability properties of such systems and derived conditions under which the systems can be approximated by their statistical average system. The second thesis is by Mary Ann Ingram (Reference 2) considered an alternative model for maneuvering vehicles and derived approximate filtering schemes for such models that involve linear systems driven by impulsive inputs whose rates depend
on the state of the system. Since exact filtering representations are not realizable, conditions for the convergence of several detection-estimation schemes were obtained and their result validated via simulations.

This report concentrates on extensions of the results to three areas. The first area involves the simulation and analysis of the multi-model approximate filtering scheme that has been tested earlier using limited memory only and its extension to three-dimensional tracking filter for a maneuvering target and is reported in Section II. The second area, covered in Section III, addresses the analysis and control of hybrid systems when the both the state dynamics and the form index exhibit fast and slow modes of behavior. The third area considers additional work in the realization of hybrid systems and is discussed in Section IV. The body of each section is relatively short, as the results are provided in appropriate appendices.
SECTION II
FILTERING SCHEMES FOR HYBRID SYSTEMS

Several models and approximations have been considered for the filtering schemes for hybrid systems and their applications. Ingram’s thesis (Reference 2) considered a continuous state model with Markov chain input whose transition matrix depends on the state of the systems. Exact filtering schemes cannot be derived analytically or implemented numerically. A new prior penalty approach to the filtering for such systems has been proposed and analyzed by Ingram in Reference 3 and is shown in Appendix A. The resulting filter is superior to the suboptimal linear smoother when the rates of change of the Markov process are very low and when the impulsive input jumps do not take very small values.

In addition to the research reported in Ingram’s thesis, additional approximation to the nonlinear filtering structure reported in Reference 4 has been proposed. In particular, the effort has been centered at reducing the memory requirement of the multi-model filter as well as providing a more realistic simulation scenario. In particular, an extension of the memory of the filter to 4 steps has been shown to provide a substantial improvements over the single step memory filter, as shown in Reference 5 and attached as Appendix B. Furthermore, it has been shown that increasing the filter dimension does not result in reduced performance as discussed in Reference 6 and attached in Appendix C. This latest work indicates that the filter is applicable to a three dimensional tracking problem, and provides and alternative approach to the modeling of the maneuver acceleration. Analysis methods to indicate the asymptotic convergence of the filter and its performance are encouraging.

Finally Reference 7 discusses a general framework for the filtering and smoothing for systems with both discrete and continuous observation models. The results are primarily analytical in nature and the implementation issues have not been resolved as yet. The representation is shown in Appendix D.
SECTION III
ANALYSIS OF HYBRID MODELS

Two avenues of research have been followed in the simulation and analysis of hybrid systems models. The first established analytical and simulation tools for the study of how well such models can be used to approximate piece-wise linear dynamic systems. Earlier results simply addressed the first order densities of such models. In this study the autocorrelation function of both the model and the original system have been simulated and compared to verify the conditions (earlier only studied in theory) for the validity of the approximation. Furthermore, an analytical approach has been developed for the analysis of the steady-state stationary probability density of the system model and its comparison to the approximating hybrid model. The results are documented in an M.S. project by K-S. Lee shown in Reference 8.

The second continued the research into hybrid systems models that involve both fast and slow dynamics. The fast and slow dynamics are involved in both the systems models and in the Markov chain that determines the transition among the various realizations. Earlier work (Reference 9) was concerned with the limiting behavior of such systems when the Markov chain was either fast or slow. More recently, the results have been extended to the case where the Markov chain can be decomposed into groups of fast transitions. Furthermore, asymptotic results for the convergence of the reduced-order models have been derived for a variety of cases of fast and slow behavior in the continuous system model and in the underlying Markov chain. The results are given in Reference 10, and are shown in Appendix E. One restriction to the resulting approximation is that system matrices of the realizations involved in each group of fast transitions have to commute. More recently, this restriction has been successfully removed as shown in Reference 11 and attached in Appendix F. However, the results still require the stability of each group of realizations. The research also provide complete analysis of the multiple-time scale approximation for such systems for both the slow and the fast dynamics of the system. The relative ratio of the time-scale of the Markov chain transition matrix to those of the continuous states is crucial to the type of the resulting approximation.

Finally, the conditions for the control and stabilization of hybrid systems using the average model constant gain controllers or switched gains controller that may depend on the correct detection of the macro-state have also been derived. Furthermore, if we assume that it is not possible to correctly identify the macro-state (the value of \( r(t) \)) of the system, conditions on the probability of detection errors have been found that will make such a controller feasible. The results are given in Reference 12 and Appendix G.

Overall, these results make it simpler to implement lower order controllers or less complex controllers for a variety of hybrid systems that either exhibit fast and slow dynamic responses or satisfy conditions that allow their robust control.
SECTION IV

REALIZATION AND CONTROL

The section addresses several issues in the realization models for hybrid systems. These models can lead to a more systematic approach to the identification and control of these systems. Canonical forms for the periodic hybrid systems have been developed in Reference 13 and shown in Appendix H. The sensitivity of various realizations of hybrid systems have been developed in Reference 14 and are shown in Appendix I. The sensitivity is crucial to the efficiency of any identification or control schemes that needs to be used in conjunction with specific realization. A special case of hybrid systems that have linear relations among its continuous states can be represented as singular hybrid system. These systems may also be considered as a limiting case of singularly perturbed systems discussed in Section III. References 15 and 16 discuss general approaches to the realization problem of such systems that have implication on their control. The results are shown in Appendix J and Appendix K. Finally, for randomly changing hybrid systems and their underlying Markov chains a novel representation for the system is given in Reference 17 and Appendix L. Similarly, a novel realization theory has been proposed in Reference 18 and shown in Appendix M for the realization of Markov chains that are crucial to the analysis and control of hybrid systems.
SECTION V

SUMMARY AND CONCLUSIONS

The research summarized in this report and supported by the Air Force Grant provides the basis for the design of estimators controllers for systems subject to random fluctuations in their models and their environments. The controllers and estimators are not optimal as it is not possible to implement and analytically derive an implementable form. Hence, approximation methods have been studied for the derivation of implementable control scheme and filtering schemes for such systems. Approximations using slow and fast dynamics separation and reduced-order modeling have been proposed for such systems to simplify the control and estimation implementation. Finally, applications to the tracking of maneuvering vehicles have been proposed, the resulting approximate filter derived and simulated for several one-dimensional and three dimensional problems.
REFERENCES


APPENDIX A

M. A. Ingram and A. H. Haddad

Smoothing for Linear Systems Excited by Point Processes
with State-Dependent Rates

Proc. 13th IMACS World Congress on Computation and Applied Mathematics

SMOOTHING FOR LINEAR SYSTEMS EXCITED BY POINT PROCESSES WITH STATE-DEPENDENT RATES

Mary Ann Ingram  Abraham H. Haddad
School of Electrical Engineering  Department of Electrical Engineering
Georgia Institute of Technology  and Computer Science
Atlanta, Georgia 30332-0250  Northwestern University
Evanston, Illinois 60208-3118

Abstract

Smoothing for a linear system driven by a point process with a rate that depends on the state of the system is considered. The observation model is the integrated version of a linear combination of the states in additive white Gaussian noise. A smoother that uses estimation and detection is compared with the optimal linear smoother and filter. The comparison is in terms of the mean squared error (MSE) of the state. The false alarm rate of the detector is shown to depend strongly on the region of support of the mark distribution. When false alarms are low, the estimation/detection scheme has lower MSE than the optimal linear smoother.

I. Introduction

We consider the state estimation problem for the following single input system:

\[ \dot{x}_t = Ax_t + B M_t, \quad t \geq 0 \]

with the scalar observation process

\[ z_t = \int_0^t C x_s ds + v_t \]

where \( v_t \) is a Wiener process with \( E\{v_t v^*_t\} = \int_0^t \sigma_v^2 \, ds \). The \( n \times n \) matrix \( A \) such that the solution to \( \dot{x} = Ax \) is exponentially stable. The scalar process \( M_t \) is a random jump process with jump heights, or marks, that are independent and identically distributed and with jump times that occur with an instantaneous average rate \( \mu(x_t) \). Thus the rate of jumps depends on the system state. An example that motivates this model is a maneuvering vehicle where a jump represents an abrupt change in acceleration. The likelihood of acceleration commands can depend on the position and velocity of the vehicle. Another example is an electromechanical system where the jumps represent failures with a likelihood of occurrence that increases under conditions of excessive heat or current. We are interested in cases where the rate of jumps is low compared to the bandwidth of the system; under this condition, the state is not well approximated by a Gaussian distribution.

The process \( x_t \) is easily shown to be in the class of "piecewise-deterministic Markov processes," defined by Davis [1]. Filtering and smoothing for systems driven by Poisson processes have been considered by Kwakernaak [2, 3] and Au [4], and for a related process by Blom [5].

II. The Prior Penalty Detector

This scheme uses observations over an interval to detect the number of jumps within the interval and estimate the times and marks of the jumps. The state estimate is constructed by superimposing the system responses to the
detected jumps. In order to reduce computational complexity and memory requirements, new observations are used to detect new jumps and update only recently detected jumps. Specifically, the observations over the interval \([A, A + T]\), denoted by \(Z_{A,A+T}\), are used to detect the number \(N_{A,A+T}\) of jumps in the interval and to estimate the vector of jump times \(\tau_{N_{A,A+T}}\) and marks \(y_{N_{A,A+T}}\) of the jumps. Thus fixed interval smoothing is performed on the observations in \([A, A + T]\). Then the interval is moved forward to \([A + \Delta, A + \Delta + T]\), and fixed-interval smoothing is performed over the new interval. A detected jump that is left behind by the moving interval is called a “finalized detection.” Here, \(\Delta\) is small enough such that \(\Pr\{N_{A,A+\Delta} > 1\} \ll 1\). The system responses to the finalized detections are superimposed to construct an estimate of the state with a fixed lag. For the sake of notational simplicity, the following expressions assume that the estimation and detection is performed on the interval \([0, T]\), and that the initial state \(z_0\) is known. In a sequential implementation, the interval is changed to \([A, A + T]\) and \(z_0\) is replaced by the smoothed estimate, \(\hat{z}_A\).

It is noted that the maximum a posteriori (MAP) estimate of \(N_{0,T}\) can be expressed as

\[
L_{\text{MAP}}(n|Z_{0,T}, z_0) = \begin{cases} 
\Pr\{N_{0,T} = 0|z_0\}, & n = 0 \\
\mathbb{E}_{\tau_n, u_n}(A\{Z_{0,T}|\tau_n, u_n, N_{0,T} = n, z_0\}) \times \Pr\{N_{0,T} = n|z_0\}, & n > 0 
\end{cases}
\]

(1)

where \(A\{Y_{0,T}|\tau_n, u_n, N_{0,T} = n, z_0\}\) is the likelihood functional. The detector in the present scheme replaces the averaged likelihood functional in (1) with the likelihood functional evaluated at the MAP estimates of \(\tau_n\) and \(u_n\), given that \(N_{0,T} = n\). Therefore the decision variable is

\[
L(n|Z_{0,T}, z_0) = \begin{cases} 
\Pr\{N_{0,T} = 0|z_0\}, & n = 0 \\
A\{Z_{0,T}|\hat{\tau}_n, \hat{u}_n, N_{0,T} = n, z_0\}\Pr\{N_{0,T} = n|z_0\}, & n > 0
\end{cases}
\]

(2)

We call this scheme the Prior Penalty Detector (PPD) because the a priori probability \(\Pr\{N_{0,T} = n|z_0\}\) serves as a penalty for overfitting and can be computed offline for the desired range of values for \(z_0\).

III. Simulation Results

Four examples are used to compare the performances of the optimal linear filter, the PPD, and the optimal linear smoother with the same lag. The performance is measured in terms of the mean squared error (MSE), normalized by state variance, and the average number of false detections per true pulse as a function of noise variance. The MSE for the optimal filter and smoother is computed using the Bode-Shannon method [6]. The MSE for the PPD is found by time averaging the computer-simulated output.
All examples have the same scalar system model of $dx_t = -5x_t dt + dM_t$. The examples differ in the rate function $\mu(z)$ and the mark pdf $p_M(u)$, as shown in Figure 1. This type of rate function was chosen to yield "bursty" behavior in the sample trajectories. If the detector succeeds in detecting the first few pulses that move the state into a high rate region, then the detector changes its characteristics to allow more detections. The pdf's were chosen to illustrate the effect the mark pdf has on the number of false detections.

The MSE results are shown in Figures 2 through 5 for Examples 1 through 4, respectively. The false alarm rates are shown in Figure 6. We observe that for Example 1, the PPD has a lower MSE than the optimal filter and smoother, and has very few false alarms. This is because the region of support of the mark pdf is confined to the positive axis and does not permit arbitrarily small pulses. In Example 2, the MSE of the PPD is only slightly lower than that of the linear smoother. The degradation in PPD performance relative to Example 1 is due to the increased number of false alarms with small marks. As the noise variance increases, the PPD makes about the same number of false alarms, but with larger marks. In Example 3, there is a rather dramatic correlation between MSE and the false alarm rate, as both increase with the noise variance. The mark pdf for this example allows the false alarms to have large positive and negative marks that nearly cancel. However, time quantization in the simulation does not allow such overlapping false alarms to approach perfect cancellation as the noise variance decreases. The mark pdf for Example 4 was selected to give the worst case performance of the PPD because it allows arbitrarily small false alarm marks as well as large false alarms that nearly cancel. Again the false alarm rate is independent of the noise variance, but the rate is larger than for Example 2 because there is no penalty for arbitrarily small marks. Also the optimal linear smoother is consistently better in terms of MSE.

IV. Conclusions

The simulation results indicate that if the PPD false alarm rate remains below 3 per true pulse, the PPD yields an MSE lower than the optimal linear smoother with the same lag. It is noted that these results are somewhat biased in favor of the linear estimators. One reason is that the MSE for the linear smoother and filter are evaluated using a formula that assumes an infinite observation interval, rather than an interval of length equal to the lag as in the PPD. Another reason is that although the ratios of (jump rate times mean squared mark value)-to-(system bandwidth times noise variance) for our examples were useful for studying false alarm behavior, they guaranteed good performance for the linear estimators. To see poorer performance by the linear estimators and better performance by the PPD, one should reduce this ratio.
Figure 1. The rate function and mark pdf for the four examples.

Figure 2. Normalized mean squared error for Example 1.

Figure 3. Normalized mean squared error for Example 2.

Figure 4. Normalized mean squared error for Example 3.

Figure 5. Normalized mean squared error for Example 4.

Figure 6. The number of false alarms per true pulse for the four examples using the PPD.

References


APPENDIX B

P. D. West and A. H. Haddad

Approximate Switched-Markov Filtering for Nonlinear Systems

Proc. 1990 American Control Conference

APPROXIMATE SWITCHED-MARKOV FILTERING FOR NONLINEAR SYSTEMS

P.D. West
Georgia Tech Research Institute
Georgia Institute of Technology
Atlanta, GA 30332

A.H. Haddad
Department of Electrical Engineering
and Computer Science
Northwestern University
Evanston, IL 60208-3118

ABSTRACT

The Kalman filter provides optimal state estimates for completely known linear systems. Unfortunately, many physical systems are neither exactly known, nor linear. Numerous filtering schemes for nonlinear systems have been introduced over the years; general theories for nonlinear systems tend to be complex, and, due to their generality, are of little practical use to the design engineer. On the other hand, solutions for specific nonlinearities usually apply only to a single nonlinearity, and thus are limited in their applications. This paper, however, presents a methodology whereby the nonlinearity is first approximated by a piecewise linear model, and then a common filtering scheme is applied. The efficacy of this approach is that the same filtering algorithm may be applied to a broad class of nonlinear stochastic systems.

I INTRODUCTION

Specifically, the problem at hand assumes that, given nonlinear observations \(y(k)\), it is desired to estimate the state \(x(k)\) of the system

\[ x_{k+1} = f(x_k) + w_k \]

(1)

where \(x_k\) is the \(n\) vector representing the system state at time \(k\), \(w_k\) is a white, discrete time \(d\)-dimensional (Gaussian) random process with covariance matrix \(Q\), and \(f\) is an \(n \times d\) dimensional matrix. The observation model is assumed to be given by

\[ y_k = h(x_k) + v_k \]

(2)

where \(y_k\) is an \(m\)-dimensional vector which represents the observation at time \(k\), and \(v_k\) is an \(m\)-dimensional white Gaussian measurement noise process with covariance \(R\). The state propagation function \(f(\cdot)\) and the observation function \(h(\cdot)\) are non-linear.

The nonlinearities in (1) and (2) are approximated by the continuous piecewise linear approximation given by the following model:

\[ g(x) = Gx + g, \quad \text{for} \quad x \in \Omega, \quad i = 1, \ldots, M \]

(3)

and

\[ h(x) = Hx + h, \quad \text{for} \quad x \in \Omega, \quad i = 1, \ldots, M \]

(4)

where \(\Omega_i\) and \(\Omega_j\) partition the state and measurement spaces. For simplicity of notation, the cross product of the two partitions may be formed to yield one partition, \(\Omega = \{\Omega_i, \Omega_j\} \), \(i = 1, \ldots, M; \quad M \leq M \times M\). At each time step \(k\), the system is assumed to be governed by one of the \(M\) models. These \(M\) system models are called system macro-states.

The final model assumption is that the system jumps from macro-state \(i\) to macro-state \(j\) according to a finite state Markov process, \(J\). In order to maintain a "memory" of the last \(r\) time steps, a parameter \(J(k)\) is introduced, where \(J(k)\) represents the set of the \(M\) macro-states, i.e.,

\[ J(k) = \{J_1, \ldots, J_r, J_{r+1}\} \]

(5)

where

\[ J_i \in \{1, 2, \ldots, M\} \]

(6)

Conceptually, the new filter is based on two assumptions for the system model: 1) The nonlinearities may be approximated by continuous piecewise linear functions and 2) This \(N\)-segment piecewise linear model may be approximated by \(N\) separate affine systems driven by the same process, with the true system output being approximated by randomly selecting one of the \(N\) outputs.

Conditions supporting the validity of these assumptions are derived in [1]. The second assumption allows application of the theory of switching systems. The optimal (albeit non-realizable) filter for switching systems was introduced by Ackerson and Fu in 1970 [2], and consists of a likelihood-weighted sum of Kalman filters "tuned" to each possible switching sequence, and, hence, involves exponentially increasing complexity with time. Their paper did not consider the additional structure present in the hybrid-state model where the dynamics of the system macro-state are independent of the system state, but where the system state is not independent of the macro-state. The optimal filter for this hybrid-state model was presented by Bruneau and Tenney in [3]. It is shown that this filter is also infinite dimensional and non-realizable. Numerous schemes have been introduced to reduce the filter complexity by causing away unlikely trajectories or combining similar estimates (see e.g. [4] [5]).

A primary difference between this work and that of [2] is that here, in the underlying piecewise linear model, the system macro-state is a function of the system state. This fact is exploited in the filtering algorithm through the consistency update stage. A consistency update occurs when the state estimate for the \(i^*\) model, \(z_i\), is compared with the domain of the \(r^*\) line segment (in the scalar case). If the state estimate produced by a given filter is not within the domain of the \(r^*\) line segment, the state estimate is said to be inconsistent with its macro-state, and less weight is placed on that estimate.

II FILTERING SCHEME

The filtering scheme applied here consists of maintaining one Kalman filter "tuned" to each of the \(M\) macro-state trajectories. Thus, for each new observation, an entire set of \(M\) innovators will be formed — one for each filter. Next, the usual measurement and time updates will be performed for each filter, producing \(M\) covariance matrices and individual state estimates. Next, the overall combined estimate, \(\hat{z}\), is formed from the likelihood weighted sum of these \(M\) individual estimates. Finally, the filters are aggregated, and the conditional probabilities and likelihood functions are modified according to the consistency update stage. A detailed summary of these steps is provided here. Additional details of the algorithm may be found in [5].

Before describing the individual filtering steps, some additional definitions are required. The Markov transition matrix, \(\Pi\), specifying the transition probabilities from macro-state \(S_i\) to \(S_j\), is obtained from:

\[ \Pi_{ij} = \text{Pr} \{x_{k+1} \in S_j \mid x_k \in S_i\} \]

(7)

The marginal steady-state probabilities \(p_i\), of macro-state \(S_i\), are defined by the solution to:

\[ p_i = p \Pi \]

(8)

where \(p\) is a row vector with components \(p_i\). The \(a\ posteriori\) probability that the system is in macro-state \(i\) at time \(k\) may be expressed as

\[ p_i(k) = \sum_{j=1}^{r} \lambda_j R_j(k) \]

(9)

where \(J(k)\) denotes all \(M^{r+1}\) sequences at time \(k\) which end in macro-state \(i\).

Consistency Update

If the variance of any individual estimate, \(P_{S_i}\), is small then the information provided by \(p_i\) may be neglected. In this case, these values are changed based on the position of the estimate \(z_i\) in the appropriate region \(\Omega_j\), and used to update the a posteriori macro state probabilities \(p_i(J(k-1))\). In turn, these are used in the next stage for updating \(p_i(k)\).
The covariance is updated using the Gaussian sum approximation,
\[ \beta(k | J(k - 1)) = \alpha(P_{r_{j-1}}) \beta(k | J(k - 1)) + \{1 - \alpha(P_{r_{j-1}})\} U_j \{x_{r_{j-1}}(k)\} \]
(10)

Here, \( \alpha(P) \) is a function of the norm of \( P \) which tends to zero as \( P \) becomes small, and which tends to unity as \( P \) becomes large. The operation \( U_j(x) \) is an indicator function that is equal to unity if \( x \in \Omega_j \), and is zero otherwise.

**Time Update**

The macro state probabilities are updated by using the consistency updated values \( \beta \), together with the transition probabilities,
\[ R_{1a_r_{j+1}}(k + 1 | k) = \beta(k | J(k - 1)) \lambda_{1a_r_{j+1}}(k) \Pi_j \]
(11)

Time updates of the individual state and covariance estimates are achieved via the standard Kalman filter equations for the appropriate models.

**Measurement Update**

As above, the individual state estimate, the innovations, and the covariance may be calculated using the Kalman filter for the appropriate model under consideration. The question now is concerned with the measurement update of the macro state probability estimates. This can be accomplished by using the standard likelihood function for a switched-Markov model, which, it should be noted, is only an approximation in this case. The expression for the \( a \) posteriori probabilities in this case will be proportional to the likelihood functions \( \lambda_{1a_r_{j+1}}(k) \). The update equation is
\[ \lambda_{1a_r_{j+1}}(k) = \beta \lambda_{1a_r_{j+1}}(k) \times \exp \{- \frac{1}{2} R_{1a_r_{j+1}}(k + 1 | k) x_{1a_r_{j+1}}(k + 1) \} \]
(12)

where \( \beta \) is a normalization coefficient, the \( R_{1a_r_{j+1}}(k + 1 | k) \) are the innovations processes arising from the Kalman filter tuned to the \( J(k + 1) \) model, and \( \Lambda \) represents the consistency updated likelihood value.

**Combined Estimate**

The combined estimate \( \hat{x}(k) \) is obtained by using the likelihood-weighted sum of the individual estimates, i.e.
\[ \hat{x}(k) = \sum_{i=0}^{m} \Lambda_{i}(k) \hat{x}_{i}(k) \]
(13)

**Aggregation**

To avoid expanding memory, it is necessary to reduce the number of filters at each time step. This may be achieved in a number of ways including casting away unlike sequences, merging similar sequences, or, the approach taken here, systematically aggregating at the earliest time.

The technique developed for reducing the number of filters required to \( M \) is, as follows: consider the collection of macro-state sequences \( J(k) \) to be the sequences (of length \( n \) at time \( k \) that began in macro-state \( i \) and progressed to macro-state \( j \) at the next time step. Similarly, the notation \( J(k) \) indicates the set of all sequences that began in macro-state \( k \). The aggregation step involves forming likelihood weighted sums over the index \( i \) for each \( j \), thus reducing the filter memory by one.

\[ \hat{x}_{i}(k) = \sum_{i=0}^{m} \Lambda_{i}(k) \hat{x}_{i}(k) \]
(14)

The covariance is updated using the Gaussian sum approximation,
\[ P_{i}(k) = \sum_{i=0}^{m} \Lambda_{i}(k) P_{i}(k) + \hat{x}_{i}(k) \hat{x}_{i}(k) - \hat{x}_{i}(k) \hat{x}_{i}(k)^{T} \]
(15)

and
\[ \Lambda_{i}(k) = \sum_{i=0}^{m} \Lambda_{i}(k) \]
(16)

**III ANALYSIS**

Since the filter is complex and nonlinear, it has yet to succumb to any closed-form performance analysis techniques. Hence, Monte-Carlo simulation techniques were used to assess its performance. In the simulation, a scalar version of the proposed filter with memory, \( r \), of length one (PF1) and four (PF4) is compared to a standard Extended Kalman Filter (EKF). Both the system function (1) as well as the measurement function (2) are defined by 3 segment affine maps \( g(x) = Gx + k \) and \( h(x) = Hx + \beta \), where:
\[ G = \begin{bmatrix} 2.0, & |x| < 1 \\ -0.2, & |x| \geq 1 \end{bmatrix}, \quad H = \begin{bmatrix} 5.0, & |x| < 0.5 \\ -0.1, & |x| \geq 0.5 \end{bmatrix} \]
(17)

\[ k_{1} = \begin{cases} 2.5 \text{sgn}(x), & |x| > 1 \\ 0, & |x| \leq 1 \end{cases}, \quad k_{2} = \begin{cases} 2.5 \text{sgn}(x), & |x| > 0.5 \\ 0, & |x| \leq 1 \end{cases} \]
(18)

Both \( Q \) and \( R \) were set to unity and \( \beta \) was varied from 1 to 10. Figure 1. is a graph which depicts the relative error variance as a function of \( \beta \) parameterized by the filter type (PF1, PF4, or the EKF). As can be seen, the improvements between either of the proposed filters and the EKF is striking. Additional simulations (not presented here) indicate similar trends, with the best performance increases being seen for non-injective nonlinearities.

**IV SUMMARY AND CONCLUSIONS**

This paper presents a new sub-optimal filter to be used for the nonlinear estimation problem in systems with piecewise linear models in both the system and observation equations. The approximations used are based on utilizing the switched-Markov model for the system as well as on modifying the resulting filter with the physical constraints of the states of the model. Not all facets of the filter are in final form, and work remains in the area of the exact formulation of the consistency update, as well as filter aggregation. Nonetheless, preliminary results show that the filter may work well in a broad class of nonlinear filtering problems.

**REFERENCES**

APPENDIX C

P. D. West and A. H. Haddad

Switched Markov Filtering for Tracking Maneuvering Targets

Proc. 1991 American Control Conference

SWITCHED-MARKOV FILTERING FOR TRACKING MANEUVERING TARGETS

P.D. West  
Georgia Tech Research Institute  
Georgia Institute of Technology  
Atlanta, GA 30332

A.H. Haddad  
Department of Electrical Engineering and Computer Science  
Northwestern University  
Evanston, IL 60208-3118

ABSTRACT

A new filtering concept is presented for tracking maneuvering targets. A conventional Markov switching process is used to model the target maneuver process, but a new filtering scheme is employed. The filter uses a traditional track-splitting approach, with one Kalman filter tuned to each branch of the tree. To limit filter complexity, aggregation is performed over the earliest timestep of an arbitrary filter memory length. Before aggregation, a unique consistency update stage is produced combining similar estimates (see Fig. 1). The modified filter is tuned to each branch of the tree. To limit filter complexity, aggregation is produced for each filter. Next, the usual measurement and time updates will be performed for each filter, producing $M$ covariance matrices and individual state estimates. Finally, the overall combined estimate, $x^c$, is formed from the likelihood weighted sum of these $M$ individual estimates. Finally, the filters are aggregated, and the conditional probabilities and likelihood functions are modified according to the consistency update stage. A detailed summary of these steps is provided here. Additional details of the algorithm may be found in [4].

Before describing the individual filtering steps, some additional definitions are required. The Markov transition matrix, $P^M$, specifying the transition probabilities from maneuver state $S_i$ to $S_j$ is obtained from:

$$
\Pi_{i,j} = P^M \{ u_{i+1} \in \Omega_j | u_i \in \Omega_i \}
$$

The a posteriori probability that the system is in maneuver state $i$ at time $k$ may be expressed as

$$
\Pi_i(k) = \sum_{j=1}^M \Pi_{i,j} a_{i,j}(k)
$$

where $\Pi_i(k)$ denotes all $M^{-1}$ sequences at time $k$ which end in macro-state $i$.

Target Model

The explicit state equations, assuming a 1 second update rate, are as follows:

$$
\begin{pmatrix}
\dot{p} \\
\dot{v} \\
\dot{a}
\end{pmatrix}
= \begin{pmatrix}
\frac{1}{2} & 0 & 0 \\
0 & 1 & 0 \\
0 & 0 & 1
\end{pmatrix}
\begin{pmatrix}
p \\
v \\
a
\end{pmatrix}
+ \begin{pmatrix}
\frac{1}{2} & 0 & 0 \\
0 & 1 & 0 \\
0 & 0 & 1
\end{pmatrix}
\begin{pmatrix}
w_1 \\
w_2 \\
w_3
\end{pmatrix}
+ \begin{pmatrix}
0 \\
0 \\
0
\end{pmatrix}
\begin{pmatrix}
u_1 \\
u_2 \\
u_3
\end{pmatrix}
$$

where $p$, $v$, and $a$ represent the target's position, velocity, and acceleration, respectively. Although numerous coordinate systems exist for target tracking, polar-spherical (measurement) coordinates have been selected for this analysis. Using well-known [5] approximations, this formulation leads to three loosely coupled state equations—one each for range, elevation, and bearing. This leaves us with three $3 \times 3$ systems rather than $9 \times 9$. The impact of this is more important, though, when the complexity introduced through multi-model approach is considered. If $N$ different maneuver commands are modeled for each axis, then there are $N^3$ possible systems for the coupled filter, and only $3N$ for the decoupled case. When a memory of $M$ timesteps is admitted then we have $M^3$ possible systems for the coupled case and only $3M$ for the decoupled case. If $S$ maneuver commands, and a memory of 3 timesteps are considered, the decoupled filter requires less than two percent of the complexity of the fully coupled system. Thus, for the same computational complexity, many more maneuver commands could be added to the decoupled filter.

In summary, then, data are measured in the spherical coordinates, range, elevation, and bearing. Independent filtering is performed in each coordinate dimension by the new filtering scheme. A single estimate is produced in each of the three dimensions, at each timestep.
Consistency Update

If the variance of any individual estimate, $P_{i ightarrow j}$, is small then the information provided by $\beta$ may be neglected. In this case, these values are changed based on the position of the estimate $\xi_{i ightarrow j}(k)$ in the appropriate region $\Omega_i$, and used to update the \textit{a posteriori} macro state probabilities $\beta_i(k | J(k - 1))$. In turn, these are used in the next stage for updating $\beta_i(k + 1 | k)$. If, on the other hand, the individual estimate covariance is large, the macro state information is weighted more heavily in determining the macro state probabilities. In this work, this updating stage was achieved through the following equation:

$$
\beta_i(k | J(k - 1)) = \alpha_i(P_{i ightarrow j}) \beta_i(k | J(k - 1)) + \\
(1 - \alpha_i(P_{i ightarrow j})) U_i(\xi_{i ightarrow j}(k))
$$

Here, $\alpha_i(P)$ is a function of the norm of $P$ which tends to zero as $P$ becomes small, and which tends to unity as $P$ becomes large. The operator $U_i(x)$ is an indicator function that is equal to unity if $x \in \Omega_i$, and is zero otherwise.

Time Update

The macro state probabilities are updated by using the consistency updated values $\beta_i$ together with the transition probabilities,

$$
\bar{\lambda}_{i ightarrow j}(k + 1 | k) = \beta_i(k | J(k - 1)) \lambda_{i ightarrow j}(k) I_i
$$

Time updates of the individual state and covariance estimates are achieved via the standard Kalman filter equations for the appropriate models.

Measurement Update

As above, the individual state estimate, the innovations, and the covariance may be calculated using the Kalman filter for the appropriate model under consideration. The question now is concerned with the measurement update of the macro state probability estimates. This can be accomplished by using the standard likelihood function for a switched-Markov model. The expression for the \textit{a posteriori} probabilities in this case will be proportional to the likelihood functions $\lambda_{i ightarrow j}(k)$. The update equation is

$$
\lambda_{i ightarrow j}(k + 1) = \beta \bar{\lambda}_{i ightarrow j}(k + 1 | k) \times \\
\exp \left[ -\frac{1}{2} r_{i ightarrow j}(k + 1) \right]
$$

where $\beta$ is a normalization coefficient, the $\lambda_{i ightarrow j}$ are the innovations processes arising from the Kalman filter tuned to the $J(k + 1)$ model, and $\bar{\lambda}$ represents the consistency updated likelihood value.

Combined Estimate

The combined estimate $\hat{x}(k)$ is obtained by using the likelihood-weighted sum of the individual estimates, i.e.

$$
\hat{x}(k) = \sum_{i=0}^{M} \lambda_{i ightarrow j}(k) \hat{x}_{i ightarrow j}(k)
$$

Aggregation

To avoid expanding memory, it is necessary to reduce the number of filters at each time step. This may be achieved in a number of ways including casting away unlikely sequences, merging similar sequences, or, the approach taken here, systematically aggregating at the earliest time.

The technique developed for reducing the number of required filters to $M'$ is as follows: consider the collection of macro-state sequences $J_i(k)$ to be the sequences of length $r$ at time $k$, that began in macro-state $i$, and progressed to macro-state $j$ at the next time step. Similarly, the notation $J_i(k)$ indicates the set of all sequences that began in maneuver state $k$. The aggregation step involves forming likelihood weighted sums over the index $i$ for each $j$, thus reducing the filter memory by one.

$$
\hat{x}_{j}(k) = \sum_{i=1}^{M} \Lambda_{i ightarrow j} \hat{x}_{i}(k)/\Lambda_{i ightarrow j}
$$

The covariance is updated using the Gaussian sum approximation, i.e.

$$
P_{j}(k) = \sum_{i=1}^{M} \Lambda_{i ightarrow j} P_{i}(k) + \hat{x}_{j}(k) \hat{x}_{j}(k)^T - \hat{x}_{j}(k) \hat{x}_{j}(k)
$$

and

$$
\Lambda_{j}(k) = \sum_{i=1}^{M} \Lambda_{i ightarrow j}
$$

III RESULTS

The filter was implemented on a digital computer using 5 maneuver commands per axis and 3 times steps of memory. Further, a simple 3-state, 3-axis Kalman filter was implemented as a first-cut benchmark. The Kalman filter noise covariance parameters and the parameters of the new filter was set equal. Realistic flightpath data were generated using the Air Force BLUERMIX II flightpath generator program, with the F-16A aircraft characteristics file. Space here allows inclusion of only a single tracking performance example. For this example, the target flies along the $x$-axis at a speed of 500 fps until it reaches $x=2000$ feet. At this point, the afterburner is turned on and the target initiates a strong climbing left-hand turn. In figure 1, the actual trajectory is highlighted with an $x$ on each data point. The trajectory highlighted with diamonds shows the performance of the simple Kalman filter, while the trajectory indicated with circles shows the performance of the new filter. As can be seen, the transient behavior of the new filter is superior.

IV SUMMARY AND CONCLUSIONS

This paper presents a new filter to be used for tracking maneuvering targets. Not all facets of the filter are in final form, and work remains in the area of the exact formulation of the consistency update, as well as filter aggregation and filter tuning, or parameter selection. Nonetheless, preliminary results are promising. Clearly, the next stage of the research should address benchmarking the proposed filter against some other well accepted multi-model filter such as the interacting multiple model filter [6].

REFERENCES

APPENDIX D

D. R. Shin and E. I. Verriest

The General Formulas for Smoothing and Prediction Problems of Mixed-Type Observations

Proc. 29th IEEE Conf. on Decision and Control

THE GENERAL FORMULAS FOR SMOOTHING AND PREDICTION PROBLEMS OF MIXED-TYPE OBSERVATIONS

D. R. Shin and E. L. Verrier

School of Electrical Engineering
Georgia Institute of Technology
Atlanta, Georgia 30332

Abstract—The filtering, smoothing, and prediction problems for mixed-type states and observations (continuous plus discontinuous) are considered. Normalized and unnormalized forms for the corresponding three types of estimates are obtained, which are formulated in the unified frameworks.

1 Introduction

Elliott and Antonelli [1,2] discussed the smoothing and prediction problems for Wiener-type observations in terms of semimartingale decompositions and measure transformations. Analogous problems for counting observations were treated in [3] and [4] independently. We extend these ideas to general estimation problems where both the signal and the observation consist of Wiener processes and counting processes, and moreover there exists dependence between signal and observation noise.

For the nonlinear filtering for these types of general problems, Gertner [5] obtained optimal conditional expectations based on a measure transformation and a Fubini-type theorem. He constructed a new equivalent reference measure under which he derived normalized and unnormalized distributions directly from the definition of conditional expectations and Bayesian formulas, not relying on innovation methods and appropriate transformations.

Based on Gertner’s general model, we will obtain normalized and unnormalized estimates for the smoothing and prediction problems in a different way. Our approach is first to derive the normalized forms using innovation methods, then the unnormalized forms indirectly, and finally direct derivations of the unnormalized forms will be made. This paper shows a unified approach to general nonlinear estimation for smoothing and prediction problems.

One example of such a mixed type observation was given by Hoversten et al. [6] who considered that in optical communication receivers, the detector output currents could be modelled as stochastic processes. Those processes contain doubly stochastic Poisson processes due to photoelectron and dark current and a Wiener process due to thermal noise.

General terminology and assumptions are presented in the next section.

2 Notation and Preliminaries

Let \((\Omega, \mathcal{F}, P)\) be a complete probability space and let \(\{\mathcal{F}_t, t \in [0,T]\}\) be a nondecreasing family of \(\sigma\)-fields of \(\mathcal{F}\) such that \(\mathcal{F}_t\) is right continuous and \(\mathcal{F}_0\) contains all null sets. All stochastic processes are defined on \((\Omega, \mathcal{F}, P)\) and a finite time interval \([0,T]\), and are scalar-valued. We denote the class of square integrable (local) martingales by \(\mathcal{M}(\mathcal{F}, P)\) (resp. \(\mathcal{M}_{\text{loc}}(\mathcal{F}, P)\)). Assume that the signal and observation processes have the following representation

\[ X_t = X_0 + \int_0^t f_s \, ds + M_t + \tilde{H}_t, \]

\[ y_t = \int_0^t h_s \, ds + \tilde{W}_t + \tilde{N}_t, \]

where \(f_s\) and \(h_s\) are square-integrable \(\mathcal{F}_t\)-adapted processes, \(M_t, W_t \in \mathcal{M}(\mathcal{F}, P)\) with \(\langle \tilde{W}, \cdot \rangle = t\), and \(\tilde{H}_t\) and \(\tilde{N}_t\) are integrable counting processes such that

\[ \hat{Q}_t = \tilde{H}_t - \int_0^t \gamma_u \, du \in \mathcal{M}(\mathcal{F}, P), \]

\[ \hat{Q}_t = \tilde{N}_t - \int_0^t \lambda_u \, du \in \mathcal{M}(\mathcal{F}, P), \]

where \(\gamma_u, \lambda_u\) are square-integrable non-negative \(\mathcal{F}_t\)-predictable processes. We assume that there exist \(\mathcal{F}_t\)-predictable processes \(\alpha_u, \beta_u\) such that

\[ < \hat{M}_t, \tilde{W} >_t = \int_0^t \alpha_u \, du, \]

\[ [\hat{M}_t, \tilde{N}] = [\hat{Q}_t, \hat{Q}_t] = \int_0^t \beta_u \, du \in \int_0^t \beta_u \, d\tilde{N}_u. \]

Next we describe a measure transformation which plays a key role in obtaining the unnormalized distribution. Define a measure \(P_0\) equivalent to \(P\) on \((\Omega, \mathcal{F})\) by

\[ \frac{dP_0}{dP} = \exp\left[ - \int_0^T h_u \, dW_u + \frac{1}{2} \int_0^T h_u^2 \, du + \int_0^T \ln(\lambda_u) \, dN_u + \int_0^T (\lambda_u - 1) \, du \right]. \]

By Girsanov’s theorem, \(P\) is absolutely continuous with respect to \(P_0\) with Radon-Nikodym derivative

\[ \frac{dP}{dP_0} = \left( \frac{dP_0}{dP} \right)^{-1} = \exp\left[ \int_0^T h_u \, dW_u - \frac{1}{2} \int_0^T h_u^2 \, du \right. \]

\[ + \int_0^T \ln(\lambda_u) \, dN_u - \int_0^T (\lambda_u - 1) \, du \right]. \]

where \(W_t = \tilde{W}_t + \int_0^t h_u \, du\). It can be shown that

\[ \lambda_t = \exp\left[ \int_0^t h_u \, dW_u - \frac{1}{2} \int_0^t h_u^2 \, du + \int_0^T \ln(\lambda_u) \, dN_u - \int_0^T (\lambda_u - 1) \, du \right] \]

is a \((\mathcal{F}_t, P_0)\)-martingale such that

\[ \lambda_t = E_0[\frac{dP}{dP_0} \mid \mathcal{F}_t]. \]

Furthermore by the differential rule, \(\lambda_t\) satisfies the integral equation

\[ \lambda_t = 1 + \int_0^t \lambda_u h_u \, dW_u + \int_0^t \lambda_u (\lambda_u - 1) \, d(N_u - u). \]

Then under the measure transformation the observation processes are simplified.
We get where and 9(t)

Note that under both measures. Similarly, the signal process is also changed under the new measure $P_0$

Thus under $P_0$ the signal and observation process are of the form

$$X_t = X_0 + \int_0^t f_u du - \int_0^t h_u du + M_t + \int_0^t b_u du$$

(5)

$$Y_t = W_t + N_t$$

(6)

It is also known that $X_t$ has the same distribution under either measure [7].

The following theorem (Fubini-type theorem for stochastic integral) is of critical importance in the derivation of the optimal conditional distribution. It provides conditions for interchanging conditional expectation and stochastic integration [5].

**Theorem 1:** Let

$$M_t, V_t \in \mathcal{M}^{2}(\mathcal{F}_t, P), \quad \mathcal{G}_t = \sigma(V_t, s \leq t)$$

Then

$$E[M_t | \mathcal{G}_s] = \int_0^t a_u dV_u$$

(7)

where

$$a_t = \begin{cases} d & < M, V \geq \downarrow \\
-d & < V \geq \downarrow \end{cases}$$

is an $\mathcal{F}_t$-predictable process, and $a_t$ is a $\mathcal{G}_t$-predictable projection of $a_t$ onto $\mathcal{G}_t$ such that

$$E\int_0^t a^2 d < V > < \infty$$

As special cases, (1) if $M \perp V$ then $E[M_t | \mathcal{G}_s] = 0$.

(2) if $f_t$ is a square-integrable $\mathcal{F}_t$-predictable process, then $\mathcal{G}_t$ and $\mathcal{G}_t(t > s)$ are conditionally independent given $\mathcal{G}_s$, then

$$E\int_0^s f_t dV_t | \mathcal{G}_s] = \int_0^s f_t dV_t$$

(8)

where $\int_t$ is the predictable projection of $f_t$ on $\mathcal{G}_t$.$\square$

Since independent increment processes $W_t, N_t$ satisfy (2), we get known results [5,6]:

$$E\int_0^t h_u dW_u | F_s = \int_0^t E[h_u | F_s] dW_u$$

$$E\int_0^t \lambda_u dW_u | F_s = \int_0^t E[\lambda_u | F_s] dW_u$$

$$E\int_0^t \alpha_u du | F_s = \int_0^t E[\alpha_u | F_s] du$$

for integrable, $\mathcal{F}_t$-adapted processes $h_t$ and $\alpha_t$, and an $\mathcal{F}_t$-predictable process $\lambda_t$.

Next is the general martingale representation when the filtration is generated by both Poisson and Wiener processes [7, p246].

**Theorem 2:**

Let $M_t \in \mathcal{M}_{\infty}(\mathcal{Y}_t, P), \quad \mathcal{Y}_t = \sigma(W_t, N_t, s \leq t) \cup (P \text{ null set}).$

Then there exist $\mathcal{Y}_t$-predictable processes $K_t$ and $R_t$ with

$$\int_0^t K_t^2 + |R_t|^2 ds < \infty,$$

such that

$$M_t = M_0 + \int_0^t K_t dW_t + \int_0^t R_t d\mathcal{Y}_t$$

where $\mathcal{Y}_t = N_t - t.$ \(\square\)

We are now in a position to derive the following optimal estimates, i.e. the conditional expectation of $X_t$ given the observation $\sigma$-fields:

- For the filtering case $\hat{X}_t = \Pi_t(X_t) = E[X_t | \mathcal{Y}_t]$
- For the smoothing case $\Pi_t(X_t) = E[X_t | \mathcal{Y}_t](s < t)$
- For the prediction case $\Pi_t(X_t) = E[X_t | \mathcal{Y}_t](s < t)$

For the filtering problem, Gerstner [5] derived the normalized and unnormalized equations directly from the definition of conditional distribution and Bayesian formulas under the transformed measure. We use alternative approaches through two steps, i.e., the innovation approach with semimartingale decomposition method for normalized forms [1] and then the measure transformation approach for unnormalized forms (indirect derivations) [2]. For smoothing and prediction problems we also include direct derivations of unnormalized forms [3,5]. All detailed derivations are included in the appendix (A-1 through A-9).

### 3 Filtering

Rewriting (1) and (2), we have the $\mathcal{F}_t$-semimartingales for a signal and an observation process

$$X_t = X_0 + \int_0^t f_u du + \int_0^t \gamma_u du + (\tilde{M}_t + \tilde{Q}_t),$$

(9)

$$Y_t = W_t + N_t = \int_0^t h_u du + \int_0^t \lambda_u du + (\tilde{W}_t + \tilde{q}_t).$$

(10)

The filtering problem is to derive

$$\hat{X}_t = \Pi_t(X_t) = E[X_t | \mathcal{Y}_t]$$

where $\mathcal{Y}_t = \sigma(Y_t, s < t) = \sigma(W_t, N_t, s \leq t)$. Noting that $\hat{X}_t - \int_0^t \lambda_u du - \int_0^t \gamma_u du$ is a $\mathcal{Y}_t$-innovation martingale, it can be represented as the stochastic integral with respect to the innovation process which has two components (for a unique representation refer to [7, p264])

$$\nu_t = W_t - \int_0^t \tilde{h}_u du$$

(11)

$$\mu_t = N_t - \int_0^t \tilde{\lambda}_u du.$$

(12)

Elliott [1] observed that for continuous-type state and observation

$$E[ X_t | Y_t ] = \pi_t E[ X_t | Y_t ] = \hat{X}_t Y_t$$

is a $\mathcal{Y}_t$-semimartingale and so its decomposition, as the sum of a martingale and a bounded predictable process, is unique.
Then he obtained two representations for $\tilde{X}_t$, compared bounded variations parts, and finally got filtering formulas. Applying this idea to the general stochastic equations (9) and (10), the resulting normalized filtering formula is of the form (see A-1)

$$
\Pi_t(X_t) = \Pi_0(X_0) + \int_0^t \Pi_u(X_u) d\omega_u + \int_0^t \Pi_u(\gamma_u) d\omega_u
$$

(13)

Then he obtained two representations for $\tilde{X}_t$, compared bounded variations parts, and finally got filtering formulas. Applying this idea to the general stochastic equations (9) and (10), the resulting normalized filtering formula is of the form (see A-1)

$$
\Pi_t(X_t) = \Pi_0(X_0) + \int_0^t \Pi_u(X_u) d\omega_u + \int_0^t \Pi_u(\gamma_u) d\omega_u
$$

(13)

where $\Pi_u(\gamma)$ heuristically implies the predictable version of that conditional expectation. This approach is a little different from ones taken in [7,8], and will be used in later developments.

This equation can be simplified by introducing the reference probability measure $P_0$ under which the observation process become simple. The resulting equation is the linear unnormalized conditional expectation. From (4) and the Bayesian rule, we have

$$
\Pi_t(X_t) = E[X_t | y_t] = E_0[A_tX_t | y_t] = \frac{\sigma_t(X)}{\sigma_t(1)}
$$

where $\sigma_t(X)$ is the unnormalized conditional expectation, $E_0$ is the expectation with respect to $P_0$, and $\sigma_t(1)(=\bar{\Lambda}_t)$ satisfies (see A-2)

$$
\sigma_t(1) = 1 + \int_0^t \sigma_u h_u dW_u + \int_0^t \sigma_u(-\lambda_u - 1) du_u.
$$

(14)

Applying the product rule to $\sigma_t(X)(=\sigma_t(1)\Pi_t(X_t))$ gives

$$
\sigma_t(X_t) = \sigma_0(X_0) + \int_0^t \sigma_u f_u + \gamma_u) du_u + \int_0^t \sigma_u(h_u X_u + a_u) dW_u
$$

(15)

+ \int_0^t \sigma_u(-\lambda_u - 1) X_u + b_u X_u + a_u) dW_u.

An alternative approach obtains (15) directly from the definition $\sigma_t(X_t) = E_0[A_tX_t | y_t]$ by representing $A_tX_t$ by a stochastic integral and taking conditioning [5].

Remark: In [5], Gertner first obtained $\sigma_t(X)$ and then derived $\Pi_t(X)$ by the product rule. Thus he could avoid the proof of the existence of martingale representation theorem.

### 4 Smoothing

The smoothing problem can be solved in a manner similar to the filtering problem. We consider the conditional expectation, $\Pi_t(X_t) = E[X_t | y_t]$, where $0 \leq s \leq t \leq T$. Notice that for fixed $s$, $\Pi_t(X_t)$ is an $y_t$-martingale and so it has a martingale representation with respect to the innovation martingale such that

$$
\Pi_t(\Phi_s) = \Pi_t(\Phi_s) + \int_s^t K_u dW_u + \int_s^t R_u d\omega_u
$$

(16)

where $K_u$, $R_u$ are $y_t$-predictable processes to be determined. The same procedure as for filtering results in a recursive smoothing equation (see A-4)

$$
\Pi_t(X_t) = \Pi_t(X_s) + \int_s^t [\Pi_u(X_u) - \Pi_u(X_s)\Pi_u(h_u)] d\omega_u
$$

(17)

+ \int_s^t [\Pi_u(\lambda_u)]^{-1} [\Pi_u(X_s, \lambda_u) - \Pi_u(X_s)\Pi_u(\lambda_u)] d\omega_u.

For the unnormalized form, by the Bayesian formula

$$
\Pi_t(X_t) = E[X_t | y_t] = \frac{E_0[A_tX_t | y_t]}{E_0[A_t | y_t]} = \frac{\sigma_t(X_t)}{\sigma_t(1)}
$$

is obtained. Again, applying the product rule to $\sigma_t(X_t)$, we obtain the unnormalized form for smoothing (see A-5):

$$
\sigma_t(X_t) = \sigma_t(1) + \int_0^t \sigma_s u(X_s) dW_u + \int_0^t \sigma_s u((\lambda_s - 1) X_s) d\omega_u.
$$

(18)

Notice that this is a linear equation for $\sigma_t(X)$. Alternatively, (18) can be derived from the definition $\sigma_t(X_t) = E_0[A_tX_t | y_t]$ by representing $X_t$ by stochastic integrals and taking conditional expectations under $E_0$ (see A-6).

### 5 Prediction

The same procedure can be applied to derive the conditional expectation of the form

$$
\Pi_t(X_t) = E[X_t | y_t]
$$

(19)

where $0 \leq s \leq t \leq T$. For fixed $t$, $\Pi_t(X_t)$ is a $y_t$-martingale, so it has a representation of the form

$$
\Pi_t(X_t) = \Pi_t(X_0) + \int_0^t K_u dW_u + \int_0^t R_u d\omega_u
$$

(20)

where $K_u$, $R_u$ are integrable $y_t$-predictable processes. For convenience of computation of the gains $K_u$, $R_u$, an auxiliary process $Z_t = E[X_t | Y_t]$, $0 \leq s \leq t \leq T$, is introduced, which produces $\Pi_t(X_t) = E[Z_t | y_t]$. Computing the gains, we can show (see A-7) that

$$
\Pi_t(X_t) = \Pi_t(X_0) + \int_0^t [\Pi_u(h_u X_t) - \Pi_u(X_t)\Pi_u(h_u)] dW_u
$$

(21)

+ \int_0^t [\Pi_u(\lambda_u)]^{-1} [\Pi_u(X_t, \lambda_u) - \Pi_u(X_t)\Pi_u(\lambda_u)] d\omega_u.

Using the Bayesian formula to get a simpler equation yields

$$
\Pi_t(Z_t) = E[Z_t | y_t] = \frac{E[Z_t | y_t]}{E_0[A_t | y_t]} = \frac{\sigma_t(Z_t)}{\sigma_t(1)}
$$

(22)

By a similar technique (see A-8),

$$
\sigma_t(X_t) = \sigma_t(1) + \int_0^t \sigma_s u(h_u X_t + a_u) dW_u
$$

(23)

is obtained, where $<Z, W>_t = \int_0^t \sigma_u dW_u, [Z, N]_t = \int_0^t \sigma_u dN_u$.

A direct derivation of (22) is also possible from the definition $\sigma_t(X_t) = E_0[Z_t, A_t | y_t]$ using the same approach (see A-9).

### 6 Discussion

Starting with the general models of stochastic systems (mixed-type of states and observations), we derived the normalized equations for filtering, smoothing, and prediction of general stochastic equations using semimartingale
decomposition techniques. We also showed that the corresponding unnormalised equations could be derived in two different ways (indirect and direct derivation). Notice that the same derivation methods result in similar structure in forms. Thus the optimal estimation problems are unified in these frameworks.

7 Appendix

A-1: the derivation of equation (12)
Conditioning (9) with respect to $y_1$, [7],

$$\Pi_t(X_t) = \Pi_0(X_0) + \int_0^t \Pi_u(f_u + \gamma_u)du + m_t \quad (24)$$

where $m_t$ is a $y_{1-}$martingale and thus it has the unique representation with respect to innovation processes $\nu$ and $\mu$:

$$m_t = \int_0^t K_u du + \int_0^t R_u du$$

for $y_{1-}$predictable processes $K_t, R_t$.

Remarks: the existence of $m_t$ can be easily shown. And in equation (24), the optional or predictable projections onto observations, denoted by $\Pi_t(X_t)$ and $\Pi_{-\infty}(X_t)$ respectively, should be taken. However, it can be shown that the following equality holds except on the set of $dtP(\omega)$ measure zero:

$$\int_0^t \Pi_u(f_u + \gamma_u)du = \int_0^t \Pi_{-\infty}(f_u + \gamma_u)du$$

Thus (24) can be identified under either projection. For details, refer to [7, p253]. In what follows, conditioning on $y_t$ implies the optional projection or predictable one as the case may be. By using Theorems 1 and 2, this martingale representation theorem with respect to innovation processes can be proved. Although $\nu$ and $\mu$ are $y_{1-}$martingales, we cannot derive (24) by a direct application of Theorem 2 because in general, $y_t \not\supset \sigma(\nu_u, \mu_u, u \leq t)$. In [7], one way around this difficulty is to take a measurement transformation and under the new measure to get the representation with respect to observations and then innovations. For details refer to [7, p264].

To find $K_u$, applying the product rule to $X_tW_t$,

$$X_tW_t = \int_0^t W_u(f_u du + d\tilde{M}_u + d\tilde{V}_u) + \int_0^t X_u(h_u du + d\tilde{V}_u) + \int_0^t a_u du.$$

Conditioning on $y_t$ and using the Fubini-type theorem,

$$E[X_tW_t | y_t] = \Pi_t(X_t)|W_t = \int_0^t \Pi_u(X_u)du + \int_0^t \Pi_u(a_u)du$$

$$(y_t - \text{martingale}).$$

This equation shows that $\Pi_t(X_t)y_t$ is a special semimartingale which is the sum of a $y_{1-}$ martingale and a predictable bounded variation process, and furthermore the decomposition is unique [1]. Whereas, from (11) and (24)

$$\Pi_t(X_t)W_t = \int_0^t \Pi_u(X_u)du + \int_0^t W_u d\Pi_u(X_u) + \int_0^t K_u du$$

Noting that $d < \Pi(X), W > = K dt$,

$$\Pi_t(X_t)W_t = \int_0^t \Pi_u(X_u)\Pi_u(h_u)du + \int_0^t K_u du + (y_t - \text{martingale}).$$

Since the decomposition is unique, comparing the bounded variation terms, we have

$$K_u = \Pi_u(X_u) - \Pi_u(X_u)\Pi_u(h_u) + \Pi_u(a_u).$$

Similarly, for $R_t$, applying the product rule to $X_tN_t$,

$$X_tN_t = \int_0^t X_u(h_u du + d\tilde{M}_u + d\tilde{H}_u) + \int_0^t h_u\lambda_u du.$$

Conditioning on $y_t$ and using the Fubini-type theorem,

$$\Pi_t(X_t)N_t = \int_0^t \Pi_u(h_u X_u)du + \int_0^t N_u\Pi_u(f_u + \gamma_u)du$$

$$+ \int_0^t \Pi_u(h_u) + (y_t - \text{martingale}).$$

On the other hand, from (12) and (24)

$$\Pi_t(X_t)N_t = \int_0^t \Pi_u-(X_u)N_u + \int_0^t N_u d\Pi_u(X_u) + \int_0^t d[\Pi(X), N_u].$$

Observing that $d[\Pi(X), N]$ is $R_t dN_t$,

Finally, substituting $K_t$ and $R_t$ into $m_t$ results in (13).

A-2: the derivation of $\sigma_t(1)$
Taking conditional expectations of $\lambda_t$ under $P_0$ and applying the Fubini-type theorem,

$$E_0[\lambda_t | y_t] = 1 + \int_0^t E_0[\lambda_u | y_t]dW_u + \int_0^t E_0[\lambda_u - 1 | y_u]d\tilde{M}_u.$$

The result follows.

A-3: the derivation of $\sigma_t(X_t)$
From (14) and (24),

$$\sigma_t(X_t) = \Pi_t(X_t)\tilde{A}_t$$

$$= \Pi_0(X_0) + \int_0^t \Pi_u(X_u)[\Pi_u(h_u)\lambda_u dW_u + (\Pi_u-(\lambda_u - 1)\lambda_u d\tilde{M}_u]$$

$$+ \int_0^t \Pi_u(h_u)\lambda_u du + \int_0^t K_u du + \lambda_u X_u + R_u dN_u.$$

With straightforward calculations, we get the equation (15).

A-4: the derivation of equation (17)
Taking conditional expectations of $X_tW_t$ with respect to $y_t$,

$$\Pi_t(X_t)W_t = \int_0^t \Pi_u(X_u)du + (y_t - \text{martingale}).$$

However, from (11) and (18)

$$\Pi_t(X_t)W_t = \int_0^t \Pi_u(X_u)[\Pi_u(h_u)du + \int_0^t K_u du + (y_t - \text{martingale})].$$

By the unique decomposition of a special semimartingale,

$$K_u = \Pi_u(X_u) - \Pi_u(X_u)\Pi_u(h_u).$$

Similarly, conditioning $X_tN_t$ on $y_t$,

$$\Pi_t(X_t)N_t = \int_0^t \Pi_u-(X_u\lambda_u) + (y_t - \text{martingale}).$$

However, from (12) and (16)
\[
\Pi_t(X_t) N_t = \Pi_t(X_t) N_t + \int_0^t \Pi_u(X_t) \Pi_u(\lambda_u) du
\]
\[
+ \int_0^t R_u \Pi_u(\lambda_u) du + (Y_t - \text{martingale}).
\]
Equating the bounded variation terms of two representations of \(\Pi(X_t) N_t\),
\[
R_{\infty} = [\Pi_u(\lambda_u)]^{-1} [\Pi_u(\lambda_u Z_u) - \Pi_u(X_t) \Pi_u(\lambda_u) + \Pi_u(\lambda_u)].
\]
Then the result directly follows.

A-5: the derivation of equation (18)
From (14) and (16),
\[
\sigma_{\pi}(X_t) = \Pi_t(X_t) \hat{\lambda}_t
\]
\[
= \Pi_t(X_t) \hat{\lambda}_t + \int_0^t \Pi_u(X_t) d\hat{\lambda}_u
++ \int_0^t \lambda_u d\Pi_u(X_t) + \int_0^t d[\Pi(X_t), \hat{\lambda}_u].
\]
Noticing that
\[
d[\Pi(X_t), \hat{\lambda}_u] = \Pi_t(\lambda_u) \hat{\lambda}_u K_{X, u} \hat{d}u + (\Pi_t(\lambda_u) - 1) R_{\infty} du N_u,
\]
With some manipulations,
\[
\sigma_{\pi}(X_t) = \Pi_t(X_t) \hat{\lambda}_t + \int_0^t \hat{\lambda}_u d\Pi_u(X_t, h_u) dW_u
++ \int_0^t \lambda_u - [\Pi_u((\lambda_u - 1) X_u)] du.
\]
The result immediately follows.

A-6: the direct derivation of (18)
\[
X_t \Delta_t = X_t \Delta_t + \int_0^t X_u d\Delta_u = \Delta_t \Delta_t + \int_0^t X_u \Delta_u dW_u + \Delta_u (\lambda_u - 1) dW_u.
\]
Therefore
\[
\sigma_{\pi}(X_t) = E_0[X_t \Delta_t | Y_t]
\]
\[
= E_0[X_t \Delta_t | Y_t] + \int_0^t E_0[\Delta_u, X_u | Y_u] dW_u
++ \int_0^t E_0[X_t \Delta_u (\lambda_u - 1) | Y_u] du.
\]
This completes the proof.

A-7: the derivation of equation (21)
Applying the product rule to \(Z_u N_u\), and conditioning on \(Y_t\) produce
\[
\Pi_t(X_t) W_t = \int_0^t \Pi_u(Z_u \Delta_u) du + \int_0^t \Pi_u(\alpha_u) du + (Y_t - \text{martingale}).
\]
From (11) and (20),
\[
\Pi_t(X_t) W_t = \int_0^t W_u d\Pi_u(Z_u) + \int_0^t \Pi_u(Z_u) dW_u + \int_0^t K_u dW_u dW_u
\]
\[
= \int_0^t \Pi_u(Z_u) \Pi_u(\Delta_u) du + \int_0^t K_u dW_u + (Y_t - \text{martingale}).
\]
Comparing the predictable processes,
\[
K_u = \int_0^t \Pi_u(Z_u \Delta_u) du - \int_0^t \Pi_u(Z_u) \Pi_u(\Delta_u) du + \int_0^t \Pi_u(\alpha_u) du.
\]
Similarly, applying the product rule to \(Z_u N_u\), and projecting this onto the observation \(\sigma\)-fields produce
\[
E[Z_u N_u | Y_t] = \Pi_t(X_t) N_t
\]
\[
= \int_0^t E[Z_u \Delta_u | Y_u] du + \int_0^t \Pi_u(\alpha_u) du + (Y_t - \text{martingale}).
\]
From (12) and (20),
\[
\Pi_t(X_t) N_t = \int_0^t \Pi_u(X_t) dN_u + \int_0^t N_u d\Pi_u(X_t) + \int_0^t R_u dN_u
\]
\[
= \int_0^t \Pi_u(X_t) \Pi_u(\lambda_u) du + \int_0^t R_u \Pi_u(\lambda_u) du + (Y_t - \text{martingale}).
\]
The unique decomposition theorem gives
\[
R_{\infty} = [\Pi_u(\lambda_u)]^{-1} [\Pi_u(\lambda_u Z_u) - \Pi_u(X_t) \Pi_u(\lambda_u) + \Pi_u(\lambda_u)].
\]
Now,
\[
\Pi_t(\lambda_u Z_u) = E[\lambda_u E[X_t | Z_u] | Y_u] = E[\lambda_u X_t | Y_u] = \Pi_t(\lambda_u X_t).
\]
Therefore, from \(K_u\) and \(R_u\), equation (21) follows.

A-8: the derivation of equation (23)
From (14) and (20),
\[
\sigma_{\pi}(X_t) = \Pi_t(X_t) \hat{\lambda}_t
\]
\[
= \Pi_t(X_t) \hat{\lambda}_t + \int_0^t \hat{\lambda}_u d\Pi_u(X_t, h_u) dW_u
++ \int_0^t \hat{\lambda}_u - [\Pi_u((\lambda_u - 1) X_u)] du.
\]
Observing that
\[
d[\Pi(X_t), \hat{\lambda}_u] = \hat{\lambda}_u \Pi_t(\lambda_u) K_{X, u} \hat{d}u + (\Pi_t(\lambda_u) - 1) R_{\infty} du N_u,
\]
with some computations, we get the equation (23).

A-9: the direct derivation of equation (23)
From (22) and the product rule,
\[
Z_t A_t = Z_0 A_0 + \int_0^t Z_u A_t dW_u + \int_0^t A_t Z_u dW_u + \int_0^t d[A, Z]_u.
\]
Note that under \(P_0\), \(A, Z\) = 0. Applications of the Fubini-type theorem gives (23).

References

APPENDIX E

C. C. Tsai and A. H. Haddad

On Singularly Perturbed Hybrid Systems

Proc. 24th Annual Conf. on Information Sciences and Systems

ON SINGULARLY PERTURBED HYBRID SYSTEMS

C. C. Tsai and A. H. Haddad

Department of Electrical Engineering and Computer Science
Northwestern University
Evanston, IL 60208

ABSTRACT

This paper considers a singularly perturbed hybrid system whose state equations depend on a near decomposable finite state Markov process with fast transitions. The limiting behavior of the fast and slow mode subsystem for the duration of intervals of fast transitions within each group is investigated. The results are shown to hold when the process is near decomposable, ergodic, time-reversible and stationary, and the values of the system matrices within each group commute.

I. INTRODUCTION AND PROBLEM FORMULATION

This paper examines the limiting behavior of a singularly perturbed hybrid system with switched parameters which depend upon a near decomposable finite state Markov process (FSMP). The state space of a hybrid system is a cross-product space of a Euclidean space and a finite-state space. Basically, hybrid systems are linear, piece-wise constant, time-varying systems, which are switching among a finite number of constant realizations. Hybrid systems have already been considered, and their properties are well documented[1]. The use of the aggregated models to describe global features of singularly perturbed FSMPs is studied in [2]. Preliminary investigation of singularly perturbed hybrid systems and singularly perturbed FSMPs is reported in [3]. Here we extend the latter results, but consider the decoupled case where the switching is near decomposable FSMP as discussed in [4].

The system models under consideration are assumed to have the following state equations:

\[ x(t) = A(t)[r(t)]x(t) \]  
\[ \dot{z}(t) = A_{\tau}(t)[z(t)] \]  

where \( \mu > 0 \) is a small parameter, \( x(t) \in \mathbb{R}^n \) represents the slow mode of the system and \( z(t) \in \mathbb{R}^m \) the fast mode, the process \( r(t) \) is the form index which takes values in \( S = \{1, 2, ..., M\} \) and determines the system model at a particular time. The process \( r(t) \) is modeled as an FSMP which contains \( N \) groups of strongly interacting states, where group \( j \) consists of \( n_j \) fast states and \( \sum_{j=1}^{M} n_j = M \). Matrices \( A_{\tau}(t) \) and \( A_{\mu}(t) \) are random through their dependence on the values of the process \( r(t) \). The current values of \( A_{\tau}(t) \) is denoted by an index, for example, for \( i = 1, 2, ..., M \), \( A_{\tau,m} \) will denote \( A_{\tau}(t) \) when \( r(t) = m \). Let \( \mu \) and \( \epsilon \) be small constants and \( \epsilon \) tend to 0. Furthermore, if we express \( \mu = \frac{1}{\epsilon^2} \) we obtain

\[ \frac{dP(t)}{dt} = P(t)(-F + G) \]  

where \( P(t) \) is an \( M \times M \) transition probability matrix at time \( t \). It is assumed that \( 0 \leq \epsilon \leq 1 \), matrices \( F, G \) and \( \frac{1}{\epsilon}F + G \) are generators such that the process \( r(t) \) has a single ergodic class. Furthermore, let each of the \( N \) groups be a FSMP with a single ergodic class[5]. The generator of the \( j \)-th group \( E_j \) is the \( j \)-th block in the block-diagonal matrix \( F \). This paper considers the stationary case, namely that the process \( r(t) \) has reached its steady state.

The following is the outline of the paper. Section II summarizes asymptotic behavior of a near decomposable FSMP. In Section III, the approximate model \( r(t) \) for the slow mode subsystem \( z_i(t) \) for the duration of intervals of transitions among the fast states of group \( j \) is derived. The probabilistic averaging procedure is adopted. The mean-squared error between \( z_i(t) \) and \( z_j(t) \) is studied. In Section IV, we study the limiting behavior of the fast mode subsystem when both \( \mu \) and \( \epsilon \) tend to 0. Section V considers an example to illustrate the methods used in Section III and IV. Section VI concludes the paper.

II. ASYMPTOTIC BEHAVIOR OF A NEAR DECOMPOSABLE FSMP

In this section we summarize the asymptotic behavior of a near decomposable FSMP. Since the process \( r(t) \) satisfies the MSST (multiple semistability) condition[2], it has only two time scales. In order to construct an asymptotic approximation of the process \( r(t) \), one establishes the following proposition:

Lemma 1 Let \( \Pi = \lim_{\epsilon \to 0} \exp(F \epsilon) \). Then

\[ \Pi = \text{diag}[\Pi_1, ..., \Pi_N] \]

With \( \Pi_j = I \) for \( j = 1, ..., N \), for some \( n_j \)-dimensional row vector \( E_j \) such that \( E_j = [e_{j1}, ..., e_{jn_j}] \) and \( E_j \cdot 1 = 1 \).

Furthermore, define the \( M \times N \) matrix \( V \) and the \( N \times M \) matrix \( U \) as follows:

\[ V = \text{diag}[1, 1, ..., 1] \]
\[ U = \text{diag}[E_1, ..., E_N] \]

then

\[ VU = \Pi \]

where \( I \) is an \( N \times N \) identity matrix.

We now use \( \Pi, V \) and \( U \) to construct a uniform asymptotic approximation of the process \( r(t) \) as shown in the following lemma:

Lemma 2 Assume that \( 0 < \epsilon < 1 \), then

\[ P(t) = \exp(UGV(t-t_0)U - \Pi \epsilon + o(1)) \]

uniformly valid for \( t \geq t_0 \).

Obviously, when \( \epsilon \to 0 \), the process \( r(t) \) can be replaced by an aggregated process \( \hat{r}(t) \) taking values in \( \hat{S} = \{1, ..., N\} \). Let \( \hat{P}(t) \) be the transition probability matrix of the process \( \hat{r}(t) \). Then

\[ \hat{P}(t) = \exp(UGV(t-t_0)), f \geq t \geq t_0 \]

and \( UGV \) is the generator of the process \( \hat{r}(t) \) and

\[ P(t) = V \hat{P}(t)U, f \geq t \geq t_0 \]

Note that the process \( r(t) \) is stochastically continuous and has a single ergodic class such that

\[ \lim_{\epsilon \to 0} \hat{P}(t) = \Pi \cdot e_{i_1, ..., e_{i_N}} \]

Thus Equation (7) can be interpreted as follows:

\[ \text{Pr}(r(t) = m \in \text{group } j \mid \text{r}(t) = k \in \text{group } i) = e_{jm} \cdot \text{Pr}(\hat{r}(t) = j \mid \hat{r}(t) = i) + o(1). \]

for \( t \geq t_0 \) and

\[ \lim_{\epsilon \to 0} \text{Pr}(r(t) = m \in \text{group } j \mid \text{r}(t) = k \in \text{group } i) = e_{j1} \cdot e_{jm} \]

where \( e_{jm} \) is the component of the ergodic probability vector \( e_m \) corresponding to state \( m \).

Furthermore, if we express (5) in the fast time scale \( r = (t-t_0) / \epsilon \), we obtain

Lemma 3 Let \( r = (t-t_0) / \epsilon \). We have

\[ \lim_{\epsilon \to 0} P(t) = \Pi \cdot e_{i_1, ..., e_{i_N}} \]
The positivity of \( \phi \) becomes
\[
\frac{\partial P(r)}{\partial r} = P(r)(F + rG)
\]
From (11) and (12) it follows that the influence of weak interaction \( rG \) in (12) will become significant after a long period of time \( r \). Assume that \( r(0) \in \text{group } j \), then we have
\[
P(r) = \sum_{m} p_{mm}^j + O(r)
\]
where \( p_{mm}^j \) is the \((m, n)\)th element of \( \exp(Fr) \).

In this work, the steady-state distribution of the process \( r(t) \) will be needed. This probability, from (10), is defined by \( e_j e_m, j = 1, \ldots, N, m = 1, \ldots, n_j \). Henceforth, it is assumed that the process \( r(t) \) has reached its steady state, i.e.,
\[
P(r(t) = m \in \text{group } j) = e_j e_m
\]
With the above assumptions, we conclude that in the fast time scale \( r, 0 < r \ll 1 \), jumps among different groups occur after a very long period of time \( r \) so that the following results are obtained:

**Lemma 4** Since group \( j \) is a finite ergodic stationary time reversible Markov chain, then the expectation of \( A[r(r)] \) in group \( j \) is given by
\[
\bar{A}_j^1 = \mathbb{E}[A_j[r(r)]|r(r) \in \text{group } j]
\]
and its autocovariance function
\[
c_j(r) = \mathbb{E}[A_j[r(r)]A_j[r(s)+r]r(s)]
\]
and \( r(s) + r \) is in group \( j \) - \( \bar{A}_j^1 \)
\[
= \sum_{m=1}^{n_j} A_j^1 e_m\{p_{mm}(r) - e_m\}A_j^1
\]
We next consider the additive process
\[
Y_j(r) = \int_0^r A_j[r(s)] ds.
\]
For large \( r \), we would expect that the process \( Y_j(r) \) is asymptotically normal in distribution, and satisfies a central limit theorem. Before doing so, the next lemma is required:

**Lemma 5** Let \( \mathcal{F}_P, F_j \) the fundamental matrix of group \( j \) and \( C_j \) be defined as follows
\[
\mathcal{F}_P = \text{diag}(e_1, \ldots, e_{n_j}), F_j = \{F_j^m \}
\]
\[
F_j^m = \int_0^{\infty} \{p_{mm}(r) - e_m\} dr
\]
and
\[
C_j = \int_0^{\infty} c_j(r) dr = [A_j^1, \ldots, A_j^m]_x
\]
\[
= [\mathcal{F}_P \otimes I_{n_j}][A_j^1]_x, \ldots, [A_j^m]_x]
\]
where \( \otimes \) denotes the Kronecker product [6].

Then \( C_j \) is strictly positive definite unless \( A_j^m \) is independent of \( m \) and \( C_j \) is zero if \( A_j^m \) is independent of \( m \).

The positivity of \( C_j \) has a natural meaning in that it corresponds to the variance of \( Y_j(r) \) in the central limit theorem.

**Lemma 6** Let \( Y_j(r) = \int_0^r A_j[r(s)] ds \). Then \( Y_j(r) \) is asymptotically normal in distribution for large \( r \) and satisfies the central limit theorem, i.e.,
\[
Y_j(r) \xrightarrow{d} \mathcal{N}[\bar{A}_j^1 r, 2 C_j r], \text{ as } r \to \infty.
\]

Meanwhile, \( Y_j(r) \) also satisfies a weak law of large number such that
\[
\frac{Y_j(r)}{r} \sim \frac{\bar{A}_j^1}{r}
\]
For proof of the above Lemmas one can see see [2], [p. 118, [5]] and [7].

### III. SLOW MODE SUBSYSTEM

This section considers the limiting behavior of the slow mode subsystem. Let \( x_j(t) \) denote the state of the subsystem when \( r(t) \) takes values in group \( j \). The trajectory of \( x_j(t) \) is based on writing the solution of \( x_j(t) \) as a function of the fast states in group \( j \) for the duration of the intervals of transitions among the fast states of group \( j \). The solution \( x_j(t) \) can be considered as a standard state equation solution of a time varying linear system. The time varying nature stems from the dependence of the system matrices on the different values of the fast states of group \( j \). In what follows we use probabilistic averaging to derive the approximate model for \( x_j(t) \). Given the group of the process \( r(t) \) (i.e., group \( j \)) and the state of the systems (1) at time \( t \) (i.e., \( x_j(t) \)), the expected state at time \( t + \Delta \) where \( \Delta \) is of order \( r \) is to be computed. Thus
\[
\hat{x}_j(t + \Delta) = \mathbb{E}[x_j(t + \Delta)|x_j(t) \in \text{group } j]\]
\[
= \sum_{m=1}^{n_j} \exp(A_j^m \Delta)(e_j e_m)/x_j(t)
\]
From Equation (21) one deduces the dynamics of \( \hat{x}_j(t) \) as
\[
\hat{x}_j(t) \equiv \lim_{\Delta \to 0} \mathbb{E}[x_j(t + \Delta) - x_j(t)|x_j(t) \in \text{group } j]/\Delta
\]
\[
= \{\sum_{m=1}^{n_j} A_j^m e_m\} \hat{x}_j(t)
\]
where \( \hat{x}_j(t) \) is the statistical average value of \( A_j^m [r(t)] \) when \( r(t) \) switches among the fast states of group \( j \). In what follows we shall show that the mean-squared error between \( \hat{x}_j(t) \) and \( x_j(t) \) tends to zero as \( r \) tends to zero.

**Theorem 1** Suppose that \( A_j^m A_j^m = A_j^m A_j^m, m, n \in [1, \ldots, n_j] \) and \( \hat{x}_j(t) \) is stable. Then
\[
\lim_{r \to \infty} \mathbb{E}[\|\hat{x}_j(t) - x_j(t)\|^2]
\]
\[
= 0, \text{ if } r \in \text{group } j, s \in [t_0, t] \]
where \( \| \cdot \| \) denotes the Euclidean norm.

**Proof:** Assume that the process \( r(t) \) switches to group \( j \) at time \( t_0 \). We express the systems (1) in the fast time scale \( r = (t - t_0)/\epsilon \) to obtain
\[
x_j(r) = c A_j^m [r(r)] X_j(r), X_j(0) = x_j(t_0)
\]
With the assumption that \( A_j^m A_j^m = A_j^m A_j^m \), we have
\[
x_j(r) = e^{r A_j^m} x_j(t_0)
\]
On the other hand, we express \( \hat{x}_j(t) \) in the fast time scale as
\[
\hat{x}_j(r) = e^{r \hat{A}_j^1} x_j(t_0)
\]
To prove the theorem, there are two cases to consider:
**Case 1:** Let \( A_j^m \) depends on the value of \( m \). The difference between \( \hat{x}_j(r) \) and \( x_j(r) \) is given by
\[ X_i(t) = X_i(t) = e^{e^{(t-t_0)}A} X_i(t_0) \quad (25) \]

For a given \( t, \) \( t \to \infty \) as \( \epsilon \to 0 \). It follows that

\[
E[\epsilon(V(t) - A^2_i(t) - 1)]
\]

\[
= \text{trace} E[\epsilon(V(t) - A^2_i(t)) - 2 \epsilon(V(t) - A^2_i(t) - 1)]
\]

\[
= \sum_{m=0}^{\infty} (e^{t/2})((4 C_i\sqrt{\epsilon})^{m} - \frac{C_i}{\epsilon^{1/2}})/m(2)!
\]

\[ = O(\epsilon) \]

and \( ||\epsilon A^2_i|| < 0 \). Hence,

\[
E[||X_i(t) - X_i(t)||^2 | e \in \text{group } j, s \in [t_0, t]] = O(\epsilon) \quad (26)
\]

Taking the limit \( \epsilon \to 0 \), we observe that the right-hand side of Equation (28) tends to zero. A change of variable yields the result.

Case II: Let \( A_{jm}^j \) be independent of \( m \) and \( A_{jm}^j = A_{jm}^i, m = 1, \ldots, n \). Lemma 3 then yields

\[ A_{jm}^j = A_{jm}^i \quad \text{and } C_j = 0 \]

Obviously, \( X_i(t) = X_i(t) \), i.e., there is no error between \( X_i(t) \) and \( X_i(t) \).

The resulting approximation implies that the slow mode subsystem can be approximated by a hybrid system depending on the aggregated Markov chain \( r(t) \) taking values in \( S \). The system model of the hybrid system at a particular time \( t \) is a statistical average of the system matrices over their values based on the group that the process \( r(t) \) takes.

IV. FAST MODE SUBSYSTEM

In this section, we study the limiting behavior of the fast mode subsystem:

\[ \mu(t) = A_m^i \mu(t) \mu(t) \quad (27) \]

Assume that \( ||A_{jm}^j|| = O(1), j = 1, \ldots, N, m = 1, \ldots, n_j \) and that \( r(t) = m \in \text{group } j \). To analyze the state \( z(t) \), the stretched time-scale \( \theta \), where \( \theta = (t - t_0)P \), is used. Hence, expressing the systems (27) in the stretched time-scale \( \theta \) yields

\[ \dot{Z}(t) = A_m^j(r(t))Z(t), \quad Z(0) = z(t_0). \quad (28) \]

Similarly, the equation for the transition matrix of \( r(t) \) becomes

\[ \frac{dP(\theta)}{d\theta} = P(\theta)(A_{zm}^i F + G), \quad P(0) = I. \quad (29) \]

The limiting behavior of the systems (28) depends on the relative size of \( \mu \) and \( \epsilon \) as they both tend to zero. There are three cases to consider:

Case I. \( \mu = O(\epsilon) \) and \( \lim_{t \to \infty} \epsilon = 0 \).

In this case we have

\[ \frac{dP(\theta)}{d\theta} = O(\epsilon), \quad P(0) = I. \quad (30) \]

Thus, in the stretched time-scale \( \theta \), the transition probability matrix tends to a constant. Since \( P(\theta) \) is an identity matrix, we observe that transitions among states of the process \( r(t) \) are very slow, i.e., \( r(t) = m \in \text{group } j \) implies \( r(t) = m \in \text{group } j \) for all \( \theta \) when \( \epsilon \to 0^+ \). Under this condition the system matrix for the state variables \( r(t) = r(t) \) when \( r(t) = m \in \text{group } j \). Similar to Section II, the solution of \( Z_i(t) \) of (28) is

\[ Z_i(t) = \exp(A_{jm}^j r(t), t_0) \quad (31) \]

which when transformed to the normal time scale \( t \) yields

\[ z_i(t) = \exp(A_{jm}^j (t - t_0)/\mu) \mu(t_0), \quad \text{for } t \geq t_0. \quad (32) \]

If all the values of \( A_{jm}^j \) are stable, then \( z_i(t) \) can be approximated by (32). The following theorem summarizes the case.

Theorem 2: Assume that \( \mu = O(\epsilon) \) and \( \lim_{t \to \infty} \epsilon = 0 \), all the values of \( A_{jm}^j \) are stable, and \( r(t_0) = m \in \text{group } j \). Then as \( \epsilon \to 0 \)

\[ z_i(t) = \exp(A_{jm}^j (t - t_0)/\mu) \mu(t_0), \quad \text{for } t \geq t_0. \quad (33) \]

Case II. \( \mu = O(\epsilon) \) and \( \lim_{t \to \infty} \epsilon = k \) where \( k = O(1) \). In this case we have

\[ \frac{dP(\theta)}{d\theta} = P(\theta)(A_{zm}^i F + O(\epsilon)), \quad P(0) = I. \quad (34) \]

From (31) it follows that, in the stretched time-scale \( \theta \), \( z_i(t) \) is approximated as an autonomous hybrid system (29) depending only on the fast states, of group \( j \), whose transition probability matrix is given by

\[ P(\theta) = \exp(F_{zm}^{jm} \theta) \]

1. If the values of \( A_{jm}^j \) are stable, then the solution of \( z_i(t) \) is uniformly asymptotically stable.

Note that logarithmic norm[1] can be used to check the stability condition of \( A_{jm}^j \).

Case III. \( \mu = \epsilon \) where \( 0 < \epsilon < 1 \) and \( \lim_{t \to \infty} \epsilon = \infty \). Denote

\[ A_{jm}^j = \sum_{m=1}^{n} A_{jm}^j \mu \]

Then

1. \( Z_i(t) \) can be approximated as an averaged system whose dynamical equations are described by

\[ \dot{Z}_i(t) = A_{jm}^j Z_i(t), \quad Z_i(0) = z_i(t_0) \quad (35) \]

2. The mean-squared error between \( Z_i(t) \) and \( \dot{Z}_i(t) \) tends to zero as \( \epsilon \) tends to zero.

3. The solution of \( Z_i(t) \) is uniformly asymptotically stable.

V. AN EXAMPLE

An example to illustrate the method in Section II and III is demonstrated here. The system is given by (1)-(2) where

\[ A_{m1}^i = \begin{bmatrix} -2 & 2 \\ 1 & -4 \end{bmatrix}, \quad A_{m2}^i = \begin{bmatrix} -3 & 2 \\ 1 & -5 \end{bmatrix} \]

\[ A_{m1}^j = \begin{bmatrix} -7 & 1 \\ 2 & -5 \end{bmatrix}, \quad A_{m2}^j = \begin{bmatrix} -8 & 1 \\ 2 & -4 \end{bmatrix} \]

\[ A_{m1}^j = \begin{bmatrix} -6 & 3 \\ 1.5 & -12 \end{bmatrix}, \quad A_{m2}^j = \begin{bmatrix} -2 & 4 \\ 2 & -10 \end{bmatrix} \]
and the FSM P r(t), shown in Fig.1, which consists of two groups, each of which contains two strongly interacting states, has the following generators:

\[
\begin{align*}
\dot{\lambda}^1_2 & = \begin{bmatrix} -7 & 1.5 \\ 8 & -4 \end{bmatrix}, \quad \dot{\lambda}^2_2 = \begin{bmatrix} -10 & 2 \\ 4 & -6 \end{bmatrix}
\end{align*}
\]

Hence, from the definition of \( \dot{\lambda}^j_t \) and \( \dot{\lambda}^j_t \), j=1,2, we have

\[
\begin{align*}
\dot{\lambda}^1_1 & = \begin{bmatrix} -2.5 & 2 \\ 1 & -4.5 \end{bmatrix}, \quad \dot{\lambda}^1_2 = \begin{bmatrix} -6.4 & 1 \\ 2 & -4.4 \end{bmatrix} \\
\dot{\lambda}^2_1 & = \begin{bmatrix} -4.5 & 3.5 \\ 1.75 & -11.2 \end{bmatrix}, \quad \dot{\lambda}^2_2 = \begin{bmatrix} -8.8 & 1.8 \\ 3.5 & -5.2 \end{bmatrix}
\end{align*}
\]

Fig.2 and Fig.3 show the sample trajectories of \( x_1(t) \) and \( 2_1(t) \), when the initial conditions \( x_1(0) = [2.0, 3.0] \), under three cases: (i) \( \epsilon = 0.1 \), (ii) \( \epsilon = 0.01 \), (iii) \( \epsilon = 0.001 \). Obviously, the error between \( x_1(t) \) and \( 2_1(t) \) is smaller as \( \epsilon \) becomes smaller. Similarly, Fig.4 shows the sample behavior for \( z_1(t) \) and \( 2_1(t) \), with \( z_1(0) = [2.0, 3.0] \) and \( \epsilon = 0.1 \). The error do become smaller when \( \epsilon \) goes decreases.

VI. CONCLUSIONS

This note considered the limiting behavior of a singularly perturbed decoupled hybrid system whose state equations depend on a near-decomposable finite state Markov process. The limiting system behavior of the slow mode subsystem for the duration of intervals of fast transition within each group can be approximated by an averaged value of the system matrix over all their values based on the fast states of the group. The limiting behavior of the fast mode subsystem for the duration of intervals of fast transitions within each group depends on the relative size of \( \alpha \) and \( \epsilon \) when both \( \alpha \) and \( \epsilon \) tend to zero: (i) the system can be approximated as a time invariant system with the constant system matrix held to the value at the initial transition when \( \mu = o(\epsilon) \), (ii) the system can be modeled as a hybrid system depending only on the fast states of the group when \( \mu = O(\epsilon) \), (iii) the system can be approximated as an averaged value of the system matrix over all their values based on the fast states of the group when \( \mu = \epsilon^\alpha \) where \( 0 < \alpha < 1 \). The results need two crucial assumptions, the ergodicity, stationary distribution and time reversibility of the process, and the fact that the values of the system matrices within each group commute. The results hold even if the aggregated process has an absorbing state.

Additional work is needed concerning the relaxation of the restrictions of commutation of the system matrices within each group. The Baker-Campbell-Hausdorff (BCH) formula [1,2] appears to be a promising feature in this direction.

Finally adapting the results of the note to the more general single perturbated (stochastic) hybrid systems where the switching is a near-decomposable finite state Markov chain may be useful.

References


Fig. 3 The Sample Trajectory of $z_{12}(t)$ and $\tilde{z}_{12}(t)$.

Fig. 4 The Sample Trajectory of $z_{11}(t)$ and $\tilde{z}_{11}(t)$.
APPENDIX F

C. C. Tsai and A. H. Haddad

Analysis of Singularly Perturbed Stochastic Hybrid Systems

Proc. 1991 American Control Conference

ANALYSIS OF SINGULARLY PERTURBED STOCHASTIC HYBRID SYSTEMS

C. C. Tsai and A. II. Iladdad

Department of Electrical Engineering and Computer Science
Northwestern University, Evanston, IL 60208-3118

ABSTRACT

This paper considers a singularly perturbed hybrid system whose state equations are governed by a stochastic switching process, which is singularly perturbed and is modeled as a near decomposable continuous time finite state Markov chain (FSMC). The decomposition of the system and the switching process together into slow and fast subsystems is investigated. An approximate model for the slow subsystem over the interval of fast transitions within each group is derived and the mean-squared error between the model and the actual subsystem is quantified. The stability of the slow mode subsystem is studied and two stability criteria are introduced. The behavior of the fast subsystem depending on the perturbation of the system and the switching process is investigated. Finally an example is used to illustrate the aforementioned techniques.

I. INTRODUCTION AND PROBLEM FORMULATION

1.1 Introduction

This paper studies the asymptotic behavior of the trajectory of a singularly perturbed hybrid system whose state equations depend on a near decomposable continuous time finite state Markov chain (FSMC). The state space of a stochastic hybrid system is a cross product of an Euclidean space and a finite discrete state space. Basically, stochastic hybrid systems are a special type of linear, piecewise constant, time varying systems which switch randomly among a finite number of linear time invariant models. The switching behavior is similar to that of an FSMC. Such systems have been successfully used to model pilot commands in target tracking, isolation levels of solar receivers, and systems subject to sudden changes in their structure and parameters which are caused by phenomena, such as component/ sensor failures or repairs, abrupt environmental disturbances and changing system interconnections in manufacturing systems and large scale flexible structures [1]. This paper is concerned with the asymptotic approximation of singularly perturbed stochastic hybrid systems whose slow continuous states and the switching process are singularly perturbed. The study of systems of this type is motivated by new applications, such as analysis of singularly perturbed systems containing quantized elements or on-off control [2], and simplified filtering schemes for singularly perturbed switched parameter systems [3].

Singular perturbation methods in [4-6] are used in the paper to decompose the continuous states and the switching process together into slow and fast mode subsystems. The methods alleviate the problems of stiffness difficulties resulting from the interaction of slow and fast dynamics. Singularly perturbed FSMCs and singularly perturbed stochastic hybrid systems have been investigated by several researchers in [3,7-10]. Aggregation methods were used to describe global features of singularly perturbed FSMCs [7,8]. In [9] the authors developed aggregation and averaging ideas to deal with approximation of stochastic hybrid systems in which the switching process depends on its current discrete state and the continuous states. Another study [3] examined the limiting behavior of a class of singularly perturbed stochastic hybrid systems where the switching process is singularly perturbed and independent of the continuous states. The authors in [10] have generalized the results of [3] by allowing a much broader class of the switching process which consists of many groups of strongly interacting discrete states. In [10] an approximate model for the slow mode subsystem within each group is derived based on the probabilistic averaging procedure. Its accuracy is quantified with the restriction of commutability of the system matrices of all realizations within each group.

Our aims here are to apply the aggregation method to derive an approximate model for the slow mode subsystem over the interval of fast transitions within each group of the switching process, to quantify the accuracy of the approximate model without the restriction imposed by the proofs derived in [3,10], and, finally, to analyze more general singularly perturbed stochastic hybrid systems.

1.2 Problem Formulation

The system models under consideration are assumed to have the following state equations:

\[ x(t) = A_1[x(t)]z(t) + A_{12}[x(t)]y(t), \quad z(t_0) = x_0, \] (1)

\[ \dot{y}(t) = A_2[y(t)]z(t) + A_{21}[y(t)]x(t), \quad z(t_0) = y_0. \] (2)

where \( \mu > 0 \) is a small parameter, \( x(t) \in \mathbb{R}^n \) represents the slow mode of the system and \( y(t) \in \mathbb{R}^m \) the fast mode, and the process \( r(t) \) is the form index (or plant mode) which takes values in \( S = \{1, 2, \ldots, m\} \), and determines the system model at a particular time. All matrices are of proper dimensions and are random through their dependence on the values of the process \( r(t) \).

We further assume that \( r(t) \) can be modeled as a near decomposable FSMC which contains \( n \) groups of strongly interacting states, where the \( i \)th group consists of \( n_i \) fast states which form the subset \( \mathcal{S}_i = \{ j \in \mathbb{N} | 1 \leq j \leq n_i \} \), where \( \mathbb{N} = \{1, 2, \ldots\} \) and \( n_i \geq 2 \) and \( N = 0 \). Note that \( \sum_{i=1}^n n_i = n \) and \( \sum_{i=1}^n \mathcal{S}_i = \mathcal{S} \). The current values of the system models are denoted by an index, for example, \( kA_i \) will denote \( A_i[x(t)] \) when \( r(t) = j, j \in \mathcal{S}_i \), where \( k = 1, 2, 12, 21. \) Let the evolution of the process \( r(t) \) satisfy

\[ \frac{dP(t)}{dt} = P(t) \begin{pmatrix} F & G \\ 0 & F \end{pmatrix} \] (3)

where \( P(t) \) is an \( m \times m \) transition probability matrix at time \( t \). It is assumed that \( 0 < \mu < 1 \), matrices \( F, G \) and \( F + G \) are generators such that the process \( r(t) \) is irreducible. Furthermore, let each of the \( n \) groups be an irreducible and time reversible FSMC with generator \( F_i \). The generator \( F_i \) is the \( i \)th block in the block-diagonal matrix \( F \). This paper considers the stationary case, namely that the process \( r(t) \) has reached its steady state. It is also assumed that \( x(t), y(t) \) and \( r(t) \) are perfectly observed.

The behavior of the overall system depends on the relative size of \( \mu \) and \( \tau \) as they both are sufficiently small. There are three cases to consider: (i) \( \mu = o(\epsilon) \), (ii) \( \mu = o(\epsilon) \), (iii) \( \epsilon = o(\mu) \). Furthermore, due to the possibility of transforming the system with Eq. (1)-(2) into a decoupled system, we shall focus on the decoupled case (i.e., \( A_{12}[r(t)] = 0 \), and \( A_{21}[r(t)] = 0 \)), and then extend the results to the coupled case.

The following is the outline of the paper. Section II presents the basic mathematical tools for the behavior of a near decomposable FSMC and the properties of additive processes. Section III studies the limiting behavior of the decoupled slow mode subsystem. An approximate model for the slow mode subsystem over the interval of fast transitions among the fast states of each group is derived in Section 3.1 based on the aggregation method. Section 3.2 quantifies the mean-squared error between the approximate model and the actual system. The asymptotic stability of the decoupled slow mode subsystem is explored in Section 3.3. In Section IV, the limiting behavior of the decoupled fast mode subsystem is investigated when both \( \mu \) and \( \epsilon \) tend to zero. Section V examines the coupled case. Section VI considers an example to illustrate the methods used in Section V. Section VII concludes the paper.

II. MATHEMATICAL PRELIMINARIES

This section surveys some notations and results concerning the asymptotic behavior of near decomposable FSMCs, and the properties of additive processes. These results play important roles in studying the asymptotic behavior of the trajectory of the system given by Eq. (1)-(3).

2.1 Near Decomposable FSMCs

In the following the asymptotic behavior of a near decomposable FSMC is summarized. There have been a number of studies in the
literature concerned with the asymptotic approximation and aggregation of a singularly perturbed FSCM [7-8]. The fast transient of the process \( r(t) \) is formed of separate transients within the strongly coupled groups. Over a longer period, each group of the strongly coupled states can be treated as an aggregate state.

Let the switching process \( r(t) \) satisfy the MSST (multiple semi-stability) condition [6]. Thus, two time scales is sufficient to describe the global evolution of the process \( r(t) \). To analyze the process \( r(t) \), one needs the following notations [6,10] :

\[
\text{Let } \{ t \} \xrightarrow{\tau} \text{exp}(\mathcal{F} \cdot \tau) \text{ Then }
\]

\[
\{ t \} = \text{diag} \{ \tau_{i} \}, i = 1, \ldots, n_1,
\]

where \( \tau_{i} = 1, \ldots, n_1 \), \( E_{i} \) is an \( n \times n \) dimensional row vector and \( E_{i,j} = [c_{ij}]_{j \in S_{i}} \), and \( 1, \ldots, i \), is an \( n \times n \) dimensional column vector with the same elements \( E_{i} \). Furthermore, define an \( m \times n \) matrix \( V \) and an \( n \times m \) matrix \( U \) as follows:

\[
V = \text{diag} \{ 1, \ldots, 1 \}, \quad U = \text{diag} \{ E_{1}, \ldots, E_{n} \}.
\]

(4)

then \( UV = \{ t \} \), \( UV = I \), where \( I \) is an \( n \times n \) identity matrix.

For fast transient analysis of \( r(t) \), Eq. (3) can be expressed in the stretched time scale \( \tau = \frac{t}{\tau} \) to obtain

\[
P(\tau) = \mathcal{F}^{	au}.
\]

(5)

For aggregate analysis of the process, \( r(t) \) can be replaced by an aggregate process \( \hat{r}(t) \) taking values in \( \hat{S} = \{ 1, \ldots, n \} \). Let \( \hat{P}(\tau) \) be the transition probability matrix of the process \( \hat{r}(t) \). Then the evolution of the aggregate process \( \hat{r}(t) \) is governed by the following equation,

\[
\hat{P}(t) = \hat{P}(0) UGV, \text{ for } t \geq t_0, \tag{6}
\]

and \( UGV \) is the infinitesimal generator of the process \( \hat{r}(t) \). Further, the relation between \( P(t) \) and \( \hat{P}(t) \) is given by

\[
P(t) = V \hat{P}(t) U, \text{ for } t \geq t_0. \tag{7}
\]

Note that the process \( \hat{r}(t) \) is stochastically continuous and has a single ergodic class such that

\[
\lim_{t \to \infty} \hat{P}(t) = 1 - [\hat{\alpha}_{1}, \ldots, \hat{\alpha}_{n}], \tag{8}
\]

where \( \hat{\gamma} \) is the ergodic probability of the aggregate process \( \hat{r}(t) \) corresponding to the state \( \gamma \).

In this paper, the steady-state distribution of the process \( r(t) \) will be needed, i.e.,

\[
\text{Prob.}[r(t) = \gamma] = \hat{\alpha}_{\gamma}.
\]

(2.2) Additive Processes

In what follows additive processes on each group of the process \( r(t) \) in the stretched time scale \( \tau \) are introduced [10]. Earlier major works on additive processes appeared in [11,12]. Let \( r(t) \in S_{i} \). Define the function \( a_{ik}^{1}[r(t)] \) as the \( (k,l) \)th element of the matrix \( A_{1}^{1}(r(t)) \in S_{i} \). The current value of \( a_{ik}^{1}[r(t)] \) is denoted by \( a_{ik}^{1, < \tau \rangle} \) when \( r(t) = j \in S_{i} \). An additive process is defined as follows:

\[
Y_{i,k}(r(t)) = \int_{0}^{r(t)} a_{ik}^{1}[s(t)] ds.
\]

(9)

For a sufficiently large \( \gamma \), one would expect the process \( Y_{i,k}(r(t)) \) to be asymptotically normal distribution and to satisfy a central limit theorem:

\[
Y_{i,k}(r(t)) \sim N \left( \frac{\hat{\alpha}_{i}^{1}}{\sqrt{\tau}}, \frac{\sigma_{i,k}^{1}}{\sqrt{\tau}} \right), \text{ as } \tau \text{ is sufficiently large}, \tag{10}
\]

where \( \hat{\alpha}_{i}^{1} = \sum_{j \in S_{i}} \hat{\alpha}_{j} a_{kj}^{1, < \tau \rangle} \).

The values of \( \sigma_{i,k}^{1} \), \( i = 1, \ldots, n \), are positive [10]. For proof of the above limit theorems see [10], [11], [12].

The central limit theorems are useful in that we exploit them to derive the accuracy of the approximate model for the decoupled slow mode subsystem.

III. DECOUPLED SLOW MODE SUBSYSTEM

This section considers the asymptotic behavior of the decoupled slow mode subsystem given by

\[
\dot{z}(t) = a_{1}[r(t)]z(t), \quad z(t_0) = z_0.
\]

(11)

An approximate model for the subsystem over the interval of fast transitions within each group is derived based on the aggregation method. Its accuracy, mean-squared error, is quantified by using of the Magnus Expansion from Lie algebra which is known as the continuous analogue of the Baker-Campbell-Hausdorff (BCH) formula. Finally, the stability of the subsystem is discussed and two stability criteria are introduced.

3.1 An Approximate Model

In general, the state variables \( z(t) \) are not Markovian. However, the joint process \( \{ z(t)^{T}, r(t) \} \) is a Markov process whose state space is \( R^{p} \times S \). Here we formulate the joint probability density function \( p.d.f \) of \( \{ z(t)^{T}, r(t) \} \) which is denoted by \( p_{ij}(t) \),

\[
p_{ij}(z(t), r(t)) = \text{Prob.} \{ z(t) \leq z + dz, r(t) = j \in S_{i} \}. \tag{12}
\]

Recall that the process \( r(t) \) remains in the ergodic distribution for all \( t \geq t_0 \). Let

\[
p_{ij}(z(t), t) = \hat{\alpha}_{i} \hat{\alpha}_{j} p_{ij}(z, t), \tag{13}
\]

so that by our assumption

\[
\int_{S_{i}} p_{ij}(z(t), t) dz = 1, \quad i \in S_{i}, j \in S_{i}. \tag{14}
\]

Then define \( \bar{p}(z, t) = (\bar{p}_{1}(z, t), \ldots, \bar{p}_{m}(z, t))^{T} \). Obviously, the evolution of \( \bar{p}(z, t) \) is governed by a forward Kolmogorov's equation (master equation) [13 ch. 3, 14, 15]. To simplify notation we introduce the matrix operator

\[
\mathcal{L}^{*} = \text{diag} \{ \mathcal{L}_{ii}^{*} \}, \tag{15}
\]

where each diagonal entry is described by

\[
\mathcal{L}_{ij}^{*} p_{ij}(z, t) = - \sum_{k=1}^{\infty} \hat{\gamma}_{k} \frac{\partial}{\partial z_{k}} [a_{ij}^{1}(1)^{T} a_{ij}^{1}(z(t), t)], \tag{16}
\]

where \( (\cdot)^{T} \) denotes the kth component of the vector ( ). Hence, the forward Kolmogorov's equation is given by

\[
\frac{\partial \bar{p}(z, t)}{\partial t} = \mathcal{L}^{*} \bar{p}(z, t) + \left( -\mathcal{F}^{T} + C^{T} \right) \bar{p}(z, t), \tag{17}
\]

\[
\bar{p}(z_0, t_0) \text{ given}. \tag{18}
\]

Because the exact solution of Eq. (17) is difficult to find, the singular perturbation method is applied to derive asymptotic representations of Eq. (17). This approach does not require the explicit solution of Eq. (17), but rather, leads directly to asymptotic expansions of \( \bar{p}(z, t) \). The basic idea is to let \( \bar{p}(z, t) \) have the outer expansion

\[
\bar{p}(z, t) = \sum_{k=0}^{\infty} \bar{p}_{k}(z, t), \tag{19}
\]

where \( \bar{p}_{k}(z, t) = (\hat{\alpha}_{1} \hat{\alpha}_{2} \cdots \hat{\alpha}_{m} \bar{p}_{k}(z, t))^{T} \). Substituting this in Eq. (17) and equating the coefficients of like powers of \( \hat{\gamma} \), we obtain
where the symbol \([\ldots]\) is the commutator product or the Lie product.

To find the limiting behavior of the state transition matrix \(\Phi(t_0, t_0 + \delta)\) as \(\epsilon\) tends to zero, we treat the terms in the exponent separately. Then, combining all these together yields

\[
\Phi(t_0, t_0 + \delta) = \exp\{t_0 \hat{A}_0 \epsilon + \delta ([1 - t_0 \hat{A}_0 \epsilon]^{-1} - \hat{A}_0) + O(\epsilon^2)\} \tag{28}
\]

From Lemma 1 it follows that all elements of \(Y_{\epsilon}(t_0, t_0)\) are mutually independent, asymptotically normal in distribution with variance \(O(\epsilon)\) as \(\epsilon\) tends to zero. Therefore, it is easy to verify

\[
\lim_{\epsilon \to 0} E(\|z(t) - \hat{z}(t)\|^2) = 0, \quad t \in [t_0, t_0 + \delta]. \tag{29}
\]

To finish the approximation over the interval \([t_0, T_i]\) where \(T_i - t_0\) is the sojourn time of \(r(t)\) in the group \(S_i\), we choose a sufficiently positive small number \(\delta'\) and a positive integer \(K\) such that \(\delta' \leq \delta\) and \(T_i - t_0 = K\delta'\). The limiting behavior of the state transition matrix \(\Phi(t_0, T_i)\) is computed by

\[
\lim_{\epsilon \to 0} \Phi(t_0, T_i) = \lim_{\epsilon \to 0} \Phi(t_0 + K\delta', t_0 + (K - 1)\delta') \tag{30}
\]

in the mean-squared sense. This implies that

\[
\lim_{\epsilon \to 0} E(\|z(t) - \hat{z}(t)\|^2) = 0, \quad t \in [t_0, T_i]. \tag{31}
\]

Note that if \(T_i - t_0 = \infty\), the mean-squared error at time \(T_i\) may be unbounded even if \(\epsilon\) tends to zero. However, it is well known [20] that the irreducible FSMC, \(r(t)\), has a finite return time for each state, i.e., \(T_i - t_0\) is bounded with probability one. The following summarizes the result.

**Theorem 1** Suppose that all values of \(\hat{A}_{ij}, \quad j \in S_i\), are bounded. Then the solution \(z(t)\) to the problem Eq. (11) converges in the mean-squared sense to the solution \(\hat{z}(t)\) of the approximate model Eq. (13) as \(\epsilon\) tends to zero, i.e.,

\[
\lim_{\epsilon \to 0} E(\|z(t) - \hat{z}(t)\|^2) = 0, \quad z(t_0) = \hat{z}(t_0) \tag{32}
\]

for \(t \in [t_0, T_i]\) where \(T_i - t_0\) is the sojourn time of the \(r(t)\) process in the group \(S_i\).

Note that Theorem 1 does not show that the approximate model is valid over the entire interval \([t_0, \infty)\).

### 3.3 Stability

In the subsection the stability of the decoupled slow mode subsystem is studied. The stability criteria are based on the logarithmic norm [21]. In order to derive the stability criteria, a brief introduction to the notion of a logarithmic norm is given as follows:

**Definition:** The logarithmic norm associated with the induced matrix norm \(\|\cdot\|\) is defined by

\[
\mu(A) = \lim_{\lambda \to 0^+} \|I + \lambda A\| - 1 \tag{33}
\]

Two stability criteria are introduced to the decoupled slow mode subsystem.

**Theorem 2** The zero solution of the slow mode hybrid system to be almost sure exponentially stable, it is sufficient to have

\[
\sum_{i=1}^n \sum_{j=k+1}^{N_{ij}} e_{ij} \mu(\hat{A}_{ij}) < 0, \tag{34}
\]

and necessary to have

\[
\sum_{i=1}^n \sum_{j=k+1}^{N_{ij}} e_{ij} \mu(-\hat{A}_{ij}) > 0. \tag{35}
\]

We next show how the Magnus Expansion and Lemma 1 can be used to prove that \(z(t)\) can be approximated by \(\hat{z}(t)\) over an interval \([t_0, T_i]\) where \(T_i - t_0 \in (0, c)\) is finite.

Assume that \(r(t) \in S_i\) at the time interval \([t_0, T_i]\). Let \(0 \leq \delta_0 < 1\) and \(0 < \epsilon < \delta \ll T_i - t_0\). With the Magnus Expansion, the state transition matrix of the slow mode system over the interval \([t_0, t_0 + \delta]\) is given by

\[
\begin{align*}
\Phi(t_0, t_0 + \delta) &= \exp\int_{t_0}^{t_0 + \delta} \hat{A}(r(s)) ds \\
&= \frac{1}{2} \int_{t_0}^{t_0 + \delta} [\hat{A}'(r(s))] ds \\
&= \frac{1}{4} \int_{t_0}^{t_0 + \delta} [\hat{A}'(r(s))] ds. \int_{t_0}^{t_0 + \delta} \hat{A}(r(s^2)) ds ds ds ds dx \\
&+ \ldots 
\end{align*}
\]
IV. DECOUPLED FAST MODE SUBSYSTEM

This section examines the limiting behavior of the decoupled fast mode subsystem given by

\[ \mu \tilde{x}(t) = A_j \tilde{x}(t), \quad x(t_0) = x_0. \]  

(35)

Assume that all values of \( A_j \), are bounded, \( j \in J \), \( j \in S \), and \( r(t) = j_0 \) \( j \in S \). To analyze the state \( x(t) \), the standard time scale \( \theta = (t - t_0)/\mu \) is used. Hence, expressing the system given by Eq. (35) in the standard time scale \( \theta \) yields

\[ \tilde{Z} (\theta) = A_j [r(\theta)] Z(\theta), \quad Z(0) = Z_0. \]  

(36)

Similarly, the equation for the evolution of \( r(\theta) \) becomes

\[ \frac{dP(\theta)}{d\theta} = \Phi'(\theta) \left( \frac{\mu}{\epsilon} F + \mu G \right) P(\theta) = I. \]  

(37)

Similar to Section 3.1, it is known that the joint process \( (Z(\theta))^T, r(\theta)) \) is a Markov process whose state space is \( R^* \times S \). Then we denote by \( p_{ij}(Z, \theta) \),

\[ p_{ij}(Z, \theta) = \text{Prob} \left\{ \frac{Z \leq Z(\theta)}{\leq Z + dZ}, r(\theta) = j \in S \right\}. \]  

(38)

The joint p.d.f. of the process \( (Z(\theta))^T, r(\theta))^T \). Defining \( p_1(Z, \theta) = (p_{11}(Z, \theta), \ldots, p_{1n}(Z, \theta))^T \), then the \( \Pi(Z, \theta) \) is governed by the forward Kolmogorov's equation shown below.

\[ \frac{\partial \Pi(Z, \theta)}{\partial \theta} = L^* \Pi(Z, \theta) + \left( \frac{\mu}{\epsilon} F^T + \mu G^T \right) \Pi(Z, \theta), \]  

(39)

where \( \Pi(Z, \theta) \) is given.

In the sequel singular perturbation methods are used to derive the asymptotic expansion of \( \Pi(Z, \theta) \). The limiting behavior of the solution of Eq. (39) depends on the relative size of \( \mu \) and \( \epsilon \) as both tend to zero. There are three cases to be considered:

**Case I.** \( \mu = O(\epsilon) \) and \( \lim_{\epsilon \rightarrow 0} \frac{\mu}{\epsilon} = 0 \).

In this case we introduce the symbol \( \mu' = \frac{\mu}{\epsilon} \) which tends to zero as \( \epsilon \) tends to zero. \( \Pi(Z, \theta) \) has the outer expansion process together into slow and fast subsystems depends on the relative size of \( \mu \) and \( \epsilon \). Theorem 4 addresses this problem for the case.

**Theorem 4** Assume that \( \mu = O(\epsilon) \) and \( \lim_{\epsilon \rightarrow 0} \mu \epsilon = k \) where \( k = O(1) \). Then \( Z(\theta) \) is approximately modeled as a hybrid system Eq. (36) depending only on the fast states of the \( j \)th group with generator \( F_k \) in the \( \theta \) time scale.

Case III. \( \epsilon = o(\mu) \) and \( \lim_{\mu \rightarrow \infty} \mu \epsilon = \infty \).

In this case we define the symbol \( \mu' = \mu \) where \( \mu' \) tends to zero as \( \epsilon \) tends to zero. Let \( \Pi(Z, \theta) \) have the outer expansion given by Eq. (40). Repeating the same procedure discussed before yields the same results as in Section III.

**Theorem 5** Suppose that \( \epsilon = o(\mu) \) and \( \lim_{\mu \rightarrow \infty} \mu \epsilon = \infty \) and \( r(t_0) \in S_i \) and all values of \( z_i \) are finite. Then, for \( t \in [t_0, T_i] \)

1. \( z_i(t) \) can be approximated as an average system whose dynamical equations are described by

\[ \mu \tilde{z}_i(t) = A_i \tilde{z}_i(t), \quad \tilde{z}_i(t_0) = z_i(t_0). \]  

(46)

2. The mean-squared error between \( z_i(t) \) and \( \tilde{z}_i(t) \) tends to zero as \( \epsilon \) tends to zero.

V. THE COUPLED CASE

This section considers the singularly perturbed stochastic hybrid system with Eq. (1) - (3). The purpose of this section is to define a slow mode subsystem that describes the slow dynamics and a fast subsystem that describes the fast dynamics. Their solutions are then used to approximate \( z_i(t) \) and \( z(t) \) over the interval of fast transitions within each group. The decomposition of the system and the \( \tau(r) \) process together into slow and fast subsystems depends on the relative size of \( \mu \) and \( \epsilon \). First, consider the case \( \mu = O(\epsilon) \) and \( \mu = O(\epsilon) \). The slow mode subsystem over the interval of fast transitions within each group is approximated by a linear time-invariant system model. The fast subsystem must be redefined over the interval of each fast transition of the \( \tau(r) \) process.

Firstly, the case \( 0 < \epsilon < \mu < 1 \) is considered. In this case the system can be regarded as a slow mode subsystem with symmetric matrix

\[ A(r(t)) \]  

with the initial conditions \( z_i(t_0) = x_i \) and \( \tilde{z}_i(t_0) = x_i \). Since all the submatrices are bounded, the solution to the system converges in the mean square sense to an approximate model with averaged system matrix

\[ A_i = \begin{bmatrix} A_i[r(t)] & A_i[r(t)] \\ A_i[r(t)] & A_i[r(t)] \end{bmatrix} \]  

(47)

Secondly, the case \( 0 < \mu < 1 \) is considered. In this case the system can be regarded as a slow mode subsystem with averaged system matrix

\[ A_i = \begin{bmatrix} \frac{1}{2} A_i & \frac{1}{2} A_i \\ \frac{1}{2} A_i & \frac{1}{2} A_i \end{bmatrix} \]  

(48)

when \( r(t) \) takes values in \( S_i \) for \( t_0 \leq t \leq T_i \).

To analyze the behavior of the system, it is assumed that all the values of system matrices \( \frac{1}{2} A_i \) are invertible and stable. Thus, a desired slow dynamics is given by

\[ \tilde{z}_i(t) = \left( t \tilde{A}_i - \frac{1}{2} A_i \tilde{A}_i - 1 \right) \tilde{z}_i(t), \]  

(49)

and a desired fast dynamics given by

\[ \mu \tilde{x}_f(t) = A_i \tilde{x}_f(t), \]  

(50)

with the initial condition \( \tilde{z}_f(t_0) = x_0 - \left( \frac{1}{2} A_i \right)^{-1} \tilde{z}_i(t_0) \) and \( \tilde{z}_i(t_0) = x_i \). Finally, the solutions of Eq. (49) and Eq. (50) are used to approximate the original slow and fast states.
Theorem 6 Assume that $0 < \epsilon \ll \mu < 1$, $r(t)$ takes values in $S_i$, for $t \in [0, T]$, and all the values of system matrices $\overline{\Lambda}_i$ are stable and invertible. Then
\[ \xi_j(t) = \tilde{z}_j(t) + O(\epsilon), \tag{51} \]
\[ z_i(t) = \tilde{z}_j(t) - (\overline{\Lambda}_i)^{-1}(\overline{\Lambda}_i)^{-1} \tilde{z}_i(t) + O(\epsilon) \tag{52} \]
where $\tilde{z}_j(t)$ and $z_i(t)$ are respective states of the slow model Eq. (49) and the fast model Eq. (50).

VI. AN EXAMPLE
An example to illustrate the method in Section V is demonstrated here. The systems are given by Eq. (1)-(3) where $z(t) \in R^1$, $\xi(t) \in R^1$, and the switching process $r(t)$, shown in Fig. 1, consists of two groups. Each group contains two strongly interacting states. Its generators are given by
\[ F = \begin{bmatrix} -2 & 0 & 0 \\ 3 & -3 & 0 \\ 0 & 0 & -2 \end{bmatrix}, \quad G = \begin{bmatrix} 0 & 0 & 0 \\ 0 & -1 & 1 \\ 0 & 0 & 0 \end{bmatrix} \]

The corresponding system matrices given by Eq. (47) of the states for the two groups are
\[ A_{11} = \begin{bmatrix} -2 & 1 \\ 2 & -3 \end{bmatrix}, \quad A_{21} = \begin{bmatrix} -3 & 2 \\ 3 & -4 \end{bmatrix} \]
\[ A_{11} = \begin{bmatrix} 2 & 1 \\ 2 & -3 \end{bmatrix}, \quad A_{22} = \begin{bmatrix} 3 & 2 \\ 3 & -4 \end{bmatrix} \]
respectively. Thus, aggregate models obtained from Eq. (48) for both groups are
\[ \tilde{A}_1 = \begin{bmatrix} -2.4 & 1.4 \\ 2.4 & -3.4 \end{bmatrix}, \quad \tilde{A}_2 = \begin{bmatrix} 2.4 & 1.4 \\ 2.4 & -3.4 \end{bmatrix} \]
The example considers two cases: (i) $\mu = 0.01$, $\epsilon = 0.001$, (ii) $\mu = 0.01$, $\epsilon = 0.0001$. Given the initial conditions $z_1(0) = [2.0, 3.0]^T$ and $[0, T] = [0, 0.3]$, the trajectories of $z_1(t)$ and $z_2(t)$ and $\tilde{z}_1(t)$ within the group $S_1, t = 1, 2$, are shown in Fig. 2 to Fig. 5 for the two cases. The simulation results illustrate that the approximate models are valid when $0 < \epsilon < \mu < 1$.

VII. CONCLUSIONS
This paper considered the asymptotic trajectory of a singularly perturbed hybrid system whose state equations depend on a near-decomposable finite state Markov chain. The limiting behavior of the decoupled slow mode subsystem over the interval of fast transitions within a group can be approximated by an averaged value of the system matrices over all their values depending on the fast states of the group. The mean-squared error between the approximate model and the fast model is $O(\epsilon)$. The stability of the decoupled slow mode subsystem is discussed and two criteria are introduced. The limiting behavior of the fast mode subsystem over the interval of fast transitions within each group depends on the relative size of $\mu$ and $\epsilon$ when both $\mu$ and $\epsilon$ tend to zero. The subsystem can be approximated as a time-invariant system with the constant system matrix held at the value at the initial transition when $\mu = O(\epsilon)$. The subsystem can be modeled as a hybrid system depending only on the fast states of each group when $\mu = O(\epsilon)$. The subsystem can be approximated as an averaged value of the system matrix over all their values depending on the fast states of each group when $\mu = O(\epsilon)$. In the coupled case approximate models for reduced order systems are investigated according to the relative size of the two perturbation parameters. The results are shown to hold when the switching process is stationary and irreducible, each group of strongly interacting states is irreducible and time-reversible.

Additional work remains to be done in the analysis of the limiting behavior of singularly perturbed hybrid systems with control, or with noise, or both by using the preceding methodology. This approach permits a unified treatment of approximate models. The results presented in this paper provide an initial step in facilitating the analysis of the behavior of singularly perturbed hybrid systems with control or noise.

References
Fig. 1 The Stochastic Switching Process \( r(t) \)

Fig. 2 \( \dot{X}_1(t) \) and \( X_1(t) \)

Fig. 3 \( \dot{X}_2(t) \) and \( X_2(t) \)

Fig. 4 \( \dot{Z}_1(t) \) and \( Z_1(t) \)

Fig. 5 \( \dot{Z}_2(t) \) and \( Z_2(t) \)
APPENDIX G

C. C. Tsai and A. H. Haddad

Stabilization of Stochastic Hybrid Systems

Submitted to
Proc. 30 IEEE Conference on Decision and Control

STABILIZATION OF STOCHASTIC HYBRID SYSTEMS

C. C. Tsai and A. H. Haddad

Department of Electrical Engineering and Computer Science
Northwestern University, Evanston, IL 60208-3118

ABSTRACT

This paper studies the stabilization of a stochastic hybrid system whose state equations are governed by a stochastic switching process, which is modeled as a continuous time finite state Markov chain (FSMC). First, Linear feedback laws with non-switching gains are proposed. The non-switching gains are computed based on the sufficient conditions derived for the definition of non-switching stochastic stabilizability. Secondly, Linear feedback laws with imperfect detectors are studied. The range of the detection probability for the detectors are computed. The results are shown to hold when the Markov chain is irreducible and the system states are perfectly observed.

I. INTRODUCTION AND PROBLEM FORMULATION

1.1 Introduction

The present paper is concerned with the stabilization of a class of stochastic hybrid systems. The state space of a stochastic hybrid system is a cross product of an Euclidean space and a finite discrete space. Basically, stochastic hybrid systems are a special type of linear, piecewise constant, time varying systems which switch randomly among a finite number of linear time invariant models. The switching behaves like a continuous time finite state Markov chain (FSMC). Such systems have been successfully used to model pilot commands in target tracking, isolation levels of

\[1\] This research was supported by the U. S. Air Force under grant AFOSR-89-0241
solar receivers, abrupt variation in the parameters of economic systems, and systems subject to sudden component/sensor failures or repairs, abrupt environmental disturbances and changing subsystem interconnections [1]. Preliminary work established the optimal control solutions for stochastic hybrid systems [2-7]. For a quadratic performance, the optimal linear feedback control law with switching gains has already been proposed. On the other hand, several schemes to stabilize stochastic hybrid systems are investigated in [1]. Another study [7] developed the new definition of stochastic stabilizability, and then established sufficient and necessary conditions for this definition. Indeed, most of previous techniques require both the continuous states and the value of the Markov chain to be measured, in order to implement on-line the feedback laws with switching gains. In practice some information, such as the complete knowledge about the Markov chain, is often difficult, if not impossible, to obtain. Some control strategies with less knowledge of the Markov chain would be more realistic [1,8].

Our aims here are to develop new stabilization schemes for a class of stochastic hybrid systems. The schemes require less knowledge of the Markov chain. These techniques are expected to aid in the design of controllers.

1.2 Problem Formulation

The system models under consideration are assumed to have the following state equations:

\[ \dot{x}(t) = A[r(t)]x(t) + B[r(t)]u(t), \]  

(1)

where \( t \in [t_0, T] \), \( T \) may be finite or infinite, \( x(t) \in \mathbb{R}^n \) represents the system states and \( u(t) \in \mathbb{R}^m \) the control. All matrices are of proper dimensions and are random through their dependence on the values of the random process \( r(t) \), called “form index”. The form index \( r(t) \) is governed by a continuous-time FSMC taking values in a finite set \( S = \{1, 2, ..., N\} \). The evolution of the form index \( r(t) \) with time is described by the state transition probabilities of associated FSMC on \( S \).
\[ \text{Prob.} \{ r(t + \Delta) = j | r(t) = i \} = \begin{cases} \lambda_{ij} \Delta + o(\Delta) & \text{if } i \neq j \\ 1 - \lambda_{ii} \Delta + o(\Delta) & \text{if } i = j \end{cases} \]

\[ \lambda_i = \sum_{j=1, j \neq i}^{N} \lambda_{ij} \]  

(2)

where \( \Delta > 0 \), and all values of \( \lambda \)'s are finite. Let \( A \) be the generator of the \( r(t) \) process. Assume that the initial values \( x_0, z_0 \) and \( r_0 \) are independent random variables; \( x_0 \) and \( z_0 \) are also independent of the \( \sigma \)-algebra generated by \( \{ r(t), t \in (t_0, T) \} \). The current value of the system model is denoted by an index, for example, \( A_i \) and \( B_i \) will denote \( A[r(t)] \) and \( B[r(t)] \) when \( r(t) = i \). The paper assumes that the \( x(t) \) is perfectly observed and the \( r(t) \) process is irreducible. The ergodic distribution of the \( r(t) \) process is given by

\[ \lim_{t \to \infty} \text{Prob.} \{ r(t) = j | r(0) = i \} = \epsilon_j, \ i, j \in S \]

The paper is organized as follows. Section II develops linear feedback laws with non-switching gains when the controllers are allowed to feedback only the continuous states. Linear feedback laws with switching gains are considered in Section III when practical detectors are used to observe the value of the Markov chain. Three illustrate examples are given in Section VI. Section V concludes the paper.

**II. Stabilization Via Non-switching Gains**

This section considers the scheme to stabilize the system given by Eq. (1)-(2) without any knowledge about the \( r(t) \) process. In [8] the author has showed that non-switching control gains for a class of stochastic hybrid systems may be preferable, in addition to the fact that they are much easier to implement. In what follow some notations and the definition of non-switching stochastic stabilizability are introduced. Sufficient conditions for the new definition are derived. The non-switching gains are computed based on the sufficient conditions.

Let \( x(t, x_0, u) \) denote the trajectories of the random processes \( x(t) \) from the initial states \( x(0) = x_0 \), under the action of the admissible control \( u(t) \) and every sample
path of $r(t)$. A new definition of non-switching stochastic stabilizability, modified from the definition in [7], is described below.

**Definition 1**—System (1)-(2) is said to be non-switching stochastically stabilizable if, for any finite $x_0 \in \mathbb{R}^n$, there exists a linear feedback control law $L$ that is constant for all values of $r(t) \in S$:

$$u(t) = -Lx(t)$$

such that there exists a symmetric positive definite matrix $M$ satisfying

$$\lim_{T \to \infty} E\{\int_0^T x'(t, x_0, u)x(t, x_0, u)dt | x_0\} \leq x_0'Mx_0$$

where $\|L\| < \infty$, and $x'(M')$ denotes the the transpose of the vector $x(t)$ (the matrix $M$).

From the above definition, non-switching stochastic stabilizability of a system means that there exists a linear feedback law which drives the $x$ states from any finite initial states $x_0$ asymptotically to the origin in the mean square sense. Sufficient conditions for non-switching stochastic stabilizability are derived as follows:

**Theorem 1**—System (1)-(2) is non-switching stochastically stabilizable if, for each form $i \in S$, there exist a control law $u(t) = -Lx(t)$ such that for any given positive definite symmetric matrix $N_i$, the (unique) set of symmetric solutions, $M_i$, of the $N$ coupled matrix equations

$$(A_i - B_iL - \frac{1}{2}\lambda_i)'M_i + M_i(A_i - B_iL - \frac{1}{2}\lambda_i) + \sum_{j=1,j\neq i}^N \lambda_{ij}M_j = -N_i \quad (3)$$

are positive definite for each $i \in S$.

Note that Theorem 1 does not require the assumption of the irreducibility of the $r(t)$ process. The proof of the theorem is similar to the proof in [6] except the unobserved value of the $r(t)$ process. In applying Theorem 1, we choose the control law $L$, let $N_i$ be identical or simple diagonal matrices, and then solve for Eq. (3) to obtain the symmetrical and positive definite matrices $\{M_i : i \in S\}$.
The sufficient conditions in Theorem 1 are difficult to check. There exits a simple necessary condition for non-switching stochastic stabilizability of system (1)-(2). If system (1)-(2) is non-switching stochastically stabilizable, then in each form $i$, $L$ can be chosen such that all the closed-loop system matrices $(A_i - B_iL - \frac{1}{2}\lambda_i)$ are stable.

The non-switching gains can be computed based on the following procedure. First, let the control law $L$ be chosen such that all the matrices $\{A_i - B_iL - \frac{1}{2}\lambda_i : i \in S\}$ are stable. Secondly, let $\{N_i : i \in S\}$ be identical or simple diagonal matrices, and then solve for Eq. (3) to obtain a set of symmetrical matrices $\{M_i : i \in S\}$. Finally, stop the procedure if all the matrices $M_i$ are positive definite. If not, go to the first step and repeat the procedure again. An example (Example 1) to illustrate such procedure is shown in Section IV.

III. STABILIZATION VIA SWITCHING GAINS WITH IMPERFECT DETECTORS

In the previous section, stabilization of stochastic hybrid systems by a non-switching linear constant feedback law was introduced. The main advantage of this stabilization scheme is lack of the need for detection and estimation of the $r(t)$ process. However, despite the simplicity of this scheme it does not permit a large class of such systems. To alleviate the shortcoming, switching gain stabilization is considered in this section. The scheme requires a form index detector to detect the current value of the $r(t)$ process. The detector is with the following characteristics:

- All jump times of the $r(t)$ process can be detected.

- Let $r^*(t)$ denote the output of the detector. The value of the $r^*(t)$ over the interval $[t_k, t_{k+1}]$, where $t_k$ and $t_{k+1}$ are two successive jump times of the $r(t)$ process, remains constant. The relationship between $r(t)$ and $r^*(t)$ at the jump time $t_k$ is given by

$$\text{Prob.}\{r^*(t_k) = j | r(t_k) = i, \} = \begin{cases} p & \text{if } i = j \\ q & \text{if } i \neq j \end{cases}$$
where $p$ is the detection probability of the detector for each form and $q = \frac{1-p}{N-1}$.

In the scheme the linear feedback control law depends on the system states and the value of the detector, i.e.,

$$u(t) = -L[r^*(t)]x(t),$$

where $\|L_i\| < \infty$. Thus, the closed-loop system becomes

$$\dot{x}(t) = \bar{A}[r(t), r^*(t)]x(t)$$

(4)

where $\bar{A}[r(t), r^*(t)] = A[r(t)] - B[r(t)]L[r^*(t)]$.

In what follows the stability of the system given by Eq. (4)-(2) is studied and two stability criteria are introduced. To have the stability criteria, a brief review to the notation of logarithmic norm is given. The logarithmic norm (also called the measure of matrix) was investigated in 1958 separately by Dahlquist [9] and Lozinskij [10]. The properties of the norm have been well documented in [11]. The norm has been applied extensively to study the growth of the solution of linear, time varying systems. Below is the definition of the logarithmic norm.

**Definition 2**: The logarithmic norm associated with the induced matrix norm $\| \cdot \|$ is defined by

$$\mu(A) = \lim_{\theta \to 0} \frac{\| I + \theta A \| - 1}{\theta}$$

With the norm $\mu$ and the irreducibility of the $r(t)$ process, we derive the following important lemma in that we use it to find the conditions for the stability criteria of the system given by Eq. (4)-(2).

**Lemma 1** Let the $r(t)$ process be irreducible. Then

$$\lim_{T \to \infty} \frac{1}{T} \int_0^T \mu(\bar{A}[r(t), r^*(t)]) dt = \sum_{i=1}^N e_i \{ p\mu(\bar{A}_{ii}) + q \sum_{j=1, j \neq i}^N \mu(\bar{A}_{ij}) \}, \text{ w.p.} 1$$

where $\bar{A}_{ij}$ denote the current value of $\bar{A}[r(t), r^*(t)]$ when $r(t) = i$, $r^*(t) = j$.

**Proof**: Let $T_i$ denote the total sojourn time over the interval $[0, T]$ for each form $i$ of the form index $r(t)$. Since $r(t)$ is an irreducible FSMC, it is well known [12] that
Further, $T_{ij}$ is defined as the total time when $r(t) = i$ and $r^*(t) = j$ over interval $[0, T]$. It is easy to show that

$$
\lim_{T \to \infty} \frac{\mu(\bar{A}_{ij}) T_{ij}}{T} = \begin{cases} 
pe_i \mu(\bar{A}_{ii}) & \text{if } i = j, \ w.p.1 \\
qe_i \mu(\bar{A}_{ij}) & \text{if } i \neq j, \ w.p.1
\end{cases}
$$

This completes the proof.

**Theorem 2** The null solution of the system given by Eq. (4)-(2) is almost sure exponentially stable if it is sufficient to have

$$
\sum_{i=1}^{N} e_i \{ p \mu(\bar{A}_{ii}) + q \sum_{j=1, j \neq i}^{N} \mu(\bar{A}_{ij}) \} < 0
$$

and necessary to have

$$
\sum_{i=1}^{N} e_i \{ p \mu(-\bar{A}_{ii}) + q \sum_{j=1, j \neq i}^{N} \mu(-\bar{A}_{ij}) \} > 0
$$

**Proof:** From the theorem in [11, pp. 89] it follows that

$$
\|x_0\| e^{-\int_0^T \mu(\bar{A}(t), r^*(t)) dt} \leq \|x(T)\| \leq \|x_0\| e^{\int_0^T \mu(\bar{A}(t), r^*(t)) dt}
$$

As $T$ tends to infinity, Lemma 1 yields the results.

If $p=1$, i.e., the detector is perfect, the results are shown in [1]. In other words, if the system states and the value of the $r(t)$ process are perfectly observed, and if the linear feedback law with switching gains satisfies

$$
\sum_{i=1}^{N} e_i \mu(\bar{A}_{ii}) < 0 \tag{5}
$$

then the null solution of the closed-loop system is almost sure exponentially stable. There arises a interesting problem: if the control law with switching gains is designed to satisfy the condition given by Eq. (5), what is the range of $p$ such that the closed system is almost sure exponentially stable when the imperfect detector is used? The following corollary answers the problem.
Corollary 1 if \( \sum_{i=1}^{N} e_i \mu(\bar{A}_{ii}) < 0 \) and \( \sum_{i=1}^{N} e_i \{ \sum_{j=1,i \neq j}^{N} \mu(\bar{A}_{ij}) \} < 0 \), and \( 0 \leq p \leq 1 \), then the system given by Eq. (4)-(2) is almost sure exponentially stable.

Corollary 2 if \( \sum_{i=1}^{N} e_i \mu(\bar{A}_{ii}) < 0 \) and \( \sum_{i=1}^{N} e_i \sum_{j=1,j \neq i}^{N} \mu(\bar{A}_{ij}) > 0 \), and

\[
p > 1 + \frac{\sum_{i=1}^{N} e_i \mu(\bar{A}_{ii})}{\sum_{i=1}^{N} e_i \sum_{j=1,j \neq i}^{N} \mu(\bar{A}_{ij}) - \sum_{i=1}^{N} e_i \mu(\bar{A}_{ii})}
\]

then the system given by Eq. (4)-(2) is almost sure exponentially stable.

IV. EXAMPLES

In this section three examples are provided to illustrate the methods derived in Section II and III. The first example demonstrates that a system given by Eq. (1)-(2) can be stabilized by using only a linear feedback law with nonswitching gains, i.e., any control law with switching gains does not satisfy the sufficient condition (or the necessary condition) given by Eq. (5). The second example examines a system which can be stabilized via both two methods discussed before. The range of the detection probability for the detector is computed. Finally a system which can not be stabilized by the method in Section II but can be stabilized by the method of Section III is considered.

Example 1: Consider a system with the form index \( r(t) \) taking values in a finite set \( S = \{1, 2\} \) with the generator

\[
A = \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix}
\]

The system and input matrices are given by

\[
A_1 = \begin{bmatrix} 0 & 1 \\ 2 & 3 \end{bmatrix}, B_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}
\]

\[
A_2 = \begin{bmatrix} 0 & 1 \\ 4 & 5 \end{bmatrix}, B_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}
\]

This example chooses the linear feedback laws \( L \) and \( \{ N_i : i \in S \} \) as follows:
Solving for Eq. (3) yields

\[ M_1 = \begin{bmatrix} 0.7634 & 1.2792 \\ 1.2792 & 7.9064 \end{bmatrix}, \quad M_2 = \begin{bmatrix} 2.2798 & 0.2418 \\ 0.2418 & 0.4156 \end{bmatrix} \]

Since the solutions \( M_1 \) and \( M_2 \) are symmetric and positive definite, the system is non-switching stochastically stabilizable. However, this system cannot be stabilized by the method in Section III. Given the logarithmic norm associated with the induced matrix norm \( \| \cdot \|_1 \) (or \( \| \cdot \|_2 \) or \( \| \cdot \|_\infty \)), for any set of designed switching gains \( L_1 \) and \( L_2 \), it is shown that

\[ e_1 \mu(A_1 - B_1 L_1) + e_2 \mu(A_2 - B_2 L_2) > 0 \]

\[ e_1 \mu(A_1 - B_1 L_2) + e_2 \mu(A_2 - B_2 L_1) > 0 \]

The sufficient condition given in Theorem 2 does not hold.

**Example 2:** Consider a system with

\[ A = \begin{bmatrix} -2 & 2 \\ 3 & -3 \end{bmatrix}, \]

\[ A_1 = \begin{bmatrix} 0 & 1 \\ 2 & 3 \end{bmatrix}, B_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, A_2 = \begin{bmatrix} 0 & 1 \\ 4 & 5 \end{bmatrix}, B_2 = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \]

The stationary distribution vector of the \( r(t) \) process is given by

\[ [e_1, e_2] = [\frac{3}{5}, \frac{2}{5}] \]

Given the logarithmic norm associated with the induced matrix norm \( \| \cdot \|_1 \) and \( L_1 = [2, 4], L_2 = [2, 5] \), we have

\[ e_1 \mu(A_1 - B_1 L_1) + e_2 \mu(A_2 - B_2 L_2) = -0.4 > 0 \]

\[ e_1 \mu(A_1 - B_1 L_2) + e_2 \mu(A_2 - B_2 L_1) = 1.2 > 0 \]
Thus, from Corollary 2 it follows that $p > 0.75$. In other words, the detector must have the detect probability $p$ where $p > 0.75$.

On the other hand, this system can be stochastically stabilized by using a non-switching feedback controller. Let $L = [2, 3]$. Solving for Eq. (3) yields two symmetrical and positive definite matrices $M_1$ and $M_2$ where

$$M_1 = \begin{bmatrix} \frac{1}{4} & -\frac{4}{239} \\ -\frac{4}{239} & \frac{1}{239} \end{bmatrix}, M_2 = \begin{bmatrix} \frac{1}{38} & -\frac{7}{144} \\ -\frac{7}{144} & \frac{1}{144} \end{bmatrix}$$

**Example 3:** Consider the two-form system with

$$A_1 = \frac{1}{3}, \quad B_1 = 1, \quad A_2 = \frac{4}{3}, \quad B_2 = -1$$

and the generator of the $r(t)$ process

$$A = \begin{bmatrix} -2 & 2 \\ 3 & -3 \end{bmatrix}$$

Obviously, this system is not nonswitching stochastically stabilizable. We design $L_1 = 2, L_2 = -3$ such that the sufficient condition

$$\frac{3}{5} \mu(A_1 - B_1 L_1) + \frac{2}{5} \mu(A_2 - B_2 L_2) = -\frac{5}{3} < 0$$

is satisfied. Then we compute

$$\frac{3}{5} \mu(A_1 - B_1 L_2) + \frac{2}{5} \mu(A_2 - B_2 L_1) = \frac{10}{3} > 0$$

From Corollary 2 it follows that $p > \frac{2}{3}$. Thus, to stabilize the system, the detector must have the detect probability $p$ where $p > \frac{2}{3}$.

**V. CONCLUSIONS**

This paper considered the stabilization of stochastic hybrid systems whose state equations depend on continuous time finite state Markov chains. Non-switching feedback laws have already been studied when the controllers are allowed to feedback only the system states. The non-switching gains are calculated based on the sufficient condition for non-switching stochastic stabilizability. Furthermore, linear feedback laws
with the practical detectors are proposed. The range of the detection probability of the detector is computed according to the sufficient conditions for the almost sure exponential stability of the closed-loop systems.

Additional work remains to be done in stabilizing singularly perturbed stochastic hybrid systems which have been studied in [13,14]. The results presented in the paper provide a initial step in facilitating the work.

Finally these results may be extended to the optimal control problems for singularly perturbed stochastic hybrid systems.

References


APPENDIX H

B. Park and E. I. Verriest

Canonical Forms on Discrete Linear Periodically Time-Varying Systems
and a Control Application

Proc. 28th IEEE Conf. on Decision and Control

Canonical Forms on Discrete Linear Periodically Time-varying Systems and A Control Application

B. Park and E.I. Verriest
School of Electrical Engineering
Georgia Institute of Technology
Atlanta, GA 30332

ABSTRACT

Canonical forms for discrete linear N-periodically time-varying (LP) completely reachable systems \( x(k) = A(k)x(k) + B(k)u(k) \) and \( y(k) = C(k)x(k) + D(k)u(k) \) are presented, which generalize the linear time-invariant (LTI) case. The derivation is first accomplished through an equivalent LTI-quadruple to 4N-tuple \([(A_k, B_k, C_k, D_k)]_{k=1,...,N-1}\). This LTI system is revealed to be a subcomponent in a decomposition of a given discrete LP system represented by the 4N-tuple. Finally, an application of the obtained canonical forms is demonstrated in a control problem: eigenvalue assignment of the monodromy matrix.

I. INTRODUCTION

Linear periodically time-varying (LP) systems have been studied by many researchers [6, and references therein]. LP systems are suitable models for some periodic behaviors such as seasonal phenomena and rhythmic biological movement. It has also been noted that LP controllers give linear time-invariant (LTI) plants more robust control in the maximum attainable gain margin sense [1][3].

Motivated by the above, we have investigated more precise mathematical descriptions of LP systems [7][8]. In this work, we primarily show a state-space canonical form for discrete LP systems \( \pi(k-1) = A(x(k)) + B(u(k), y(k) = C(x(k)) + D(u(k) \) where the quadruple is N-periodic \( (A_k, B_k, C_k, D_k) = (A_{k+N}, B_{k+N}, C_{k+N}, D_{k+N}) \). As for LTI systems, we believe that the canonical forms for the discrete LP systems play the same role in such problems as realization, control and identification.

In section II, a reachable canonical form is derived using a representative \( (A_k, B_k, C_k, D_k) \). In section III, a system decomposition \( S^*L^*S \) is derived for a given discrete LP system, and the equivalent quadruple is revealed as the realization matrices for the LTI system \( L \). In section IV, a typical application is demonstrated using the reachable canonical form along with the stability analysis and feedback connection rule developed in section III: eigenvalue assignment of the monodromy matrix for a completely reachable discrete MIMO LP system.

Throughout this work, LP systems are assumed to be of dimension \( n \) with \( m \) inputs, \( p \) outputs and period \( N \) such that \( (A_k, B_k, C_k, D_k) \in \mathbb{R}^{n\times n} \times \mathbb{R}^{n\times m} \times \mathbb{R}^{m\times n} \times \mathbb{R}^{m \times m} \) for \( k \in \{0, 1, \ldots, N-1\} \). Bold face characters are reserved for matrices and vectors of 'big dimension' (a multiple of \( N \)) such as \( (A_k, B_k, C_k, D_k) \in \mathbb{R}^{n\times n} \times \mathbb{R}^{n\times m} \times \mathbb{R}^{m\times n} \times \mathbb{R}^{m \times m} \) and \( f(k) \in \mathbb{R}^m \). The superscript '*' is used to indicate a set such as \{0,1,2,\ldots, N-1\} = \( N \).

II. STATE-SPACE CANONICAL FORMS

Before deriving the canonical forms, let us recall the significance in general setting. Let an 'objective' function \( F \) on a parameter set \( X \) be given. In many situations, the function \( F \) is many-to-one so that we can find an invariant partitioning \( X \subset X : \bigcup_{i \neq j} X_i = X \) for some index set \( I, X_i \cap X_j = \varnothing \) for \( i \neq j \)

\[ F(x) = F(y) \iff x \equiv y \in X \] (1)

The parameter set \( X \) is then too 'redundant' with respect to the function \( F \). Therefore, it is natural to select a representative or a canonical element \( x \in X_i \) for the subset \( X_i \). This selecting process would be understood as a map \( A \) from \( X \) into \( X_i \). Such map \( A \) is called a canonical form for the parameter set \( X \) under the partitioning \( \{X_i : i \in I\} \) with respect to the objective function \( F \). The function \( F \) on \( X \) is 'simplified' by restricting its domain to the subset \( \{x : i \in I\} \) without missing the original objective

\[ F = F|_{\{x_i\}} \] (2)

More specifically, consider a parameter space \( \{(A_k, B_k, C_k, D_k)\}_{k \in \mathbb{N}} : (A_k, B_k)_{k \in \mathbb{N}} \) is completely reachable, an equivalence relation \( R \), and an objective function \( H \) such that \( H \) is the vector valued pulse response or its one sided z-transformation of a discrete linear periodically time-varying (LP) system \( LP((A_k, B_k, C_k, D_k))_{k \in \mathbb{N}} \)

\[ H(z, i) = \left[ C_{A1}A_{i1} \cdots C_{A_N}A_{iN} \cdots C_{A1}z^{-N+1} + \cdots + C_{A1}z^{-1} \right] \]

\[ \left[ (z^{N-1} - A_1 \cdots A_{N-1} \cdots A_{N-1}z^{-N+1} + \cdots + C_{A1}z^{-1})D_{kN} \right] \text{ for } i \in \mathbb{N}^* \] (3)

\[ LP((A_k, B_k, C_k, D_k))_{k \in \mathbb{N}} : u(k) \rightarrow y(k) \]

\[ x(k+1) = A_kx(k) + B_ku(k), y(k) = C_kx(k) + D_ku(k) \]

\[ (A_k, B_k, C_k, D_k) = (A_{k+N}, B_{k+N}, C_{k+N}, D_{k+N}) \text{ for all } k \in \mathbb{N} \] (4)

\[ (A_k, B_k)_{k \in \mathbb{N}} \text{ is called completely reachable if each reachability matrix } R((A_k, B_k))_{k \in \mathbb{N}} \text{ for } i \in \mathbb{N}^* \text{ is full rank} \]

\[ R((A_k, B_k))_{k \in \mathbb{N}} = [E_{i} A_{i1} A_{i2} A_{i3} \cdots] \] (5)

This matrix relates the state \( x(i-1) \) to past inputs \( u(k), k \leq i \).

\[ LP((A_k, B_k, C_k, D_k))_{k \in \mathbb{N}} : u(k) \rightarrow y(k) \]

\[ \{A_k, B_k, C_k, D_k\}_{k \in \mathbb{N}} : R((A_k, B_k, C_k, D_k))_{k \in \mathbb{N}} \text{ is completely reachable } \] in (6) under \( Rp \) in (7) with respect to \( H \) in (4). Since a system (5) is involved, such \( A \) is simply called a reachability canonical form for discrete LP systems.

In our derivation, it will be very convenient to consider an equivalent quadruple \( (A_k, B_k, C_k, D_k) \) to a 4N-tuple \( ((A_k, B_k, C_k, D_k))_{k \in \mathbb{N}} \) induced by a map \( E \)

\[ E : ((A_k, B_k, C_k, D_k))_{k \in \mathbb{N}} \rightarrow (A_k, B_k, C_k, D_k) \] (8)
The map \( E \), which is called hereafter the extended form, preserves the complete reachability of \( (A, B_1, C, D) \) in (6) and the equivalence relation \( R_p \) in (7) in the following sense.

**LEMMA 1:** \( (A, B_1) \) is completely reachable iff \((A, B_1)\) is reachable.

**PROOF:** Observe the reachability matrix

\[
R(A, B_1) = \begin{bmatrix} B_1 & A \end{bmatrix} A \begin{bmatrix} B_1 & A \end{bmatrix} & \cdots & \begin{bmatrix} B_1 & A \end{bmatrix} A^{n-1} \begin{bmatrix} B_1 & A \end{bmatrix} & \cdots & \begin{bmatrix} B_1 & A \end{bmatrix} A^{m-1} \begin{bmatrix} B_1 & A \end{bmatrix} \\
\end{bmatrix}
\]

where \( T_i \) denotes those column change operations. Since \( (A, B_1) \) is the equivalence relation defined in (6), we proved the lemma.

**LEMMA 2:** Let \( R_p \) be an equivalence relation defined by:

\[
(R((A_k, B_k), C_k, D_k))_k \in \mathbb{N} - \text{complete reachability}
\]

iff there exists a diagonal nonsingular matrix \( T_D = \begin{bmatrix} T_{11} & \cdots & T_{1n} \\
0 & \cdots & 0
\end{bmatrix} \) such that

\[
A_1T_D = T_D A_1, B_1 = T_D B_1, C_1T_D = C_1, D_1 = D_1.
\]

Then,

\[
(A_1, B_1, C_1, D_1) \text{ is a canonical form as in } (11) \iff (A_1, B_1, C_1, D_1) \text{ is reachable.}
\]

**PROOF:** By direct calculation of (11), we obtain (7). \( \therefore \)

For showing our main result, let \( (A_1, B_1, C_1, D_1) \) be a canonical form as in (11) \( \Rightarrow \)

\[
E((A_k, B_k), C_k, D_k) \text{ and } (A_1, B_1, C_1, D_1) \text{ are respectively reachable and completely reachable.}
\]

**THEOREM 1:** If a map \( T_D \) is a canonical form for \( (A_1, B_1, C_1, D_1) \) under \( R_p \), then a composite map \( E^{-1}T_D \) is also a canonical form for \( (A_1, B_1, C_1, D_1) \) under \( R_p \).

**PROOF:** (i) The map \( E^{-1}T_D \) should satisfy (3).

\[
E((A_k, B_k), C_k, D_k) = E((A_1, B_1, C_1, D_1)) \Rightarrow E^{-1}T_D E((A_k, B_k), C_k, D_k) = E^{-1}T_D E((A_1, B_1, C_1, D_1)).
\]

(ii) The map \( E^{-1}T_D \) should satisfy (3).

\[
E((A_k, B_k, C_k, D_k)) = E((A_1, B_1, C_1, D_1)) \Rightarrow E^{-1}T_D E((A_k, B_k, C_k, D_k)) = E^{-1}T_D E((A_1, B_1, C_1, D_1)).
\]

We proved the theorem. \( \therefore \)

Although it is relatively easy to find a canonical form \( E((A_k, B_k, C_k, D_k)) \) such as the 'reachability' canonical form with the Scheme \( II \) [2] (since the notions of reachability and controllability are different for discrete time systems), we need to find a canonical form \( E \), which is of the extended form as in (8). This is because we can apply the inverse map \( E^{-1} \) to obtain a canonical \( 4N \)-tuple according to **THEOREM 1**.
By assumption of the complete reachability, \( R(W, \Phi) \) is full rank for each \( i \), and the number of independent columns is now the dimension \( n \) of the monodromy matrix \( \Phi \). In step (ii), we might order independent cells according to the ordering procedure of the usual Scheme II. The resulting transformation would yield the reachability canonical form with the Scheme II \( (A_{12n}, B_{12n}, C_{12n}, D_{12n}) \), which is similar to the canonical quadruple \( (A_{1m}, B_{1m}, C_{1m}, D_{1m}) \) in (16) by column and row permutation.

**EXAMPLE 1:** If \( (C_i, D_i) \) is arbitrary, and \( ((A_i, B_i)) \) is given as

\[
\begin{pmatrix}
48 & 192 & -4 & -42 & 30 \\
3 & 12 & 0 & -1 & 4 \\
-16 & -63 & 0 & 6 & -16 \\
22 & -88 & 1 & 15 & -19 \\
10 & 40 & 0 & -4 & 12 \\
& 0 & -3 & -4 & -1 & -8 \\
& 0 & 20 & -2 & -4 & -19 \\
& 0 & 21 & 9 & -1 & 5 \\
& 0 & -38 & 7 & 9 & 45 \\
\end{pmatrix}
\]

(i) Searching dependent columns:

\[
R_0((A_i, B_i))_{\mathbb{C}^L} = [B_0 | A_0 B_2 | A_0 A_2 B_3 | A_0 A_3 A_1 | B_0 |]
\]

\[
= \begin{pmatrix}
-1 & -11 \\
0 & 0 \\
0 & 1 \\
0 & 4 \\
0 & 1 \\
-2 & 0 \\
0 & 0 \\
0 & 1 \\
0 & 1 \\
-1 & -2 \\
-1 & 1 \\
-1 & 0 \\
\end{pmatrix}
\]

(ii) Ordering independent columns for each \( i \):

\[
T_0 = [B_0 | A_0 B_2 | A_0 A_2 B_3 | A_0 A_3 A_1 | B_0 ]
\]

(iii) Form \( T_D \) in (15) with (27) to (29). Applying (16) and utilizing (21) to (26), we obtain (there is no need for numerically calculated \( T_D \) yet):

\[
A_{12n} =
\]

\[
\begin{pmatrix}
0 & -4 & 0 & 0 & 2 \\
1 & -3 & 0 & 0 & 0 \\
0 & 2 & 0 & 0 & -2 \\
0 & 3 & 1 & 0 & -5 \\
0 & 0 & 0 & 1 & -4 \\
\end{pmatrix}
\]

Calculate \( T_D \) in (15) with (27) to (29) and (16) to (20). Applying (16), we obtain \( C_{12n} \) and \( D_{12n} \).

(iv) Read out nonzero block in \( (A_{12n}, B_{12n}, C_{12n}, D_{12n}) \) according to \( E^{-1} \), and we finally obtain \( ((A_{12n}, B_{12n}, C_{12n}, D_{12n}))_{\mathbb{C}^L} \).

In fact, each value of a canonical form is completely specified by two things, a structural index \( \kappa \) and a parametric quantity \( \theta \): \( \kappa \) tells which positions of all matrices \( E^{-1}T_{12n}E((A_i, B_i, C_i, D_i))_{\mathbb{C}^L} \) are fixed to 0 or 1, and \( \theta \) gives real values for the nonfixed positions. In the canonical quadruple \( (A_{12n}, B_{12n}, C_{12n}, D_{12n}) \) of EXAMPLE 1, all entries of \( (C_{12n}, D_{12n}) \) are arbitrarily determined, and the invariants are only found in \( (A_{12n}, B_{12n}) \) in (30). The ordered set of real numbers on the nonfixed positions in \( (A_{12n}, B_{12n}) \) in (30) and the set of indices indicating the fixed 0 and 1 entry positions are respectively values of the parametric complete invariant \( \theta \) and structural invariant \( \kappa \). Specifically, since the reachability index \( (2,3,2,3,2,3) \) determines the fixed 0 and 1 entry positions, we can write

\[
\alpha T_{12n}(A_i, B_i) = (2,2,3,2,3,3)
\]

The range space of \( \kappa \), however, is a proper subset

\[
\alpha T_{12n}((A_i, B_i)) \subset \{(k_1, k_2, \ldots, k_{Nn}) : \sum_{j=1}^{Nn} k_j = Nn \}
\]

An index \( (2,2,3,2,3,1,4) \), for instance, does not correspond to \( \kappa \) for reachable \( (A_i, B_i) \) in the case \( n = 5, m = 2 \) and \( N = 3 \), since there are respectively 3, 6 and 6 independent cells (columns) in groups of 0, 1 and 2 numbered cells (in matrices of \( R_0((A_i, B_i))_{\mathbb{C}^L} \), \( R_1((A_i, B_i))_{\mathbb{C}^L} \) and \( R_0((A_i, B_i))_{\mathbb{C}^L} \)), which violates the assumption of the complete reachability as mentioned in REMARK 1. The parametric complete invariant \( \theta \) is indeed the 'simplified' parameter. All quadruples \( (A_i, B_i, C_i, D_i) \) (and hence \( ((A_i, B_i, C_i, D_i))_{\mathbb{C}^L} \)) are parameterized by \( \kappa \) and \( \theta \).

### III. A SYSTEM DECOMPOSITION

In the foregoing section, the extended form \( (A_i, B_i, C_i, D_i) \) played a key role in the derivation of the canonical form for discrete LP systems. Now, a question naturally arises what the relationship is between the discrete LP system \( LP((A_X, B_X, C_X, D_X)) \).
output transformation in (39) yields LTI system \( L(A, B, C, D) \) in (40). Moreover, the input and output transformation in (41) applied to (40) yields the quadruple in (44) is not of the extended form. But by transforming the quadruple \( (44) \) becomes as follows:

\[
M_s^{-1}A_sM_s A_s \to \tilde{A}_s, M_s^{-1}B_sM_s A_s \to \tilde{B}_s, M_s^{-1}C_sM_s A_s \to \tilde{C}_s, M_s^{-1}D_sM_s A_s \to \tilde{D}_s
\]

(40)

Observe that matrix \( M_s \) is the generator of a cyclic group of order \( N \) and \( \tilde{A}_s \) is its I/O-equivalent from the other invariant models. The extended invariant system of \( L(A, B, C, D) \) is called the extended invariant system of \( L(A, B, C, D) \).

LEMMA 6: Assume the dimensions of all subsystems are compatible. The parallel (addition) or serial (multiplication) connection of the extended invariant systems as in (35) again yields an extended invariant system.

PROOF: Let \((A_1, B_1, C_1, D_1)\) and \((A_2, B_2, C_2, D_2)\) denote the subsystems, and \((V_1(k), z_1(k))\) and \((V_2(k), z_2(k))\) variable sets of the input, state, and output of the two extended invariant systems \( L_1 \) and \( L_2 \), respectively.

(i) Let \( V(k) = V_1(k) + V_2(k) \), \( z(k) = z_1(k) + z_2(k) \), and \( w(k) = w_1(k) + w_2(k) \), and the parallel connected system is

\[
\begin{bmatrix}
  z_1(k+1) \\
  z_2(k+1)
\end{bmatrix} = \begin{bmatrix}
  A_1 & 0 \\
  0 & A_2
\end{bmatrix}
\begin{bmatrix}
  z_1(k) \\
  z_2(k)
\end{bmatrix} + \begin{bmatrix}
  B_1 & B_2
\end{bmatrix}
\begin{bmatrix}
  v_1(k) \\
  v_2(k)
\end{bmatrix}
\]

(44)

The quadruple of \( (44) \) is not of the extended form. But by transformation \( T_k \),

\[
T_k = \begin{bmatrix}
  I_n & 0 & 0 & \ldots & 0 \\
  0 & I_n & 0 & \ldots & 0 \\
  0 & 0 & I_n & \ldots & 0 \\
  \vdots & \vdots & \vdots & \ddots & \vdots \\
  0 & 0 & 0 & \ldots & I_n
\end{bmatrix}
\]

(45)

where \( n_1 \) and \( n_2 \) are dimensions of \( L_1 \) and \( L_2 \), the quadruple is

\[
E \left( \left[ \begin{array}{c}
  A_1 \\
  B_1
\end{array} \right], \left[ \begin{array}{c}
  B_1 \\
  B_1
\end{array} \right], \left[ \begin{array}{c}
  C_1 \\
  C_1
\end{array} \right], \left[ \begin{array}{c}
  D_1 + D_2
\end{array} \right] \right)
\]

(46)

(ii) Let \( V(k) = w_1(k) \), \( V(k) = v(k) \), and \( w = w(k) \), and the serial connected system is

\[
\begin{bmatrix}
  z_1(k+1) \\
  z_2(k+1)
\end{bmatrix} = \begin{bmatrix}
  A_1 & 0 \\
  B_1C_1 & A_2
\end{bmatrix}
\begin{bmatrix}
  z_1(k) \\
  z_2(k)
\end{bmatrix} + \begin{bmatrix}
  B_1 & B_2
\end{bmatrix}
\begin{bmatrix}
  v(k)
\end{bmatrix}
\]

(45)

Again, by \( T_k \) in (45), the quadruple is

\[
E \left( \left[ \begin{array}{c}
  A_1 \\
  B_1
\end{array} \right], \left[ \begin{array}{c}
  B_1 \\
  B_1
\end{array} \right], \left[ \begin{array}{c}
  C_1 \\
  C_1
\end{array} \right], \left[ \begin{array}{c}
  D_1 + D_2
\end{array} \right] \right)
\]

(46)

Therefore, \( r(k) = L_s u(k) \), which means the identity (43).

LEMMA 5: The following identity holds

\[
p_s(z) = det(zI - A_sM_s) = det(\lambda I - A_1 \cdots A_{n-1} A_n) = p(\lambda), \text{ for } \lambda = z^N
\]

(42)

PROOF: By direct calculation of det \((zI - M_s^{-1}A_sM_s)\). The subsystems of \( S^*\) and \( S \) have unique properties which are useful for the feedback connection.
Theorem 4: The following two feedback systems are the same:

\[
\begin{bmatrix}
S_1 & S_2 & S_3 & S_4
\end{bmatrix}
\begin{bmatrix}
L_1 & L_2 & L_3 & L_4
\end{bmatrix}
\]

\[
\begin{bmatrix}
S_1 & S_2 & S_3 & S_4
\end{bmatrix}
\begin{bmatrix}
L_1 & L_2 & L_3 & L_4
\end{bmatrix}
\]

Proof: By using Lemma 5 and 6.

IV. Application to Eigenvalue Assignment

A typical control application of the canonical form derived in section II is the eigenvalue assignment of the monodromy matrix \( \Phi_0 \) as in (17) of a completely reachable discrete MIMO LP system

\[
x(k+1) = A_x x(k) + B u(k), \quad y(k) = L x(k)
\]
where \( L \) is a \( n \times n \) identity matrix. The eigenvalue assignment problem is stated as: By what cyclo-static state feedback \( u(k) = L y(k) \) is the monodromy matrix \( \Phi_0 \) of the plant \( \Phi_0 \) controlled to \( \Phi_{x0} = (A_{xN-1} - B_{xN-1} L_{N-1}) \cdots (A_1 - B_1 L_1) (A_0 - B_0 L_0) \) such that for a given desired polynomial \( p(\lambda) \)

\[
det(A I - (A_{N-1} - B_{N-1} L_{N-1}) \cdots (A_1 - B_1 L_1) (A_0 - B_0 L_0)) = p(\lambda)
\]

Although the eigenvalue assignment problem has been solved \([4]\), our approach using the canonical form is simple and generalizes the time-invariant case. The solution is eventually obtained in terms of the cyclo-static state feedback \((L_k)_{k\in\mathbb{N}}\) and input transformation \((G_k)_{k\in\mathbb{N}}\).

\[
u(k) = G_v r(k) - L_d y(k)
\]

However, the problem is first considered through the extended invariant systems: The controlled system in \((47)\) and \((49)\) can be understood as in Theorem 4 where \( L P_1 = L P((A_1, B_1, L_0))_{k\in\mathbb{N}} \), \( L P_2 = L P((0, 0, 0, L_0))_{k\in\mathbb{N}} \), and \( L P_3 = L P((0, 0, 0, G_0))_{k\in\mathbb{N}} \). The overall extended invariant system \( L_{3d} (1 + L_1 L_2) L_{1d} = (A_s - B_s G_s^{-1} L_m, B_s G_s, M_s, 0) \)

\[
L_{3d} (1 + L_1 L_2) L_{1d} = (A_s - B_s G_s^{-1} L_m, B_s G_s, M_s, 0)
\]

Let the characteristic polynomial of the controlled extended invariant system matrix \( A_s - B_s L_s M_s \) be

\[
det(A I - (A_{N-1} - B_{N-1} L_{N-1}) \cdots (A_1 - B_1 L_1) (A_0 - B_0 L_0)) = p(\lambda)
\]

then, by applying Lemma 4 to \((48)\) and \((51)\),

\[
p'(\lambda) = p(\lambda) \quad \text{for} \quad \lambda = \lambda^M
\]

Therefore, the problem is solved by finding \( L_s \) in \((51)\) with proper \( G_s \) for \( p'(\lambda) \) given from \( p(\lambda) \).

Without loss of generality, we show the procedure for finding \( L_s \) and \( G_s \), and hence \((L_s)_{k\in\mathbb{N}}\) and \((G_s)_{k\in\mathbb{N}}\) in \((49)\) with an example.

Example 2: A completely reachable LP system \( L P((A_s, B_s, L_0))_{k\in\mathbb{N}} \) as in \((47)\) of \( n = 3, m = 2 \) and \( N = 2 \) is given. Let the extended invariant system be \( L(A_s, B_s, M_s, 0) \) where \( M_s \) is as in \((31)\).

(i) Using the Algorithm in Section II, find the extended canonical quadruple \( T_{3s} \), of \((A_s, B_s, M_s, 0)\). Let us assume its structural invariance (reachability index) is \( \kappa = (3, 0, 1, 0) \)

The following \( T_D \) yields the extended canonical quadruple \((A_{3s}, B_{3s}, M_s, T_D, 0) \) of \((A_s, B_s, M_s, 0)\)

\[
T_D = \text{diag}(T_0, T_1)
\]

\[
(A_{3s}, B_{3s}) =
\]

\[
\begin{bmatrix}
B_0 & 0 \\
0 & B_1
\end{bmatrix}
\]

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

The following \( T_{3s} \) or \( Q \) yields the usual reachability canonical form with the Scheme II \((A_{3s}, B_{3s}, M_s, T_P Q, 0) \) of \((A_s, B_s, M_s, 0)\)

\[
T_{3s} =
\]

\[
\begin{bmatrix}
B_1 & 0 \\
0 & A_1
\end{bmatrix}
\]

\[
A_{3s}, B_{3s} =
\]

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

\[
Q =
\]

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

\[
(A_{3s}, B_{3s}) =
\]

\[
\begin{bmatrix}
0 & a_{0,1} & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0
\end{bmatrix}
\]

\[
R =
\]

\[
\begin{bmatrix}
1 & 0 & 0 & 0 & 0 & 0 \\
0 & -a_{1,1} & 1 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0
\end{bmatrix}
\]

\[
(A_{3s}, B_{3s}) =
\]

\[
\begin{bmatrix}
0 & a_{0,0} & 0 & a_{1,1} & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & a_{0,1} & 0
\end{bmatrix}
\]

\[
(60)
\]
(iv) Assume that a desired characteristic polynomial is given as \( p(\lambda) = \lambda^3 - \rho_3 \lambda^2 - \rho_1 \lambda - \rho_0 \). Then, by (52),
\[
p'(s) = s^3 - \rho_3 s^2 - \rho_1 s - \rho_0
\]
(v) Using the controller method [2, pp. 500-503], we now intend to control the system \( L(A_{1e}, B_{1e}, M, T, Q, R, 0) \) to \( L(A_{1e}, B_{1e}, M, T, Q, R, 0) (A_{1e}, B_{1e}) =
\[
\begin{bmatrix}
0 & p_1 & p_0 & p_0 \\
1 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 \\
0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1 \\
0 & 0 & 0 & 0
\end{bmatrix}
\]
by the input transformation \( G \), and state feedback \( G^{-1}L_{1e}M_{1e}T_{D}Q \) such that
\[
B_{1e} = B_{1e}G, A_{1e} = A_{1e} - B_{1e}G, G^{-1}L_{1e}M_{1e}T_{D}Q \]
(vi) From (60) to (62),
\[
G_e = \begin{bmatrix}
1 & -b_{0} & 0 & 0 \\
0 & 1 & 0 & 0 \\
0 & 0 & 1 & -b_{1} \\
0 & 0 & 0 & 1
\end{bmatrix}
\]
From (60) to (62),
\[
G^{-1}_{e}L_{1e}M_{1e}T_{D}Q = \begin{bmatrix}
0 & k_0 & k_1 & 0 \\
0 & 0 & 0 & 0 \\
0 & 0 & 0 & k_0 \\
0 & 0 & 0 & 0
\end{bmatrix}
\]
where
\[
k_1 = a_{0,1} - p_0, \quad k_0 = a_{1,1} - p_1
\]
From (53),(57),(59),(63) and (64),
\[
L_e = \begin{bmatrix}
k_0 & k_1 & a_{1,1} k_0 + k_1 \\
0 & 0 & 0 \\
0 & 0 & 0
\end{bmatrix} T_1 0
\]
Observe that \( G_e \) in (63) and \( L_e \) in (65) are always block-diagonal (hence the subscript 'e' is justified). This is clear from as follows: \( (A_{1e}, B_{1e}) \) in (60) is similar to \( (A_{2e,e}, B_{2e,e}) \) in (54)
\[
A_{2e,e} = QR^{-1}A_{2e}QR, B_{2e,e} = QR^{-1}B_{2e}
\]
Moreover, since \( (A_{2e,e}', B_{2e,e}') = (A_{2e}, B_{2e}) \) except that \( a_{0,1} = 1, b_{0,0} = 0, \) and \( b_{0,1} = 0, \) there exist \( (A_{1e,e}', B_{1e,e}') \) which is similar to \( (A_{2e,e}', B_{2e,e}') \), and is of the extended form. Specifically,
\[
A_{1e,e} = QR^{-1}A_{1e}QR, B_{1e,e} = QR^{-1}B_{1e}
\]
(67) \( G_e \) and \( L_e \) are now diagonal respectively from (62),(66) and (67).
(vii) Finally, the controlled extended invariant system is obtained as \( L(A_{1e} - B_{1e}L_{1e}M_{1e}T_{D}Q R_{1}G_{e}, M_{1e}T_{D}Q R_{0}) \), which is similar to \( L(A_{1e} - B_{1e}L_{1e}M_{1e}T_{D}Q R_{0}) \). This is the overall extended invariant system in (50) and THEOREM 4, and hence the input transformation \( G \), and the state feedback \( L_e \) is realized with the cyclo-static input transformation and state feedback \( L_e = LP(0,0,0,G_e) \) and \( L_e = (0,0,0,L_e) \), respectively.

V. CONCLUDING REMARKS

Another reachability canonical form can be obtained using the Scheme I instead of Scheme II in the ALGORITHM. Observability canonical forms can also be found by using the same steps in the ALGORITHM and the obvious duality. These canonical forms are useful mathematical tools, applicable to the modeling, robust control and multi-rate systems.

APPENDIX

For simple notation, let \( A = A_{3e}e, e_{11} = |B_{3e}|, \) and \( e_{11} = |B_{3e}|. \) From (58),
\[
A^e_{11} = a_{2,1}A^e_{11} - a_{1,1}A_{11} + a_{0,1}e_{11}
\]
(68)
\[
Ae_{11} = a_{0,1}e_{11}
\]
(69)
Let
\[
A^e_{11} = a_{2,1}A^e_{11} - a_{1,1}A_{11} = e_{11}
\]
(70)
then, from (68) and (70),
\[
Ae_{11} = a_{0,1}e_{11}
\]
(71)
Let
\[
A^e_{11} = a_{2,1}A^e_{11} - a_{1,1}e_{11} = e_{11}
\]
(72)
then, from (70) and (72),
\[
Ae_{11} = a_{1,1}e_{11} = e_{11}
\]
(73)
Let
\[
A^e_{11} = a_{2,1}A^e_{11} = e_{11}
\]
(74)
then, from (72) and (74),
\[
Ae_{11} = e_{11}
\]
(75)
Let
\[
A^e_{11} = e_{11}
\]
(76)
then, from (74) and (76),
\[
Ae_{11} = a_{0,1}e_{11} + e_{11}
\]
(77)
Form a transformation \( R = [e_{11} e_{11} e_{11} e_{11} e_{11}] \), and we obtain (69) and (60) by applying the formula \( A_{2e,e}R = R,A_{2e}, B_{2e} = B_{2e} \) with (69),(71),(73),(75),(76) and (77).

REFERENCES


APPENDIX I

W. S. Gray and E. I. Verriest

On the Sensitivity of Generalized State-Space Systems

Proc. 28th IEEE Conf. on Decision and Control

ON THE SENSITIVITY OF GENERALIZED STATE-SPACE SYSTEMS

W. Steven Gray  
Department of Electrical and Computer Engineering  
Drexel University  
Philadelphia, Pennsylvania 19104

Erik I. Verriest  
School of Electrical Engineering  
Georgia Institute of Technology  
Atlanta, Georgia 30332

Abstract

The synthesis of minimum sensitivity state-space realizations of linear time-invariant systems is a well understood problem. Such realizations have been linked to balanced realizations. In this paper, the theory is extended to the synthesis of minimum sensitivity generalized state-space models for singular linear systems. A scalar sensitivity measure is first defined. Then the minimization of the measure is considered over all admissible realizations. Since minimal realizations are not required to be related by similarity transformation, the optimization problem is more complex. A criterion is given for determining optimal sensitivity structures. In the nonsingular case, the criterion reduces to the familiar result. The simple example of the right-shift operator is considered.

I. Introduction

The sensitivity of state-space realizations of linear time-invariant systems has been a subject of considerable interest to researchers during the past few years [8,9,10,14,15,16,17,18]. Much of the motivation for this research has been the desire to design digital filters and analog networks with minimum parameter sensitivity. At present, the sensitivity theory for this class of systems is well established. The purpose of this paper is to study the sensitivity of generalized state-space realizations for singular linear systems. An example from this class is the system described by the set of difference equations of the form

\[ E_{x_{n+1}} = Ax_n + b_0 \]  
\[ y_0 = c x_n \]  

where \( E \) is a singular \( n \times n \) matrix such that the semi-state vector \( x_n \in \mathbb{R}^n \) is defined implicitly. We restrict the discussion here to the discrete-time, single-variable case for brevity, but much of the following development can be extended to the continuous-time and multivariable cases.

In the usual linear system case, where \( E = I \), it is well known that state-variable models are not unique. Any two minimal state-space realizations of a given transfer function are related by a similarity transformation, i.e., by a change in basis for the state-space. In the singular case, however, minimal realizations are not necessarily related by similarity. The set of admissible transformations is much larger. Thus, determining optimal sensitivity realizations for the singular case is a non-trivial extension of existing theory for the usual linear case.

An outline of the paper is as follows. We begin by defining sensitivity measures in a manner analogous to the usual linear case [14,15,16,17]. Using these measures, the problem of identifying optimal generalized state-space structures is considered. The theory is then applied to the example of the right-shift operator.

II. SENSITIVITY MEASURES

Given the system (1)-(2), the corresponding transfer function is

\[ h(z) = c(zE - A)^{-1}b. \]  

(3)

Any matrix pair \( (U, V) \in \text{GL}_n(\mathbb{R}) \) will provide another \( n \times n \) order realization of the transfer function by applying the action

\[ \phi : \text{GL}_n(\mathbb{R}) \times \Sigma \rightarrow \Sigma \]  
\[ (U, V) \times (E, A, b, c) \rightarrow (U^{-1}EV, U^{-1}AV, U^{-1}b, cV), \]  

(4)

where \( \Sigma \) is defined to be the set of all \( n \times n \) order realizations. In general \( \phi \) is not a similarity action since \( U \) and \( V \) may be distinct. When \( n \) corresponds to the minimal order required to realize \( h(z) \), then all minimal realizations of \( h(z) \) lie on an orbit of \( \phi \) [7].

The sensitivity of the transfer function to the realization parameters is described by the sensitivity functions:

\[ \frac{\partial h(z)}{\partial e} = f(z); \quad \frac{\partial h(z)}{\partial c} = g(z). \]  

(5)

Analogous to the results for the usual linear system case reported in [14,15,16,17], we have the following lemma.

Lemma 2.1 Define the vectors

\[ f(z) = (zE - A)^{-1}b \]  
\[ g(z) = (zE^T - A^T)^{-1}c^T. \]  

(6)

(7)

then

\[ \frac{\partial h(z)}{\partial \epsilon} = f(z); \quad \frac{\partial h(z)}{\partial c} = g(z). \]  

(8)

Proof. The expressions in (8) are obvious generalizations of those in [14,15,16,17]. The expressions in (9) are proven using Kronecker matrix algebra [3,4] as follows:

\[ \frac{\partial h(z)}{\partial A^T} = (L_e \otimes c)(zE - A)^{-1}(L_e \otimes b) \]  
\[ = (L_e \otimes c)(zE - A)^{-1}U_{nn}(I_e \otimes (zE - A)^{-1}b), \]  

where \( U_{nn} \) is the \( n \times n \) identity matrix.
where \( U_{saa} = \sum_{s} \sum_{a} E_{s} \otimes E_{a} \), and \( E_{s} \) is an \( n \times n \) matrix with unity at the \((i, k)\) position and zeroes elsewhere. Thus,

\[
\begin{align*}
\frac{\partial h(z)}{\partial A} &= \left( \frac{\partial h(z)}{\partial A^T} \right)^T \\
&= (L_s \otimes b^T(zE^T - A^T)^{-1}) C_{saa}^T (L_s \otimes (zE^T - A^T)^{-1} c^T) \\
&= (L_s \otimes b^T(zE^T - A^T)^{-1}) (zE^T - A^T)^{-1} c^T \otimes I_s \\
&= (zE^T - A^T)^{-1} c^T b^T(zE^T - A^T)^{-1} \\
&= g(z)f'(z)
\end{align*}
\]

Thus, it follows from the Cauchy-Schwarz inequality that

\[
\begin{align*}
\frac{\partial h(z)}{\partial A} &= \left( \frac{\partial h(z)}{\partial A^T} \right)^T \\
&= (L_s \otimes b^T(zE^T - A^T)^{-1}) C_{saa}^T (L_s \otimes (zE^T - A^T)^{-1} c^T) \\
&= (L_s \otimes b^T(zE^T - A^T)^{-1}) (zE^T - A^T)^{-1} c^T \otimes I_s \\
&= (zE^T - A^T)^{-1} c^T b^T(zE^T - A^T)^{-1} \\
&= g(z)f'(z)
\end{align*}
\]

where \( I(z) = z \). Letting \( \{f_s\} \) and \( \{g_s\} \) denote the inverse \( z \)-transform of \( f(z) \) and \( g(z) \), respectively, it follows from Parseval’s theorem that

\[
\begin{align*}
\|g_s\|^2 &= \frac{1}{2\pi j} \int \|g(z)\|^2 \, z^{-1} \, dz \\
&= \sum_{\alpha=0}^{\infty} \tilde{g}_\alpha \tilde{g}_\alpha \tag{15}
\end{align*}
\]

and

\[
\begin{align*}
\|f(z)\|^2 &= \frac{1}{2\pi j} \int \|f(z)\|^2 \, z^{-1} \, dz \\
&= \sum_{\alpha=0}^{\infty} f_{\alpha} f_{\alpha} \tag{16}
\end{align*}
\]

(17)

(18)

(19)

(20)

(21)

where \( \mu \geq 1 \) is the index of nilpotency of the matrix pencil \((zE - A)\). The fundamental matrix \((\phi_\alpha)\) satisfies the recursion relations

\[
\begin{align*}
E \phi_\alpha - A \phi_{\alpha-1} &= \delta_\alpha I \tag{22} \\
\phi_{\alpha} E - \phi_{\alpha-1} A &= \delta_\alpha I. \tag{23}
\end{align*}
\]

(\( \delta_\alpha \) is the Kronecker delta.) In terms of the fundamental matrix, \( \{f_s\} \) and \( \{g_s\} \) are expressed as

\[
\begin{align*}
f_s &= \phi_\alpha b \, u_{\alpha} \\
g_s &= \phi_\alpha c \, u_{\alpha}.
\end{align*}
\]

where \( u_0 \) is the unit step function. Hence, another expression for \( M^* \) is

\[
M^* = 2 \sum_{\alpha=0}^{\infty} b^T \phi_\alpha b \cdot \sum_{\alpha=0}^{\infty} c \phi_\alpha c^T
\]

\[
+ \sum_{\alpha=0}^{\infty} b^T \phi_\alpha b \cdot \sum_{\alpha=0}^{\infty} c \phi_\alpha c^T. \tag{26}
\]

For the usual linear system case, \( \phi_\alpha = A^k \) for \( k \geq 0 \) (zero otherwise) and \( \mu = 1 \). In this case, the summation involving \( b \) in (26) can easily be shown to be equal to the trace of the reachability grammian. Likewise, the summation involving \( c \) is equal to the trace of the observability grammian. The set of bound-optimal state-space structures are those that minimize

\[
M^*_e \triangleq M^* \mid \frac{\partial h}{\partial A} = 0 \tag{27}
\]
over all possible realizations. The bound-optimal set is characterized by the property that each member has its reachability grammian equal to its observability grammian. Such realizations are said to be essentially balanced [8,18]. In order to establish an analogous result for singular systems, we wish to express $M^*$ in terms of grammian matrices. There are at least two ways to define the reachability and observability grammians for singular system. Each definition ultimately depends on how the concepts of reachability and observability are extended from the nonsingular to the singular case. This is still an active area of research [1,2,6,7,11,13]. Consider first the reachability and observability matrices defined by Lewis in [11,13]. The subscript $\sigma$ indicates that these definitions follow from the notion of reachability and observability in the symmetric sense.

Definition 2.1 [11] For the system (1)-(2) with $zE - A \neq 0$, define the symmetric reachability and observability grammians as

$$R_{\sigma}(E, A, b) = \begin{bmatrix} \phi_0 b & \phi_1 b & \cdots & \phi_\sigma b \\
\sigma_{-1} b & \phi_{-1} b & \cdots & \phi_{-\sigma} b \end{bmatrix}$$

and

$$O_{\sigma}(E, A, c) = \begin{bmatrix} c_0 & \cdots & c_{-\sigma} \\
c_{\sigma+1} & \cdots & c_{2\sigma} \\
c_{\sigma+1} & \cdots & c_{2\sigma} \\
\vdots & \vdots & \vdots \\
c_{\sigma+1} & \cdots & c_{2\sigma} \end{bmatrix},$$

respectively, for some non-negative integer $\sigma$.

In terms of the matrices in the definition above, the natural definitions for reachability and observability grammians are as follows.

Definition 2.2 For the system (1)-(2) with $zE - A \neq 0$, define the symmetric reachability and observability grammians as

$$P_{\sigma} = R_{\sigma} R_{\sigma}^T$$

and

$$Q_{\sigma} = C_{\sigma}^T O_{\sigma},$$

respectively.

It is easily shown using (22) and (23) that the steady-state grammians (when the limits exist)

$$P_s \triangleq \lim_{t \to \infty} P_s = \begin{bmatrix} \sum_{\sigma=0}^{\infty} \phi_\sigma b \phi_\sigma^T & 0 \\
0 & \sum_{\sigma=0}^{\infty} \phi_\sigma b \phi_\sigma^T \end{bmatrix}$$

satisfy the matrix equations

$$E_s P_s E_s^T = A_s P_s A_s^T + E_s b \phi b^T E_s^T$$

and

$$E_s^T Q_s E_s = A_s^T Q_s A_s + E_s^T \phi^T c E_s,$$

where

$$E_s = \begin{bmatrix} I & 0 \\
0 & I \end{bmatrix}$$

and

$$\phi = \begin{bmatrix} \phi_0 & \phi_1 & \cdots & \phi_\sigma \\
\sigma_{-1} & \phi_{-1} & \cdots & \phi_{-\sigma} \\
\vdots & \vdots & \vdots & \vdots \\
\phi_{\sigma+1} & \cdots & \phi_{2\sigma} \end{bmatrix},$$

$$\phi^T = \begin{bmatrix} \phi_0^T & \cdots & \phi_{\sigma}^T \\
\sigma_{-1} & \phi_{-1}^T & \cdots & \phi_{-\sigma}^T \\
\vdots & \vdots & \vdots & \vdots \\
\phi_{\sigma+1} & \cdots & \phi_{2\sigma}^T \end{bmatrix},$$

and

$$c = \begin{bmatrix} c_0 \\
c_{\sigma+1} \\
c_{\sigma+1} \\
\vdots \\
c_{\sigma+1} \\
c_{\sigma+1} \end{bmatrix}. $$

Another useful reachability-observability grammian matrix pair comes from the following definition.

Definition 2.3 [1,13] For the system (1)-(2) with $zE - A \neq 0$, define the forward reachability and observability grammians as

$$R_s(E, A, b) = \begin{bmatrix} \phi_0 b & \phi_1 b & \cdots & \phi_\sigma b \\
\sigma_{-1} b & \phi_{-1} b & \cdots & \phi_{-\sigma} b \end{bmatrix}$$

and

$$O_s(E, A, c) = \begin{bmatrix} c_0 \\
c_{\sigma+1} \\
c_{\sigma+1} \\
\vdots \\
c_{\sigma+1} \\
c_{\sigma+1} \\
c_{\sigma+1} \end{bmatrix},$$

respectively, for some non-negative integer $\sigma$.

In terms of the forward reachability and observability grammians, the natural definitions of the reachability and observability grammians are given below.

Definition 2.4 For the system (1)-(2) with $zE - A \neq 0$, define the forward reachability and observability grammians as

$$P_s = R_s R_s^T$$

and

$$Q_s = C_s^T O_s,$$

respectively.

It is easily verified that the steady-state forward grammians (when the limits exist)

$$P \triangleq \lim_{t \to \infty} P_t = \sum_{k=0}^{\infty} \phi_k b \phi_k^T$$

satisfy the equations

$$E P E^T = A P A^T + (E \phi b \phi^T E^T - A \phi b \phi^T A)$$

and

$$E^T Q E = A^T Q A + (E \phi^T c c^T E - A^T \phi^T c c^T A).$$

For the usual linear case, equations (45) and (46) reduce to the familiar Lyapunov equations. (Note that Definition 2.4 is closely related to, yet distinct from, that given by Bender in [1]. The main difference is that $P$ and $Q$ as defined above will always be non-negative definite.) The main result of this section is now given in the following lemma.

Lemma 2.2 For the system in (1)-(2), the sensitivity measure $M$ satisfies the inequality $M \leq M^*$, where

$$M^* = 2 Tr P T r Q + Tr P + Tr Q$$

and

$$P_s = T r P_s + T r Q_s.$$

Proof. Substitute the definitions for $P$ and $Q$ into (47). Equation (48) follows directly from (47), since $Tr P_s = Tr P$ and $Tr Q_s = Tr Q$.  

1339
III. MINIMUM SENSITIVITY REALIZATIONS

We first consider the problem of minimizing $M^e$ over all minimal realizations of a fixed transfer function $h(z)$. In view of (4), the effect of a transformation pair $(U, V)$ on the grammians $P_2$ and $Q_2$ is given by the action

$$\psi : GL_2^2(R) \times (\mathbb{R}^{a \times a})^2 \rightarrow (\mathbb{R}^{a \times a})^2$$

$$(U, V) \times (P_2, Q_2) \rightarrow (V^{-1}P_2V^{-T}, U^TQ_2U).$$

(49)

where $C = I_2 \otimes U$ and $V = I_2 \otimes V$. Similarly, the effect of a transform pair on $P$ and $Q$ is given by

$$\varphi : GL_2^2(R) \times (\mathbb{R}^{a \times a})^2 \rightarrow (\mathbb{R}^{a \times a})^2$$

$$(U, V) \times (P, Q) \rightarrow (V^{-1}PV^{-T}, U^TQU).$$

(50)

Since $(U, V)$ can be selected arbitrarily from $GL_2^2(R)$, then for some given $(P, Q)$ apply the transformation

$$U_i = e_i; \quad V_i = \frac{1}{\epsilon} I_a$$

(51)

with $\epsilon > 0$. Then,

$$M^e(\epsilon) = 2\epsilon^2Tr(P + T^2P + \epsilon^2 Tr Q).$$

(52)

Thus, the sensitivity measure $M$ can be made arbitrarily small since $0 \leq M \leq M^e(\epsilon)$. This is a phenomenon that does not occur in the usual linear system case, primarily because the set of similarity transformations

$$S = \{(U, V) \in GL_2^2(R) \mid U = V\}$$

(53)

is much more restrictive. Observe that $(U_i, V_i)$ is not an element of $S$.

From a practical point of view, the transformation $\phi_{(u,v)}$ is not very useful. The resulting realizations for small values of $\epsilon$ would have poor quantization properties since the components of $E$ and $A$ could be made arbitrarily large relative to the components of $b$ and $c$. The problem herein is to sufficiently restrict the set of admissible transformations such that the problem has a meaningful answer. There are many possibilities. We shall consider the problem of minimizing $M^e$ over the set of generalized state-space transformations having the same fixed $E$ matrix. For example, in studying the solvability of singular systems or in realizations methods researchers often use the so-called semi-explicit form of (1)-(2), where $E$ is fixed to be

$$E = \begin{bmatrix} I_a & 0 \\ 0 & 0 \end{bmatrix}.$$  

(54)

and $0 \leq \tau \leq a$ (see [5,19]).

Define for arbitrary fixed $E$, the set of transformations leaving $E$ invariant:

$$\tau_E = \{(U, V) \in GL_2^2(R) \mid EU = UE\}.$$  

(55)

When $E = I$ then clearly $\tau_E = S$. To perform the minimization of $M^e \equiv M^e_E$ over $\tau_E$, we employ a generalization of the Lagrange multiplier technique given in [8,18]. The method allows one to adjoin a matrix constraint to a scalar valued performance index by using the following lemma.

**Lemma 3.1** For arbitrary matrices $X$ and $Y$,

$$\text{Tr} \Lambda(X - Y) = 0$$

(56)

for all orthogonal matrices $\Lambda$, if and only if $X = Y$.

**Theorem 3.1** $M^e_E$ is minimized over $\tau_E$ for arbitrary fixed $E$ only if

$$EP = QE.$$  

(57)

**Proof.** First consider minimizing the product $Tr P \cdot Tr Q$ over $\tau_E$. Define a Hamiltonian

$$H = Tr(V^{-1}PV^{-T}) + \lambda Tr(AU' - UE),$$  

(58)

where $\lambda$ is the Lagrange multiplier and $A$ is any orthogonal matrix. A necessary condition for an extremal is

$$\frac{\partial H}{\partial U} = Tr(V^{-1}PV^{-T}) - 2\lambda UE = 0$$

(59)

$$\frac{\partial H}{\partial V} = (-2V^{-1}PV^{-T}) + Tr(U^TQU) + \lambda A = 0.$$  

(60)

Eliminating $\lambda$ and $A$ by combining (59) and (60) gives

$$Tr(U^TQU) \cdot EV^{-1}PV^{-T} = Tr(V^{-1}PV^{-T}) \cdot U^TQE.$$  

(61)

Thus, the optimal transformations $(U, V)$ will generate a matrix pair $(P, Q)$ such that

$$Tr(Q) \cdot EP = Tr(P) \cdot QE.$$  

(62)

Now observe that for non-negative definite $P$ and $Q$

$$Tr P + Tr Q \geq 2(Tr P \cdot Tr Q)^2,$$  

(63)

with equality if and only if $Tr P = Tr Q$. Hence, minimizing the full performance index

$$M^e_E = Tr(P^TQ + P + Tr Q)$$  

(64)

requires $Tr(Q) \cdot EP = Tr(P) \cdot QE$ and $Tr P = Tr Q$.

In the event that $E = I$, it follows that those realizations satisfying (57) are the essentially balanced realizations of [8,18] or those specified in [9,10,15,16,17]. If $E$ has the form given in (54), then it follows easily that any bound-optimal realization $(E, A, b, c)$ has a corresponding matrix pair $(P, Q)$ with the structure

$$P = \begin{bmatrix} \Sigma & 0 \\ 0 & P_D \end{bmatrix}; \quad Q = \begin{bmatrix} \Sigma & 0 \\ 0 & Q_D \end{bmatrix},$$  

(55)

where $\Sigma \in \mathbb{R}^{axa}$ and $P_D, Q_D \in \mathbb{R}^{(a-x)x(a-x)}$. Choosing $(U, V) \in \tau_E$ such that

$$U = \begin{bmatrix} T & 0 \\ 0 & \epsilon I_{x \tau} \end{bmatrix}; \quad V = \begin{bmatrix} T & 0 \\ 0 & \epsilon I_{x \tau} \end{bmatrix}$$  

(66)

gives for small values of $\epsilon > 0$ the approximation

$$M^e_E \approx Tr T^{-1} \Sigma T^{-T} \cdot Tr T^T \Sigma T + Tr T^{-1} \Sigma T^{-T} + Tr T^T \Sigma T.$$  

(67)

Letting $\lambda_i$ denote the $i$th eigenvalue of $\Sigma$ and using the methods in [14,15,16,17], it follows directly that

$$M^e_E \geq \left( \sum_{i=1}^{x} \lambda_i \right)^2 + 2 \sum_{i=1}^{x} \lambda_i.$$  

(68)

Furthermore, there always exist $T \in GL_2^2(R)$ such that this lower bound is achieved arbitrarily closely.
IV. EXAMPLE

Consider the problem of determining the minimum sensitivity generalized state-space realization of the right-shift operator in semi-explicit form. It is easily verified, for example, that \( h(z) = z \) can be minimally realized by

\[
E = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}; \quad A = \begin{bmatrix} 1 & 1 \\ -1 & 0 \end{bmatrix}
\]

\[
b = \begin{bmatrix} -2 \\ 1 \end{bmatrix}; \quad c = \begin{bmatrix} -1 & 1 \end{bmatrix}.
\]

The corresponding fundamental matrix is given by

\[
\phi_k = \begin{cases} \begin{bmatrix} 0 & 1 \\ -1 & -1 \end{bmatrix} : k = -1 \\ \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} : k = -2 \\ 0 & \text{otherwise} \end{cases}
\]

Thus,

\[
P = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}; \quad Q = \begin{bmatrix} 1 & 2 \\ 2 & 5 \end{bmatrix}
\]

and \( M^*_2 = 27 \). Observe that \( (P, Q) \) above does not satisfy the bound-optimality criterion. Now apply the transformation pair

\[
U_1 = \begin{bmatrix} 1 & 2 \\ 0 & -1 \end{bmatrix}; \quad V_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}
\]

The new realization is

\[
E_1 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}; \quad A_1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}
\]

\[
b_1 = \begin{bmatrix} 0 \\ -1 \end{bmatrix}; \quad c_1 = \begin{bmatrix} 0 & 1 \end{bmatrix}
\]

with

\[
P_1 = Q_1 = I_2.
\]

The bound-optimality criterion is now satisfied, but \( M^*_2 = 8 \) is not the minimum value of the measure. (The condition in Theorem 3.1 is only a necessary condition.) Apply a final transformation pair

\[
U_2 = \begin{bmatrix} 1 & 0 \\ 0 & 0.1 \end{bmatrix}; \quad V_2 = \begin{bmatrix} 1 & 0 \\ 0 & 10 \end{bmatrix}
\]

as suggested by (66) such that

\[
E_2 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}; \quad A_2 = \begin{bmatrix} 0 & 10 \\ 10 & 0 \end{bmatrix}
\]

\[
b_2 = \begin{bmatrix} 0 \\ -10 \end{bmatrix}; \quad c_2 = \begin{bmatrix} 0 & 10 \end{bmatrix}.
\]

Hence,

\[
P_2 = Q_2 = \begin{bmatrix} 1 & 0 \\ 0 & 0.01 \end{bmatrix}
\]

such that \( M^*_2 = 3 \) as predicted by (68).

V. CONCLUSIONS

The problem of determining minimum sensitivity generalized state-space realizations was considered by first defining a sensitivity measure analogous to that for the usual linear system case. The minimization of the measure was then shown to be a meaningful problem only if the set of admissible realizations was significantly restricted. An interesting example of such a restriction was the subset of transformations which keep the \( E \) matrix invariant. In this case, the optimality criterion developed was an extension of the known result for the nonsingular case and led to the notion of an essentially balanced generalized state-space realization. The theory was applied to the minimum sensitivity synthesis problem for the right-shift operator.

ACKNOWLEDGEMENTS

This research was supported, in part, by the U. S. Air Force under contract AFOSR-87-0308, AFOSR-89-0241 and Northwestern University.

REFERENCES


APPENDIX J

W. S. Gray, E. I. Verriest, and F. L. Lewis

A Hankel Matrix Approach to Singular System Realization Theory

Proc. 29th IEEE Conf. on Decision and Control

Honolulu, Hawaii, pp. 73-78, December 1990.
A Hankel Matrix Approach to Singular System Realization Theory

W. Steven Gray  
Department of Electrical and Computer Engineering  
Drexel University  
Philadelphia, Pennsylvania 19104

Erik I. Verriest  
School of Electrical Engineering  
Georgia Institute of Technology  
Atlanta, Georgia 30332

Frank L. Lewis  
School of Electrical Engineering  
Georgia Institute of Technology  
Atlanta, Georgia 30332

Abstract

The system Hankel matrix plays a central role in realization theory for linear time-invariant systems. For example, a complete solution to the problem of minimal realization exists for the nonsingular case [9,19,20]. Theory also exists for characterizing minimality [3,17], reachability and observability [1,12,13,14,16], and determining canonical forms [5]. A strong unifying realization theory for singular systems comparable to that of the nonsingular case has yet to be presented. A solid realization theory is certainly a prerequisite for deriving physical interpretations of systems, as well as understanding computational structures.

In this paper, a Hankel matrix approach to singular linear system realization theory is presented, which is analogous to the methods of Kung for the nonsingular case [11]. The focus is exclusively on discrete-time or descriptor systems. Motivated by definitions of reachability and observability, the notion of a system Hankel matrix is first defined. The system Hankel matrix is then shown to have reachability-observability properties which can be used to solve the realization synthesis problem. Next, the notion of a balanced realization for singular systems is derived using the singular value decomposition of the system Hankel matrix. The final section gives an application of this theory for synthesizing realizations which have minimum parametric sensitivity properties. It is demonstrated that such realizations are related to the notion of balancing. Analogous connections exist for the nonsingular case [9,19], and such a connection has been demonstrated in the singular case by a completely independent method [10].

I. Introduction

Developing a complete realization theory for singular linear systems has provided some challenging problems for researchers. There has been significant success in the development of algorithms for minimal realization synthesis from given input-output behavior [3,4,6,7,18]. Theory also exists for characterizing minimality [8,17], reachability and observability [1,12,13,14,16], and determining canonical forms [5]. A strong unifying realization theory for singular systems comparable to that of the nonsingular case has yet to be presented. A solid realization theory is certainly a prerequisite for deriving physical interpretations of systems, as well as understanding computational structures.

II. The Singular System Hankel Matrix

A singular linear system of difference equations

\[ \begin{align*}
Ez_{k+1} &= Az_k + Bu_k \\
y_k &= Cz_k,
\end{align*} \]

where \( E \) is a singular \( n \times n \) matrix, is said to realize a given \( p \) by \( m \) rational transfer function matrix \( H(z) \) when

\[ H(z) = C(zE - A)^{-1}B. \]

In the usual linear system case, where \( E = I \), it is well known that state variable models are not unique. Any two minimal state space realizations of a given transfer function matrix are related by a similarity transformation, i.e., by a change in basis for the state space. In the singular case, however, minimal realizations are not necessarily related by similarity. The set of admissible transformations is considerably larger. Any matrix pair \((U,V) \in GL_n(\mathbb{R}) \otimes GL_n(\mathbb{R}) \times GL_n(\mathbb{R})\) will provide another \( n^{th} \) order realization of the transfer matrix by applying the action

\[ \phi : GL_n(\mathbb{R}) \times \mathcal{L}_n(\mathbb{R}) \to \mathcal{L}_n(\mathbb{R}) \]

\[ : (U,V) \times (E,A,B,C) \to (U^{-1}EV,U^{-1}AV,U^{-1}B,CV), \]

where \( \mathcal{L}_n(\mathbb{R}) \) is defined as the set of all \( n^{th} \) order realizations. In general \( \phi \) is not a similarity action since \( U \) and \( V \) may be distinct.

For a given realization \((E,A,B,C)\), the corresponding fundamental matrix \( \{\phi_k\}_{k=1}^{\infty} \) is defined in terms of the unique Laurent expansion about the point at infinity of the resolvent matrix \((zE - A)^{-1}\)

\[ (zE - A)^{-1} = z^{-\mu} \sum_{m=-\mu}^{\infty} \phi_m z^{m}, \]

where \( \mu \geq 1 \) is the index of nilpotency of the matrix pencil \((zE - A)\) [12,13]. The fundamental matrix is known to satisfy the recursion relations

\[ E\phi_k - A\phi_{k-1} = \delta_{k0} I \]

\[ \phi_kE - \phi_{k-1}A = \delta_{k0} I. \]

(\( \delta_{k0} \) is the Kronecker delta.) Furthermore, equation (4) implies that the corresponding transfer matrix, \( H(z) \), has a series representation

\[ H(z) = \sum_{i=0}^{\infty} H_i z^{-i}, \]

where

\[ H_i = C\phi_{i-1}B. \]
Hence, we shall refer to the sequence \( \{H_i\}_{i=-\infty}^{-1} \) defined by (8) as the generalized Markov parameters for the system (1)-(2) [15].

The notions of reachability and observability for singular systems is an active area of research [1,8,12,13,14,16]. The possible noncausality of (1)-(2) makes the extension of these concepts nontrivial. For the purpose of factoring the system Hankel matrix, it useful to define the notions of forward reachability and observability as given below.

**Definition 2.1** [14] A regular system (1)-(2) is said to be forward reachable if, for every \( z \in \mathbb{R}^n \), there exists an integer \( j > 0 \) and an input \( \{u_k\}_{k=-\infty}^{-1} \) such that \( x_j = x \) when \( x_0 = 0 \).

**Lemma 2.1** [14] A regular system (1)-(2) is forward reachable if and only if the zero input response \( \{y_k\}_{k=-\infty}^{-1} \) is precisely zero then \( E_{x_0} = 0 \).

The following tests can be used for determining forward reachability and observability.

**Lemma 2.1** [14] A regular system (1)-(2) is forward reachable if and only if the forward reachability matrix

\[
R_r(E, A, B) = \begin{bmatrix}
\phi_{r-1} & \phi_{r-1}B & \cdots & \phi_1B & \phi_1 \phi_1B & \cdots & \phi_1B
\end{bmatrix}
\]

has rank \( n \) for \( j = \deg(|zE - A|) \).

**Lemma 2.2** [14] A regular system (1)-(2) is forward observable if and only if the forward observability matrix

\[
O_r(E, A, C) = \begin{bmatrix}
C\phi_0 & C\phi_1 & \cdots & C\phi_{r-1} & C\phi_r & \cdots & C\phi_{r+1}
\end{bmatrix}
\]

has at least the rank of \( E \) when \( i = \deg(|zE - A|) \).

Consider the following definition.

**Definition 2.3** The system (block) Hankel matrix for a given rational transfer matrix \( H(z) \) is defined as

\[
H[i,j] = \begin{bmatrix}
0 & 0 & \cdots & 0 \\
0 & 0 & \cdots & -H_{r+1} & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
0 & -H_{r+1} & \cdots & -H_{i-1} & H_1 & H_2 & \cdots & H_{i+1} & 0 \\
& & & & H_2 & H_3 & \cdots & H_{i+2} & \vdots & \vdots \\
& & & & & & & H_{i+1} & H_{i+2} & \cdots & H_{i+j+1}
\end{bmatrix}
\]

**Theorem 2.1** The system Hankel matrix for a given rational transfer matrix \( H(z) \) has a finite dimensional generalized state-space realization if and only if there exist non-negative integers \( r \) and \( \nu \) such that the rank \( \rho(H[i+j, r+i]) = \nu \) for all \( i = 0, 1, 2, \ldots \) and \( j = 0, 1, 2, \ldots \).

Proof. Viewing \( H[i,j] \) as a \( 2 \times 2 \) block partitioned matrix, it is clear that the block matrix in the lower right position is the Hankel matrix for the strictly proper portion of \( H(z) \), say \( H_{s+}(z) \). It is well known that \( H_{s+}(z) \) has a finite dimensional state-space realization if and only if its corresponding Hankel matrix \( H_{s+}[i,j] \) has the property that there exists positive integer \( r \) such that \( \rho(H_{s+}[r+i, r+j]) = \rho \) for all \( i, j = 0, 1, 2, \ldots \). In view of the Weierstrass form [13], if \( \rho \) is finite the necessity condition follows directly. The sufficiency condition follows from the fact that when both nontrivial submatrices of \( H[z, +\infty, +\infty] \) have finite rank then a finite dimensional generalised state space realization can be constructed by known algorithms (see for example [6]).

**Lemma 2.3** Every realization \((E, A, B, C)\) of a given transfer matrix is related to the system Hankel matrix via the equality

\[
H[i,j] = C_iE_i.
\]

Proof. The result follows directly from the definitions using the property

\[
C_iE_i = \begin{cases}
\phi_{i+j} : i < 0, j < 0 \\
\phi_{i+j} : i \geq 0, j \geq 0 \\
0 : \text{otherwise}
\end{cases}
\]

which is proven in [14].

This factorization is the natural extension of the result known for the usual linear system case \((E = I)\). A surprising property, however, of the generalised Hankel matrix is that it only specifies the system uniquely modulo a feedforward component. That is, the parameter \( H_0 \) does not appear in \( H[i,j] \). The consequences of this fact will be discussed in the next section. Also, note in reference to Theorem 2.1 that

\[
\rho(E) \geq \rho(H[+\infty, +\infty]) \geq \nu
\]

for any realization \((E, A, B, C)\) of \( H(z) \). Furthermore, it follows directly that

\[
(\mu - 1)\rho(H_{s+1}) + r \leq \nu \leq (\mu - 1)\min(p, m) + r.
\]

When \((E, A, B, C)\) is a minimal realization with \( n \) \( \geq \) \( n_{max} \) then

\[
r \equiv \deg(|zE - A|)
\]

From the characterization of \( n_{max} \) in [6] a simple calculation gives

\[
\mu(\rho(H_{s+1})) + r \leq \mu \leq \mu(\min(p, m)) + r.
\]

By assumption \( H_{s+1} \) is not identically zero. So in the case where \( \min(p, m) = 1 \), the inequalities (11) and (13) combine to give

\[
n_{max} = \nu + 1.
\]

**III. Hankel Matrix Realization Theory**

In this section, a realization theory and algorithm is presented which is analogous to that given by Kung in [11] for classical linear systems. That is, we wish to consider the problem
of extracting generalized state space realizations from appropriate factorizations of a given system Hankel matrix as suggested by Lemma 2.3. The general algorithm presented herein is not viewed as being particularly efficient or stable, but rather as a theoretical tool to exhibit some of the structure of the realization theory. The extended theory is considerably more complex due to the singular nature of $E$. For example, unlike the non-singular case, not every factorization maps to a corresponding realization. Consider the following definition.

**Definition 3.1** A factorization $\mathcal{K}(i, j) = Q_i E R_j$ of a given rank $\nu$ system Hankel matrix, where $E \in \mathbb{R}^{n \times n}$, is said to be consistent if

1. $n \geq n_{\text{rin}}$, 
2. rank$(H[i, j]) = \text{rank}(E)$, 
3. $Q_i E R_j - \Gamma_{ij} = Q_i^t E R_j - \Gamma_{ij}$,

where $\Gamma_{ij}$ is defined as

$$\Gamma_{ij} = \begin{bmatrix} H_{n+1} & \vdots & H_1 & H_0 & \cdots & 0 \\ 0 & \ddots & 0 & \vdots & \vdots & \vdots \\ \vdots & \ddots & \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & 0 & H_0 & \cdots & 0 \\ H_1 & \cdots & H_0 & \cdots & \vdots & \vdots \\ \vdots & \cdots & \vdots & \ddots & \vdots & \vdots \\ \vdots & \cdots & \vdots & \cdots & \ddots & \vdots \\ \vdots & \cdots & \vdots & \cdots & \cdots & \ddots \end{bmatrix}$$

with the $\mu$th block column being nonzero, and $L_{ii}$ similarly defined with the $\mu$th block row being nonzero. The notations $[\cdot]^t$ and $[\cdot]^r$ represent the block column left shift and the block row up shift operators, respectively.

A consistent factorization $(Q_i, E, R_j)$ has the property that at least one realization can be synthesized from it. Observe, from equations (5)-(6) it follows directly that

$$ER_j = AR_j + [0 \ldots 0 B 0 \ldots 0]$$

(16)

$$Q_i E = QA + [0 \ldots 0 C^* 0 \ldots 0]^T.$$  

(17)

Premultiplying equation (16) by $Q_i$ and postmultiplying (17) by $R_j$ gives

$$Q_i E R_j = Q_i A R_j + \Gamma_{ij}$$

(18)

$$Q_i^t E R_j = Q_i A R_j + L_{ij}.$$  

(19)

Thus, the remaining realization matrices are given by

$$C = \text{the } \mu\text{th block row of } L_{ij} R_j^t$$

(20)

$$B = \text{the } \mu\text{th block column of } Q_i \Gamma_{ij}$$

(21)

$$A = Q_i^t (E - [0 \ldots 0 C^* 0 \ldots 0]^T)$$

(22)

$$= (E R_j^t - [0 \ldots 0 B 0 \ldots 0]) R_j^t$$

(23)

for sufficiently large $i$ and $j$ and where $[\cdot]^r$ denotes a pseudo-inverse. These realization matrices are uniquely specified if $\rho(Q_i) = \rho(R_j) = n$. That is, when $(E, A, B, C)$ is both forward reachable and forward observable.

It is of interest to note that since the system Hankel matrix is not a function of $H_0$, the direct feed term, one has the option of setting $H_0 = 0$ in the assignment of $\Gamma_{ij}$ and $L_{ij}$ above and then compensating by adding the known direct feed term to the output equation, i.e.,

$$y_a = C x_a + H_0 u_a.$$  

(24)

In fact, this interplay between the direct feed term in the semi-state equation and that in the output equation is arbitrary. If one sets the direct feed parameter in $\Gamma_{ij}$ and $L_{ij}$ to any compatible matrix $\Delta$ and then compensates by adding the direct feed term $(H_0 - \Delta) u_a$ to the output equation, the transfer matrix is invariant.

### IV. Balanced Generalized Realizations

It is well known in linear system theory that a balanced realization of a given transfer matrix can be extracted from the singular value decomposition (SVD) of its system Hankel matrix [11]. More specifically, the SVD can be used as a tool for computing a special set of factorizations of the Hankel matrix which has the property that all corresponding realizations yield equal and diagonal reachability and observability grammians. In this section, the extension of this idea is considered for the singular system case. It should be mentioned that the notion of balancing for singular systems has been defined in [10], but in a quite different context. The following approach is consistent with this earlier definition, but is considerably more direct.

In terms of the forward reachability and observability matrices, the natural definitions of the reachability and observability grammians are given below.

**Definition 4.1** For a regular system (1)-(8), define the forward reachability and observability grammians as

$$P = E R E^T$$

(25)

and

$$Q = Q_i^t Q_i$$

(26)

respectively.

It is easily verified that the steady-state forward grammians (when the limits exist)

$$P = \lim_{t \to \infty} P_t = \sum_{i=0}^{\infty} \phi_i \phi_i^T$$

(27)

$$Q = \lim_{t \to \infty} Q_t = \sum_{i=0}^{\infty} \phi_i \phi_i^T$$

(28)

satisfy the equations

$$E P E^t = A P A^t + (E \phi_i B B^t \phi_i^* E^T - A \phi_i B B^t \phi_i^* A^T)$$

$$E^t Q E = A^T Q A + (E^T \phi_i C^T C \phi_i E - A^T \phi_i C^T C \phi_i A).$$

For the usual linear system case, these equations reduce to the familiar Lyapunov equations.

**Theorem 4.1** The infinite system Hankel matrix $\mathcal{K}(+\infty, +\infty)$ with rank $\nu$ corresponding to a given rational transfer matrix $H(s)$ has a consistent factorization of the form

$$O = U \Sigma_0^{1/2}; \quad R = \Sigma_0^{1/2} V^T$$

(29)

$$E = \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix},$$

(30)

where $\Sigma_0^{1/2}$ and $\Sigma_0^{1/2}$ are full rank diagonal matrices and $U^T U = V^T V = I_n$.  

75
Proof. The singular value decomposition of \( \mathcal{K}[+\infty, +\infty] \) is
\[
\mathcal{K}[+\infty, +\infty] = U\Sigma V^T = U^\dagger\Sigma E V^T \quad (31)
\]
where \( \Sigma = \text{diag} (\sqrt{\lambda_1}, \ldots, \sqrt{\lambda_r}, \ldots, \sqrt{\lambda_m}) \) and \( \Sigma_{\Theta} = \text{diag} (\sqrt{\lambda_{1,\Theta}}, \ldots, \sqrt{\lambda_{r,\Theta}}, \ldots, \sqrt{\lambda_{m,\Theta}}) \) with the \( \alpha_i \)'s and \( \beta_i \)'s as nonzero free parameters. \( \Theta \) denotes a compatibly sized zero matrix. This decomposition leads directly to the factorization in equations (29)-(30). This factorization is consistent since conditions (i) and (ii) are satisfied by design, and condition (iii) can be shown to be satisfied in the limit by direct substitution of equations (29)-(30).

Corollary 4.1 Any factorization of the form given in Theorem 4.1 has the property that all corresponding realizations have forward reachability and observability grammians such that
\[
EP = QE = E\Sigma = \Sigma_0 E. \quad (33)
\]

Definition 4.2 Any realization \( (E, A, B, C) \) satisfying equation (33) is said to be a balanced realization.

V. Minimum Sensitivity Realizations

In this section, we consider the problem of finding minimum sensitivity generalized state space realizations of a given rational (possibly noncausal) transfer matrix \( H(z) \). This problem has been studied in a purely algebraic context [10], but the following approach is geometric in nature and provides a natural application of the realization theory described above. First, an abstract geometric approach is briefly described for solving generic minimum sensitivity synthesis problems [9,19]. Then the approach is applied for the singular linear system case.

A generic realization space, say \( \tilde{\Theta} \), is defined to be an affine space of admissible realizations (usually some subset of \( \mathbb{R}^n \)) with the structure of a smooth Riemannian manifold. Each point \( \tilde{\Theta} \in \tilde{\Theta} \) corresponds to an admissible realization. In every modeling problem, there are invariants which are related to the observed behavior one is trying to model, e.g., Markov parameters in the linear systems case or the Volterra kernels for more general Volterra type systems. Using these invariants, called observables, it is possible to partition a realization space into equivalence classes. Two realizations are in the same equivalence class if their observables assume the same values.

Assume that \( f : \tilde{\Theta} \rightarrow \mathbb{R} \) is a smooth function that maps each realization to a corresponding scalar observable. Furthermore, assume that \( f \) has no critical points in \( \tilde{\Theta} \). Then if \( f \) is a smooth function on a given realization, the metric tensor \( g \) on \( \tilde{\Theta} \) induces the norm \( \| \cdot \| : T_{\tilde{\Theta}} \tilde{\Theta} \rightarrow \mathbb{R} : v \mapsto g(v,v) \) on the tangent space \( T_{\tilde{\Theta}} \tilde{\Theta} \).

Definition 5.1 A realization \( \tilde{\Theta} \in \mathcal{M}_f(f) \) is an extremal sensitivity point of \( \mathcal{M}_f(f) \) if and only if \( \tilde{\Theta} \) is an extremal of the performance index \( I(\Theta) = \frac{1}{2} \| \nabla f \|^2 \) over the manifold \( \mathcal{M}_f(f) \).

The following Theorem (see [9,19]) gives a necessary condition satisfied by all extremal points in the event that the metric tensor \( g \) on \( \tilde{\Theta} \) is taken to be uniform (a typical assumption in the analysis of fixed point arithmetic).

Theorem 5.1 If a realization \( \tilde{\Theta} \in \tilde{\Theta} \) is an extremal sensitivity point then
\[
(\nabla f(\tilde{\Theta})) - \lambda I(\nabla f(\tilde{\Theta})) = 0, \quad (34)
\]
where \( \nabla f(\tilde{\Theta}) \) and \( \nabla^2 f(\tilde{\Theta}) \) are the gradient vector and the Hessian operator, respectively, at \( \tilde{\Theta} \), \( I \) is the identity operator on the tangent space \( T_{\tilde{\Theta}} \tilde{\Theta} \) and \( \lambda \in \mathbb{R} \).

In other words, the gradient vector of \( f \) at \( \tilde{\Theta} \) is an eigenvector of the Hessian matrix at \( \tilde{\Theta} \). The stated condition is the Euler-Lagrange equation for the constrained optimization problem. The type of extremum is easily determined by the definiteness of the second variation.

To cast the minimum sensitivity synthesis problem for singular systems in the geometric context described above, we must first identify the relevant realization spaces. There are in fact two general realization spaces we shall consider: the space of all \( n \)th order generalized state space realizations \( (E, A, B, C) \), \( \Sigma_0(\mathbb{R}) \), and a space related to all consistent factorizations of all possible system Hankel matrices with rank \( r \leq n \). \( \Sigma_0 \) is clearly isomorphic to a closed subset of \( \mathbb{R}^{2n+r-n} \). To define the second realization space, consider the mapping
\[
\omega : \Sigma_0(\mathbb{R}) \rightarrow \mathbb{R}^{n+i(i+1)+...+n} \times \mathbb{R}^{n+i(i+1)+...+n} \times \mathbb{R}^{n+i(i+1)+...+n}
\]
\[
: (E, A, B, C) \rightarrow (\xi(E, A, C)E, ER(E, A, B)), \quad (35)
\]
where \( i, j \geq r - 1 \) are assumed to be fixed. For fixed \( E \), the marginal map \( \omega_{\tilde{\Theta}} \) defines the following subsets of \( \mathbb{R}^{n+i(i+1)+...+n} \times \mathbb{R}^{n+i(i+1)+...+n} \times \mathbb{R}^{n+i(i+1)+...+n} \):
\[
\Sigma_{\tilde{\Theta}}(\mathbb{R}) = \omega_{\tilde{\Theta}}(\Sigma_0(\mathbb{R})),
\]
\[
\Sigma_{\tilde{\Theta}}^0(\mathbb{R}) = \omega_{\tilde{\Theta}}(\Sigma_0^0(\mathbb{R})),
\]
\[
\Sigma_{\tilde{\Theta}}^{\epsilon,\epsilon}(\mathbb{R}) = \omega_{\tilde{\Theta}}(\Sigma_0^{\epsilon,\epsilon}(\mathbb{R})),
\]
where
\[
\Sigma_{\tilde{\Theta}}^{\epsilon,\epsilon}(\mathbb{R}) = \{ s \in \Sigma_0(\mathbb{R}) : \| sE - A \| \neq 0, \rho(Z_{\tilde{\Theta}}E(A, B)) = n \}
\]
\[
\Sigma_{\tilde{\Theta}}^{\epsilon}(\mathbb{R}) = \{ s \in \Sigma_0(\mathbb{R}) : \| sE - A \| \neq 0, \rho(Z_{\tilde{\Theta}}(A, E, C)) = n \}
\]
\[
\Sigma_{\tilde{\Theta}}^{\epsilon,\epsilon}(\mathbb{R}) = \Sigma_{\tilde{\Theta}}^{\epsilon}(\mathbb{R}) \cap \Sigma_{\tilde{\Theta}}^{\epsilon}(\mathbb{R}).
\]
If \( f \) and \( f_0 \) refer to forward reachable and forward observable, respectively. Note that \( \Sigma_{\tilde{\Theta}}^{\epsilon}(\mathbb{R}) \) is only a subset of all possible forward observable realizations.

In general, a group action on a manifold is said to be a foliated action if the orbit form leaves of a foliation. A foliated action is characterized by the property that the dimension of the isotropy subgroup at any point on the manifold is fixed. The actions
\[
\phi : GL(\mathbb{R}) \times \Sigma_0(\mathbb{R}) \rightarrow (U,V) \times (E, A, B, C) \rightarrow (U^{-1}EV, U^{-1}AV, U^{-1}B, CV)
\]
and
\[
\psi : GL(\mathbb{R}) \times \Sigma_0(\mathbb{R}) \rightarrow (\mathbb{R}) \rightarrow (U, V) \times (O(E, E), R) \rightarrow (O(UE, U^{-1}E))
\]
It follows then

with these representations, the gradient vector

or as

defines the observable functions

family

An orbit in either realization space \( \Sigma^s_i(E) \) or \( \Omega^s_i(E) \) under its group action is characterized not by one observable but by several, namely the entries from the corresponding Hankel matrix. Hence, we must slightly generalize the geometric method described above. Define the following families of scalar-valued observable functions

\[
\hat{f}_a(E, A, B, C) = \text{Tr } \Delta (Y - O(E, A, C) R(E, A, B)) \tag{36}
\]

\[
f_a(OE, ER) = \text{Tr } \Delta (Y - O ER). \tag{37}
\]

where \( \Delta \) is any compatible matrix which has all of its singular values precisely equal to unity. Note that if \( f_a = 0 \) for all such \( \Delta \), then \( Y = O ER \) (proof in [9,19]). Thus, it follows that for a given singular linear system characterized by a rank \( \nu \) Hankel matrix \( H \), the corresponding orbit in either \( \Sigma^s_i(E) \) or \( \Omega^s_i(E) \) is uniquely characterized by \( f_a = 0 \) or \( f_a = 0 \), respectively, for all admissible \( \Delta \). If we express \( f_a \) in component form

\[
f_a = \sum_{i} \Delta_{ii} (Y - O ER)_{ii}, \tag{39}
\]

it is apparent that this family of observables is defined by the set of constraints on the components of the \( O E \) and \( E R \) matrices with each component of \( \Delta \) playing the role of a Lagrange multiplier.

To characterize extremal sensitivity realizations of singular linear systems, we should apply Theorem 5.1 to the observable \( f_a \). This turns out to be a formidable problem. So instead we shall work in the realization (factorization) space \( \Omega^s_i(E) \) with the goal of relating the solutions of the two problems by other means. Applying Theorem 5.1 to the observable \( f_a \) is a relatively simple problem because it is a linear function of the components of \( (O E, E R) \). Consider the following Theorem.

Theorem 5.2 Given a singular linear system characterized by a square Hankel matrix \( H[i, i] \), then extremal sensitivity points on the corresponding leaf of the foliation induced by the observable family

\[
f_a(OE, ER) = \text{Tr } \Delta (Y[i, i] - O ER). \tag{39}
\]

have the property that

\[
ER, R_0^T = O^T O E. \tag{40}
\]

Proof. Use the optimality equation (34). In this case

\[
\theta = \begin{bmatrix}
\text{vec}(R_0^T E^T) \\
\text{vec}(O E)
\end{bmatrix}. \tag{41}
\]

The observable function \( f_a \) can be expressed in terms of a quadratic form in \( \theta \) via Kronecker product algebra [2] as either

\[
f_a(\theta) = \text{Tr } \Delta (Y[i, i] - \text{vec}(R_0^T E^T))^T (I, \otimes \Delta) \cdot \text{vec}(O E) \tag{42}
\]

or as

\[
f_a(\theta) = \text{Tr } \Delta (Y[i, i] - \text{vec}(R_0^T))^T (I, \otimes \Delta) \cdot \text{vec}(O E). \tag{43}
\]

With these representations, the gradient vector is computed as

\[
\nabla f_a(\theta) = \begin{bmatrix}
-(I, \otimes \Delta) \text{vec}(O E) \\
-(I, \otimes \Delta)^T \text{vec}(R_0^T)
\end{bmatrix}. \tag{44}
\]

It is interesting to note that this computation is possible even though \( O \) and \( ER \) cannot be determined uniquely from \( O E \) and \( ER \) due to the singularity of \( E \).

The computation of the Hessian matrix is not so obvious since the gradient vector is not an explicit function of \( \theta \). Consider, however, that by the product rule (see [2], T4.3)

\[
\frac{\partial \text{vec}(OE, ER)}{\partial \text{vec}(OE)} = \frac{\partial}{\partial \text{vec}(OE)}(E^T \otimes I)v\text{ec}(O E) = E^T \otimes I,
\]

\[
\frac{\partial \text{vec}(ER)}{\partial \text{vec}(ER)^T} = \frac{\partial}{\partial \text{vec}(ER)^T}(E \otimes I)v\text{ec}(ER^T) = E \otimes I.
\]

Thus, it follows that

\[
(E^T \otimes I) \frac{\partial \text{vec}(O E)}{\partial \text{vec}(OE)^T} = I \omega, \tag{45}
\]

\[
(E \otimes I) \frac{\partial \text{vec}(ER)}{\partial \text{vec}(ER)^T} = I \omega. \tag{46}
\]

For brevity, let

\[
D_o = \frac{\partial \text{vec}(O E)}{\partial \text{vec}(OE)^T}, \quad D_R = \frac{\partial \text{vec}(ER)}{\partial \text{vec}(ER)^T}. \tag{47}
\]

Then, by the chain rule (see [2], T4.6), it follows that the Hessian matrix is

\[
\hat{\Phi}^2 f_a(\theta) = \begin{bmatrix}
0 & -(I, \otimes \Delta) D_o \\
-(I, \otimes \Delta)^T D_R & 0
\end{bmatrix}. \tag{48}
\]

The optimality condition, then, is as follows:

\[
\hat{\Phi}^2 f_a(\theta) - \lambda I_{2n} \nabla f_a(\theta) = 0,
\]

\[
\begin{bmatrix}
-(I, \otimes \Delta)^T D_R & -\lambda I_{2n} \\
-(I, \otimes \Delta) D_o & -(I, \otimes \Delta)^T \text{vec}(R_0^T)
\end{bmatrix} = 0. \tag{50}
\]

Equation (50) gives directly that

\[
\lambda \text{vec}(O E) + D_o (I, \otimes \Delta)^T \text{vec}(R_0^T) = 0 \tag{51}
\]

\[
D_R (I, \otimes \Delta) \text{vec}(O E) + \lambda \text{vec}(R_0^T) = 0. \tag{52}
\]

Premultiplication of equations (51) and (52) by \( (E^T \otimes I) \) and \( (E \otimes I) \), respectively, and application of the properties given in (45) and (46), gives

\[
(I, \otimes \Delta) \text{vec}(O E) + \lambda (E^T \otimes I) \text{vec}(R_0^T) = 0 \tag{53}
\]

\[
(I, \otimes \Delta)^T \text{vec}(O E) + \lambda (E \otimes I) \text{vec}(R_0^T) = 0. \tag{54}
\]

It follows then that

\[
\lambda O E + \lambda R_0^T E^T = 0 \tag{55}
\]

\[
\lambda O E + \lambda D^T R_0 = 0. \tag{56}
\]

Hence, the conclusion follows immediately using the fact that \( \Delta \) is an orthogonal matrix and \( \lambda \neq 0 \).

Thus, we concluded that minimum sensitivity factorizations with \( E \) fixed in the form given in (30) have corresponding realizations that are nearly balanced (letting \( i \to \infty \)) in the sense that they are only an orthogonal transformation (rotation) away from being balanced. This conclusion is analogous to that reached via the earlier algebraic approach in [10] which worked directly in the first realization space, rather than in this intermediate factorization space. When \( E = I \), the optimality condition reduces to the usual result for linear systems.
VI. Conclusions

In this paper, a definition of a system Hankel matrix is given for singular linear systems and used to develop a realization theory based on a factorization approach. The singular value decomposition was then applied to the system Hankel matrix to define a balanced generalized state-space realization. Such realizations were then used to characterize those realizations which have minimum parametric sensitivity properties.

Acknowledgements

This research was supported, in part, by the U. S. Air Force under contract AFOSR-87-0308, AFOSR-89-0241 and Northwestern University.

References


APPENDIX K

E. I. Verriest

Representations and Realizations of Singular Systems: The Tortoise and the Hare Revisited

Proc. 29th IEEE Conf. on Decision and Control

Honolulu, Hawai'i, pp. 61-66, December 1990.
REPRESENTATIONS AND REALIZATIONS
OF SINGULAR SYSTEMS:
THE TORTOISE AND THE HARE REVISITED

Erik I. Verriest
School of Electrical Engineering
Georgia Institute of Technology
Atlanta, Ga 30332

ABSTRACT

The notion of system description or representation is reconciled with system realization or implementation. By the latter a causal construct for obtaining the solution will be understood. The realization starts from a particular description, where dynamical and algebraic equations are separated, and the dynamical states correspond to integrator outputs. Clearly such a description is nonunique. The optimal realization problem is then to find the particular realization in the orbit of a particular group H, which minimizes the sensitivity measure in [7]. This group H is a subgroup, leaving the (separation) structure invariant, of the group characterizing the orbits under restricted system equivalence. A practical implementation is one for which asymptotically (in a well defined sense) the solution is obtained with a causal regular realization. The asymptotics are obtained by considering sequences of mathematical tools we propose, via the use of nonstandard analysis methods. The behavior of the regularizing parameter is determined by the structure at infinity of the original singular system. As examples, a differentiator and a purely algebraic set of equations are discussed.

Introduction

In previous work on the sensitivity minimization of singular systems [7], the problem of deriving the realization of the singular system was attacked. This problem is well understood for regular systems. Its practical significance is that in the presence of parameter disturbances the response of the (perturbed) optimal realization is close to the nominal or desired response. Now in following a similar programme for singular systems one is faced with an additional problem of interpretation. What is an implementation of a singular system? Realization in the pure sense always means an implementation with integrators. As is well known, singular systems may exhibit net differentiation. How is one to implement this? In discrete time, the situation seems even worse, since noncausal behaviour (i.e. advances) may result. On the other hand, any simulation or computation is inherently causal.

In this paper we shall try to reconcile the notion of system representation or description, and that of system realization or implementation. The latter will always mean a causal construction, for instance using a universal Turing machine. Only when this problem is satisfactorily answered, and practical ways of computing solutions to so-called singular systems have been found, will it make sense to optimize the computation, and speak of minimum sensitivity implementations.

The purpose of this presentation is to provide some ideas towards the solution of the above sketched problem, with practical implementations in mind. Some new tools will be given, in particular the nonstandard analysis. While still a young (a little more than twenty years) branch of mathematics, its presence has already been felt in the theory of differential equations. And while its name may insinuate abstractions, it is not true that it lies outside the "classical" domain of mathematics, nor is it in conflict with it. As expressed by Diener and Reeb [5] in their introduction, the nonstandard analysis adds "new possibilities to one's toolbox by giving the existing tools more power. This paper will then also only be a rather modest exploration of a potential use of nonstandard analysis in a branch of system theory. Perhaps more inspired researchers will smooth out the corners.

In this spirit, the differentiator is explored first in the next two sections, the last of which considers the infinite frequency behaviour more closely. The following section presents some general ideas distilled from this case study. In turn this is followed by a system consisting of pure algebraic equations. An effort is made to compute (in a causal way), solutions, or approximations of solutions to singular systems. Some reflections are collected in the conclusions, and an appendix gives a short "tutorial" on nonstandard analysis.

Case Study of a Singular System: The Differentiator

The differentiator has a singular representation

\[
\begin{bmatrix}
0 & 1 \\
0 & 0
\end{bmatrix}
\begin{bmatrix}
\dot{x}_1 \\
\dot{x}_2
\end{bmatrix}
= \begin{bmatrix}
1 & 0 \\
0 & 1
\end{bmatrix}
\begin{bmatrix}
x_1 \\
x_2
\end{bmatrix}
+ \begin{bmatrix}
0 \\
1
\end{bmatrix} u
\]

\[y = [-1, 0]' [x_1, x_2]'\]

Symbolically, we shall represent the system in an "open" form, as indicated in figure 1. Here two new symbols are introduced, a destructor or black hole (or sink), indicated by the big dark arrow, and a creator or white hole (a source), indicated by the big white arrow. The idea is to let these symbols represent the algebraic constraints in the above "system". Thus the black hole MUST have a zero signal going INTO it, whereas at the "white hole", a signal is CREATED, here \(x_1\), which forces the black hole input to be zero.
So far we have only a new picture, but nothing essentially new yet. Something new is obtained, at least heuristically, if we think of the black and white holes respectively as input and output of some fast system, as indicated in figure 1, sitting "behind the scene", and then relax the constraint that the black hole input must be zero.

![Figure 1. The "open form" representation.](image1)

Figure 2. The "fast" system.

Given a small deviation at the input of the "fast" system, it then quickly changes the signal at the white hole port, its output. Then one can hope that with this feedback structure, the black hole input will be driven to zero. This sounds plausible in words, but too familiar from the construction of, for instance, a new structure.

![Figure 2. The "fast" system.](image2)

Practically, this means now that if $u(t)$ has highest frequency components $\omega_0 = 2\pi f_0$, the implementation must be "faithful" up to frequency $2\pi f_0$. From the above discussion, this requires

$$\epsilon < 1/(4\pi^2 f_0^2)$$

The implementation is then

$$\dot{x}_2 = x_1$$

$$\dot{x}_1 = L(x_2 + u)/\epsilon$$

which has transferfunction

$$H_s(s) = aL/\{1 - a^2\}$$

Faithfulness yields then that

$$H_s(s) = s$$

The above heuristic ideas thus seem to work. However some more quantitative work and precise definitions will be needed. In particular it was learned from the above example that in singular problems, a regularization should be defined for which the system becomes purely dynamical. As usual this is done with some $\epsilon$-parameter which one lets tend to zero. It is here however that problems arise. The behaviour when $s$ goes to infinity is highly dependent on how $\epsilon$ approaches to zero in relation to $s$. In normal mathematical parlance, when one considers extensions, one considers sequences, introduces a notion of equivalence, and then considers the equivalence classes as the extended set, wherein the original set is nicely imbedded by mapping an element $x$ from the original set to the sequence $(x,x,x,x,\ldots)$. The construction is all too familiar from the construction of, for instance, the reals from the rationals. Here one identifies Cauchy sequences approaching the same limit, i.e. the equivalence class of all such Cauchy sequences defines the real number. However, this equivalence is too coarse for some applications. Indeed, using Cauchy sequences, the notion of "rate of convergence" is lost, as the sequences $(1/n)$ and $(1/n^2)$ define the same real (0), but are clearly remarkably distinguishable.

It is in this sense also that the singular systems are limit points of Cauchy sequences of regular systems. Such ideas have already been used by Hinrichsen and O'Halloran in [8], generalizing the idea of Young, Kokotovic and Utkin in [14] on high gain feedback. Only the behavior at infinity is ambiguous. This need not be so, if one takes the rates of convergence of these Cauchy sequences into account, as was clearly shown in the above differentiator example. Indeed for some rates the behaviour at infinity will be identical to that of the singular systems as described in [11]. Implementations (i.e. Cauchy sequences of regularized realizations) with such implied rates of convergence will be called faithful.

Clearly then we are faced with the problem of characterizing a singular system as a Cauchy sequence of regular systems, while retaining the information regarding the rate of convergence, pertaining to the structure at infinity. We need clearly more structure in our equivalence than is usually implied. It is now known that also the real line contains much more structure than is usually implied. The reals can be imbedded in the "hyperreals". Hyperreals can be
identified with the equivalence classes of sequences of reals \( a_1, a_2, \ldots \). Two sequences are considered equivalent if they agree a.e. (in a well specified measure, based on the notion of an ultrafilter). The hyperreals contain the infinitesimally small and their inverses, the infinitely large numbers. Nonstandard analysis seems therefore to be the right framework to analyze the realization problem for singular systems. But before galloping too far, it must be mentioned that everything that can be shown with nonstandard analysis can also be shown using conventional tools, but at the expense of some more work. Nonstandard analysis merely provides a convenient language. First let us consider again the example of the differentiator in order to shed some more light on this idea.

**Structure at infinity of the differentiator.**

Consider an equation decomposition form of the differentiator

\[
\begin{bmatrix}
1 & 0 \\
0 & 1 \\
0 & \varepsilon \\
0 & 0
\end{bmatrix}
\begin{bmatrix}
x_1 \\
x_2 \\
x_1 \\
x_2
\end{bmatrix} = \begin{bmatrix}
0 & 1 \\
1 & 0 \\
0 & L \\
L & 0
\end{bmatrix}
\begin{bmatrix}
x_1 \\
x_2 \\
0 \\
0
\end{bmatrix} + \begin{bmatrix}
0 \\
1 \\
0 \\
0
\end{bmatrix} u
\]

\[
y = [0 -1] \begin{bmatrix} x_1 \ x_2 \end{bmatrix}^	op
\]

and its associated regularized representation

\[
\begin{bmatrix}
1 & 0 \\
0 & \varepsilon \\
0 & 1 \\
0 & 0
\end{bmatrix}
\begin{bmatrix}
x_1 \\
x_2 \\
x_1 \\
x_2
\end{bmatrix} = \begin{bmatrix}
0 & 1 \\
1 & 0 \\
0 & L \\
L & 0
\end{bmatrix}
\begin{bmatrix}
x_1 \\
x_2 \\
0 \\
0
\end{bmatrix} + \begin{bmatrix}
0 \\
1 \\
0 \\
0
\end{bmatrix} u
\]

\[
y = [0 -1] \begin{bmatrix} x_1 \ x_2 \end{bmatrix}^	op
\]

Their polar structures are respectively given by [11]

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

and since \( L = 0 \), the latter has the same (finite) zero structure as \( \varepsilon \rightarrow 0 \). The zero-structure at infinity is determined by the zero structure at \( \lambda = 0 \) of respectively

\[
(1/\lambda) X - A = \begin{bmatrix}
(1/\lambda) & -1 \\
-1 & 0
\end{bmatrix}
\]

and

\[
(1/\lambda) \hat{X} - \hat{A} = \begin{bmatrix}
(1/\lambda) & -1 \\
-L & \varepsilon/\lambda
\end{bmatrix}
\]

The first one has the Smith-McMillan form

\[
\begin{bmatrix}
1/\lambda & 0 \\
0 & \lambda
\end{bmatrix}
\]

thus displaying a zero at \( \lambda = 0 \). The \( \varepsilon \)-implementation has the Smith-McMillan form

\[
\begin{bmatrix}
1/\lambda & 0 \\
0 & (1^2 - \varepsilon L)/\lambda
\end{bmatrix}
\]

Faithfulness at \( \lambda = 0 \) requires that \( \lambda^2 - \varepsilon L \) "behaves" as \( \lambda^2 \), implying in turn that \( \varepsilon \) goes to zero faster than \( \lambda^2 \).

The zero structure of the singular representation is determined by the matrix

\[
\begin{bmatrix}
s \varepsilon - A & B \\
C & 0
\end{bmatrix}
\]

For instance for the input decoupling zeros, we find

\[
\begin{bmatrix}
s \varepsilon - A & B \\
C & 0
\end{bmatrix}
\]

which has full rank for all finite \( s \), implying finite controllability and reachability. At infinity, using a Möbius transformation \( s = 1/\lambda \), the invariant factors are found to be \( 1/\lambda \) and \( 1 \), indicating the absence of an input decoupling zero. Incidentally, also note that \( [s, B] \) has full rank, so that the realization is also reachable at infinity [9][13].

Now let us turn to the regularized representation. The controllability pencil is now

\[
\begin{bmatrix}
s \varepsilon - A & \hat{A} \\
L & s \varepsilon L
\end{bmatrix}
\]

and it has obviously the same finite zero structure as the singular pencil, as long as \( \varepsilon \) converges to zero. The infinite zero structure is now obtained from the zero structure at \( \lambda = 0 \) of the reduced form

\[
1/\lambda \begin{bmatrix}
1 & 0 & 0 \\
0 & \lambda & \varepsilon
\end{bmatrix}
\]

Since we already established that \( \varepsilon = 1^2 \), this shows that the \( \lambda \)-realization has no input decoupling zero at infinity.

In fact note that at infinity, the system matrix of the singular system and the realization have the same zero structure (i.e. no zero at infinity), however their singular structure differs.
Representations and Realizations.

Motivated by the analysis of the differentiator in our previous section, we now turn to the discussion of representations and realizations. Simply stated a realization or implementation should mean causal implementation, say with integrators. Therefore, the derivatives of all state-variables must appear on the left hand side of an equation with coefficient one, i.e., realization = 1. A representation or description is nothing but a set of equations (dynamic and algebraic) that must be satisfied by the variables in the discourse: inputs, outputs and the x's. We refer to these also as the generalized system equations.

Obviously representations (descriptions) carry the same information under restricted system equivalence [8]. On the other hand restricted equivalence cannot be allowed for realizations in the above sense. A pseudo realization form is obtained by writing the generalized system equations in the so-called Second Equivalent Form [4], which decouples the dynamical equations from the algebraic constraints. This decomposition reflects the physical meaning of interconnected regular subsystems.

\[
\begin{bmatrix}
  \dot{x}_1 \\
  0
\end{bmatrix} = \begin{bmatrix}
  A_{11} & A_{12} \\
  A_{21} & A_{22}
\end{bmatrix} \begin{bmatrix}
  x_1 \\
  x_2
\end{bmatrix} + \begin{bmatrix}
  B_1 \\
  B_2
\end{bmatrix} u
\]

In contrast, the Weierstrass-form or First-Equivalent Form seems to be more useful for characterizing the solutions of the generalized system, and its associated observability and reachability properties. An equivalence leaving the above structure invariant is obtained by left and right multiplication by respectively

\[
\begin{bmatrix}
  U_{11} & U_{12} \\
  0 & U_{22}
\end{bmatrix} \quad \text{and} \quad \begin{bmatrix}
  V_{11} & 0 \\
  V_{12} & V_{22}
\end{bmatrix}
\]

where \((U_{11}, V_{22}) \in \text{GL}(n_1) \times \text{GL}(n_2)\), and \(V_{11} U_{11}^{-1}\) and \(V_{12} U_{22}\) arbitrary.

Now obtain an implementable realization from the pseudo realization by regularization. Because of the form we refer to this as the \(\epsilon\)-realization. This simply means that an \(x_\epsilon\) is placed where the zero appears in the pseudo realization. In view of the equivalence described above we shall just consider that the algebraic set of equations is replaced by

\[
x_\epsilon = L(A_\epsilon x_\epsilon + A_{21} x_1 + B_2 u)/\epsilon
\]

But notice that this is nothing but an observer, with infinitely large gain \(L/\epsilon\), for the system with state \(x_\epsilon\) and \((A_\epsilon, C) = (0, -A_{22})\), receiving an "output-measurement" \(y_\epsilon = A_\epsilon x_\epsilon + B_2 u\). This "fast" system is observable if \(r(A_{22}) = n_2\), i.e., if \(A_{22}\) has full rank. Since this rank is invariant under the above described equivalence, lack of observability of the fast system cannot be overcome, unless regularization is brought into the picture, e.g., \(A_{22} \rightarrow A_{22} + \mu I\), thus making the system \((0, A_{22} + \mu I)\) observable. Notice that generalized systems for which \(A_{22}\) has full rank are in fact redundant representations of regular systems, since we can always solve for

\[
x_\epsilon = -A_{22}^{-1}(A_{21} x_1 + B_2 u)/\epsilon
\]

and backsubstitute

\[
x_1 = (A_{11} - A_{12} A_{22}^{-1} A_{21}) x_1 + (B_1 - A_{12} B_2) u
\]

\[
y = (C_1 - C_2 A_{22}^{-1} A_{21}) x_1
\]

Application to Purely Algebraic Equations

In this section we describe the dynamical solution of a set of linear equations \(y = Ax\), where \(A\) is square and invertible. This is indeed a special case of the representation, having no dynamical elements, but containing \(2n\) "states". However, we can model this by some pseudo-realization

\[
\begin{bmatrix}
  y \\
  0
\end{bmatrix} = \begin{bmatrix}
  Fy + Gx \\
  0
\end{bmatrix}
\]

and thus the algebraic subsystem has dimension \(n\).

The \(\epsilon\)-realization yields for the fast subsystem the "observer"

\[
x = -Lx/\epsilon + Ly/\epsilon
\]

which has the solution

\[
x(t) = A_{\epsilon}^{-1}y + \exp(-Lt/\epsilon) \left( x_0 - A_{\epsilon}^{-1}y \right)
\]

For instance, the choice \(L = A'\) yields a balanced realization [CV], if one lets \(\epsilon = 1\). In this case we actually have no reference for how fast "fast" really is, since there is no dynamical equation. This balancedness is with respect to the disturbability due to the measurement error, and the reconstructability of \(y\) from \(x\), thus respectively the equations

\[
x = -A'Ax - A'\gamma
\]

and

\[
x = -A'Ax ; \quad y = Ax
\]

Their associated reachability and observability grammians equal \(I\).

Conclusions

This paper characterized the implementability of a generalized system in terms of the ability to build an implementation that asymptotically remains faithful to the properties of the original description. The practicality of this lies in the fact that all physical signals are inherently bandlimited, and the description, more particularly its structure at infinity, gives then an idea of how "fast" practical implementations should work. This way the tortoise can outrun the hare! Rates of convergence are important, and therefore we propose to use the language of nonstandard analysis. This has actually not been done yet in this preliminary version of the paper, as we were merely interested in demonstrating the feasibility of an idea. For the purpose of orientation in this field, an appendix on nonstandard analysis is included.
Nonstandard analysis is a modern approach to using infinitesimals in analysis to express limits and its derived notions. The theory is originated by Abraham Robinson and modeled after Leibnitz's theory of infinitesimals. Its notion of "infinitely close" is useful in representing limits, not just on the real axis, but also in a topological sense, and even in contexts where the notion is not exactly topological. This is the notion which will make it useful to study the theory of singular systems.

The essential ingredient of the nonstandard theory is the observation that the real line allows for a much richer structure than it is usually endowed with. Whereas classically the reals are defined as Dedekind cuts or Cauchy sequences, the richer structure is obtained by a similar procedure, but using instead the notion of a free ultrafilter. Hence, besides the usual reals, which will be called standard, the new set of hyperreals will also contain additional "nonstandard" elements. Intuitively speaking the new elements build up a universe of infinitessimals near each standard real. Every element in this universe is infinitely close to the given real. Infinitely close means that the distance is smaller than any nonzero ordinary standard real. The fact that this is all brought on firm logical foundations, makes the rules for manipulating infinitesimals rigorous. But not only are "infinitely small" numbers brought in, the hyperreals also contain the "infinitely large" numbers, the inverses of the infinitesimals. Once the structure of the hyperreals is defined, it is possible to speak of nonstandard functions, operators, and other mathematical objects in the same vain. Moreover, all the theorems of the ordinary standard mathematics apply in this enriched universe, provided of course that they are appropriately interpreted. This property is referred to as the Transfer Principle, or Leibnitz's Principle, as he proposed that all infinitesimals should obey the same rules as ordinary standard numbers.

Transfer Principle
The weak form of the transfer principle postulates that for every standard formula P(x) having no other free variables than x, we have

\[ \forall^* x \; P(x) = \forall x \; P(x) \]

or, equivalently,

\[ 2 \forall x \; P(x) = 3 \forall x \; P(x) \]

where \( \forall^* x \; P(x) \) means \( \forall x \; [s(x) \equiv P(x)] \) and \( 3 \forall x \; P(x) \) means \( 3 \forall [s(x) \; P(x)] \). Here s(x) indicates that x is standard, i.e. an element in the usual discourse of mathematics.

Of course it is also possible to project the hyperreals on the reals. The operation is referred to as taking the Standard Part. This standard part is also defined for other mathematical objects. A typical approach in nonstandard analysis is to obtain a continuous standard object from a discrete nonstandard object. Infinitary mathematics is obtained from finitary mathematics (the nonstandard construction). The complete nonstandard solution to a standard problem consequently involves first of all a "lifting" of the given standard problem to a nonstandard one, of which it is the standard part. Next comes the nonstandard solution which is usually finitary. Then one shows that the standard part to the solution exists, and that this solves also the standard problem. Free movement between the standard world and the nonstandard one is allowed by the "Transfer Principle", and the "Standard Part" map. Such an approach eliminates much of the burden of modern mathematical rigor, since it deals rather simply and in a more naive way with the infinitesimals and the infinitely large. In order to illustrate its power, consider the criterion for continuity in nonstandard analysis:

Given \( f: \mathbb{R} \to \mathbb{R} \), a standard function and \( x_0 \) a standard real number. Then \( f \) is continuous at \( x_0 \) if and only if

\[ \forall \varepsilon > 0 \; \forall x \in \mathbb{R} \setminus \{x_0\} \; [f(x) - f(x_0)] < \varepsilon \]

The equivalence identifies numbers that are infinitesimally close, i.e. \( x \sim y \) is equivalent to stating that \( x - y \) is infinitesimal.

Of course it has the disadvantage that the language is not yet common, especially to the nonspecialist who should be using it.

A very readable introduction to nonstandard analysis is the recent book [in French] by Diener and Reeb [5]. An approach via nonstandard analysis to probability is for instance described in [10]. It contains a remarkably simple and concise introduction to nonstandard analysis. Applications to singular perturbation theory of ordinary differential equations are described in [2] and [16]. As far as this author knows, the only attempt to introduce nonstandard analysis in systems theory was [3], in the context of instantaneous stabilizability and its robustness properties. The recent book [1] on nonstandard analysis has a chapter devoted to differential operators. Its introduction deals with calculus, topology and linear spaces in nonstandard mathematics.

It is this authors aim to see if by rephrasing the singular system problem in the language of nonstandard analysis, some of the obscurity presently cloaking the theory cannot be eliminated, and thus its applicability enhanced to a wider community.

References


APPENDIX L

E. I. Verriest

On a Hyperbolic PDE Describing the Forward Evolution of a Class of Randomly Alternating Systems

Proc. 29th IEEE Conf. on Decision and Control

ON A HYPERBOLIC PDE DESCRIBING THE FORWARD EVOLUTION
OF A CLASS OF RANDOMLY ALTERNATING SYSTEMS

Erik I. Verriest
School of Electrical Engineering
Georgia Institute of Technology
Atlanta, GA 30332

Abstract

For a class of randomly switched linear systems, a transition functional is introduced. It is shown that the expectation of this functional at time t satisfies a HYPERBOLIC partial differential equation, which plays a similar role as the backwards Kolmogorov equation for diffusions. Its formal adjoint leads to the forward equation, and their complexity is determined by the Lie algebra associated with the set of values assumable by the dynamic matrix A(t). The usual PARABOLIC Kolmogorov equation is derived from this as a limiting case. The result leads to Monte-Carlo simulation methods for solving hyperbolic PDE's.

Introduction

In this paper, the class of stochastically switched systems is considered, where the system parameters are piecewise constant, and assumes only a finite number of values, i.e. 2(t) E {2, ..., 2}. The applications of such models are widespread: from target tracking, where the parameter change occurs as changes in acceleration, bank angle etc., to fault tolerant control, the different modes being associated with different failure nodes. These hybrid systems have also been used as approximations for certain nonlinear systems [1]. In this paper, N = 2, but the generalization is straightforward although of increased complexity. The switching phenomenon is assumed to occur at purely random times, and its stochasticity is determined by the probability of a switch in an infinitesimal interval of length At be At. If N(t) denotes the number of switches in the finite interval [0, t], then it is well known that the probability that N(t) equals k is given by the ubiquitous Poisson formula

\[ \text{Prob} \{ N(t) = k \} = e^{-\lambda t} \frac{\lambda^k}{k!} \]

Furthermore, if the switching times are ordered, \( \tau_1 < \tau_2 < \cdots < \tau_n \), then the increments \( N(t_{\tau_k}) - N(t_{\tau_{k-1}}) \) are independent. Our method generalizes a result by Kac [2], where the one-dimensional motion of a particle with constant speed \( v \), but with random (at Poisson times) reversing direction was considered, i.e. the first order affine system \( dx/dc \in (v, -v) \). In his paper, the order of any function of the position was interpreted as a solution to the wave equation, but with a "random path time" substituted for the real time, followed by averaging over all paths. This led to Monte-Carlo methods for solving the related Klein-Gordon equation, which is significant in quantum electrodynamics, and factors to the Dirac equation [3].

First we define the notion of a general transition functional: \( F_{\phi,u} \in \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R} \) where \( \phi \) is a smooth map (linear or nonlinear) from the state space \( \mathbb{R}^n \) to \( \mathbb{R} \), and \( F_{\phi,u} \) implicitly defined by

\[ F_{\phi,u}(x(t), u(t), t) = \phi(x(t)) \]

This research is supported by a grant from AFOSR and Northwestern.

\[ \frac{dx(t)}{dt} = f(x(t), u(t), t) ; x(0) = x \]

Main Results

Consider the alternating undriven linear system

\[ \frac{dx(t)}{dt} = A(t)w(t), \]

where the random matrix \( A(t) \) assumes the values \( A_0 \) and \( A_1 \) according to a random switching process. For instance the matrix process \( A(t) \) is patterned after a telegraph signal, based on the number \( N(t) \) of random points in the interval \([0, t)\). If this number is even, then \( A(t) = A_0 \), whereas if it is odd, then \( A(t) = A_1 \).

Define also the averaged dynamic matrix and its "excursion" respectively by \( A_+ = (A_0 + A_1)/2 \) and \( A_- = (A_0 - A_1)/2 \). An approximate solution of the random system is first obtained by discretizing the time in steps of length \( \Delta t \), and letting the switching be commensurate with these sampling times. The matrix \( A(\Delta t) \) can be expressed as

\[ A_0 + \epsilon_n \delta_n, \epsilon_1 

where \( \epsilon_n \) is a Bernoulli (1, -1) process, with \( \text{Pr}(\epsilon_n = 1) = 1/2 \). A switching at step \( k \) corresponds then to \( \epsilon_k = -1 \). Hence the state \( x(t) \), if \( A(0) = A_0 \), is the limit for \( \Delta t/n \to 0 \)

\[ \left[ A_0 + \epsilon_n \delta_n \right] [x_1, x_2, \ldots, x_{n-1}, x_n]_{\text{mod} 1} \]

The state transition matrices \( F_+ \) and \( F_- \) follow

\[ F_+ = [A_0 + \epsilon_n \delta_n] [x_1, x_2, \ldots, x_{n-1}, x_n]_{\text{mod} 1} \]

The expected transition functionals are now obtained by

\[ F_{\phi,u}(x) = \mathbb{E} \phi(F_{\phi,u}^+), \]

and by taking the expectations over \( \epsilon_1 \) separately one derives the recursions

\[ F_{\phi,u}^+(x) = (1 - \lambda) F_{\phi,u}^+(x) + \lambda F_{\phi,u}^-(x) \]

\[ F_{\phi,u}^-(x) = (1 - \lambda) F_{\phi,u}^-(x) + \lambda F_{\phi,u}^-(x) \]

By reorganizing the terms, and taking limits for \( \Delta t \to 0 \), one obtains the partial differential system

\[ \frac{\partial \phi(x, t)}{\partial t} = \frac{\partial F_{\phi,u}^+(x, t)}{\partial x} \]

\[ \frac{\partial \phi(x, t)}{\partial t} = \frac{\partial F_{\phi,u}^-(x, t)}{\partial x} \]
Upon setting
\[ G = \left[ F^* + F \right]/2 \]
\[ H = \left[ F^* - F \right]/2 \]
this is equivalent to the system of PDE's
\[ \delta G/\delta t = (\delta G/\delta x) A_0 x + (\delta H/\delta x) \partial x \]
\[ \delta H/\delta t = (\delta H/\delta x) A_0 x + (\delta G/\delta x) \partial x - 2A \]
Its initial conditions follow directly from the definition of \( G(x,t) \). Indeed, \( G(x,0) = -\nu(x) = -\psi(x) \), since randomization over \( A(0,\omega) \) just gives the identity. The second set of conditions follows similarly from \( H(x,0) = 0 \). This then proofs the following theorem.

Theorem 1. The evolution \( F \) of the randomly switched system (3) is given by the following pair of PDE's
\[ \delta F/\delta t = (\delta F/\delta x) A_0 x + (\delta H/\delta x) \partial x \]
\[ \delta H/\delta t = (\delta H/\delta x) A_0 x + (\delta F/\delta x) \partial x - 2A \]
with initial conditions \( F(x,0) = \psi(x) \) and \( H(x,0) = 0 \).

The coupled set of PDE's plays the role of the backwards Kolmogorov equation (5). Introducing the first order differential forms \( \delta_\omega = \delta/\delta t + \varphi(x) \), \( \delta_\omega = \delta/\delta x + \psi(x) \), and \( \delta_\omega = \delta/\delta x, \delta_\omega = \delta/\delta x \), the following special case can be deduced.

Corollary 2. If \( A_\omega \) and \( A_\omega \) commute, then the evolution of the system (3) is governed by the hyperbolic PDE
\[ (\delta_\omega + 2\delta_\omega)(\delta_\omega - \delta_\omega)F = (\delta_\omega)^2 F \]
with initial conditions \( F(x,0) = \psi(x) \) and \( \delta F(x,0) = 0 \).

Proof: Indeed, if \( A_0 \) and \( A_0 \) commute, then so do \( A_\omega \) and \( \partial \). But then the differential operators \( \delta_\omega \) and \( \delta_\omega \) commute, and upon elimination of \( H(x,t) \) one obtains the PDE (13). That the initial conditions are as stated follows also from the main theorem.

If \( \rho^+ \) and \( \rho^* \) represent the conditional density (assuming it exists) of \( x(t) \) given \( x(0) = x \), and \( A(0) \) respectively \( A_\omega \) and \( A_0 \), then it follows from
\[ \delta /\delta t \rho^+, \phi = -\rho^+, \left[ x' A_\omega (\delta/\delta x) - \lambda \phi \right] + \rho^+, \lambda \phi + \rho^+, \phi \]
that (in the weak sense)
\[ \delta \rho^+/\delta t = -V(A_\omega x) - \lambda (\rho^+ - \rho^-) \]
\[ \delta \rho^-/\delta t = -V(A_\omega x) - \lambda (\rho^- - \rho^+) \]
The density of \( x(t) \) is finally obtained by \( \rho(x,t) = \rho^+(x,t) + \rho^*(x,t)/2 \).

Using arguments similar to the ones in [1], it can be shown [6] that this equation has an interesting asymptotic form for \( \lambda \to \infty \) and \( \lambda(t,\omega) = A_\omega (t,\omega) \) with \( \mu \) such that \( \mu /\lambda \) is kept constant (Q say). Indeed, a PARABOLIC PDE results, which is equivalent to the Ito-differential system where \( w(t) \) is a Wiener process with \( E(w(t)w(s)\mid s) = \sigma^2 \).

\[ dx = \left( A_0 + Q \right) x \partial t + \partial_0 x \partial w(t) \]
Clearly, the "jittering" caused by very fast switching over very large amplitudes in the direction \( \partial_0 \) has the same effect as a diffusion. The drift is however NOT the one given by the averaged dynamics \( A_\omega \), but an additional drift \( \partial_0 x \) is present. This can be stabilizing or destabilizing, depending on \( \partial_0 \). For instance if \( \partial_0 \) has imaginary eigenvalues, stabilization may occur, since \( \partial_0^2 \) has then negative eigenvalues. If on the other hand one has high frequency switching, but \( \mu \) remains finite, then the stochastic energy (the integral of \( |A(t)cA(t+c)|^2 \) is zero, and the dynamics of the averaged system is all that remains.

In the noncommutative case a higher order PDE for \( F \) is obtained. Its structure depends on the dimension of the Lie algebra, generated by \( A_\omega \) and \( A_\omega \). For instance if the commutator of \( A_0 \) and \( A_\omega \) is nonzero, but commutes with both, then it is known [6] that \( G \) satisfies a third order PDE.

For \( \lambda = 0 \), the equations are readily solved in terms of the characteristics which are exactly the deterministic evolutions according to the different modes. For \( \lambda = 0 \), the solutions of the PDE are still interpreted in terms of the characteristics, but via a "random time operator" [6].

Summary and Extensions
It was shown that for randomly alternating systems, a hyperbolic system of first order PDE's describes the behavior of the system. From this, a single higher order PDE results through elimination of the auxiliary variables. The mechanization of this elimination process and its ensuing complexity is determined by the Lie algebra generated by the \( A(t) \) values. By using the formal adjoint, this can be interpreted that under some smoothness assumptions the density satisfies a type of forward Kolmogorov or Fokker-Planck equation, which in this case is also of hyperbolic type. It was shown that asymptotically, the parabolic equations of diffusion type result, if the limits are taken in such a way that the stochastic energy is conserved in the limiting system. The results presented here were for linear autonomous systems, but extend easily to the nonlinear driven case with markovian switching between a countable number of models.

Finally, one can reverse the ideas and develop stochastic solution methods for hyperbolic PDE's as was done for parabolic and elliptic ones based on Dynkin's equation [5]. Indeed, such Monte Carlo simulation methods are not based on the stochastic evolution in the (narrow) Ito sense (i.e. based on an underlying Brownian process), but on a counting process.

References
APPENDIX M

J. A. Ramos and E. I. Verriest

A 2-D Realization Theory for Markov Chains

Proc. 29th IEEE Conf. on Decision and Control
Honolulu, Hawaii, pp. 853-858, December 1990.
A 2-D Realization Theory for Markov Chains

Jose A. Ramos
United Technologies Optical Systems
P.O. Box 109660, MS-9
West Palm Beach, FL 33410-9660

Erik I. Verriest
Department of Electrical Engineering
Georgia Institute of Technology
Atlanta, GA 30332

Abstract

In this paper we study the dynamics of time-homogeneous Markov chain models from a state-space modeling point of view. It is shown that a Markov chain model can be embedded in a 2-D realization theory where markov parameters correspond to higher-order transition probabilities. The implications of formulating a Markov chain model in this state-space domain is that many equivalent representations may exist, some of which may have better robustness properties. A modified Hankel approximation algorithm is presented which exactly matches all the Markov parameters. The algorithm is an extension of the 2-D harmonic retrieval algorithm introduced in [6].

1. Introduction

Markov chain models have been used extensively to model random phenomena with a particular type of dependence; the Markov dependence. A stationary, finite state Markov chain is defined as a stochastic process having a finite number of states, the Markovian dependence, stationary transition probabilities, and an initial set of probabilities \( \pi(0) \). Such a process is said to be memoryless; the future behavior depends only on its present state and not on its past history. Hence, only a limited amount of information is required to propagate the conditional distribution of a Markov process. Such Markov structure arises in connection with decision making under uncertainty [1], queueing theory [2], hidden Markov models [3], stochastic dynamic programming [4], and the solution of linear algebraic, integral, and differential equations [5], to name only a few.

Although Markov chains have the concept of state and Markov propagation property embedded in it, there does not seem to be any connection with the state-space formulation of linear dynamical systems. Having an equivalent linear systems theory for Markov chain models, one can select a canonical representa-
are called one-step transition probabilities, and are said to be stationary if
\[ p_{ij}(1) = p(\xi(n+1) = j | \xi(n) = i) = p(\xi(1) = j | \xi(0) = i) \quad \forall \ n = 0, 1, 2, \ldots \]

so that the transition probabilities remain unchanged through time. These values may be displayed in a matrix \[ P(1) = [p_{ij}(1)] \]
called the one-step transition matrix. The \((N \times N)\) matrix \[ P(1) \]
satisfies
\[ 0 \leq p_{ij}(1) \leq 1 \quad (4a) \]

\[ \sum_{j=0}^{N-1} p_{ij}(1) = 1 \quad \text{for} \ 0 \leq i \leq N-1 \quad (4b) \]

The existence of one-step, stationary transition probabilities implies the existence of higher-order transition probabilities, which can be computed from the Chapman-Kolmogorov equations, i.e.,
\[ p_{ij}(k+s) = \sum_{m=0}^{N-1} p_{im}(k)p_{mj}(s) \quad (5) \]

\[ \forall \ k, s = 0, 1, 2, \ldots \ \text{and} \ 0 \leq i, j \leq N-1. \]

Here \[ p_{ij}(k) = p(\xi(n+k) = j | \xi(n) = i) \]
called \( k \)-th step transition probabilities, and may be displayed in a \( k \)-th step transition matrix \[ P(k) = [p_{ij}(k)] \]
where, in general
\[ 0 \leq p_{ij}(k) \leq 1 \ \forall \ k = 0, 1, 2, \ldots \]
and \( 0 \leq i, j \leq N-1 \)
\[ \sum_{k=0}^{N-1} p_{ij}(k) = 1 \ \forall \ k = 0, 1, 2, \ldots \]
and \( 0 \leq i \leq N-1 \)

It should be noted that \[ p_{ij}(0) = \delta_{ij} \]
(Kronecker delta), thus, \( P(0) = I_N \)
\((N \times N)\) identity matrix).

The unconditional probability of \( (\xi(n)) \)
being in state \( j \) at time \( n = k \) is given by
\[ \pi_{ij}(k) = p(\xi(k) = j | \xi(0) = i) = \sum_{k=0}^{N-1} \pi_i(0)p_{ij}(k) \]
\[ \forall \ k = 1, 2, \ldots \ \text{and} \ 0 \leq j \leq N-1 \]
and in row vector form
\[ \pi(k) = [\pi_1(k) \pi_2(k) \pi_3(k) \ldots \pi_N(k)] \]

In general, for irreducible, ergodic Markov chains, the steady-state probabilities \( \pi_j \)
are independent of \( i, \) i.e.,
\[ \lim_{k \to \infty} p_{ij}(k) = \lim_{k \to \infty} \pi_j(k) = \pi_j \]

and satisfy the following conditions:
\[ 0 < \pi_j \leq 1 \]
\[ \sum_{j=0}^{N-1} \pi_j = 1 \]

Finally, for \[ \overline{P} = \lim_{k \to \infty} P(k), \lambda = 1 \]
is the only nonzero eigenvalue and \( \pi \)
and \( 1 = [1 \ 1 \ 1 \ldots 1] \]
are its left and right eigenvectors, respectively.

The interested reader may consult references [7,8] for further details on Markov chains.

3. 2-D realization Theory for Markov Chains

Consider a 2-D state-space model such as
\[ x_{i,j}(k) = A_1x_{i,j}(k) \]
\[ x_{i,j}(k) = A_2x_{i,j}(k) \]
\[ p_{ij}(k) = Cx_{i,j}(k) \]

where \( x_{i,j}(k) \)
is an \((N \times 1)\) state vector, \( A_1 \) and \( A_2 \)
are \((N \times N)\) constant matrices, \( c \)
is a \((1 \times N)\) vector, and \( p_{ij}(k) \)
is a scalar measurement corresponding to the \((i,j)\)th element of the \(k\)-th order transition matrix. The dynamics of the Markov chain can be incorporated by allowing the state vector to vary with transitions, i.e,
\[ x_{i,j}(k+1) = Wx_{i,j}(k) \]
where \( W \)
is an \((N \times N)\) transition matrix. In addition we assume that \( A_1 \) and \( A_2 \)
are stability matrices and the pairs \((A_1, c)\)
and \((A_2, x_{i,j}(0))\)
are observable and controllable, respectively. If we recursively solve the state equations (10a), (11), and (10b), and substitute them in (10c), we find that
\[ p_{ij}(k) = cA_I W^k A_2x_{i,j}(0) \]
corresponds to the Markov parameters of the 2-D model (10)-(11). However, it should be clear from our 2-D model that the matrices \( A_1, A_2, \) and \( W \)
must commute with each other. As we will see later, the constraints imposed by \( P(0) \)
does not allow this commutativity property to hold. This imposes a constraint on the order in which the state equations can be updated. In
order to avoid any confusion with this partial ordering, we rewrite state equations (10b) and (11) as

\[ x_{j+1}(k) = A_2 x_{j}(k) \]  \hspace{1cm} (10b')
\[ x_j(k+1) = W x_j(k) \]  \hspace{1cm} (11')

Notice that (11') implies \( x_{j}(k+s) = W^k A_2 x_{j+s}(s) \). Also, the order of state updates is j-k-i (column, time, row), which leads to the Markov parameters (12).

The solution to the Chapman-Kolmogorov equations yield the higher-order summation (p(O) = p(C(O)) = N), and by the observability and controllability assumption (\( \rho(O) = \rho(C(O)) = N \)), \( O = C(O)^{-1} \). Furthermore, \( Ox_{x_0} = e_1 \) (the first element of the standard basis in \( \mathbb{R}^N \)) and \( Ox_{x_0} = 0 \). Thus, \( A_2 x_{x_0} = 0 \) implies that \( x_{x_0} \) is an element of \( N_{sp}(A_1) \). In fact, \( \text{span}\{x_{x_0}\} = N_{sp}(A_1) \).

We remark that (15) requires \( N \) to be large or \( A_1^n = A_1^m = 0 \) \( \forall m \geq N \). In the following theorem we prove the latter case, along with other properties from (A_1, A_2, W, c, x_{x_0}(0))_n.

**Theorem 1:** Given an Nth-order 2-D realization (A_1, A_2, W, c, x_{x_0}(0))_n, the following properties have to be satisfied in order for it to characterize a Markov chain:

i) \( x_{x_0}(0) \in N_{sp}(A_1) \) and \( c^T \in N_{sp}(A_2^T) \) such that \( c x_{x_0}(0) = 1 \)

ii) \( A_1^n = A_2^m = 0 \) \( \forall m \geq N \)

iii) \( A_2 A_1 = I_N - x_{x_0}(0)c \)

iv) \( \rho(A_1) = \rho(A_2) = \rho(A_1 A_2) = N \)

v) \( \lambda(A_1) = \lambda(A_2) = 0 \) ; \( k = 1, 2, \ldots, N \) and \( \lambda(A_1 A_2) = (1, 1, 1, \ldots, 1) \)

vi) \( A_2 A_1 A_1 = A_1 \) and \( A_2 A_2 A_2 = A_2 \)

vii) \( A_1, A_2, \) and \( W \) cannot commute with each other

**Proof:** Recall that \( P(0) = OC(0) = I_N \), where

\[ O = \begin{bmatrix} c \\ c A_1 \\ \vdots \\ c A_1^{N-1} \end{bmatrix} \]

\[ C(0) = [ x_{x_0}(0) \ A_2 x_{x_0}(0) \ A_2^2 x_{x_0}(0) \ \ldots \ A_2^{N-1} x_{x_0}(0) ] \]

and by the observability and controllability assumption (\( \rho(O) = \rho(C(O)) = N \)), \( O = C(O)^{-1} \). Furthermore, \( Ox_{x_0} = e_1 \) (the first element of the standard basis in \( \mathbb{R}^N \)) and \( Ox_{x_0} = 0 \), thus \( A_2 x_{x_0} = 0 \) implies that \( x_{x_0} \) is an element of \( N_{sp}(A_1) \). In fact, \( \text{span}\{x_{x_0}\} = N_{sp}(A_1) \).

A dual argument implies \( c A_2 C(0) = 0^T \) and thus \( c^T \) is an element of the left null space of \( A_2 \). The normalization comes from \( p_{x_0}(0) = c x_{x_0}(0) = 1 \). To prove property (ii) we need to make use of the Cayley-Hamilton Theorem which states that \( \Delta(A_1) = \Delta(A_2) = 0 \) (characteristic polynomial), i.e., \( \forall 0 \leq j \leq N-1 \)

\[ A_1^N = \alpha_{N:1} A_1^{N-1} - \alpha_{N:2} A_1^{N-2} \ldots - \alpha_1 A_1 \]

\[ c A_1^{N} A_2 x_{x_0}(0) = c A_1^{N} A_2 x_{x_0}(0) = c A_1^{N} A_2 x_{x_0}(0) = \ldots = c A_1^{N} A_2 x_{x_0}(0) = c A_1^{N} A_2 x_{x_0}(0) \]

\[ = -c_1 = 0 \]

thus, \( c A_1^N C(0) = 0^T \) and since \( c^T \in N_{sp}(A_2^T) \), it follows that \( A_1^N = 0 \). A dual argument can be used to show that \( A_2^N = 0 \). Property (iii) follows from the fact that \( W_{x_0}(0) = C(O) = I_N \) and the use of (15). To prove property (iv) we need the following identities from [10, pp. 140 - 141]:

\[ 1 - c x_{x_0}(0) = | I_N - x_{x_0}(0) c | \]
\[ | x_{x_0}(0) c - \mu I_N | = (\mu)^{N-1} | c x_{x_0}(0) - \mu | \]

which implies that \( \mu = 1 \) is the only nonzero eigenvalue of \( c x_{x_0}(0) c \), therefore, the eigenvalues of \( A_2 A_1 \) are \( \lambda = (\mu - 1) \) with multiplicity (N-1) and \( \lambda = 0 \). The same holds true for \( A_1 A_2 \), hence, \( \text{rank}(A_1 A_2) = N-1 \). Now, since

\[ \rho(A_1 A_2) \leq \min(\rho(A_1), \rho(A_2)) \leq N \]

\[ N-1 \leq \min(\rho(A_1), \rho(A_2)) \leq N \]

we know that the lower bound is satisfied since the dimension of the null spaces of \( A_1 \) and \( A_2^T \) is at least one (it is indeed one), therefore,
the left by m columns. Then by the previous eigenvalue-eigenvector properties, one can show that \( N(N-1) \) rows of \( \mathbf{O}(\mathbf{A}, \mathbf{A}, c) \) and columns of \( \mathbf{C}(\mathbf{A}, \mathbf{A}, x_\infty(0)) \) are repeated. This proves properties (i) and (iii) since \( \text{rank}(\mathbf{O}) = \text{rank}(\mathbf{C}(\mathbf{C})) = N \) by definition of \( \mathbf{O}(\mathbf{0}) \). To prove property (ii) it is easy to show that \( \overline{\mathbf{H}} \) can be constructed so that column block\((j+1) = \mathbf{P}(1)\)column block\((j) \) and row block\((i+1) = \mathbf{P}(1)\)row block\((i) \). Therefore, \( \mathbf{P}(0) \) is the only block in \( \overline{\mathbf{H}} \) that is independent of the others, and is of full rank.

4. Markov Chain Realization Algorithm

Given \( \mathbf{P}(1) \), we can form the Hankel matrix using 2K-2 Markov parameter matrices from \( \mathbf{P}(k) = \mathbf{P}(1)^k \), i.e.,

\[
\overline{\mathbf{H}} = \begin{bmatrix}
\mathbf{P}(0) & \mathbf{P}(1) & \ldots & \mathbf{P}(K-1) \\
\mathbf{P}(1) & \mathbf{P}(2) & \ldots & \mathbf{P}(K) \\
\vdots & \vdots & \ddots & \vdots \\
\mathbf{P}(K-1) & \mathbf{P}(K) & \ldots & \mathbf{P}(2K-2)
\end{bmatrix}
\]

\[
= \overline{\mathbf{O}} \mathbf{C}
\]

(20)

where \( \mathbf{P}(k) \) is given by

\[
\mathbf{P}(k) = \begin{bmatrix}
c \mathbf{W}^k x_\infty(0) \\
c \mathbf{W}^k A_2 x_\infty(0) \\
\ldots \\
c \mathbf{W}^k A_2^{N-1} x_\infty(0) \\
c \mathbf{A}_1 \mathbf{W}^k x_\infty(0) \\
\ldots \\
c \mathbf{A}_1 \mathbf{W}^k A_2 x_\infty(0) \\
\ldots \\
c \mathbf{A}_1 \mathbf{W}^k A_2^{N-1} x_\infty(0) \\
\ldots \\
c \mathbf{A}_1 \mathbf{W}^k A_2^{N-1} A_1 x_\infty(0) \\
\ldots \\
\ldots
\end{bmatrix} = \mathbf{OC}(k)
\]

(21)

and represents a lower level set of Markov parameters. Similarly, the upper level observability and controllability matrices are

\[
\overline{\mathbf{C}} = [\mathbf{C}(0) \mathbf{W} \mathbf{C}(0) \ldots \mathbf{W}^{K-1} \mathbf{C}(0)]
\]

(22a)

\[
\overline{\mathbf{O}} = \begin{bmatrix}
\mathbf{O} \\
\mathbf{O} \mathbf{W} \\
\mathbf{O} \mathbf{W}^2 \\
\vdots \\
\mathbf{O} \mathbf{W}^{K-1}
\end{bmatrix}
\]

(22b)

The Markov chain realization algorithm consists of a pair of upper/lower level steps to determine \( [\mathbf{O}(0), \mathbf{C}(0), \mathbf{W}] \) and \( [\mathbf{A}_1, \mathbf{A}_2, \mathbf{c}, x_\infty(0)] \), respectively. Both steps can be achieved through a singular value decomposition (svd) of \( \overline{\mathbf{H}} \), i.e.,

\[
\overline{\mathbf{H}} = \mathbf{U} \Sigma \mathbf{V}^T = \overline{\mathbf{O}} \mathbf{C}
\]

\[
\overline{\mathbf{O}} = \mathbf{U} \Sigma^{12}
\]

\[
\mathbf{C} = \Sigma^{12} \mathbf{V}^T
\]

(23)

where \( \mathbf{U} \) is a \((K N \times N)\) orthonormal matrix, \( \Sigma \) is an \((N \times N)\) diagonal matrix containing the Hankel singular values, and \( \mathbf{V} \) is a \((K N \times N)\) orthonormal matrix. The parameters are obtained from

Upper Level Parameters: \([\mathbf{O}(0), \mathbf{C}(0), \mathbf{W}]\)

\[
\mathbf{O} = \text{first} (N \times N) \text{ block of } \overline{\mathbf{O}}
\]

\[
\mathbf{C}(0) = \text{first} (N \times N) \text{ block of } \overline{\mathbf{C}}
\]

\[
\mathbf{W} = [\overline{\mathbf{O}}_1 \mathbf{O}_2 \ldots \overline{\mathbf{O}}_N \overline{\mathbf{O}}_1] = \mathbf{C}_1 \mathbf{C}_2 \ldots \mathbf{C}_K^{-1}
\]

where \( \overline{\mathbf{O}}_1 \) consists of the first \( N(K-1) \) rows of \( \overline{\mathbf{O}} \) and \( \overline{\mathbf{O}}_N \) of the last \( N(K-1) \) rows of \( \overline{\mathbf{O}} \). The same definition applies to \( \overline{\mathbf{C}} \).

Lower Level Parameters: \([\mathbf{A}_1, \mathbf{A}_2, \mathbf{c}, x_\infty(0)] \)

\[
\mathbf{c} = \text{first} (1 \times N) \text{ row of } \mathbf{O}
\]

\[
x_\infty(0) = \text{first} (N \times 1) \text{ column of } \mathbf{C}(0)
\]

(24)

(25)

\[
\mathbf{A}_1 = \mathbf{C}(0) \mathbf{O}^T
\]

(26)

\[
\mathbf{A}_2 = \mathbf{C}(0) \mathbf{O}
\]

(27)

where \( \mathbf{O}^T \) is equal to \( \mathbf{O} \) shifted upwards by one row, the last one being a \((1 \times N)\) row of zeros since \( \mathbf{A}_1^N = 0 \); and \( \mathbf{C}(0) \) is \( \mathbf{C}(0) \) shifted to the left by one column.

We should point out that since \( \rho(\overline{\mathbf{H}}) = N \), \( \overline{\mathbf{O}} \overline{\mathbf{C}} \) has Hankel structure, therefore, the following Hankel norm property is satisfied

\[
|| \overline{\mathbf{H}} - \overline{\mathbf{O}} \overline{\mathbf{C}} || = \sigma_{N+1}
\]

(26)

which is of the order of machine precision.
\[ p(A_i) = p(A_2) = N-1. \] Property (v) follows from properties (ii) and (iv) since every nilpotent matrix has all its eigenvalues equal to zero. Property (vi) follows from properties (i) and (iii), which implies that \( A_1 \) and \( A_2 \) are g-inverses of one another. To prove property (vii) we need

\[
O_{A_2} = \begin{bmatrix}
    c & \\
    cA_2 & cA_2^2 & \\
     \vdots & \\
    cA_2^{N-1}
\end{bmatrix}
\]

and, if we recall, \( O_{A_1} = O \) and \( C_{A_2}(0) = C(0) \). Now, suppose \( P(0) = O_{A_2}C_{A_1}(0) = I_N \), then one can show that \( cA_2C_{A_1}(0) = 0^T = c \) since \( c \in N_{sp}(A_2^T) \), therefore, \( A_1A_2 = A_2A_1 \). Furthermore, it can be shown that \( cA_1A_2 = c \) and \( A_1A_2x_{0,0}(0) = x_{0,0}(0) \), thus, we have \( cWA_1A_2x_{0,0}(0) = cWx_{0,0}(0) \), for instance. A similar argument shows that \( W \) cannot commute with \( A_2 \) either.

We now establish the equivalence between Markov chains characterized by \( [\pi(0), P(1)]N \) and a state-space realization characterized by \( [A_1, A_2, W, c, x_{0,0}(0)]N \).

**Theorem 2:** A Markov chain defined by \( [\pi(0), P(1)]N \) is equivalent to a 2-D state-space realization \( [A_1, A_2, W, c, x_{0,0}(0)]N \) provided this one satisfies the properties of Theorem 1.

**Proof:** Since we know that \( P(k) = P(1)^k \), we can use these as Markov parameters. Let us now form the Hankel matrix from these higher-order transition matrices, i.e.,

\[
\overline{H} = \begin{bmatrix}
    P(0) \\
    P(1) \\
    \vdots \\
    P(K-1)
\end{bmatrix}
\]

Then \( O = P(0), C(0) = P(0), W = P(1), c = [1, 0, 0, \ldots, 0], x_{0,0}(0) = c^T, A_1 = I_N, \) and \( A_2 = \tilde{I}_N \) (arrows denote shifted identity matrices) can be shown to satisfy the properties of Theorem 1. Suppose there is another \( N \)th dimensional realization \( [\hat{A}_1, \hat{A}_2, \hat{W}, \hat{c}, \hat{x}_{0,0}(0)]_N \) that satisfies Theorem 1, i.e., then it can be brought to the above canonical form by a similarity transformation, i.e.,

\[
T = C(0)\hat{O}. \]

Hence, the two realizations are equivalent in the sense of \( [A_1, A_2, W, c, x_{0,0}(0)]_N = [A_1, A_2, \hat{W}, \hat{c}, \hat{x}_{0,0}(0)]_N \). If we recall from the previous section, the initial probabilities are used in a state equation such as

\[
\pi(k) = \pi(0)P(k) = \pi(0)O\hat{W}C(0)
\]

or

\[
\hat{z}(k) = \hat{z}(0)\hat{W}
\]

where \( \hat{z}(k) = \pi(k)\hat{O} \). Then if we apply the similarity transformation to (19), i.e., \( W^* = TW T^{-1} \), we get \( \hat{z}(k)T^{-1} = z(k) = \pi(k)O = \pi(k) \). This shows that the two type of systems \( [\pi(0), P(1)]N \) and \( [A_1, A_2, \hat{W}, \hat{c}, \hat{x}_{0,0}(0)]_N \) carry the same information.

**Theorem 3:** Given an \( N \)th order 2-D realization \( [A_1, A_2, W, c, x_{0,0}(0)]_N \) that satisfies the properties of Theorem 1, the following properties are equivalent:

i) \( (A_1, A_2, c) \) and \( (A_1, A_2, x_{0,0}(0)) \) are observable and controllable

ii) \( \rho(\overline{H}) = \rho(P(0)) = N \)

iii) \( \rho(O(A_1, A_2, c)) = \rho(O(A_1, c)) = N \)

\( \rho(C(A_1, A_2, x_{0,0}(0))) = \rho(C(A_2, x_{0,0}(0))) = N \)

**Proof:** One can show that the global observability and controllability matrices [11] have the following structure

\[
\begin{bmatrix}
    O \\
    OA_2 \\
    \vdots \\
    OA_2^{N-1}
\end{bmatrix}
\]

\[
\begin{bmatrix}
    O \\
    O_1 \\
    \vdots \\
    O_{N-1}
\end{bmatrix}
\]

\[
C(A_1, A_2, x_{0,0}(0)) = [C(0) A_1 C(0) \ldots A_1^{N-1} C(0)]
\]

\[
= [C(0) C_1(0) C_2(0) \ldots C_N(0)]
\]

where \( O \) denotes the observability matrix shifted downwards by \( m \) rows (padded with zero rows). Similarly, \( C(0) \) denotes \( C(0) \) shifted to
5. Conclusions

We have presented a 2-D realization theory for Markov chains which yields an exact representation. It was shown that the Markov parameters of the 2-D realization exactly match the higher-order transition probability matrices of the Markov chain. Since the model is obtained from a "balanced" type (in this case optimal) realization algorithm, one should expect the robustness properties inherent in these algorithms. Moreover, a parametrization of the 2-D realizations presented here may lead to canonical structures for certain probability matrices, i.e., birth-death chains, queueing chains, etc. Another potential application is in the identification of Markov chains from given data. These issues and other extensions are currently being investigated and will be reported elsewhere.

References


