# ORDER REDUCING APPROXIMATION OF TWO-TIME-SCALE STOCHASTIC DISCRETE LINEAR TIME-VARYING SYSTEMS

by

N. Nordstrom

Memorandum No. UCB/ERL M85/72

30 August 1985

# ORDER REDUCING APPROXIMATION OF TWO-TIME-SCALE STOCHASTIC DISTRETE LINEAR TIME-VARYING SYSTEMS

by

Niklas Nordström

Memorandum No. UCB/ERL M85/72

30 August 1985

ELECTRONICS RESEARCH LABORATORY

College of Engineering
University of California, Berkeley
94720

# Order Reducing Approximation of Two-Time-Scale Stochastic Discrete Linear Time-Varying Systems

*Niklas Nordström*

Department of Electrical Engineering and Computer Science
Electronics Research Laboratory
University of California
Berkeley CA 94720, USA

In this paper we study a two-time-scale stochastic discrete time linear time varying system. We heuristically find a reduced order approximation to its asymptotic behavior as the time scale separation tends to infinity. This approximation results in a white noise representation of the fast state vector, and a corresponding approximation error in the slow state vector. After introducing the approximate system we define concepts of continuity and rate of variation, which are needed in the discrete time analysis. We then prove, that as the time scale separation increases, the state of the reduced order system asymptotically coincides with the slow state of the original system in the mean square sense on compact time intervals. We also find the order of, the slow state approximation error covariance.

## 1. Introduction

Singularly perturbed systems have been studied extensively in the engineering litterature during the last fifteen years. Much or all of the motivation for this recent activity seems to have come from the need to deal with systems evolving at two or more time scales. Such systems are e.g. of interest in estimation and control problems.

(1) Dynamics of observers used in estimation are typically designed to be much faster than the process dynamics. In practice one hopes that the time scale separation between the dynamics of the observer and those of the rest of the system is so good, that the situation is almost the same as in the ideal case, where the observer dynamics is neglected.

(2) In control systems the actuators providing the control inputs are usually designed to give quick or instant response, resulting in a similar multiple time scale system with a corresponding order reducing approximation.

This approximation is exactly one which can be made in the analysis of a singularly peturbed system, suggesting that this class of systems has to be analysed and understood, in order to mathematically justify the heuristic approximations, which are desirable in the estimation- and control- problems such as those mentioned above.

The deterministic continuous time varying control problem has been studied e.g. by Kokotović and Yackel (1972) [1], who show, that the Riccati equations corresponding to the full and reduced systems asymptotically agree on compact time intervals, as the singular perturbation parameter $\mu$ tends to zero. A similar result concerning the state and costate trajectories has been obtained by Kokotović and Wilde (1973) [2].

Haddad (1976) [3] has studied the stochastic continuous time varying estimation problem, and shown that the Kalman filter can be decomposed into slow and fast filters, still giving estimates, which are asymptotically correct on compact time intervals. This problem is essentially the dual of the deterministic control problem, but differs in a significiant way due to the presence of a white noise input.

For discrete time systems, on the other hand, work to describe multiple time scale behavior and the corresponding order reducing approximations, has been very limited until the last five years. Given recent advances in micro processor technology, most sophisticated filters and controllers are

implemented with digital computers. Hence the analysis of singularly perturbed discrete time systems is of just as much importance as its continuous time counterpart. This is the major motivation for the work presented in this paper. Phillips (1980) [4] has introduced a notion of two-time-scale discrete time invariant systems, and exhibited a class of systems having this property. Since then Naidu and Rao (1982) [5] have shown, that the approximations resulting from decomposition of some of these systems are asymptotically correct. This approach however does not represent proper time scaling in the sense, that the parameter $\mu$ can be eliminated from the system equations by a change of time scale. Instead it focuses on the eigen values of the discrete system. Fernando and Nicholson (1983) [6] have proposed a model for properly time scaled discrete time invariant systems with fixed time step, and Kimura (1983) [7] has studied the control problem for a similar time varying system. These papers consider however only the comparison between the degenerate full system ($\mu$=0) and the reduced system, without giving any conditions under which their approximations are asymptotically accurate. Although Kokotović and others have proven, that these heuristic approaches yield asymptotically correct approximations in continuous time, this does not seem to be the case in discrete time due to the catastrophic instability, that occurs when the parameter $\mu$ tends to zero faster, than the time step does. These instability problems have clearly been observed by Blankenship (1981) [8], who let the fast dynamics determine the stepsize of the discrete system, and show that the optimal control of such systems asymptotically decomposes into fast and slow components.

In this paper we develop appropriate concepts and techniques, which extend the domain of singular perturbation analysis to properly time scaled two-time-scale discrete time varying systems. Furthermore our systems are stochastic in contrast to those in the foregoing papers. As in [8], we consider a stepsize proportional (for simplicity equal) to the parameter $\mu$. This brings up the fact, that as the fast dynamics of our sampled data system becomes faster, a shorter sampling time is needed, to capture the stability of the fast subsystem. We furthermore propose a heuristically obvious reduced order approximation, give conditions under which this approximation is valid, and carry out a careful error analysis, showing that our approximation is asymptotically correct.

The paper is organized as follws. In section 2 we present the system to be studied along with a heuristically derived reduced order approximation to the slow part of its dynamics. In section 3 we define concepts of continuity, rate of variation and boundedness, and present a few related propositions. Section 4 and section 5 are devoted to results about the fast and the slow state tranxition matrix respectively. In section 6 we show that the heuristic approximation introduced in section 2 is asymptotically correct for block triangular systems.

Finally in section 7 we indicate how this result can be extended to the general case by means of a nonsingular block triangularizing transformation.

Throughout the entire paper we will use the following notation:

$D^o = \{z \in \mathbb{C} : |z| < 1\}$ = open unit disk in $\mathbb{C}$

$\lfloor \alpha \rfloor$ = integer part of the number $\alpha$

$A^H$ = Hermitian transpose of the matrix $A$

$\sigma(A)$ = spectrum of the matrix $A$
$\delta$ is the forward difference operator

## 2. System Description

The system we consider is given by:

$$\delta\begin{bmatrix} x_\mu(i) \\ z_\mu(i) \end{bmatrix} = \begin{bmatrix} \mu A_{1,\mu}(i) & \mu A_{12,\mu}(i) \\ A_{21,\mu}(i) & A_{2,\mu}(i) \end{bmatrix} \begin{bmatrix} x_\mu(i) \\ z_\mu(i) \end{bmatrix} + \begin{bmatrix} \mu B_{1,\mu}(i) \\ B_{2,\mu}(i) \end{bmatrix} u_\mu(i) \tag{2.1}$$

where $\delta$ is the difference operator, $i \in \mathbb{N}_0 := \{0\} \bigcup \mathbb{N}$ is the time index, $\mu \in (0,\infty)$ is a small parameter, $x_\mu \in \mathbb{C}^{n_s}$ is the slow state, $z_\mu \in \mathbb{C}^{n_f}$ is the fast state, $u_\mu \in \mathbb{C}^{n_i}$ is a driving white noise and $A_{1,\mu} \in \mathbb{C}^{n_s \times n_s}$, $A_{12,\mu} \in \mathbb{C}^{n_s \times n_f}$, $A_{21,\mu} \in \mathbb{C}^{n_f \times n_s}$, $A_{2,\mu} \in \mathbb{C}^{n_f \times n_f}$, $B_{1,\mu} \in \mathbb{C}^{n_s \times n_i}$ and $B_{2,\mu} \in \mathbb{C}^{n_f \times n_i}$ are given system parameters. The vectors $x_\mu$, $z_\mu$ and $u_\mu$ are assumed to have the following statistical properties:

$$E\{\begin{bmatrix} x_\mu(0) \\ z_\mu(0) \\ u_\mu(i) \end{bmatrix}\} = 0 \qquad \forall\, i \in \mathbb{N}_0 \ \forall\, \mu>0 \tag{2.2a}$$

and

$$E\{\begin{bmatrix} x_\mu(0) \\ z_\mu(0) \\ u_\mu(j) \end{bmatrix} \begin{bmatrix} x_\mu^H(0) & z_\mu^H(0) & u_\mu^H(i) \end{bmatrix}\} = \begin{bmatrix} \Pi_{x,\mu} & \frac{1}{\sqrt{\mu}}\Pi_{xz,\mu} & 0 \\ \frac{1}{\sqrt{\mu}}\Pi_{xz,\mu}^H & \frac{1}{\mu}\Pi_{z,\mu} & 0 \\ 0 & 0 & \frac{1}{\mu}Q_\mu(j)\delta_{ij} \end{bmatrix} \tag{2.2b}$$

$\forall\, i, j \in \mathbb{N}_0 \ \forall\, \mu>0$

where the matrices $\Pi_{x,\mu} \in \mathbb{C}^{n_s \times n_s}$, $\Pi_{xz,\mu} \in \mathbb{C}^{n_s \times n_f}$, $\Pi_{z,\mu} \in \mathbb{C}^{n_f \times n_f}$ and $Q_\mu(i) \in \mathbb{C}^{n_i \times n_i}$ are known $\forall\, i, j \in \mathbb{N}_0$ and $\forall\, \mu>0$.

The interpretation of (2.1) is simple. If we identify t with i$\mu$, it is just a discrete time version of the two-time-scale continuous time system.

$$\begin{bmatrix} \dot{\tilde{x}}(t) \\ \mu\dot{\tilde{z}}(t) \end{bmatrix} = \begin{bmatrix} \tilde{A}_1(t) & \tilde{A}_{12}(t) \\ \tilde{A}_{21}(t) & \tilde{A}_2(t) \end{bmatrix} \begin{bmatrix} \tilde{x}(t) \\ \tilde{z}(t) \end{bmatrix} + \begin{bmatrix} \tilde{B}_1(t) \\ \tilde{B}_2(t) \end{bmatrix} \tilde{u}(t) \tag{2.3}$$

In fact in our proofs concerning properties of the model (2.1), we will assume the existence of at least parts of a continuous time approximation of the form (2.3). The continuous time system parameters will be distiguished from those of the discrete time system by a tilde on top of the letter as in (2.3) above. The way the descrete time system (2.1) is parameterized by its associated sample period is just $\mu$. We therefore define the time index $k_\mu(t) \in \mathbb{N}_0$ for each time $t \geq 0$ by

$$k_\mu(t) := \left\lfloor \frac{t}{\mu} \right\rfloor \tag{2.4}$$

In order to simplify notation we also define the matrices:

$$D_\mu(i) := I + \mu A_{1,\mu}(i) \quad \in \mathbb{C}^{n_s \times n_s} \tag{2.5a}$$

$$F_\mu(i) := I + A_{2,\mu}(i) \quad \in \mathbb{C}^{n_f \times n_f} \tag{2.5b}$$

The aim of this paper is to approximate the system (2.1) by a lower order system and show that when these two systems are subject to the same initial slow state, their slow states are asymptotically equal on compact time intervals as $\mu \downarrow 0$. Hence we shall be content with results valid for $\mu$ small enough, i.e. whenever there exists $\overline{\mu}>0$ such that the result under consideration holds $\forall\ \mu \in (0,\overline{\mu})$.

If we formally set $\mu=0$ in the continuous time equation (2.3), and solve for z in terms of u, we obtain the white noise approximation

$$\tilde{\zeta}(t) = -\tilde{A}_2^{-1}(t)[\tilde{A}_{21}(t)\tilde{\xi}(t) + \tilde{B}_2(t)\tilde{u}(t)] \tag{2.6}$$

and the approximate reduced order system

$$\dot{\tilde{\xi}}(t) = [\tilde{A}_1(t) - \tilde{A}_{12}(t)\tilde{A}_2^{-1}(t)\tilde{A}_{21}(t)]\tilde{\xi}(t) + [\tilde{B}_1(t) - \tilde{A}_{12}(t)\tilde{A}_2^{-1}(t)\tilde{B}_2(t)]\,\tilde{u}(t) \tag{2.7}$$

The approximations, that will be considered in this paper, are given by the discrete time versions of these equations, namely

$$\delta\xi_\mu(i) = \mu\,[A_{1,\mu}(i) - A_{12,\mu}(i)A_{2,\mu}^{-1}(i)A_{21,\mu}(i)]\,\xi_\mu(i) \tag{2.8a}$$

$$+\ \mu\,[B_{1,\mu}(i) - A_{12,\mu}(i)A_{2,\mu}^{-1}(i)B_{2,\mu}(i)]\,u_\mu(i)$$

$$\zeta_\mu(i) = -A_{2,\mu}^{-1}(i)\,[A_{21,\mu}(i)\xi_\mu(i) + B_{2,\mu}(i)u_\mu(i)] \tag{2.8b}$$

## 3. Basic Definitions

In order to be able to describe the behavior of a singularly perturbed discrete time linear system, as the small parameter $\mu$ tends to zero, we need concepts similar to continuity, boundedness and Lipschitz continuity. In this section we introduce a few such concepts, and state a few intuitively rather obvious related facts.

**Definition 3.1:** We say that the map

$$A : (0,\infty)\times\mathbb{N}_0 \to \mathbb{C}^{m\times n} : (\mu,i) \mapsto A_\mu(i) \tag{3.1}$$

is *discrete continuous (d.c) on the time set $T$*, if there exists a continuous map

$$\tilde{A} : T \to \mathbb{C}^{m\times n} : t \mapsto \tilde{A}(t) \tag{3.2}$$

called the limit of $A$ such that

$$A_\mu(k_\mu(t)) \underset{\mu\to 0}{\to} \tilde{A}(t) \tag{3.3}$$

uniformly on every compact subset of $T$. We say that the map is *discrete continuous* if it is d.c. on $[0,\infty)$.

**Definition 3.2:** For $\rho>0$ we say that the matrix valued map $A_\mu(i)$ has $O(\mu^\rho)$ *-rate of variation (r.v.) on the time set $T$* if there exists a constant $L_A<\infty$ such that

$$\sup_{t\in T}\|A_\mu(k_\mu(t) + 1) - A_\mu(k_\mu(t))\| \le L_A\mu_\rho \tag{3.4}$$

for $\mu$ small enough. We say that it has $O(\mu^\rho)$ *-rate of variation* if it has $O(\mu^\rho)$-r.v. on $[0,\infty)$. $O(\mu)$-r.v. is called *bounded rate of variation (b.r.v.)*.

**Definition 3.3:** We say that the matrix valued map $A_\mu(i)$ is *bounded on T*, if there exists a constant $M_A < \infty$ such that

$$\sup_{t \in T} \|A_\mu(k_\mu(t))\| \leq M_A \tag{3.5}$$

for $\mu$ small enough.

In general, the properties defined above are closed under finite products and inversion. Two such results are stated in the following propositions, whose simple proofs are omitted.

**Proposition 3.4:** Let $\{A_{j,\mu}(i) : j=1, \ldots, J\}$ be a finite set of matrix valued maps, which are bounded with $O(\mu^\rho)$-r.v. on $T$,

then $\prod_{j=1}^{J} A_{j,\mu}(i)$ has $O(\mu^\rho)$-r.v.

**Proposition 3.5:** Assume that:

(1) The map $A: (0,\infty) \times \mathbb{N}_0 \rightarrow \mathbb{C}^{n \times n}: (\mu,i) \mapsto A_\mu(i)$ is d.c. with limit $\tilde{A}(t)$ on $T$.

(2) $\tilde{A}(t)$ is nonsingular $\forall\, t \in T$.

Then

(i) $A_\mu^{-1}(i)$ and $\tilde{A}^{-1}(t)$ are bounded on every compact subset of $T$.

(ii) $A_\mu^{-1}(i)$ is d.c. with limit $\tilde{A}^{-1}(t)$ on every compact subset of $T$.

(iii) If in addition to (1) and (2), $A_\mu(i)$ has $O(\mu^\rho)$-r.v. on $T$, then so does $A_\mu^{-1}(i)$ on every compact subset of $T$.

## 4. Properties of the Fast State Transition Matrix

One of the crucial reasons, for which the white noise approximation to the fast state, that was suggested in the introduction, is valid, is that the fast part of the system is stable at each time, and as $\mu$ tends to zero, it becomes "infinitely stable". This stiff behavior kills off the dependence of the fast state on the past. Hence the fast state dynamics becomes almost memoryless, and since it is driven by a white input, the fast state becomes almost white. In order to find out how good this heuristic approximation is, we clearly need to characterize the stability of the fast dynamics in terms of the parameter $\mu$. This leads us to the analysis of the state transition matrix $\Phi_{F_\mu}(j,i)$ associated with the difference equation

$$z_\mu(i+1) = F_\mu(i) z_\mu(i) \tag{4.1}$$

when $F_\mu(i)$ is d.c. with given r.v. and a stable limit $\tilde{F}(t)$. The state transition matrix $\Phi_{F_\mu}(j,i)$ is defined in the usual way i.e.

$$\Phi_{F_\mu}(j,i) = F_\mu(j-1) F_\mu(j-2) \cdots F_\mu(i) \qquad j \geq i \tag{4.2}$$

Roughly speaking, the results of this section assert that in the limit as $\mu \downarrow$, the matrices $F_\mu^i(k_\mu(t))$, $\Phi_{F_\mu}(j+i,j)$, $\Phi_{F_\mu}(j+i,j) - F_\mu^i(j)$ and $\Phi_{F_\mu}(j+i,j) - F_\mu^i(j+i)$ can all be bounded on any compact time interval by exponentially decaying expressions proportional to $\lambda^i$, where $\lambda$ is arbitrarily close to, but strictly greater than the largest eigenvalue of $\tilde{F}(t)$. The precise results are stated in the following lemmas. For most of the proofs we refer to [9].

**Lemma 4.1:** Consider the matrices $F \in \mathbb{C}^{n\times n}$ and $F(i) \in \mathbb{C}^{n\times n}$, $i=0,\ldots,I$. Assume $\sigma(F) \subseteq D^\circ$ and $\|F(i) - F\| < \alpha$ $i=0,\ldots,I$.
Then there exist constants $M \in [1,\infty)$ and $\lambda \in [0,1)$ such that

(i) $\|\Phi_F(i,0)\| < M\beta^i$ $\quad i=0,\ldots,I+1$

(ii) $\|\Phi_F(i,0) - F^i\| < M\beta^i$ $\quad i=0,\ldots,I+1$

where $\beta := \lambda + M\alpha$

(iii) Moreover if $\beta \leq 1$, then

$$\|\Phi_F(i,0) - F^i\| < \frac{M^2\alpha}{1-\lambda} \leq M \tag{4.3}$$

**Lemma 4.2:** Let $T$ be compact, and let the map $F : (0,\infty)\times\mathbb{N}_0 \to \mathbb{C}^{n\times n} : (\mu,i) \mapsto F_\mu(i)$ be d.c. with limit $\tilde{F}(t)$ on $T$. Assume that $\sigma(\tilde{F}(t)) \subseteq D^\circ$ $\forall\, t \in T$.
Then there exist constants $M = M(F,T) \in [1,\infty)$ and $\lambda = \lambda(F,T) \in [0,1)$ such that

$$\sup_{t\in T} \|F_\mu^i(k_\mu(t))\| \leq M\lambda^i \qquad \forall\, i \in \mathbb{N}_0 \tag{4.4}$$

for $\mu$ small enough.

**Lemma 4.3:** Let the map $F : (0,\infty)\times\mathbb{N}_0 \to \mathbb{C}^{n\times n} : (\mu,i) \mapsto F_\mu(i)$ be d.c. with b.r.v. and limit $\tilde{F}(t)$ on $T=[T_1,T_2]$. Assume that $\sigma(\tilde{F}(t)) \subseteq D^\circ$ $\forall\, t \in T$
Then there exist constants $\hat{M} = \hat{M}(F,T) \in [1,\infty)$ and $\hat{\lambda} = \hat{\lambda}(F,T) \in [0,1)$ such that

$$\|\Phi_{F_\mu}(l,k)\| = \|F_\mu(l-1)\cdots F_\mu(k)\| \leq \hat{M}\hat{\lambda}^{l-k} \tag{4.5}$$

whenever $k_\mu(T_1) \leq k \leq l \leq k_\mu(T_2)$, and $\mu$ is small enough.

**Proof:** Since the hypotheses of lemma 4.2 are satisfied, there exist constants $M = M(F,T) \in [1,\infty)$ and $\lambda = \lambda(F,T) \in [0,1)$ such that

$$\max_{k=k_\mu(T_1)}^{k_\mu(T_2)-1} \|F_\mu^i(k)\| \leq M\lambda^i \qquad \forall\, i \in \mathbb{N}_0 \tag{4.6}$$

for $\mu$ small enough. Since $F_\mu(i)$ has b.r.v. on $T$, given $\epsilon>0$ there exists $\Delta t = \Delta t_\epsilon > 0$ such that

$$\sup_{t\in T} \max_{i=k_\mu(t)}^{k_\mu((t+\Delta t)\wedge T_2)-1} \|F_\mu(i) - F_\mu(k_\mu(t))\| < \epsilon \tag{4.7}$$

for $\mu$ small enough. Indeed $F_\mu(i)$ having b.r.v. on $T$ means that there exists a constant $L_F < \infty$ such that

$$\sup_{t\in T} \max_{i=k_\mu(t)}^{k_\mu((t+\Delta t)\wedge T_2)-1} \|F_\mu(i) - F_\mu(k_\mu(t))\| \leq L_F(k_\mu(t+\Delta t) - 1 - k_\mu(t))\mu \tag{4.8}$$

$$\leq L_F\left(\frac{t+\Delta t}{\mu} - 1 - \frac{t}{\mu} + 1\right)\mu = L_F\Delta t$$

for $\mu$ small enough, and hence any $\Delta t < \frac{\epsilon}{L_F}$ will do the trick. Pick $\Delta t$ accordingly.

It now follows from lemma 4.1 that for $\mu$ small enough

$$\sup_{t\in T} \max_{i=k_\mu(t)}^{k_\mu((t+\Delta t)\wedge T_2)} \|\Phi_{F_\mu}(i,k_\mu(t)) - F_\mu^{i-k_\mu(t)}(k_\mu(t))\| \leq M\beta^{i-k_\mu(t)} \tag{4.9}$$

where

$$\beta = \lambda + M\epsilon \tag{4.10}$$

By choosing $\epsilon$ and $\Delta t$ smaller if necessary, we can w.o.l.g. assume that $\beta<1$.

Now given $\Delta t \in (0,\infty)$, $\mu \in (0,\Delta t)$ and $k,l \in \{k_\mu(T_1), \ldots, k_\mu(T_2)\}$ such that $k\leq l$, let

$$J := \left\lfloor \frac{(l-k)\mu}{\Delta t} \right\rfloor \tag{4.11}$$

$$t_j := k\mu + j\Delta t \qquad j=0,\ldots,J \tag{4.12}$$

$$H_j := \Phi_{F_\mu}(k_\mu(t_{j+1}),k_\mu(t_j)), \qquad j=0,\ldots,J-1 \tag{4.13}$$

$$H_J := \Phi_{F_\mu}(l,k_\mu(t_J)) \tag{4.14}$$

$$K_j := F_\mu^{k_\mu(t_{j+1})-k_\mu(t_j)}(k), \qquad j=0,\ldots,J-1 \tag{4.15}$$

$$K_J := F_\mu^{l-k_\mu(t_J)}(k) \tag{4.16}$$

$$G_j := H_j-K_j, \qquad j=0,\ldots,J \tag{4.17}$$

Then from equations (4.6), (4.9) and (4.10) we see that for $\mu$ small enough

$$\|K_j\| \leq M\lambda^{k_\mu(t_{j+1})-k_\mu(t_j)} \leq M\beta^{k_\mu(t_{j+1})-k_\mu(t_j)}, \qquad j=0,\ldots,J-1 \tag{4.18}$$

$$\|K_J\| \leq M\lambda^{l-k_\mu(t_J)} \leq M\beta^{l-k_\mu(t_J)} \tag{4.19}$$

$$\|G_j\| \leq M\beta^{k_\mu(t_{j+1})-k_\mu(t_j)}, \qquad j=0,\ldots,J-1 \tag{4.20}$$

$$\|G_J\| \leq M\beta^{l-k_\mu(t_J)} \tag{4.21}$$

Thus for $\mu$ small enough

$$\|\Phi_{F_\mu}(l,k)\| = \|H_J\cdots H_0\| \tag{4.22}$$

$$= \|(K_J+G_J)\cdots(K_0+G_0)\| \leq (\|K_J\|+\|G_J\|)\cdots(\|K_0\|+\|G_J\|) \leq \hat{M}\hat{\lambda}^{l-k}$$

where

$$\hat{M} := (2M)^{J_{\Delta t}+1} \in [1,\infty) \tag{4.23}$$

$$\hat{\lambda} := \beta \in [0,1) \tag{4.24}$$

$$J_{\Delta t} := \left\lfloor \frac{T_2-T_1}{\Delta t} \right\rfloor + 1 < \infty \quad \blacksquare \tag{4.25}$$

**Lemma 4.4:** Let the map $F : (0,\infty)\times\mathbb{N}_0 \to \mathbb{C}^{n\times n} : (\mu,i) \mapsto F_\mu(i)$ be d.c. with b.r.v. and limit $\tilde{F}(t)$ on $T=[T_1,T_2]$, and assume that $\sigma(\tilde{F}(t)) \subseteq D^\circ \ \forall\, t \in T$.

Then for every fixed $m \in \mathbb{N}_0$, $\exists$ constants $\lambda \in [0,1)$ and $b<\infty$ such that for $\mu$ small enough:

(i) $\|\Phi_{F_\mu}(l,k) - F^{l-k}(k-m)\| \le \mu b(l-k)(l-k-1)\lambda^{l-k-1}$

whenever $\frac{T_1}{\mu} \le k-m \le k \le l \le \frac{T_2}{\mu}$.

(ii) $\|\Phi_{F_\mu}(l,k) - F^{l-k}(l+m)\| \le \mu b(l-k)(l-k-1)\lambda^{l-k-1}$

whenever $\frac{T_1}{\mu} \le k \le l \le l+m \le \frac{T_2}{\mu}$.

## 5. Properties of the Slow State Transition Matrix

For the analysis of the slow dynamics of the system (2.1) we need a fact about the state transition matrix $\Phi_{D_\mu}(j,i)$ associated with the difference equation

$$x_\mu(i+1) = D_\mu(i)\, x_\mu(i) = (I+\mu A_\mu(i))x_\mu(i) \tag{5.1}$$

As usual the state transition matrix $\Phi_{D_\mu}(j,i)$ is defined by:

$$\Phi_{D_\mu}(j,i) := D_\mu(j-1)D_\mu(j-2)\cdots D_\mu(i) \qquad j\ge i \tag{5.2a}$$

$$\Phi_{D_\mu}(j,i) := D_\mu^{-1}(j)D_\mu^{-1}(j+1)\cdots D_\mu^{-1}(i-1) \qquad j\le i \tag{5.2b}$$

**Lemma 5.1:** Let the map $A : (0,\infty)\times\mathbb{N}_0 \to \mathbb{C}^{n\times n} : (\mu,i) \mapsto A_\mu(i)$ be bounded on $T$.
Let $D_\mu(i)$ and $\Phi_{D_\mu}(j,i)$ be defined as in (5.2).
Then for every closed interval $[T_1,T_2] \subseteq T$, the map $\Phi_D : (0,\infty)\times\mathbb{N}_0^2 \to \mathbb{C}^{n\times n} : (\mu,i,j) \mapsto \Phi_{D_\mu}(j,i)$ is bounded and has b.r.v. (in both i and j) on $[T_1,T_2]^2$.

For the proof of this lemma and a few related facts we refer to [9].

## 6. Approximation of a Block Triangular Two-Time-Scale System

### 6.1. Full Order System

For the rest of this paper we study the two-time-scale stochastic discrete linear time varying system introduced in section 2 above under the following additional assumptions:

(A1) $A_{1,\mu}$, $B_{1,\mu}$, $B_{2,\mu}$ and $Q_\mu$ are bounded on $T$.

(A2) $A_{12,\mu}$ and $A_{21,\mu}$ are bounded and have $O(\mu^\rho)$-r.v. on $T$ for some $\rho>0$.

(A3) $A_{2,\mu}$ is d.c. with limit $\tilde{A}_2$ and has b.r.v. on $T$.

(A4) $\sigma(\tilde{F}(t)) \subseteq D^\circ \quad \forall\, t \in T$,
where $\tilde{F}(t) = I + \tilde{A}_2(t)$ is the limit of $F_\mu(i) = I + A_{2,\mu}(i)$.

(A5) $\Pi_{x,\mu}$, $\Pi_{xz,\mu}$ and $\Pi_{z,\mu}$ are $O(\mu^0)$.

where $T$ denotes a compact interval of the form $[0,T_2]$.

In this section we furthermore restrict attention to the case, when $A_{21,\mu}(i) \equiv 0$. This assumption considerably simplifies the analysis. In the next two sections we will indicate how the general case can be treated by introducing a nonsingular block triangularizing transformation. In general all of the discrete time quantities depend on the parameter $\mu$, so for brevity we drop the $\mu$-subscripts, where no ambiguities are to be expected.

## 6.2. Fast State Autocorrelation

Before we introduce the reduced order approximation, we derive the autocorrelation of the fast state, and prove two of its properties. Using the state transition matrix $\Phi_{F_\mu}(j,i)$ defined as in (4.2) above, from the system equation (2.1) we have, that for $l \geq k$

$$z(l) = \Phi_F(l,k)z(k) + \sum_{i=k}^{l-1}\Phi_F(l,i+1)B_2(i)u(i) \tag{6.1}$$

We now introduce the scaled fast state autocorrelation

$$\Lambda_\mu(i) := \mu E\{z_\mu(i)z_\mu^H(i)\} \tag{6.2}$$

Since $z(k)$ is uncorrelated with $u(i)$ $\forall\ i \geq k$, we have

$$\Lambda(l) = \Phi_F(l,k)\Lambda(k)\Phi_F^H(l,k) + \sum_{i=k}^{l-1}\Phi_F(l,i+1)G(i)\Phi_F^H(l,i+1) \qquad \forall\ l \geq k \tag{6.3}$$

where

$$G_\mu(i) := B_{2,\mu}(i)Q_\mu(i)B_{2,\mu}^H(i) \tag{6.4}$$

As an immediate consequence of (6.3), it follows that

$\Lambda_\mu(i)$ is bounded on $T$. Indeed by lemma 4.3 $\exists$ constants $M_\Phi \in [1,\infty)$ and $\lambda_\Phi \in [0,1)$ such that

$$\|\Phi_F(k,j)\| \leq M_\Phi \lambda_\Phi^{k-j} \tag{6.5}$$

whenever $0 \leq j \leq k \leq k_\mu(T_2)$.

Since moreover $G_\mu(i)$ is bounded on $T$, say by $M_G < \infty$

$$\|\Lambda(k)\| \leq M_\Phi \lambda_\Phi^k \|\Lambda(0)\|\, M_\Phi \lambda_\Phi^k + \sum_{i=0}^{k-1} M_\Phi \lambda_\Phi^{k-i-1} M_G M_\Phi \lambda_\Phi^{k-i-1} \tag{6.6}$$

$$\leq M_\Phi^2(\|\Pi_{z,\mu}\| + \frac{M_G}{1-\lambda_\Phi^2}) = O(\mu^0) \qquad k=0,\ldots,k_\mu(T_2) \quad \blacksquare$$

## 6.3. Reduced Order Approximation

When $A_{21,\mu}(i) \equiv 0$ the white noise representation (2.6) of $z_\mu(i)$ reduces to:

$$\varsigma_\mu(i) := -A_{2,\mu}^{-1}(i)B_{2,\mu}(i)u_\mu(i) \tag{6.7}$$

Note that

$$\sigma(\tilde{F}(t)) \subseteq D^\circ \quad \forall\ t \in [0,\infty) \implies \sigma(\tilde{A}_2(t)) \subseteq D^\circ - 1 \quad \forall\ t \in [0,\infty) \tag{6.8}$$

So $\tilde{A}_2(t)$ is nonsingular $\forall\ t \in [0,\infty)$.

Since moreover $A_{2,\mu}(i)$ is d.c., from proposition 3.5 we see that for $\mu$ small enough, $A_{2,\mu}(k_\mu(t))$ is nonsingular $\forall\ t \in T$. So $\varsigma_\mu(i)$ is well defined by (6.7). Substituting (6.7) in (2.1a) we obtain the resulting approximate slow state vector $\xi_\mu(i)$ as:

$$\delta\xi_\mu(i) = \mu A_{1,\mu}(i)\xi_\mu(i) + \mu\,[B_{1,\mu}(i) - A_{12,\mu}(i)A_{2,\mu}^{-1}(i)B_{2,\mu}(i)]\,u_\mu(i) \tag{6.9a}$$

$$\xi_\mu(0) = x_\mu(0) \tag{6.9b}$$

The initial condition (6.9b) just reflects the fact, that initially no injection of the fast state into the slower part of the system has occured, and hence no approximation error has been made. To evaluate the quality of this approximation we introduce the slow and the fast state errors:

$$d_\mu(i) := x_\mu(i) - \xi_\mu(i) \tag{6.10}$$

$$e_\mu(i) := z_\mu(i) - \varsigma_\mu(i) \tag{6.11}$$

and derive their covariances

$$R_{d_\mu}(k) := E\{d_\mu(k)d_\mu^H(k)\} \tag{6.12}$$

$$R_{e_\mu}(l,k) := E\{e_\mu(l)e_\mu^H(k)\} \tag{6.13}$$

### 6.4. Slow State Error Covariance

Using the statistical properties (2.2) and the equations (6.1) and (6.2) in a straight forward manner we can express the fast state error covariance as:

$$R_e(l,k) - \frac{1}{\mu}\, \Phi_F(l,k+1)[F(k)\Lambda(k) + G(k)A_2^{-H}(k)] \qquad \forall\, l>k \tag{6.14}$$

and

$$R_e(k,k) = \frac{1}{\mu}\, [\Lambda(k) + A_2^{-1}(k)G(k)A_2^{-H}(k)] \tag{6.15}$$

Next we define the fast to slow state impulse response:

$$H_\mu(j,i) := \Phi_{D_\mu}(j,i+1)A_{12,\mu}(i) \tag{6.16}$$

This is the impulse response from $z_\mu$ to $x_\mu$, from $\varsigma_\mu$ to $\xi_\mu$ or from $e_\mu$ to $d_\mu$.

Since $A_{1,\mu}(i)$ is bounded on $T$, it follows by lemma (5.1), that $\Phi_{D_\mu}(j,i)$ is bounded and has b.r.v. on $T^2$. By hypothesis $A_{12,\mu}(i)$ is bounded and has $O(\mu^\rho)$-r.v. on $T$. Hence $H_\mu(j,i)$ is bounded on $T^2$, and by proposition 3.4 it follows that $H_\mu(j,i)$ has $O(\mu^{\rho\wedge 1})$-r.v. on $T^2$ as well.

Subtracting (6.9a) from (2.1a) we obtain the slow state error dynamics, and from (6.9b) we get the initial value of the slow state error.

$$\delta d(i) = \mu A_1(i)d(i) + \mu A_{12}(i)e(i) \tag{6.17a}$$

$$d(0) = 0 \tag{6.17b}$$

It follows that the slow state error covariance is given by

$$R_d(k) = E\{d(l)d^H(k)\} = \mu^2 \sum_{i=0}^{k-1}\sum_{j=0}^{k-1} H(l,i)R_e(i,j)H^H(k,j) \tag{6.18}$$

For the rest of this section we write $k$ for $k_\mu(t)$ and let

$$J_\mu(i,j) := H_\mu(k_\mu(t),i)R_e(i,j)H_\mu^H(k_\mu(t),j) \tag{6.19}$$

Then defining the slow state error covariance at time $t$ in the obvious way, using (6.18) and (6.19) we get

$$\tilde{R}_{d_\mu}(t) := R_{d_\mu}(k_\mu(t)) = \mu^2 \sum_{i=0}^{k-1}\sum_{j=0}^{k-1} J(i,j) \tag{6.20}$$

Changing the summation indices from $(i,j)$ to $(i,m) := (i,i-j)$ (6.20) now yields

$$\tilde{R}_d(t) = \mu^2 \sum_{m=1}^{k-1} \sum_{i=0}^{k-1-m} [J^H(i+m,i) + J(i+m,i)] + \mu^2 \sum_{i=0}^{k-1} J(i,i) \tag{6.21}$$

For $m>0$ as in the sum above from (6.14) we see that

$$J(i+m,i) = H(k,i+m)\,\frac{1}{\mu}\,\Phi_F(i+m,i+1)L(i)H^H(k,i) \tag{6.22}$$

where

$$L(i) = L_\mu(i) := F_\mu(i)\Lambda_\mu(i) + G_\mu(i)A_{2,\mu}^{-H'}(i) \tag{6.23}$$

Since $F_\mu(i)$ is d.c. it is bounded on $T$. >From before we also know that $\Lambda_\mu(i)$ and $G_\mu(i)$ are bounded on $T$. Finally since $A_{2,\mu}(i)$ is d.c. with nonsingular limit $\tilde{A}_2(t)$ on $T$, it follows by proposition 3.5, that $A_{2,\mu}^{-1}(i)$ is d.c. and hence bounded on $T$. Thus $F_\mu(i)$, $\Lambda_\mu(i)$ and $A_{2,\mu}^{-H}(i)$ are all bounded on $T$. Therefore $L_\mu(i)$ given by (6.23) is bounded o $T$ as well.

We have now accumulated facts enough, to be able to prove that the reduced order approximation (6.9) is asymptotically correct. The precise result is stated in the following lemma.

**Lemma 6.1:** Consider the two-time-scale stochastic discrete linear time-varying system governed by (2.1) and (2.2).

Assume the conditions (A1) - (A5) are satisfied for some compact interval $T = [0,T_2]$, and that $A_{21,\mu}(i) \equiv 0$. Then the reduced order approximation

$$\delta\xi_\mu(i) = \mu A_{1,\mu}(i)\xi_\mu(i) + \mu\mathbf{B}_{1,\mu}(i)u_\mu(i) \tag{6.24a}$$

$$\xi_\mu(0) = x_\mu(0) \tag{6.24b}$$

where

$$\mathbf{B}_{1,\mu}(i) := B_{1,\mu}(i) - A_{12,\mu}(i)A_{2,\mu}^{-1}(i)B_{2,\mu}(i) \tag{6.25}$$

is asymptotically correct in the mean square sense. More precisely

$$\sup_{t\in T} \|\tilde{R}_{d_\mu}(t)\| = O(\mu^{\rho\wedge 1}) \tag{6.26}$$

**Proof:** Since $H_\mu(j,i)$ is bounded and has $O(\mu^{\rho\wedge 1})$-r.v. on $T^2$, $\exists$ finite constants $M_H$ and $L_H$ such that

$$\sup_{t\in T} \max_{i=0}^{k_\mu(T_2)} \|H(k_\mu(t),i)\| \le M_H \tag{6.27}$$

and

$$\sup_{t\in T} \max_{i=0}^{k_\mu(T_2)} \|H(k_\mu(t),i+1) - H(k_\mu(t),i)\| \le L_H\mu^{\rho\wedge 1} \tag{6.28}$$

for $\mu$ small enough.

Similarly since $L_\mu(i)$ is bounded on $T$, $\exists$ a constant $M_L<\infty$ such that

$$\sup_{t\in T} \|L(k_\mu(t))\| \le M_L \tag{6.29}$$

Moreover as we have seen before, by lemma 4.3 $\exists$ constants $M_\Phi \in [1,\infty)$ and $\lambda_\Phi \in [0,1)$ such that

$$\|\Phi_F(i+m,i+1)\| \le M_\Phi\lambda_\Phi^{m-1} \tag{6.30}$$

whenever $0 \leq i < i+m \leq k_\mu(T_2)$.

Using the bounds (6.27), (6.28), (6.29), (6.30) and the expression (6.22) we obtain

$$\|\mu^2 \sum_{m=1}^{k-1} \sum_{i=0}^{k-1-m} J(i+m,i) - \mu \sum_{i=0}^{k-1} H(k,i) \sum_{m=1}^{k-1} \Phi_F(i+m,i+1)L(i)H^H(k,i)\| \tag{6.31}$$

$$= \|\mu \sum_{m=1}^{k-1} \sum_{i=0}^{k-1-m} H(k,i+m)\Phi_F(i+m,i+1)L(i)H^H(k,i)$$

$$- \mu \sum_{m=1}^{k-1} \sum_{i=0}^{k-1} H(k,i)\Phi_F(i+m,i+1)L(i)H^H(k,i)\|$$

$$\leq \mu \sum_{m=1}^{k-1} \sum_{i=0}^{k-1-m} \|H(k,i+m) - H(k,i)\| \, \|\Phi_F(i+m,i+1)\| \, \|L(i)\| \, \|H^H(k,i)\|$$

$$+ \mu \sum_{m=1}^{k-1} \sum_{i=k-m}^{k-1} \|H(k,i)\| \, \|\Phi_F(i+m,i+1)\| \, \|L(i)\| \, \|H^H(k,i)\|$$

$$\leq \mu \sum_{m=1}^{k-1} ( \sum_{i=0}^{k-1-m} mL_H \mu^{\rho\wedge 1} M_\Phi \lambda_\Phi^{m-1} M_L M_H + \sum_{i=k-m}^{k-1} M_H M_\Phi \lambda_\Phi^{m-1} M_L M_H)$$

$$= \mu \sum_{m=1}^{k-1} [(k_\mu(t) - m) m L_H \mu^{\rho\wedge 1} M_\Phi \lambda_\Phi^{m-1} M_L M_H + m M_H^2 M_\Phi \lambda_\Phi^{m-1} M_L]$$

$$\leq \mu M_\Phi M_L M_H \sum_{m=1}^{\infty} [(\frac{t}{\mu} - m) L_H \mu^{\rho\wedge 1} + M_H] \, m \lambda_\Phi^{m-1}$$

$$\leq M_\Phi M_L M_H (T_2 L_H \mu^{\rho\wedge 1} + M_H \mu) \sum_{m=1}^{\infty} m \lambda_\Phi^{m-1} = O(\mu^{\rho\wedge 1})$$

since the series converges.

>From lemma 4.4 we know that $\exists$ constants $b<\infty$ and $\lambda_\Delta \in [0,1)$ such that

$$\|\Phi_F(i+m,i+1) - F^{m-1}(i)\| \leq \mu b(m-1)(m-2)\lambda_\Delta^{m-2} \tag{6.32}$$

whenever $0 \leq i < i+m \leq k_\mu(T_2)$.

Moreover since $\sigma(\tilde{F}(t)) \subseteq D^\circ \ \forall \, t \in T$, for $\mu$ small enough $A_{2,\mu}(i)$ is nonsingular, $i=0, \ldots, k_\mu(T_2)$. Hence

$$\sum_{m=1}^{k-1} F^{m-1}(i) = [F^{k-1}(i) - I] A_2^{-1}(i) \tag{6.33}$$

whenever $i=0, \ldots, k_\mu(T_2)$.

Using these two facts we find that

$$\|\mu \sum_{i=0}^{k-1} H(k,i) \sum_{m=1}^{k-1} \Phi_F(i+m,i+1)L(i)H^H(k,i) \tag{6.34}$$

$$- \mu \sum_{i=0}^{k-1} H(k,i)[F^{k-1}(i) - I]A_2^{-1}(i)L(i)H^H(k,i)\|$$

$$= \mu \| \sum_{i=0}^{k-1} H(k,i) \sum_{m=1}^{k-1} [\Phi_F(i+m,i+1) - F^{m-1}(i)] L(i)H^H(k,i)\|$$

$$\leq \mu \sum_{i=0}^{k-1} M_H \sum_{m=1}^{k-1} \mu b(m-1)(m-2)\lambda_\Delta^{m-2} M_L M_H \leq T_2 M_H^2 b M_L \sum_{m=1}^{\infty} m^2 \lambda_\Delta^{m-1} \mu \qquad \forall \, t \in T$$

Together with (6.31) this implies that

$$\mu^2 \sum_{m=1}^{k-1} \sum_{i=0}^{k-1-m} J(i+m,i) = \mu \sum_{i=0}^{k-1} H(k,i)[F^{k-1}(i) - I]A_2^{-1}(i)L(i)H^H(k,i) + O(\mu^{\rho\wedge 1}) \tag{6.35}$$

on $T$.

Using (6.19), (6.21) and (6.35), the slow state error covariance can now be written as

$$\tilde{R}_d(t) = \mu \sum_{i=0}^{k-1} H(k,i)R(i)H^H(k,i) + O(\mu^{\rho\wedge 1}) \tag{6.36}$$

where according to (6.15)

$$R_\mu(i) := L_\mu^H(i)A_{2,\mu}^{-H}(i)[F_\mu^{k-1}(i) - I]^H + [F_\mu^{k-1}(i) - I]A_{2,\mu}^{-1}(i)\acute{L}_\mu(i) + \Lambda_\mu(i) + A_{2,\mu}^{-1}(i)G_\mu(i)A_{2,\mu}^{-H}(i) \tag{6.37}$$

Thus

$$R_\mu(i) = R_{1,\mu}(i) + L_\mu^H(i)A_{2,\mu}^{-H}(i)[F_\mu^{k-1}(i)]^H + F_\mu^{k-1}(i)A_{2,\mu}^{-1}(i)L_\mu(i) \tag{6.38}$$

where

$$R_{1,\mu}(i) := -L_\mu^H(i)A_{2,\mu}^{-H}(i) - A_{2,\mu}^{-1}L_\mu(i) + \Lambda_\mu(i) + A_{2,\mu}^{-1}(i)G_\mu(i)A_{2,\mu}^{-H}(i) \tag{6.39}$$

Using (6.23), this can be written as

$$R_1(i) = A_2^{-1}(i)\big(\Lambda(i) - F(i)\Lambda(i)F^H(i) - G(i)\big)A_2^{-H}(i) \tag{6.40}$$

so by (6.3) (with $k$ replaced by $i$ and $l$=1)

$$R_1(i) = -A_2^{-1}(i)\delta\Lambda(i)A_2^{-H}(i) \tag{6.41}$$

By proposition 3.5, $A_{2,\mu}^{-1}(i)$ is bounded, say by $M_A<\infty$, and has b.r.v. on $T$. We also know that $\Lambda_\mu(i)$ is bounded on $T$, say by $M_\Lambda<\infty$. Moreover from lemma 4.2 we know that $\exists$ constants $M_F \in [1,\infty)$ and $\lambda_F \in [0,1)$ such that

$$\|F_\mu^{k-1}(i)\| \le M_F\lambda^{k-1} \qquad i=0,\ldots,k_\mu(T_2),\ \forall\, t \in T \tag{6.42}$$

Finally since $H(j,i)$ is bounded and has $O(\mu^{\rho\wedge 1})$-r.v. on $T^2$, it follows by proposition 3.4, that $\exists$ a finite constant $L_{HA}$ such that

$$\|H(j,i+1)A_2^{-1}(i+1) - H(j,i)A_2^{-1}(i)\| \le L_{HA}\mu^{\rho\wedge 1} \qquad \forall\, i,j \in \{0,\ldots,k_\mu(T_2)\} \tag{6.43}$$

Using these bounds together with (6.36), (6.38) and (6.41), we estimate the slow error covariance as follows.

$$\|\tilde{R}_d(t)\| = \|\mu \sum_{i=0}^{k-1} H(k,i)\big(-A_2^{-1}(i)\delta\Lambda(i)A_2^{-H}(i) + L^H(i)A_2^{-H}(i)[F^{k-1}(i)]^H + F^{k-1}(i)A_2^{-1}(i)L(i)\big)H^H(k,i)\| + O(\mu^{\rho\wedge 1}) \tag{6.44}$$

$$\le \mu\|\sum_{i=0}^{k-1} H(k,i)A_2^{-1}(i)[\Lambda(i) - \Lambda(i+1)]A_2^{-H}(i)H^H(k,i)\|$$

$$+ \mu\|\sum_{i=0}^{k-1} H(k,i)\big(L^H(i)A_2^{-H}(i)[F^{k-1}(i)]^H + F^{k-1}(i)A_2^{-1}(i)L(i)\big)H^H(k,i)\| + O(\mu^{\rho\wedge 1})$$

$$\le \mu\|\sum_{i=0}^{k-2} H(k,i+1)A_2^{-1}(i+1)\Lambda(i+1)A_2^{-H}(i+1)H^H(k,i+1)$$

$$- \sum_{i=0}^{k-2} H(k,i)A_2^{-1}(i)\Lambda(i+1)A_2^{-H}(i)H^H(k,i)\|$$

$$+ \mu\|H(k,0)A_2^{-1}(0)\Lambda(0)A_2^{-H}(0)H^H(k,0) - H(k,k-1)A_2^{-1}(k-1)\Lambda(k)A_2^{-H}(k-1)H^H(k,k-1)\|$$

$$+ \mu \sum_{i=0}^{k-1} M_H(M_L M_A M_F \lambda_F^{k-1} + M_F\lambda_F^{k-1}M_A M_L)M_H + O(\mu^{\rho\wedge 1})$$

$$\leq \mu \sum_{i=0}^{k-2}\Big( \|H(k,i+1)A_2^{-1}(i) - H(k,i)A_2^{-1}(i)\| \, \|\Lambda(i+1)A_2^{-H}(i+1)H^H(k,i+1)\|$$

$$+ \|H(k,i)A_2^{-1}(i)\Lambda(i+1)\| \, \|A_2^{-H}(i+1)H^H(k,i+1) - A_2^{-H}(i)H^H(k,i)\|\Big)$$

$$+ 2\mu M_H M_A M_\Lambda M_A M_H + 2\mu k M_H^2 M_L M_A M_F \lambda_F^{k-1} + O(\mu^{\rho\wedge 1})$$

$$\leq \mu \sum_{i=0}^{k-2}(L_{HA}\mu^{\rho\wedge 1} M_\Lambda M_A M_H + M_H M_A M_\Lambda L_{HA}\mu^{\rho\wedge 1}) + O(\mu)'$$

$$+ 2\mu M_H^2 M_L M_A M_F \sup_{j\in N} j\lambda^j + O(\mu^{\rho\wedge 1}) = O(\mu^{\rho\wedge 1})$$

on $T$. ∎

## 7. Approximation of the General Two-Time-Scale System

The results of the previous section for block triangular systems can be extended to the general case by means of the nonsingular transformation:

$$\begin{bmatrix} v_\mu(i) \\ w_\mu(i) \end{bmatrix} = \begin{bmatrix} I & 0 \\ N_\mu(i) & I \end{bmatrix} \begin{bmatrix} x_\mu(i) \\ z_\mu(i) \end{bmatrix} \tag{7.1}$$

where $N_\mu(i)$ is defined by the difference equation

$$\delta N_\mu(i) = A_{2,\mu}(i)N_\mu(i) - A_{21,\mu}(i) - \mu N_\mu(i+1)[A_{1,\mu}(i) - A_{12,\mu}(i)N_\mu(i)] \tag{7.2a}$$

$$N_\mu(0) = A_{2,\mu}^{-1}(0)A_{21,\mu}(0) \tag{7.2b}$$

It can be shown that this transformation is well defined for $\mu$ small enough, and that the transformed system is block triangular. It can furthermore be shown, that the approximation of the original system (2.1) induced by the straight forward reduced order approximation of the transformed block triangular system is asymptotically equivalent to the reduced order approximation (2.8) of the original system itself. I.e. the covariance of the difference between the slow state trajectories of these two systems tends to zero uniformly on compact time intervals as $\mu \downarrow 0$. By lemma 6.1 of the previous section the same is true for the difference between the slow state trajectories of the transformed block triangular system and its reduced order approximation. Hence lemma 6.1 extends to the following theorem:

**Theorem 7.1:** Consider the two-time-scale stochastic discrete linear time-varying system governed by (2.1) and (2.2).

Assume that the conditions (A1) - (A5) are satisfied for some compact time interval $T = [0, T_2]$. Then the reduced order approximation

$$\delta\xi_\mu(i) = \mu \mathbf{A}_{1,\mu}(i)\xi_\mu(i) + \mu \mathbf{B}_{1,\mu}(i)u_\mu(i) \tag{7.3a}$$

$$\xi_\mu(0) = x_\mu(0) \tag{7.3b}$$

where

$$\mathbf{A}_{1,\mu}(i) := A_{1,\mu}(i) - A_{12,\mu}(i)A_{2,\mu}^{-1}(i)A_{21,\mu}(i) \tag{7.4a}$$

$$\mathbf{B}_{1,\mu}(i) := B_{1,\mu}(i) - A_{12,\mu}(i)A_{2,\mu}^{-1}(i)B_{2,\mu}(i) \tag{7.4b}$$

is asymptotically correct in the mean square sense.

For the details of the block triangularizing transformation and the proof of this theorem we refer to [9].


## 8. Acknowledgements

I would like to thank my advisor Professor Shankar Sastry for many helpful suggestions.

———

The research was supported in part by:

JSEP-AFOSR-F49620-79-C0178
Ericsson Radio Systems, Sweden
Regents Fellowship, University of California


## 9. Bibliography


[1] P. V. Kokotović and R. A. Yackel, "Singular perturbation of linear regulators: Basic theorems," *IEEE Trans. Automat. Contr.*, vol. AC-17pp. 29-37, Feb. 1972.

[2] R. R. Wilde and P. V. Kokotović, "Optimal open- and closed-loop control of singularly perturbed linear system," *IEEE Trans. Automat. Contr.*, vol. AC-18, pp. 616-626, Dec. 1973.

[3] A. H. Haddad, "Linear filtering of singularly perturbed systems," *IEEE Trans. Automat. Contr.*, vol. AC-21, pp. 515-519, Aug. 1976.

[4] R. G. Phillips, "Reduce order modelling and control of two-time-scale discrete systems," *Int. J. Control*, vol. 31, no 4, pp. 765-780, April 1980.

[5] D. S. Naidu and A. K. Rao, "Singular perturbation methods for a class of initial- and boundary-value problems in discrete systems," *Int. J. Control*, vol. 36, no. 1, pp. 77-94, July 1982.

[6] K. V. Fernando and H. Nicholson, "Singular perturbational approximations for discrete-time balanced systems," *IEEE Trans. Automat. Contr.*, vol. AC-28, pp. 240-242, Feb. 1983.

[7] M. Kimura, "On the matrix Riccati equation for a singularly perturbed linear discrete control system," *Int. J. Control*, vol. 38, no. 5, pp. 959-975, Nov. 1983.

[8] G. Blankenship, "Singularly perturbed difference equations in optimal control problems," *IEEE Trans. Automat. Contr.*, vol. AC-26, pp. 911-917, Aug. 1981.

[9] N. Nordström, "Order Reducing Approximation of Two-Time-Scale Stochastic Discrete Linear Time-Varying Systems," *Master's thesis, University of California, Berkeley* May 1985.