Information Theoretic Methods for Stochastic Model Reduction Based on State Projection

Hui ZHANG, You Xian SUN

***Abstract*—Based on state projection method with two-step operations, this paper deals with the model reduction problem by analyzing the information descriptions of system states. Our basic idea in obtaining the reduced-order models is to minimize the information loss or the conditional information loss caused by truncation by eliminating the state variables with the least contribution to system information. Before truncation, an entropy preserving transformation of the original state is required. The derived minimum information loss (MIL) and minimum conditional information loss (MCIL) methods are proved to be efficient for approximating stable and unstable systems, respectively, and connected with the balanced truncation methods firmly. Illustrative examples are given.**

I. Introduction

MODEL reduction is a long-term studied problem in the field of control theory. For linear time invariant (LTI) systems, many approximation approaches are available in the literature, such as Lyapunov balanced truncation [10], LQG balanced truncation [7] and stochastic balanced truncation [3]. Most of the existing approximation methods fall into the category of state space projection with two-step operations [5]. The first step is a state transformation into a state space realization in which the state variables can be ranked according to some measure of importance. The second step is truncation of the least important state variables.

Information theoretic methods for control systems, including that dealing with the model reduction problem, are attracting more and more attention [9, 11]. For example, the Kullback-Leibler information (KLI) was adopted [9] as a measure of statistic distance between full- and reduced-order models, and was made to be minimum in obtaining the approximating models. Stochastic balanced truncation method also possesses information theoretic interpretation [3]. Based on state projection, the present paper studies the model reduction problem of stochastic LTI systems by analyzing the information and conditional information descriptions of system states. Different from [9], our basic idea is to minimize the (conditional) information loss caused by truncation by eliminating the state variables with the least contribution to system information.

H. Zhang (corresponding author to provide phone: 86-571-87951879; fax: 86-571-87951879; e-mail: zhanghui_iipc@ zju.edu.cn) and Y. X. Sun (e-mail: yxsun@iipc.zju.edu.cn) are with the National Laboratory of Industrial Control Technology, Institute of Modern Control Engineering, Department of Control Science & Engineering, Zhejiang University, Hang Zhou, 310027, China.

II. Statement of the Problem

Consider a full-order LTI stochastic system

$$\begin{aligned} \delta \boldsymbol{x}(t) &= \boldsymbol{A}\boldsymbol{x}(t) + \boldsymbol{B}\boldsymbol{w}(t) \\ \boldsymbol{y}(t) &= \boldsymbol{C}\boldsymbol{x}(t) + \boldsymbol{v}(t) \end{aligned} \quad (1)$$

where $\boldsymbol{x}(t) \in \mathbf{R}^n$, $\boldsymbol{w}(t) \in \mathbf{R}^m$, $\boldsymbol{y}(t), \boldsymbol{v}(t) \in \mathbf{R}^p$, $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C}$ are constant matrices with appropriate dimensions. δ denotes shift or derivative operator regarding we are dealing with discrete or continuous time systems, respectively. $\boldsymbol{w}(t)$ and $\boldsymbol{v}(t)$ are mutually independent zero-mean white Gaussian random vectors with covariance matrices $\boldsymbol{Q}$ and $\boldsymbol{R}$, respectively, and uncorrelated with $\boldsymbol{x}(0)$. To approximate system (1), we wish to find a reduced-order model

$$\begin{aligned} \delta \boldsymbol{x}_{\mathrm{r}}(t) &= \boldsymbol{A}_{\mathrm{r}}\boldsymbol{x}_{\mathrm{r}}(t) + \boldsymbol{B}_{\mathrm{r}}\boldsymbol{w}(t) \\ \boldsymbol{y}_{\mathrm{r}}(t) &= \boldsymbol{C}_{\mathrm{r}}\boldsymbol{x}_{\mathrm{r}}(t) + \boldsymbol{v}(t) \end{aligned} \quad (2)$$

where $\boldsymbol{x}_{\mathrm{r}}(t) \in \mathbf{R}^l$, $l < n$, $\boldsymbol{y}_r(t) \in \mathbf{R}^p$, $\boldsymbol{A}_{\mathrm{r}}, \boldsymbol{B}_{\mathrm{r}}, \boldsymbol{C}_{\mathrm{r}}$ are constant matrices. Denote realizations (1) and (2) as $\{\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}\}$ and $\{\boldsymbol{A}_{\mathrm{r}}, \boldsymbol{B}_{\mathrm{r}}, \boldsymbol{C}_{\mathrm{r}}\}$, respectively.

In this paper, we will discuss this problem based on the methodology of state space projection. Define the projection $\boldsymbol{\Psi} = \boldsymbol{V}\boldsymbol{\Lambda}$, where $\boldsymbol{V}, \boldsymbol{\Lambda}^{\mathrm{T}} \in \mathbf{R}^{n \times l}$ and $\boldsymbol{\Lambda}\boldsymbol{V} = \boldsymbol{I}_l$. $\boldsymbol{\Psi}$ is the projection onto the range space of $\boldsymbol{V}$ along the null space of $\boldsymbol{\Lambda}$. Now the idea of state projection is that at each time, the state $\boldsymbol{x}(t)$ is approximated by $\boldsymbol{\Psi}\boldsymbol{x}(t) = \boldsymbol{V}\boldsymbol{\Lambda}\boldsymbol{x}(t)$ so that

$$\{\boldsymbol{A}_{\mathrm{r}}, \boldsymbol{B}_{\mathrm{r}}, \boldsymbol{C}_{\mathrm{r}}\} = \{\boldsymbol{\Lambda}\boldsymbol{A}\boldsymbol{V}, \boldsymbol{\Lambda}\boldsymbol{B}, \boldsymbol{C}\boldsymbol{V}\}, \quad (3)$$

where we defined

$$\boldsymbol{x}_{\mathrm{r}}(t) = \boldsymbol{\Lambda}\boldsymbol{x}(t). \quad (4)$$

0-7803-9098-9/05/$25.00 ©2005 AACC

If we change the state-space coordinate basis of the system (1) by choosing

$$\bar{\boldsymbol{x}} = \boldsymbol{T}\boldsymbol{x} \tag{5}$$

with the matrix $\boldsymbol{T} \in \mathbf{R}^{n\times n}$ nonsingular, then the system (1) is equivalent to the system $\{\bar{\boldsymbol{A}}, \bar{\boldsymbol{B}}, \bar{\boldsymbol{C}}\} = \{\boldsymbol{T}\boldsymbol{A}\boldsymbol{T}^{-1}, \boldsymbol{T}\boldsymbol{B}, \boldsymbol{C}\boldsymbol{T}^{-1}\}$.

Since $\boldsymbol{\Lambda}\boldsymbol{V} = \boldsymbol{I}_l$, it is always possible to find a similarity transformation $\boldsymbol{T}$ so that the projecting matrices $\boldsymbol{\Lambda}$ and $\boldsymbol{V}$ consist of the first l rows of $\boldsymbol{T}$ and the first l columns of $\boldsymbol{T}^{-1}$, respectively. Hence $\{\boldsymbol{A}_{\mathrm{r}}, \boldsymbol{B}_{\mathrm{r}}, \boldsymbol{C}_{\mathrm{r}}\}$ is constructed by first transforming the model (1) by choosing a particular basis transformation $\boldsymbol{T}$, giving the realization $\{\bar{\boldsymbol{A}}, \bar{\boldsymbol{B}}, \bar{\boldsymbol{C}}\}$, and then truncating the transformed state-space model by restriction to the first l rows and/or columns. Based on this two-step procedure, we will discuss the model reduction problem from the viewpoint of information theory.

III. Minimizing Information Loss

In this section, we assume that the matrix $\boldsymbol{A}$ is Hurwitz and [$\boldsymbol{A}$, $\boldsymbol{B}$] is controllable.

A. State Information

System dynamics is defined by model structure and parameters. However, the "information" of the dynamics is contained in system state vectors. The dynamic information of the "full-order description" (1) is contained in $\boldsymbol{x}(t)$, while the dynamic information of the "reduced-order description" (2) is contained in $\boldsymbol{x}_{\mathrm{r}}(t)$. In information theory, the amount of information of a stochastic variable is measured by the entropy function [6]. For the full-order Gaussian system (1), the transient state entropy is defined by

$$H(\boldsymbol{x}(t)) = \frac{n}{2}\ln(2\pi e) + \frac{1}{2}\ln \det\boldsymbol{\Pi}(t)$$

where $\boldsymbol{\Pi}(t)$ is the covariance matrix of $\boldsymbol{x}(t)$. For the reduced-order model (2), the state entropy is

$$H(\boldsymbol{x}_{\mathrm{r}}(t)) = \frac{l}{2}\ln(2\pi e) + \frac{1}{2}\ln \det\boldsymbol{\Pi}_{\mathrm{r}}(t)$$

where $\boldsymbol{\Pi}_{\mathrm{r}}(t)$ is the covariance matrix of $\boldsymbol{x}_{\mathrm{r}}(t)$.

In this paper we will focus on the steady state information. The steady state covariance of the system (1) defined by

$$\boldsymbol{\Pi} = \lim_{t\to\infty} E\{\boldsymbol{x}(t)\boldsymbol{x}^{\mathrm{T}}(t)\}$$

is the unique positive definite solution to

$$\boldsymbol{A}\boldsymbol{\Pi} + \boldsymbol{\Pi}\boldsymbol{A}^{\mathrm{T}} + \boldsymbol{B}\boldsymbol{Q}\boldsymbol{B}^{\mathrm{T}} = 0 \tag{6}$$

when (1) is a continuous time system, or, the unique positive definite solution to

$$\boldsymbol{A}\boldsymbol{\Pi}\boldsymbol{A}^{\mathrm{T}} + \boldsymbol{B}\boldsymbol{Q}\boldsymbol{B}^{\mathrm{T}} = \boldsymbol{\Pi} \tag{7}$$

when (1) is a discrete time system.

Suppose the steady state covariance of system (2) is

$$\boldsymbol{\Pi}_{\mathrm{r}} = \lim_{t\to\infty} E\{\boldsymbol{x}_{\mathrm{r}}(t)\boldsymbol{x}_{\mathrm{r}}^{\mathrm{T}}(t)\}. \tag{8}$$

It can be concluded from (4) and (8) that when the reduced -order model is constructed by (3), then $\boldsymbol{\Pi}_{\mathrm{r}} = \boldsymbol{\Lambda}\boldsymbol{\Pi}\boldsymbol{\Lambda}^{\mathrm{T}}$ is the covariance matrix of the reduced state, and is the unique positive definite solution of the following l-order Lyapunov equation, for example, when system is discrete time.

$$\boldsymbol{A}_{\mathrm{r}}\boldsymbol{\Pi}_{\mathrm{r}}\boldsymbol{A}_{\mathrm{r}}^{\mathrm{T}} + \boldsymbol{B}_{\mathrm{r}}\boldsymbol{Q}\boldsymbol{B}_{\mathrm{r}}^{\mathrm{T}} = \boldsymbol{\Pi}_{\mathrm{r}}. \tag{9}$$

Then, we can define the steady state information of systems (1) and (2) as

$$H(\boldsymbol{x}) = \frac{n}{2}\ln(2\pi e) + \frac{1}{2}\ln \det\boldsymbol{\Pi}, \tag{10}$$

$$H(\boldsymbol{x}_{\mathrm{r}}) = \frac{l}{2}\ln(2\pi e) + \frac{1}{2}\ln \det\boldsymbol{\Lambda}\boldsymbol{\Pi}\boldsymbol{\Lambda}^{\mathrm{T}}, \tag{11}$$

respectively.

B. Reduced-order model with minimum information loss

From the viewpoint of information theory, the approximating performance of reduced-order model is determined by the amount of information of the full-order state $\boldsymbol{x}$ retained in the reduced-order state $\boldsymbol{x}_{\mathrm{r}}$. If $\boldsymbol{x}_{\mathrm{r}}$ retains more information of $\boldsymbol{x}$, then better performance can be expected. Based on this understanding, we propose a criterion for getting the reduced-order model: To make the system $\{\boldsymbol{A}_{\mathrm{r}}, \boldsymbol{B}_{\mathrm{r}}, \boldsymbol{C}_{\mathrm{r}}\}$ retain as much steady state information of the original system $\{\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}\}$ as possible, or to minimize the steady state information loss defined by

$$IL(\boldsymbol{x}, \boldsymbol{x}_{\mathrm{r}}) := H(\boldsymbol{x}) - H(\boldsymbol{x}_{\mathrm{r}}). \tag{12}$$

The minimum information loss principle had been applied successfully in the field of pattern recognition [8].

Under transformation (5), the steady state information becomes $H(\bar{\boldsymbol{x}}) = H(\boldsymbol{T}\boldsymbol{x})$. Before truncation, a natural requirement is that the transformation $\boldsymbol{T}$ does not change the information of the original state, i.e.

$$H(\bar{\boldsymbol{x}}) = H(\boldsymbol{x}). \tag{13}$$

Such kind of transformation can be referred to as the entropy preserving transformation. When system state vectors take values in a discrete space, equation (13) is true if $\boldsymbol{T}$ is nonsingular [6]. However, when system state vectors take values in a continuous space, we have [6]

$$H(\bar{\boldsymbol{x}}) = H(\boldsymbol{x}) + \ln|\det\boldsymbol{T}|. \tag{14}$$

Hence, the nonsingular condition is not sufficient for $\boldsymbol{T}$ to be an entropy preserving transformation when state space is continuous. There must be some further restriction on $\boldsymbol{T}$.

If the state transformation $\boldsymbol{T}$ is restricted to be unitary, then $\det\boldsymbol{T} = 1$. Such a transformation $\boldsymbol{T}$ is entropy preserving even when state space is continuous.

From equations (10) ~ (13) we have

$$\begin{aligned} IL(\boldsymbol{x}, \boldsymbol{x}_\mathrm{r}) &= H(\bar{\boldsymbol{x}}) - H(\boldsymbol{x}_\mathrm{r}) \\ &= \frac{n-l}{2}\ln(2\pi e) + \frac{1}{2}(\ln\det\boldsymbol{\Pi} - \ln\det\boldsymbol{\Pi}_\mathrm{r}). \end{aligned} \tag{15}$$

Let $\sigma_1, ..., \sigma_l$ be the eigenvalues of $\boldsymbol{\Pi}_\mathrm{r} = \boldsymbol{\Lambda\Pi\Lambda}^\mathrm{T}$, then $\det(\boldsymbol{\Pi}_\mathrm{r}) = \sigma_1 \cdots \sigma_l$. The steady state entropy of reduced -order model is $H(\boldsymbol{x}_\mathrm{r}) = \frac{l}{2}\ln(2\pi e) + \frac{1}{2}\sum_{i=1}^{l}\ln\sigma_i$. The first item in the right of (15) is independent of the choice of elements of $\boldsymbol{\Lambda}$. $\det\boldsymbol{\Pi}$ is fixed. So, minimizing $IL(\boldsymbol{x}, \boldsymbol{x}_\mathrm{r})$ is equivalent to maximizing $\sum_{i=1}^{l}\ln\sigma_i$. Because the transformation $\boldsymbol{T}$ is entropy preserving, it can be seen that $\sum_{i=1}^{l}\ln\sigma_i$ is maximized if and only if $\sigma_1, ..., \sigma_l$ are the l largest eigenvalues of the covariance matrix $\boldsymbol{\Pi}$.

To fulfill this, let $\boldsymbol{T}$ consist of all the ortho-normal eigenvectors of $\boldsymbol{\Pi}$, $\boldsymbol{\Lambda}$ consist of the first l rows of $\boldsymbol{T}$ and $\boldsymbol{V}$ consist of the first l columns of $\boldsymbol{T}^{-1}$, so that

$$\boldsymbol{\Lambda} = [\boldsymbol{\eta}_1\ \boldsymbol{\eta}_2 \cdots \boldsymbol{\eta}_l]^\mathrm{T}, \tag{16}$$

where $\boldsymbol{\eta}_1, \boldsymbol{\eta}_2, \ldots, \boldsymbol{\eta}_l$ are the l ortho-normal eigenvectors corresponding to the l largest eigenvalues of the covariance matrix $\boldsymbol{\Pi}$. Since $\boldsymbol{TT}^\mathrm{T} = \boldsymbol{I}$, then $\boldsymbol{T}^{-1} = \boldsymbol{T}^\mathrm{T}$, and $\boldsymbol{V} = \boldsymbol{\Lambda}^\mathrm{T}$. From (3) the reduced-order model is given by

$$\{\boldsymbol{A}_\mathrm{r}, \boldsymbol{B}_\mathrm{r}, \boldsymbol{C}_\mathrm{r}\} = \{\boldsymbol{\Lambda A \Lambda}^\mathrm{T}, \boldsymbol{\Lambda B}, \boldsymbol{C\Lambda}^\mathrm{T}\}. \tag{17}$$

Under the transformation $\boldsymbol{T}$ as constructed above, the steady state covariance matrix becomes

$$\bar{\boldsymbol{\Pi}} = \boldsymbol{T\Pi T}^\mathrm{T} = \mathrm{diag}[\sigma_1,\ \sigma_2,\ \ldots,\ \sigma_n],$$

where $\sigma_1 \geq \cdots \geq \sigma_n > 0$ are eigenvalues of $\boldsymbol{\Pi}$. Then

$$H(\bar{\boldsymbol{x}}) = H(\boldsymbol{x}) = \frac{n}{2}\ln(2\pi e) + \frac{1}{2}\sum_{i=1}^{n}\ln\sigma_i, \tag{18}$$

and the minimum information loss is

$$IL(\boldsymbol{x}, \boldsymbol{x}_\mathrm{r}) = \frac{n-l}{2}\ln(2\pi e) + \frac{1}{2}\sum_{i=l+1}^{n}\ln\sigma_i. \tag{19}$$

We come to the conclusion here that the minimum information loss (MIL) reduced-order model based on projection is constructed by equations (16) and (17), where the steady state covariance $\boldsymbol{\Pi}$ is the unique positive definite solution to the Lyapunov equation (6) or (7) for continuous and discrete time systems, respectively.

IV. Minimizing Conditional Information Loss

A limitation of the application of the MIL method is that the original model must be stable so that the steady state covariance $\boldsymbol{\Pi}$ exists. In this section, we will show that an unstable system can be dealt with by analyzing the conditional information of system state. For simplicity, we assume that the system discussed in this section is discrete time. The results for continuous time systems can be gotten in a similar way. In order to distinguish it from section 3, here we denote the state transformation as $\boldsymbol{S}$, and the projection as $\boldsymbol{\Omega} = \boldsymbol{K\Gamma}$, where $\boldsymbol{\Gamma}, \boldsymbol{K}^\mathrm{T} \in \mathbf{R}^{l\times n}$, $\boldsymbol{\Gamma K} = \boldsymbol{I}_l$.

A. Conditional Information

Suppose we have a certain observation sequence of the true system, $\boldsymbol{Y}^t = \{\boldsymbol{y}(1), \boldsymbol{y}(2), ..., \boldsymbol{y}(t)\}$. Suppose $\hat{\boldsymbol{x}}(t+1)$ is the one step ahead Kalman estimate of $\boldsymbol{x}(t+1)$ based on the given $\boldsymbol{Y}^t$ (t is large enough). Let $\tilde{\boldsymbol{x}}(t+1)$ and $\boldsymbol{\Sigma}(t+1)$ be the state estimate error and the error covariance matrix, respectively. For $\boldsymbol{Y}^t$ is known and fixed, from estimation theory [2] and information theory [6] we get the conditional information of model (1) as

$$\begin{aligned} H(\boldsymbol{x}(t+1)|\boldsymbol{Y}^t) &= H(\tilde{\boldsymbol{x}}(t+1)) \\ &= \frac{n}{2}\ln(2\pi e) + \frac{1}{2}\ln\det\boldsymbol{\Sigma}(t+1). \end{aligned} \tag{20}$$

For the reduced-order system (2), let $\hat{\boldsymbol{x}}_{\mathrm{r}}(t+1)$, $\tilde{\boldsymbol{x}}_{\mathrm{r}}(t+1)$ and $\boldsymbol{\Sigma}_{\mathrm{r}}(t+1)$ denote the Kalman estimate of $\boldsymbol{x}_{\mathrm{r}}(t+1)$, the estimate error and the error covariance, respectively. Then,

$$H(\boldsymbol{x}_{\mathrm{r}}(t+1)|\boldsymbol{Y}^t) = H(\tilde{\boldsymbol{x}}_{\mathrm{r}}(t+1)) = \frac{l}{2}\ln(2\pi e) + \frac{1}{2}\ln\det\boldsymbol{\Sigma}_{\mathrm{r}}(t+1). \quad (21)$$

Suppose $\tilde{\boldsymbol{x}} = \lim_{t\to\infty}\tilde{\boldsymbol{x}}(t)$, and $\boldsymbol{\Sigma} = \lim_{t\to\infty}\boldsymbol{\Sigma}(t)$. It is well known that even when the original system is unstable, the steady covariance $\boldsymbol{\Sigma}$ exists and satisfies the following Riccati equation if the pair [$\boldsymbol{A}$, $\boldsymbol{C}$] is observable and the pair [$\boldsymbol{A}$, $\boldsymbol{H}$] (where $\boldsymbol{H}\boldsymbol{H}^{\mathrm{T}} = \boldsymbol{BQB}^{\mathrm{T}}$) is stabilizable:

$$\boldsymbol{\Sigma} = \boldsymbol{A}(\boldsymbol{\Sigma} - \boldsymbol{\Sigma C}^{\mathrm{T}}(\boldsymbol{C\Sigma C}^{\mathrm{T}} + \boldsymbol{R})^{-1}\boldsymbol{C\Sigma})\boldsymbol{A}^{\mathrm{T}} + \boldsymbol{BQB}^{\mathrm{T}}. \quad (22)$$

We also suppose $\tilde{\boldsymbol{x}}_{\mathrm{r}} = \lim_{t\to\infty}\tilde{\boldsymbol{x}}_{\mathrm{r}}(t)$, $\boldsymbol{\Sigma}_{\mathrm{r}} = \lim_{t\to\infty}\boldsymbol{\Sigma}_{\mathrm{r}}(t)$. The steady conditional information of models (1) and (2) can be gotten from (20) and (21) as

$$H(\boldsymbol{x}\,|\,\boldsymbol{Y}) = H(\tilde{\boldsymbol{x}}) = \frac{n}{2}\ln(2\pi e) + \frac{1}{2}\ln\det\boldsymbol{\Sigma}, \quad (23)$$

$$H(\boldsymbol{x}_{\mathrm{r}}\,|\,\boldsymbol{Y}) = H(\tilde{\boldsymbol{x}}_{\mathrm{r}}) = \frac{l}{2}\ln(2\pi e) + \frac{1}{2}\ln\det\boldsymbol{\Sigma}_{\mathrm{r}}, \quad (24)$$

respectively, where $\boldsymbol{Y} = \{\boldsymbol{y}(1), \boldsymbol{y}(2), ..., \boldsymbol{y}(\infty)\}$. Conditional entropies $H(\boldsymbol{x}\,|\,\boldsymbol{Y})$ and $H(\boldsymbol{x}_{\mathrm{r}}\,|\,\boldsymbol{Y})$ are respectively related to the posterior error covariances $\boldsymbol{\Sigma}$ and $\boldsymbol{\Sigma}_{\mathrm{r}}$, and can be referred to as posterior information. While $H(\boldsymbol{x})$ and $H(\boldsymbol{x}_{\mathrm{r}})$ are respectively related to prior error covariances $\boldsymbol{\Pi}$ and $\boldsymbol{\Pi}_{\mathrm{r}}$, and can be referred to as prior information.

B. Reduced-order model with minimum conditional information loss

If the reduced-order model retains more conditional information contained in the full-order model, then the better approximating performance can be expected. In order to minimize the conditional information loss, similar to the consideration in section 3, we make the transformation matrix $\boldsymbol{S}$ consist of all the ortho-normal eigenvectors of $\boldsymbol{\Sigma}$, and let $\boldsymbol{\Gamma}$ and $\boldsymbol{K}$ consist of the first l rows of $\boldsymbol{S}$ and the first l columns of $\boldsymbol{S}^{-1}$, respectively. It can be seen that $\boldsymbol{S}$ is also an entropy preserving transformation, and $\boldsymbol{S}^{-1} = \boldsymbol{S}^{\mathrm{T}}$, $\boldsymbol{K} = \boldsymbol{\Gamma}^{\mathrm{T}}$. The reduced-order model is given by

$$\{\boldsymbol{A}_{\mathrm{r}}, \boldsymbol{B}_{\mathrm{r}}, \boldsymbol{C}_{\mathrm{r}}\} = \{\boldsymbol{\Gamma A \Gamma}^{\mathrm{T}}, \boldsymbol{\Gamma B}, \boldsymbol{C\Gamma}^{\mathrm{T}}\}, \quad (25)$$

with

$$\boldsymbol{\Gamma} = [\boldsymbol{\mu}_1\ \boldsymbol{\mu}_2 \cdots \boldsymbol{\mu}_l]^{\mathrm{T}}, \quad (26)$$

where $\boldsymbol{\mu}_1, \boldsymbol{\mu}_2, \cdots, \boldsymbol{\mu}_l$ are the l ortho-normal eigenvectors corresponding to the l largest eigenvalues of $\boldsymbol{\Sigma}$.

Since $\tilde{\boldsymbol{x}}_{\mathrm{r}}(t) = \boldsymbol{\Gamma}\tilde{\boldsymbol{x}}(t)$, it can also be concluded easily that $\boldsymbol{\Sigma}_{\mathrm{r}} = \boldsymbol{\Gamma\Sigma\Gamma}^{\mathrm{T}}$ is the steady state estimate error covariance matrix of model (25), and the unique positive definite solution to the following Riccati equation:

$$\boldsymbol{\Sigma}_{\mathrm{r}} = \boldsymbol{A}_{\mathrm{r}}(\boldsymbol{\Sigma}_{\mathrm{r}} - \boldsymbol{\Sigma}_{\mathrm{r}}\boldsymbol{C}_{\mathrm{r}}^{\mathrm{T}}(\boldsymbol{C}_{\mathrm{r}}\boldsymbol{\Sigma}_{\mathrm{r}}\boldsymbol{C}_{\mathrm{r}}^{\mathrm{T}} + \boldsymbol{R})^{-1}\boldsymbol{C}_{\mathrm{r}}\boldsymbol{\Sigma}_{\mathrm{r}})\boldsymbol{A}_{\mathrm{r}}^{\mathrm{T}} + \boldsymbol{B}_{\mathrm{r}}\boldsymbol{QB}_{\mathrm{r}}^{\mathrm{T}}. \quad (27)$$

Let $\rho_1, ..., \rho_n$ denote the eigenvalues of $\boldsymbol{\Sigma}$. Then the steady conditional information of model (1) is

$$H(\boldsymbol{x}\,|\,\boldsymbol{Y}) = H(\tilde{\boldsymbol{x}}) = \frac{n}{2}\ln(2\pi e) + \frac{1}{2}\sum_{i=1}^{n}\ln\rho_i,$$

while the maximum steady conditional information retained in the reduced state is

$$H(\boldsymbol{x}_{\mathrm{r}}\,|\,\boldsymbol{Y}) = H(\tilde{\boldsymbol{x}}_{\mathrm{r}}) = \frac{l}{2}\ln(2\pi e) + \frac{1}{2}\sum_{i=1}^{l}\ln\rho_i.$$

Then the minimum steady conditional information loss caused by truncation is

$$IL(\boldsymbol{x}; \boldsymbol{x}_{\mathrm{r}}|\boldsymbol{Y}) := H(\tilde{\boldsymbol{x}}) - H(\tilde{\boldsymbol{x}}_{\mathrm{r}}) = \frac{n-l}{2}\ln(2\pi e) + \frac{1}{2}\sum_{i=l}^{n}\ln\rho_i. \quad (28)$$

The realization constructed by equations (25) and (26), where the steady error covariance $\boldsymbol{\Sigma}$ of state estimate is the unique positive solution to the Riccati equation (22), can be referred to as the minimum condition information loss (MCIL) reduced-order model.

V. Analysis and Discussion

A. Properties of the reduced-order models

It can be proved easily that the reduced-order models obtained by MIL and MCIL methods preserve the stability, controllability and observability of the original model. Due to the space limitation we omit the proof.

B. On the order selection of approximating model

It is known from equation (19) that the steady state information loss $IL(\boldsymbol{x}, \boldsymbol{x}_r)$ of MIL method is defined by the eliminated eigenvalues of covariance matrix $\boldsymbol{\Pi}$ besides the factor of dimension reduction. So, smaller eliminated eigenvalues $\sigma_{l+1}, \ldots, \sigma_n$ imply better approximating performance. If there are two neighbored eigenvalues σ_m and σ_{m+1} that $\sigma_m >> \sigma_{m+1}$, then we can set $l = m$ as the order of the reduced-order model. A same conclusion can be drawn for the MCIL method. This order selection method corresponding to the eigenvalues of covariance matrix is in some sense similar to that proposed in [3] corresponding to the eigenvalues of canonical correlation matrix.

C. Comparison of MIL and MCIL methods

As we know, the MIL and MCIL methods consider the minimum loss of the prior information $H(\boldsymbol{x})$ and the posterior information $H(\boldsymbol{x} \mid \boldsymbol{Y})$, respectively. In other words, the MIL and MCIL reduced-order models are obtained on different "information basis". It was pointed out that for a stable system, the posterior error covariance $\boldsymbol{\Sigma}$ is always smaller than the prior error covariance $\boldsymbol{\Pi}$ [1]. Since $\boldsymbol{\Sigma} < \boldsymbol{\Pi}$ implies $H(\boldsymbol{x} \mid \boldsymbol{Y}) < H(\boldsymbol{x})$, we can get more "knowledge" of system dynamics from the prior information than from the posterior information. Hence, the MIL model will possess better approximating performance than the MCIL model in general. So, although the MCIL method can be applied to stable systems, it is advisable to use the MIL method when system is stable. However, the advantage of MCIL method is that it is applicable to unstable systems.

D. Connections with balanced truncations.

It was known that the Lyapunov balanced truncation which eliminates the state variables that are both difficult to control and observe can also be considered as one particular case of the state projection. Suppose the controllability and observability Gramians of the original realization $\{\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}\}$ are $\boldsymbol{M}$ and $\boldsymbol{N}$, respectively. From equations (6) or (7) we can see that the steady state covariance matrix $\boldsymbol{\Pi}$, based on which the MIL reduced-order model is gotten, becomes the controllability Gramian if the system noise $\boldsymbol{w}(t)$ is normalized, i.e. if $\boldsymbol{Q} = \boldsymbol{I}$. In this case, the steady state information $H(\boldsymbol{x})$ can also be considered as the "controllability information". In the same sense, we can suppose an "observability information" corresponding to the observability Gramian. It can also be considered as the steady state information of the dual system.

The system is called Lyapunov balanced if $\boldsymbol{M} = \boldsymbol{N} = \mathrm{diag}[\xi_1, \ldots, \xi_n]$ (where $\xi_1 \geq \cdots \geq \xi_n > 0$ are the Hankel singular values of the system). Hence, from the viewpoint of information theory, the Lyapunov balanced truncation method can also be considered as minimizing the controllability and observability information losses (or, minimizing the steady state information losses of the original and dual systems), simultaneously. However, the Lyapunov balanced truncation method requires that the original realization be minimal. In large-scale setting balancing the whole system and then truncating the balanced basis are numerically inefficient and ill-conditioned [4]. In this sense, the idea of MIL in section 3 gives not only an information theoretic interpretation but also a simple method for model reduction.

With similar consideration to the above, it can be concluded that the LQG balanced truncation also has information theoretic interpretation: It minimizes the steady conditional information losses of states of the original and the dual systems, simultaneously.

VI. ILLUSTRATIVE EXAMPLES

A. A stable system

Consider a lightly damped, simply supported beam model [9] described by equation (1) with parameters as:

$$\text{A:}\quad \boldsymbol{A} = \mathrm{diag}\{\boldsymbol{A}_1, \cdots, \boldsymbol{A}_5\},\ \boldsymbol{B} = [\boldsymbol{B}_1^{\mathrm{T}} \cdots \boldsymbol{B}_5^{\mathrm{T}}]^{\mathrm{T}},\quad \boldsymbol{C} = [\boldsymbol{C}_1 \cdots \boldsymbol{C}_5],$$

where $\boldsymbol{A}_j = \begin{bmatrix} 0 & 1 \\ -\omega_j^2 & -2\xi\omega_j \end{bmatrix}$, $\boldsymbol{B}_j = \begin{bmatrix} 0 & b_j \end{bmatrix}^{\mathrm{T}}$, $\boldsymbol{C}_j = \begin{bmatrix} b_j & 0 \end{bmatrix}$, $\omega_j = j^2$, $[b_1, \cdots, b_5] = [0.9877, 0.309, -0.891, -0.5878, 0.7071]$, $j = 1, \cdots, 5$, and $\xi = 0.005$; $w(t)$, $v(t)$ are independent Gaussian white processes with intensities 1 and 0.1, respectively. The poles of A are $\{-0.0050 \pm 1.0000i, -0.0200 \pm 3.9998i, -0.0450 \pm 8.9999i, -0.0800 \pm 15.9998i, -0.1250 \pm 24.9997i\}$. The eigenvalues of the steady state covariance matrix $\boldsymbol{\Pi}$ are $\{\sigma_i, i = 1, \ldots, 10\} = \{48.7776, 48.7776, 4.4105, 1.1935, 1.0797, 1.0000, 0.0746, 0.0544, 0.0042, 0.0016\}$. For $\sigma_2 >> \sigma_3$, it is suitable to select 2 as the order of approximating model.

Applying the MIL method to this 10-order model, we get a 2-order approximating model A_r. The poles of model A_r are $\{-0.0050 \pm 1.0000i\}$. We denote the 2-order approximating model obtained in [9] by KLI method as A_m.

The Bode plots of the reduced-order models A_m and A_r are compared with that of the full-order model A in Figure 1. It can be seen that the approximating performance of A_r obtained by the present MIL method is better than that of A_m obtained by KLI method in [9]. Further, our present algorithm is simple and easily applicable.

B. An unstable system

To illustrate the efficiency of the MCIL method, we consider a continuous time unstable model described by equation (1) with parameters as [12]:

$$\text{B:}\quad \boldsymbol{A}=\begin{bmatrix}0&0&0&0&-114\\1&0&0&0&-86\\0&1&0&0&35\\0&0&1&0&3\\0&0&0&1&-5\end{bmatrix},\quad \begin{array}{l}\boldsymbol{B}=[70\ \ 114\ \ 55\ \ 12\ \ 1]^{\mathrm{T}},\\ \boldsymbol{C}=[0\ \ 0\ \ 0\ \ 0\ \ 1].\end{array}$$

The noises w and v are assumed to be mutual independent Gaussian processes with covariance $Q=R=1$. In this example, we examine the approximating performance of the reduced-order model by comparing its singular value to that of the original model. The poles of model B are $\{2.0565\pm1.4622i,-1.2261,-3.9435\pm1.7012i\}$. The eigenvalues of the steady Kalman estimate error matrix $\boldsymbol{\Sigma}$ are $\{357.9370, 62.2093, 9.4782, 0.9637, 0.0711\}$.

The poles of the 3-order approximating model, B_r gotten by MCIL method, are $\{-9.4857, 2.6979\pm0.8941\}$. Figure 2 compares the singular value of the reduced-order model B_r to that of the full-order model B. It shows that the MCIL method is efficient for unstable model reduction.

VII. CONCLUSION

Based on state projection, the present paper studied the model reduction problem of linear time invariant systems by analyzing the information and the conditional information descriptions of system states. The basic idea in this paper is to make the information loss or the conditional information loss caused by truncation be minimum by eliminating the state variables with the least contribution to system information. Before truncation, an entropy preserving transformation of the original state was required. The derived MIL method was proved to be efficient for approximating stable model, while the MCIL method was proved to be efficient for approximating unstable model. Connections between MIL, MCIL methods and balanced truncation methods were also analyzed.

REFERENCES

[1] Baram Y, Kailath T, Estimability and regularity of linear systems. *IEEE Trans. Automatic Control*, 33(12): 1116-1121, 1988.

[2] Caines P E, *Linear Stochastic Systems, Wiley Series in Probability and Mathematical Statistics*. John Wiley & Sons, Inc. 1988

[3] Desai U B, Pal D, A realization approach to stochastic model reduction and balanced stochastic realizations. *Proc. 21st IEEE Conference on Decision & Control*, 1982: 1105-1112.

[4] Gugercin S, Sorensen D C, Antoulas A C, A modified low-rank Smith method for large-sacle Lyapunov equations. *Numerical Algorithms*, 32(1): 27-55, 2003.

[5] Hyland D C, Bernstein D S, The optimal projection equations for model reduction and the relationships among the methods of Wilson, Skelton, and Moore. *IEEE Trans. Automatic Control*, 30(12): 1201-1211, 1985.

[6] Ihara S, *Information Theory for Continuous Systems*. Singapore: World Scientific Publishing Co. Pte. Ltd., 1993.

[7] Jonckheere E A, Silverman L M, A new set of invariants for linear systems — applicatopn to reduced order compensator design. *IEEE Trans. Automatic Control*, 28(10): 953-964, 1983.

[8] Kapur J N, Kesavan H K, *Entropy Optimization Principles with Applications*. Academic Press, Inc., 1992.

[9] Leland R, Reduced-order models and controllers for continuous-time stochastic systems: an information theory approach. *IEEE Trans. Automatic Control*, 44(9): 1714-1719, 1999.

[10] Moore B C, Principle component analysis in linear systems: controllability, observability, and model reduction. *IEEE Trans. on Automatic Control*, 26(1): 17-32, 1981.

[11] Touchette H, Lloyd S, Information-theoretic approach to the study of control systems. *Physica A*, 331: 140-172, 2004.

[12] Xiao C-S, Feng Z-M, Shan X-M, On the solution of the continuous-time Lyapunov matrix equation in two canonical forms. *IEE Proceedings-D*, 139(3): 286-290, 1992.

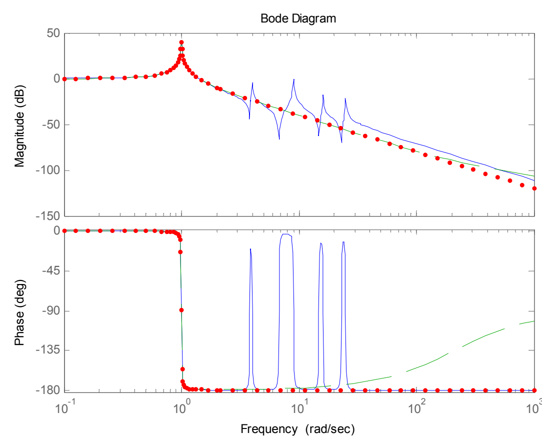

Figure 1: Bode plots of models A (solid line), A_m (dashed line) and A_r (dot line)

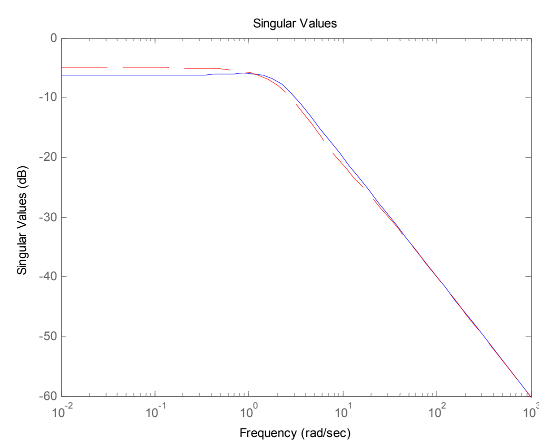

Figure 2: Singular value plots of B (solid line) and B_r (dashed line)