Gramian-based model order reduction of parameterized time-delay systems

Xiang Wang*†, Zheng Zhang‡, Qing Wang§ and Ngai Wong

Department of Electrical and Electronic Engineering, The University of Hong Kong, Hong Kong

ABSTRACT

Time-delay systems (TDSs) frequently arise in circuit simulation especially in high-frequency applications. Model order reduction (MOR) techniques can be used to facilitate the simulation of TDSs. On the other hand, many kinds of variations, such as temperature and geometric uncertainties, can have significant impact on the transient responses of TDSs. Therefore, it is important to preserve parametric dependence during the MOR procedure. This paper presents a new parameterized MOR scheme for TDSs with parameter variations. We derive parameterized reduced-order models (ROMs) for TDSs using balanced truncation by approximating the Gramians in the multi-dimensional space of parameters. The resulting ROMs can preserve the parametric dependence, making it efficient for repeated simulations under different parameter settings. Numerical examples are presented to verify the accuracy and efficiency of our proposed algorithm. Copyright © 2012 John Wiley & Sons, Ltd.

1. INTRODUCTION

In high-speed circuit design, time-delay phenomena frequently appear due to the propagation delays, for instance, caused by the transmission lines in circuit packaging and printed circuit board (PCB) design [1–5]. In many cases of packaging and PCB design/optimization, propagation delay can dominate circuit performance, whereby attenuation effects caused by transmission lines are negligible, and the transmission lines can be regarded as lossless. The problem size and computational cost make direct simulation impractical for variational analysis of time-delay systems (TDSs). Therefore, model order reduction (MOR) techniques are applied to reduce the problem size for efficient simulation [6–13].

A TDS can be a variational model when there are various parameters affecting system performance. The geometric parameters of transmission lines, such as width and thickness, can have a significant impact on system performance [14], whereas temperature is another factor that can affect transmission line characteristics [15]. In this case, the computational cost is much higher than the non-parametric case because the original TDS needs to be simulated repeatedly in order to verify a number of parameter settings. This is common for worst-case (corner) simulation at the early stage of circuit design for the purpose of performance optimization. Therefore, parameterized MOR becomes necessary [16–19]. In parameterized MOR, the MOR procedure is performed only once to

*Correspondence to: Xiang Wang, Department of Electrical and Electronic Engineering, The University of Hong Kong, Hong Kong
†E-mail: xwang@eee.hku.hk
‡Zheng Zhang is in the Department of Electrical Engineering and Computer Science, MIT, Cambridge. Part of this work was finished when Zheng Zhang was in the University of Hong Kong.
§Qing Wang is also with the School of Information Science and Technology, Sun Yat-Sen University, Guangzhou, China.

Copyright © 2012 John Wiley & Sons, Ltd.
obtain a variational reduced-order model (ROM), which is subsequently used for repeated simulations under different parameter settings. Recently, [20, 21] propose new parameterized MOR approaches for partial element equivalent circuit models which can preserve system passivity.

There are generally two main categories of MOR techniques. The first category is the moment-matching scheme [6–8]. It projects the original system onto a Krylov subspace by matching the first several moments of the transfer function. The second category of MOR, named truncated balanced realization (TBR) [9], stems from control theory and utilizes the controllability and observability Gramians. The TBR approaches construct the projection matrices by preserving the dominant states of the balanced systems. Moment-matching based approaches are numerically more efficient than TBR-based approaches; thus, they have extensive applications. Due to the fact that there does not exist an analytical error bound for moment-matching, the ROMs are not optimal, and it is difficult to automatically control the error. On the other hand, the TBR-type methods provide a better global accuracy and usually obtain nearly optimal ROMs subject to a given accuracy requirement. This motivates us to develop a TBR-based algorithm for macromodeling TDSs with variational analysis.

The major cost of TBR lies in computing of the controllability and observability Gramians. For a linear time invariant (LTI) system, these two Gramians are computed from two Lyapunov equations at the cost of $O(n^3)$, where $n$ is the size of the original system. Recently, a novel TBR-type algorithm, named the Poor Man’s TBR (PMTBR) [22], was introduced to facilitate the computation of the Gramians. By Parseval’s theorem, PMTBR approximates the controllability and observability Gramians in a limited frequency band, other than using their time-domain definitions which are infeasible to be computed directly. PMTBR can provide a global accuracy within the frequency band of interest, especially when the original system’s input signals have finite bandwidths. Since PMTBR is based on matrix-vector multiplication (as in moment-matching schemes), high computational efficiency can be obtained.

In the literature, the most widely used MOR approaches for TDSs are moment-matching based [23–26]; however, few results are reported on the Gramian-based MOR of TDSs [27], and there is also a dominant-pole algorithm for reduction of second-order TDSs [28]. The major difficulty of applying a TBR-type approach on a TDS is the complicated definitions of the Gramians, which involve the Lambert W function [29]. Because there is no existing analytical convergence criterion for the Lambert W function, one cannot determine how many branches are needed to calculate the TDS Gramians, either from their time-domain definitions or from the corresponding Lyapunov-type equations. Moreover, the direct computation of TDS Gramians in the time domain involves matrix inversion, making the computational cost even higher [30]. Therefore, the authors propose to utilize the PMTBR scheme to facilitate the Gramian calculation for TDSs.

The cost would be extremely high when design or process parameter variations need to be considered. Most methodologies for variational analysis are based on Monte Carlo simulation [31]. Specifically, one needs to perform full system simulation at every sampling point to obtain the system behaviors corresponding to different process or design variations. Therefore, it is highly desirable to use an efficient parameterized macromodeling scheme to avoid the prohibitive computational cost caused by repeated simulations. Because the parameterized ROM is generated only once and then can be repeatedly used in subsequent simulations, the computational cost can be remarkably reduced.

The proposed algorithm takes advantage of the fact that the typical TDSs usually have a finite input bandwidth and that the time-delay effects are only considered for giga-Hertz (or above) systems (because the delay values caused by the transmission lines are in the magnitude of nanoseconds). Therefore, it is justifiable to evaluate the TDS Gramians by the frequency-domain integration in a limited frequency band. To tackle the problem of variational Gramians in parameterized systems, we utilize the variational PMTBR scheme proposed in [32]. Numerical accuracy and efficiency of our proposed algorithm are shown by various numerical examples.

The rest of the paper is organized as follows. Section 2 introduces the formulations of TDSs, as well as the existing definitions of TDS Gramians. Section 3 first reviews the PMTBR and parameterized MOR, and then introduces the parameterized MOR using variational Gramians. In Section 4, we first analyze the Lyapunov-type equations of TDSs, and then the proposed algorithm is elaborated.
with a discussion of practical implementation. Numerical examples are presented in Section 5. Finally, Section 6 concludes the paper.

2. BACKGROUND OF TDSS

2.1. TDSs

A TDS is formulated as

\[
\begin{align*}
    \dot{x}(t) &= Ax(t) + A_dx(t-h) + Bu(t), & t > 0 \\
    y(t) &= Cx(t) \\
    x(0) &= x_0 \\
    x(t) &= g(t), & t \in [-h, 0]
\end{align*}
\]

(1)

where \(x(t) \in \mathbb{R}^n\) is the state vector, \(A, A_d \in \mathbb{R}^{n \times n}\) are the state matrices, \(h > 0\) is a constant time delay in the states, \(B \in \mathbb{R}^{n \times p}\) is the input matrix, \(CT \in \mathbb{R}^{q \times n}\) is the output matrix, \(y(t) \in \mathbb{R}^q\) is the output vector, and \(u(t) \in \mathbb{R}^p\) is the input vector. \(x_0\) and \(g(t)\) are the initial conditions within the delay interval \([-h, 0]\).

For simplicity, we assume that the TDS is a single-delay system; the generalization to multiple delays is straightforward and will be discussed in Section 4.3.

To analyze the controllability and observability Gramians, one needs to derive the analytical state solution for a TDS. In [33], the state solution of (1) is formulated as

\[
x(t,0,g,u) = x(t,0,g,0) + \int_0^t K(t,\tau)Bu(\tau)d\tau,
\]

(2)

where \(x(t,0,g,u)\) denotes a solution at time \(t\) corresponding to the initial time 0, initial condition \(g(t)\), and input signal \(u(t)\). The homogeneous solution \(x(t,0,g,0)\) of (1) is obtained by setting the input function \(u(t)\) to be constant 0. The fundamental matrix function \(K(t,\tau)\) of (1) satisfies the following equations

\[
\begin{align*}
    \frac{\partial K(t,\tau)}{\partial \tau} &= -K(t,\tau)A - K(t,\tau + h)A_d, & 0 \leq \tau \leq t - h \\
    K(t,t) &= I \\
    K(t,\tau) &= 0, & \text{for } \tau > t
\end{align*}
\]

(3)

Here, \(K(t,\tau)\) is essential for deriving the controllability and observability Gramians for a TDS.

2.2. Controllability and observability Gramians of TDSs

In the literature, there are many kinds of definitions for the Gramians of a TDS [33–35], such as \(M_2\) controllability, absolute controllability, point-wise controllability, and some hybrid types. In this paper, we use the point-wise controllability which is the most widely used. We choose this point-wise definition for TDS controllability/observability because in the literature, we find that the TDS Gramians can only be derived from this definition, but not the others. Using the other definitions, by now, one cannot derive such closed-form Gramians for TDSs, which are critical for the balanced-truncation approach.

Definition 1

(Controllability of a TDS) [36]: The system (1) is point-wise controllable if, for any given initial conditions \(g(t)\) and \(x_0\), there exists \(0 < t_1 < \infty\), and an admissible (i.e. measurable and bounded on a finite time interval) input \(u(t)\) for \(t \in [0, t_1]\) such that \(x(t_1, 0, g(t), u(t)) = 0\).

It is still difficult to use Definition 1 to numerically check the controllability of a TDS in practice due to the lack of a well-established analytical criterion. Weiss proposed an algebraic criterion to
check the point-wise controllability of linear TDSs [33]. In [36], the Lambert W function is used to define the Gramian matrices analytically. Algebraic Gramians are derived from the time-domain solution to system (1) using the matrix Lambert W function [36]

\[ x(t) = \sum_{k=-\infty}^{\infty} e^{S_k t} C_k^T + \int_0^t \sum_{k=-\infty}^{\infty} e^{S_k (t-\tau)} C_k^N B u(\tau) d\tau, \]  
(4)

with

\[ S_k = \frac{1}{h} W_k(A_d h Q_k) + A, \]  
(5)

where \( W_k(H_k) \) is the matrix Lambert W function of \( H_k \) defined in [36]. For more details of the convergence conditions, definitions, and calculations of \( C_1^T, C_k^N, \) and \( Q_k, \) we refer the readers to [36, 37, 30].

As discussed in Section 2.1, the Gramians can be derived directly from the analytical fundamental matrix function

\[ K(t, \tau) = \sum_{k=-\infty}^{\infty} e^{S_k (t-\tau)} C_k^N. \]  
(6)

Analogous to LTI systems, the controllability Gramian of a TDS is defined as [36]

\[ P(0, t_1) = \int_0^{t_1} \sum_{k=-\infty}^{\infty} e^{S_k (t_1-\tau)} C_k^N B B^T \left\{ \sum_{k=-\infty}^{\infty} e^{S_k (t_1-\tau)} C_k^N \right\}^T d\tau. \]  
(7)

The analytical criterion of observability is defined similarly as below.

**Definition 2**

(Observability of a TDS) [36]: The system (1) is point-wise observable in \([0, t_1]\) if the initial point \( x_0 \) can be uniquely determined from the knowledge of \( u(t), g(t), \) and \( y(t). \)

The observability Gramian [36] is defined as

\[ Q(0, t_1) = \int_0^{t_1} \left\{ \sum_{k=-\infty}^{\infty} e^{S_k (t_1-\tau)} C_k^N \right\}^T C^T C \left( \sum_{k=-\infty}^{\infty} e^{S_k (t_1-\tau)} C_k^N \right) d\tau. \]  
(8)

If there is no delay effect (e.g. setting \( A_d = 0), \) a TDS reduces to an LTI system, and the controllability and observability Gramians are exactly the same as in the LTI case. Therefore, the TBR procedure for a TDS reduces to the LTI case when \( A_d = 0. \)

These analytical Gramians and their time-domain evaluations were first proposed in [36], where only the first several branches of the Lambert W function are used to approximate the Gramians. The unavoidable matrix inversion in computing coefficients \( S_k \) and \( C_k^N \) [30] makes such an evaluation too expensive in practice. Furthermore, it is common in VLSI circuit simulation that the original TDSs are of very high order; thus, this time-domain evaluation approach is infeasible due to its extremely high computational cost. Moreover, one cannot determine how many branches should be used to approximate the Gramians without a convergence criterion for those infinite branches of the Lambert W function.
3. BALANCE TRUNCATION WITH PARAMETER VARIATIONS

3.1. Review of the PMTBR

The controllability Gramian $P$ and observability Gramian $Q$ of a stable LTI state-space model are defined as

$$P = \int_0^\infty e^{At}BB^Te^{At^T}dt, \quad Q = \int_0^\infty e^{At^T}C^Te^{At}dt,$$

(9)

It can be shown that $P$ and $Q$ are unique positive semidefinite solutions to the Lyapunov equations

$$AP + PA^T + BB^T = 0, \quad A^TQ + QA + CT^T = 0.$$  

(10)

The conventional TBR computes the Gramians directly from the Lyapunov equations (10). TBR has advantages over moment-matching because it provides an analytical error bound in terms of $H^\infty$ norm of the transfer function [38]. However, the TBR methods are generally more expensive than the moment-matching algorithms.

According to the Parseval’s theorem, Gramians (9) defined in the time domain can be reformulated as [22]

$$P = \frac{1}{2\pi} \int_{-\infty}^{\infty} (j\omega I - A)^{-1}BB^T(j\omega I - A)^{-H}d\omega,$$

(11a)

$$Q = \frac{1}{2\pi} \int_{-\infty}^{\infty} (j\omega I - A^T)^{-1}C^TC(j\omega I - A^T)^{-H}d\omega,$$

(11b)

where the superscript $H$ denotes Hermitian transpose. Note that the coefficient $\frac{1}{2\pi}$ can be ignored in numerical implementation because it does not affect the resulting projection matrices. With the weighting function $w_k$ which results from a proper sampling scheme, a finite summation is then used to approximate the integration defined in an infinite frequency band. Defining

$$z_{ck} = (j\omega_k I - A)^{-1}B, \quad z_{ok} = (j\omega_k I - A^T)^{-1}C^T,$$

(12)

the approximate Gramians can be calculated as

$$\tilde{P} = \sum_k w_kz_{ck}z_{ck}^H, \quad \tilde{Q} = \sum_k w_kz_{ok}z_{ok}^H.$$  

(13)

The detailed discussions on sampling-point selection and error estimation are presented in [22], and in practice, we can set $w_k = 1$ [39, 40]. In most practical circuits, the inputs usually have finite bandwidth, and we are interested in the system performance only within a limited frequency band. Therefore, it is reasonable to calculate the approximate Gramians (13) within the finite frequency band of interest and subsequently use the resulting finite-bandwidth Gramians to construct the projection matrices [22,40].

3.2. Parameterized MOR

Consider a general linear system in the frequency domain with parameter variations, and we denote all the varying/uncertain parameters (frequency, temperature, geometry, etc.) by a set of $k$ parameters $\hat{\lambda}_1, \ldots, \hat{\lambda}_k$.
where \( J \) denotes the system matrix in the frequency domain with parameter variations. The conventional parameterized MOR methods [41–44] use polynomial fitting technique to transfer system (14) into a tractable form

\[
J(\lambda_1, \ldots, \lambda_k) = J_0 + \sum_{i=1}^{k} \lambda_i J_i + \sum_{j=1}^{k} \lambda_j \lambda_j J_{ji} + \cdots,
\]

and then truncate (15) to an approximate parameterized system (17) by introducing some new parameter variables

\[
\tilde{J}_r = \begin{cases} J_i & i = 0, \ldots, k \\ J_{ji} & j = 1, \ldots, k; l = 1, \ldots, k \\ \tilde{\lambda}_i & i = 1, \ldots, k \\ \tilde{\lambda}_j \tilde{\lambda}_j & j = 1, \ldots, k; l = 1, \ldots, k, \end{cases}
\]

\[
\left[ \tilde{J}_0 + \tilde{\lambda}_1 \tilde{J}_1 + \cdots + \tilde{\lambda}_p \tilde{J}_p \right] x = Bu.
\]

The second step of parameterized MOR is to construct the projection matrix \( V \) which contains the variational information within the parameter space of interest. After that, the reduced system is obtained by

\[

\left[ V^T J_0 V + \tilde{\lambda}_1 V^T J_1 V + \cdots + \tilde{\lambda}_p V^T J_p V \right] \tilde{x} = V^T Bu.
\]

When parameters are considered, a TDS can be described as

\[
\dot{x}(t) = A(\lambda_1 \cdots, \lambda_k)x(t) + A_d(\lambda_1 \cdots, \lambda_k)x(t - h(\lambda_1 \cdots, \lambda_k)) + Bu(t).
\]

We assume that input/output matrices \( B \) and \( C \) are parameter independent, because using the TDS models described in [23], these matrices are only related to the mapping topology of input/output signals. Similarly, before reduction by projection, we need to first approximate the variational TDS (19) into an explicit parameterized form like (17). According to the fitting procedure (15), another assumption of the proposed algorithm is that the matrices of the original model must have the same dimensions as a function of the parameters.

3.3. Computational complexity with parameter growth

From the parameterized formulation (15), it is noted that system accuracy can be improved if the largest order of derivative grows. However, the size of the parameterized system (17) will grow quickly and become extremely high if high-order parameterization is performed. The relation between the size \( p \) of the parameterized model and the largest order \( m_p \) of derivative is given by [44]

\[
p = O \left( \frac{k^{m_p}}{(m_p - 0.5)^{m_p}} \right),
\]

where \( k \) is the number of parameters in the original system. If parameter dependencies are smooth, \( m_p = 1 \) is enough to obtain a good accuracy. In such a case, the size \( p \) of the parameterized system grows linearly with the number of parameters; therefore, this algorithm can apply on a larger number of parameters. In this paper, we focus on handling systems
with smooth parameter dependencies; therefore, only low-order derivative is needed in the parameterization step. If the parameter dependencies are less smooth where a large \( m_p \) is required, the size \( p \) of the parameterized system (17) will grow quickly even for a very small number \( k \) of parameters.

In fact, the exponential growth of the ROM size with respect to the parameter number and expansion order is an open question in parameterized MOR community. Although a recent optimization-based work [45] whose resulting ROM size is independent of the number of parameter has been proposed, no solution has been proposed for the time-delay cases. For the non-smooth parameter dependence, the optimization-based idea [45] can be used for systems without delay effects, and its ROM size is also independent of the number of parameters. We may extend this idea to TDSs based on balanced truncation (BT). However, this topic is beyond the scope of the paper, and it will be addressed in our future work.

### 3.4. Parameterized MOR via variational Gramians

Assume the LTI system (21) with a single parameter \( \lambda \) in its system matrix. The extension to multiple parameters is shown in Section 4.2.

\[
\dot{x}(t) = A(\lambda)x(t) + Bu(t), \quad y = Cx(t).
\] (21)

The variational Gramians in the frequency-parameter space are defined as [32]

\[
P_\lambda = \frac{1}{2\pi} \int_{-\infty}^{\infty} (j\omega I - A(\lambda))^{-1}BB^T (j\omega I - A(\lambda))^{-H} d\omega d\lambda, \tag{22a}
\]

\[
Q_\lambda = \frac{1}{2\pi} \int_{-\infty}^{\infty} (j\omega I - A(\lambda)^T)^{-1}C^T C (j\omega I - A(\lambda)^T)^{-H} d\omega d\lambda, \tag{22b}
\]

where \( P_\lambda \) and \( Q_\lambda \) are the variational Gramians, and \( S_\lambda \) is the space of parameter variations. Note that the coefficient \( \frac{1}{2\pi} \) can be ignored in numerical implementation because it does not affect the resulting projection matrices. From a quadrature-motivated viewpoint, one can use multi-dimensional finite summations to approximate the multi-dimensional integral Gramians. By defining

\[
z_{ckl} = (j\omega_k I - A(\lambda_l))^{-1}B, \quad z_{okl} = (j\omega_k I - A(\lambda_l)^T)^{-1}C^T,
\] (23)

the approximate variational Gramians are formulated as

\[
\tilde{P}_\lambda = \sum_k \sum_l z_{ckl}z_{ckl}^H, \quad \tilde{Q}_\lambda = \sum_k \sum_l z_{okl}z_{okl}^H. \tag{24}
\]

After computing these Gramians, the conventional TBR procedure can then be applied to reduce the parameterized system.

### 4. PROPOSED ALGORITHM

#### 4.1. Lyapunov-type equations of TDSs

Here, we briefly analyze the complexity of Lyapunov-type equations for a TDS.
By defining

\[ w^P_{ij} = e^{S_i} C^i B B^T \left( e^{S_j} C^j \right)^T, \quad w^Q_{ij} = (e^{S_i} C^i)^T C^T C e^{S_j} C^j, \]  

(25)

\[ P_{ij} = \int_0^\infty w^P_{ij} dt, \quad Q_{ij} = \int_0^\infty w^Q_{ij} dt, \]  

(26)

the controllability and observability Gramians (7) and (8) can be rewritten as

\[ P = \sum_{i=-\infty}^{+\infty} \sum_{j=-\infty}^{+\infty} \int_0^\infty w^P_{ij} dt = \sum_{i=-\infty}^{+\infty} \sum_{j=-\infty}^{+\infty} P_{ij}, \]  

(27a)

\[ Q = \sum_{i=-\infty}^{+\infty} \sum_{j=-\infty}^{+\infty} \int_0^\infty w^Q_{ij} dt = \sum_{i=-\infty}^{+\infty} \sum_{j=-\infty}^{+\infty} Q_{ij}. \]  

(27b)

For a stable TDS, we can obtain a set of Lyapunov-type equations for the Gramian components \( P_{ij} \) and \( Q_{ij} \)

\[ S_i P_{ij} + P_{ij} S_j^T + C^i B B^T \left( C^j \right)^T = 0, \]  

(28a)

\[ S_i^T Q_{ij} + Q_{ij} S_j + \left( C^i \right)^T C^T C \left( C^j \right) = 0. \]  

(28b)

Then, the Gramians (27) can be approximated by using their first several branches of solutions to (28).

However, evaluating the Gramians this way is as difficult as evaluating them by their time-domain definitions. Because of the \( O(n^3) \) complexity involved in the Lyapunov-type equations and without a convergence criterion of the branches, it is impractical to perform TBR on TDSs by computing the Gramians from these Lyapunov-type equations. Therefore, we utilize the sampling-based PMTBR scheme to evaluate the TDS Gramians in the frequency domain.

4.2 Parameterized TBR of TDSs

For a variational TDS with \( k \) parameters modeled in (19), we use the notation \( \theta \) to represent the combination of various \( k \) parameters

\[ \theta = (\hat{\lambda}_1, \ldots, \hat{\lambda}_k), \]  

(29)

and we use \( \theta^i \), to denote the \( i^{th} \) point in the parameter space. The variational TDS (19) now can be rewritten as

\[ \dot{x}(t) = A(\theta)x(t) + A_d(\theta)x(t - h(\theta)) + Bu(t). \]  

(30)

By using second-order affine model

\[ A(\theta) = A_0 + \sum_{i=1}^{k} \hat{\lambda}_i A_i + \sum_{j=1}^{k} \hat{\lambda}_j \hat{\lambda}_j A_{ij}, \]  

(31a)

\[ A_d(\theta) = A_d0 + \sum_{i=1}^{k} \hat{\lambda}_i A_{di} + \sum_{j=1}^{k} \hat{\lambda}_j \hat{\lambda}_j A_{dij}, \]  

(31b)

the variational TDS (30) now has explicit dependencies on various parameters, and it is ready for the parameterized MOR procedure. Note that we can also choose high-order affine models in the
parameterization step if the parameter dependencies are not so smooth and high accuracy is required. The delay is not parameterized because using the TDS models described in [23] the delay has explicit dependencies on parameters, and it can be evaluated directly. The matrices $A(\theta)$ and $A_d(\theta)$ are fit in the sense of matrix-wise, and if the dimension of the parameter space is high, the complexity of this fitting procedure will be also very high. By now, the proposed algorithm is only efficient for cases with a few delays, and it can not deal with cases with hundreds or thousands of delays.

By Laplace transform, the fundamental matrix function of a non-parameterized TDS becomes [36]

$$
\mathcal{L}\left\{ \sum_{k=-\infty}^{\infty} e^{Sk} C_k^N B \right\} = (sI - A - A_d e^{-sk})^{-1} B, \tag{32a}
$$

$$
\mathcal{L}\left\{ \sum_{k=-\infty}^{\infty} e^{Sk} C_k^N C^T \right\} = (sI - A^T - A_d^T e^{-sk})^{-1} C^T, \tag{32b}
$$

where $\mathcal{L}$ denotes the Laplace transform operation.

Combining (32) with (22), the variational Gramians of a TDS are calculated in the frequency-parameter space as

$$
P_{\theta} = \frac{1}{2\pi} \int_{S_\theta} \int_{-\infty}^{\infty} \left( j\omega I - A(\theta) - A_d(\theta) e^{-j\omega \theta} \right)^{-1} B B^T \left( j\omega I - A(\theta) - A_d(\theta) e^{-j\omega \theta} \right)^{-H} d\omega d\theta, \tag{33}
$$

$$
Q_{\theta} = \frac{1}{2\pi} \int_{S_\theta} \int_{-\infty}^{\infty} \left( j\omega I - A(\theta) - A_d(\theta) e^{-j\omega \theta} \right)^{-H} C^T C \left( j\omega I - A(\theta) - A_d(\theta) e^{-j\omega \theta} \right)^{-1} d\omega d\theta, \tag{34}
$$

where $S_\theta$ is the space of parameter variations. Note that the coefficient $\frac{1}{2\pi}$ can be ignored in numerical implementation because it does not affect the resulting projection matrices. By defining

$$
z_{ckl} = \left( j\omega_k I - A(\theta_l) - A_d(\theta_l) e^{-j\omega_k \theta(l)} \right)^{-1} B, \tag{35a}
$$

$$
z_{okl} = \left( j\omega_k I - A(\theta_l) - A_d(\theta_l) e^{-j\omega_k \theta(l)} \right)^{-H} C^T, \tag{35b}
$$

the approximate Gramians can be calculated as

$$
P_{\theta} = \sum_k \sum_l z_{ckl} z_{ckl}^H, \quad Q_{\theta} = \sum_k \sum_l z_{okl}^H z_{okl}, \tag{36}
$$

By experiments, we have found that usually a moderate or small set of sampling points is adequate. To provide a global accuracy within the frequency-parameter varying range, the sampling points were chosen to be distributed uniformly in logarithmic scale within the frequency band and distributed uniformly in linear scale within the parameter space. Combine all the $N$ sampling points in the frequency-parameter space, and denoting

$$
Z_c = [z_{c1}, z_{c2}, \ldots, z_{cN}], \quad Z_o = [z_{o1}, z_{o2}, \ldots, z_{oN}], \tag{37}
$$

the approximate Gramians are further expressed as
After calculating the variational Gramians of a TDS, the standard TBR procedure is then applied to construct the parameterized ROM. The proposed algorithm is summarized in Algorithm 1. Note that parameterized MOR by projection preserves the TDS structure, and the delay effects remain the same as in the original system. Similar to classical TBR-based MOR techniques, the proposed variational Gramian-based method for TDSs may not preserve system passivity. One assumption to use our proposed algorithm is that the original model should be stable. Stability and passivity preservation are necessary for stable simulation of interconnected systems; however, such an enhancement is still an open question in the EDA community. This topic is beyond the scope of this paper, and it will be addressed in our future work.

Algorithm 1: Parameterized TBR of TDSs via Variational Gramians

1. Input: \( A(\lambda_1, \ldots, \lambda_k), A_d(\lambda_1, \ldots, \lambda_k), h(\lambda_1, \ldots, \lambda_k), B, C \)
2. Denote the combination of \( k \) parameters by \( \theta = (\lambda_1, \ldots, \lambda_k) \), and use least-squares polynomial interpolation to fit a TDS in an affine model with explicit dependencies on various parameters

\[
A(\theta) = A_0 + \sum_{i=1}^{k} \lambda_i A_i + \sum_{j=1}^{k} \lambda_j A_{ij},
\]

\[
A_d(\theta) = A_{d0} + \sum_{i=1}^{k} \lambda_i A_{di} + \sum_{j=1}^{k} \lambda_j A_{dij},
\]

3. Select \( N \) sampling points within the frequency-parameter space \( (\omega_i, \theta_i), i = 1, \ldots, N \)
4. Compute \( Z = [z_{c1}, z_{c2}, \ldots, z_{cN}] \), where \( z_{ci} = (j\omega_i I - A(\theta_i))^{-1} B \)
5. Compute \( Z_0 = [z_{s1}, z_{s2}, \ldots, z_{sN}] \), where \( z_{si} = (j\omega_i I - A(\theta_i))^{-1} B \)
6. Compute the Gramians \( P = Z \tilde{Z}_c^H \) and \( Q = Z_0 \tilde{Z}_s^H \)
7. Compute Cholesky factors \( P = XX^H \) and \( Q = YY^H \)
8. Compute SVD of cross factors and partition at the significant singular-value drop according to a given threshold \( \epsilon \) (i.e., \( \Sigma_1(2,1) < \epsilon \Sigma(1,1) \))

\[
X^H Y = U \Sigma V^T = U \begin{bmatrix} \Sigma_1 & \Sigma_2 \\ \Sigma_1 & \Sigma_2 \end{bmatrix} V^T
\]
9. Compute the balanced transformation matrices

\[
T_L = \Sigma^{-1/2} \Sigma_2 \Sigma_1^{-1/2}
\]

\[
T_R = X U \Sigma^{-1/2}
\]
10. Partition the balanced realization conformal to Step 8

\[
A_b(\theta) = T_L A(\theta) T_R = \begin{bmatrix} A_{11}(\theta) & A_{12}(\theta) \\ A_{21}(\theta) & A_{22}(\theta) \end{bmatrix},
\]

\[
A_{db}(\theta) = T_L A_d(\theta) T_R = \begin{bmatrix} A_{d1}(\theta) & A_{d12}(\theta) \\ A_{d21}(\theta) & A_{d22}(\theta) \end{bmatrix},
\]

\[
B_b = T_L B = \begin{bmatrix} B_1 \\ B_2 \end{bmatrix}, \quad C_b = C T_R = \begin{bmatrix} C_1 \\ C_2 \end{bmatrix}
\]
11. Truncate the balanced realization to obtain the ROM

\[
A_r(\theta) = A_{11}(\theta), \quad A_{dr}(\theta) = A_{d1}(\theta), \quad B_r = B_1, \quad C_r = C_1
\]
12. Output: \( A_r(\theta), A_{dr}(\theta), h(\theta), B_r, C_r \)

4.3. Practical implementation

In practice, TDSs may appear like descriptor systems (DSs), and the TDSs can also have multiple delay values resulting from different signal propagation paths. Therefore, the state-space model of a variational TDS becomes

\[
E(\theta)\dot{x}(t) = A(\theta)x(t) + \sum_{i=1}^{m} A_{di}(\theta)x(t - h_i(\theta)) + Bu(t),
\]

where \( m \) is the number of delay terms, and \( h_i \) is the delay value. In this case, the formulations (35) for the components of variational Gramians become

\[
\tilde{P}_\theta = Z_c Z_c^H, \quad \tilde{Q}_\theta = Z_s Z_s^H.
\]
\[ z_{ckl} = (j\omega_k E(\theta_l) - A(\theta_l) - \sum_{i=1}^{m} A_{1i}(\theta_l) e^{-j\omega_k h_i(\theta_l)})^{-1} B, \] (40a)

\[ z_{ckl} = (j\omega_k E(\theta_l) - A(\theta_l) - \sum_{i=1}^{m} A_{1i}(\theta_l) e^{-j\omega_k h_i(\theta_l)})^{-H} C^T, \] (40b)

By replacing steps 4 and 5 in Algorithm 1 with (40), we can perform variational TBR on TDSs (39) in the form of DSs with multiple delays.

5. EXPERIMENTAL RESULTS

5.1. Proposed TBR on a non-parametric TDS with variable length

In this paper, a TDS is constructed by connecting a linear RLC network with several sets of lossless MTLs. Figure 1 shows the model structure of a TDS. When the input signals have very high frequencies, the MTLs can no longer be regarded as lumped elements, and the delay effects become significant. The coupled parameters of MTLs, including coupled capacitance and inductance, result in signal propagation delays in the range of nanoseconds. The resulting delay effects in the frequency domain emerge from around 1 GHz. This motivates us to evaluate the Gramians within the high-frequency band of interest. The detailed procedures of constructing such TDSs are elaborated in [23], and the per-unit-length (P.U.L.) parameters of MTLs are extracted utilizing the approaches in [14]. Numerical experiments in this paper are performed on a platform of Intel Core 2 Q8400 with 2.66 GHz CPU and 3.25 GB RAM.

We can include the temperature and geometric parameter variations in our MTL model such as width and thickness, which have significant impacts on the transient responses of TDSs. To efficiently evaluate variational TDSs within the parameter-varying band of interest, parameterized ROMs are constructed at the nominal point using variational Gramians by sampling in the frequency-parameter space. The ROMs can then be repeatedly used to predict system performance.

In this section, we first use a non-parametric TDS to verify the frequency-sampling scheme for Gramian approximation. A TDS of order 688 is constructed by using a two-port linear interconnect network.

![Figure 1. A TDS consisting of an RLC interconnect network connected with two sets of lossless MTLs.](image)
network connected with 24 lossless three-conductor transmission lines, and this TDS has two different delay values in its state variable. Given a threshold value of $2 \times 10^{-9}$, the Hankel singular value (HSV) matrix is truncated at the significant singular-value drop, and an ROM of order 20 is constructed by utilizing the proposed TBR approach. Because there is no analytical error bond available for this approximate Gramian-based TBR, the threshold value for truncation is determined heuristically and based on experiments. The TDS Gramians are approximated by using 20 sampling points distributed uniformly in a logarithmic scale within the frequency band of interest, i.e. from $10^6$ to $10^{12}$ Hz in our experiments. It is noted that the number of sampling points is chosen by experiments, and in practice, a small set of sampling points can provide a good accuracy. This sampling strategy is also applied in other examples. Figure 2 shows the time-domain input current source and its frequency-domain spectrum, and this input is also applied in the following experiments. Figure 3 compares the frequency response of the ROM by our proposed TBR with the original TDS. In this paper, transfer functions in the frequency domain use impedance representation. It is shown that our ROM provides good accuracy within the frequency band of interest. The time-domain transient response is also compared in Figure 4, showing that the ROM accurately captures the delay effects.

The proposed TBR uses about 14.2 s to construct the ROM, and the transient simulation times are 12.6 s for the original model and 0.26 s for the ROM, obtaining about 50× speedup. Note that the ROM construction time is comparable with the transient simulation time, which means that the proposed algorithm is relatively efficient, especially when the ROM is normally used for repeated simulations.

![Figure 2](image1.png)

*Figure 2. Input current source and its frequency-domain spectrum. (a) Input source. (b) Frequency-domain spectrum.*

![Figure 3](image2.png)

*Figure 3. Frequency responses of Example 1 and its relative errors. (a) Frequency responses. (b) Frequency-domain relative errors.*
The length of MTLs has evident effects on TDS performance, and changing the length of an MTL can lead to an evident shift of the resonances in frequency/transient responses. Using the TDS formulations presented in [23], the length does not affect the P.U.L. parameters of MTLs, and the system matrices are not functions of length. Therefore, there is no need to perform parameterization on length, and the delay values can be evaluated directly. In this experiment, we verify the proposed TBR on three length settings (i.e. 0.5 m, 1 m, and 1.5 m). The nominal model has the length of 1.2 m, and we calculate the variational Gramians by using three length-sampling points (i.e. 0.7 m, 1.2 m, and 1.7 m), while the other simulation settings are the same as in the previous example. It is noted from Figure 5(a) that, the shorter the length is, the higher frequency the ripples emerge from, because delay value is proportional to the length of MTLs. It is also noted from Figure 6(a) that, the shorter the length is, the earlier the resonances emerge. From the frequency responses plotted in Figure 5 and transient responses plotted in Figure 6, it is shown that our algorithm can capture delay effects accurately.

5.2. TDSs with temperature variation

A TDS can be a temperature-dependent model resulting from MTL parameter variation. The dependencies of P.U.L. parameters (i.e. the capacitance $C$ and inductance $L$) of MTLs on temperature $T$ are modeled using second-order relations [15]

![Figure 4. Transient responses of Example 1 and its relative errors. (a) Transient responses. (b) Time-domain relative errors.](image1)

![Figure 5. Frequency responses of Example 1 with different length settings. (a) Frequency responses. (b) Frequency-domain relative errors.](image2)
\[ L(T) = L(T_0)(1 + x_1^L(T - T_0) + x_2^L(T - T_0)^2), \]  

\[ C(T) = C(T_0)(1 + x_1^C(T - T_0) + x_2^C(T - T_0)^2), \]  

where \( T_0 \) is the nominal value of temperature, \( x_1^C \) and \( x_2^C \) are the first- and second-order coefficients for \( C \), \( x_1^L \) and \( x_2^L \) are the first- and second-order coefficients for \( L \), respectively. It is noted that \( x_1^L \), \( x_2^L \), \( x_1^C \), and \( x_2^C \) are scalars, and the P.U.L. parameters \( L \) and \( C \) can be matrices whose dimensions are equal to the number of conductors excluding the reference one. In this experiment, the ROM is constructed at \( T_0 = 25^\circ C \), and the temperature varies from \(-15^\circ C \) to \( 80^\circ C \).

In this section, a TDS of order 688 is constructed by using a two-port linear interconnect network connected with 24 lossless three-conductor transmission lines, and this TDS has two different delay values in its state variable. We calculate the variational Gramians by using ten frequency-sampling points distributed logarithmically from \( 10^6 \) to \( 10^{12} \) Hz and five temperature-sampling points distributed linearly from \(-15^\circ C \) to \( 65^\circ C \). Given a threshold value of \( 7 \times 10^{-9} \), the HSV matrix is truncated at the significant singular-value drop, and the parameterized ROM is of order 20 which can accurately predict the system performance corresponding to various parameter variations. The frequency responses in Figure 7(a) and transient responses in Figure 8(a) are verified at three temperature points (i.e., \(-10^\circ C \), \( 40^\circ C \), and \( 80^\circ C \)). The frequency-domain relative errors and time-domain relative errors are further compared in Figures 7(b) and 8(b), respectively, showing that the proposed TBR provides good accuracy to capture the temperature dependence.

Figure 6. Transient responses of Example 1 with different length settings. (a) Transient responses. (b) Time-domain relative errors.

\[
L(T) = L(T_0)(1 + x_1^L(T - T_0) + x_2^L(T - T_0)^2),
\]

\[
C(T) = C(T_0)(1 + x_1^C(T - T_0) + x_2^C(T - T_0)^2),
\]
The proposed TBR uses about 24.1 s to construct the ROM, and transient simulation times for the original model and the ROM are about 12.8 and 0.27 s, respectively, obtaining about 50/C2 speedup. Note that the ROM construction time is comparable with the original simulation time, which means the proposed algorithm is quite efficient if the ROM needs to be reused repeatedly at various temperature points.

5.3. TDSs with geometric variation

A TDS can also have geometric variation. The variation of width and thickness of MTLs will change the values of P.U.L. parameters, making the whole TDS a variational model. The dependencies of the P.U.L. capacitance and inductance on various geometric parameters are very complicated, and there are no uniform formulations to calculate them explicitly [14,46]. Therefore, a parameterization step is needed to first fit the original TDS to a parameterized model with explicit dependencies on various parameters, and then we can construct the variational ROM. After using least-squares polynomial interpolation to fit the original geometrically dependent TDS to a second-order affine model (31), the parameterized TDS now has explicit dependencies on geometric parameters, and we call this step the parameterization step. In this example, we consider a single-delay TDS, and the P.U.L. parameters $L$ and $C$ are then scalars. It is shown in Figure 9 that the error of least-squares polynomial interpolation is quite small, making it viable to use the following second-order polynomial relations (42) to extract the geometric dependencies.

Figure 9. Extraction of the geometric dependencies for capacitance and inductance using second-order polynomial interpolation. Comparisons with original values are made corresponding to width = 5, 7.5, 10, 12.5, 15 mils and thickness = 20, 25, 30, 35, 40, 45, 50 mils, respectively. (1 mil = 0.001 inch) (a) P.U.L. capacitance. (b) P.U.L. inductance.
\[ L(w, t) = L(w_0, t_0) \cdot (1 + x_{w1} \Delta w + x_{w2} \Delta w^2 + x_{t1} \Delta t + x_{t2} \Delta t^2 + x_{wt} \Delta w \Delta t), \]  
\[ C(w, t) = C(w_0, t_0) \cdot (1 + x_{w1} C \Delta w + x_{w2} C \Delta w^2 + x_{t1} C \Delta t + x_{t2} C \Delta t^2 + x_{wt} C \Delta w \Delta t), \]  
(42a)  
(42b)

It is noted that all the coefficients \( x \) are scalars. Figures 6(a) and (b) compare the original values with interpolation values for the P.U.L. capacitance and inductance, respectively.

In this section, we use a nine-port RLC network connected with 40 lossless two-conductor transmission lines to construct a TDS of order 1740, and this TDS has one constant delay value in its state variable. Given a threshold value of \( 5 \times 10^{-12} \), the HSV matrix is truncated at the significant singular-value drop, and an ROM of order 50 is constructed with parameter variations in width and thickness. The parameterized ROM is constructed at the nominal point of width = 10 mils and thickness = 35 mils (1 mil = 0.001 inch), and the variational Gramians are calculated by using ten frequency-sampling points distributed logarithmically from \( 10^6 \) to \( 10^{12} \) Hz, three width-sampling points distributed linearly from 7.5 mils to 12.5 mils and three thickness-sampling points distributed linearly from 25 mils to 45 mils. Figure 10(a) and Figure 11(a) respectively show the frequency responses and impulse responses corresponding to various values of width while the thickness is set.

![Figure 10](image1.png)

**Figure 10.** Frequency responses of transfer function \( H(1,1) \) in Example 3 corresponding to thickness = 20 mils with different width settings. (a) Frequency responses. Data 1, data 3, and data 5 are curves of original systems with width = 5, 10, 15 mils, respectively. Data 2, data 4, and data 6 are curves of reduced systems with width = 5, 10, 15 mils, respectively. (b) Frequency-domain relative errors.

![Figure 11](image2.png)

**Figure 11.** Time-domain responses of output \( y(1) \) in Example 3 corresponding to thickness = 20 mils with different width settings. (a) Time-domain responses. Data 1, data 3, and data 5 are curves of original systems with width = 5, 10, 15 mils, respectively. Data 2, data 4, and data 6 are curves of reduced systems with width = 5, 10, 15 mils, respectively. (b) Time-domain relative errors.
to be a corner value of 20 mils. Setting the thickness to be another corner value of 50 mils (note that this corner value is different from the sampling points 25 mils, 35 mils, and 45 mils), the frequency responses and impulse responses corresponding to various values of width are plotted again in Figures 12(a) and 13(a), respectively. It is noted that the variations of frequency responses are not very large. However, from the results of transient simulation, it is shown that geometric variations can have significant impact on TDS time-domain responses. The frequency-domain and time-domain relative errors of the parameterized ROM corresponding to various parameter settings are also plotted in Figures 10(b), 11(b), 12(b), and 13(b), showing that the proposed TBR provides a good accuracy to capture the geometric dependencies.

The construction time of the parameterized ROM is about 180 s, and the transient simulation times for the original model and ROM are about 40 and 0.55 s, respectively, obtaining about 70\times speedup. Note that the ROM construction time is comparable with the simulation time, which means the proposed algorithm can be quite efficient if the ROM needs to be reused repeatedly at various parameter settings.

5.4. Efficiency analysis and speed-up performance

We first show the CPU time of the most computationally expensive steps of Algorithm 1 in Table I. From the results, we can see that the most computationally expensive step is to collect the TDS

Figure 12. Frequency responses of transfer function $H(1,1)$ in Example 3 corresponding to thickness = 50 mils with different width settings. (a) Frequency responses. Data 1, data 3, and data 5 are curves of original systems with width = 5, 10, 15 mils, respectively. Data 2, data 4, and data 6 are curves of reduced systems with width = 5, 10, 15 mils, respectively. (b) Frequency-domain relative errors.

Figure 13. Time-domain responses of output $y(1)$ in Example 3 corresponding to thickness = 50 mils with different width settings. (a) Time-domain responses. Data 1, data 3, and data 5 are curves of original systems with width = 5, 10, 15 mils, respectively. Data 2, data 4, and data 6 are curves of reduced systems with width = 5, 10, 15 mils, respectively. (b) Time-domain relative errors.
variational Gramians. The cost will become high if too many sampling points are used to compute the variational Gramians. The cost of Cholesky factorization, SVD of cross factors, and computing the projection matrices will grow quickly if system dimension becomes too high.

For the CPU time of parameterization procedure, we may not be able to provide such data because in our experiments, this procedure is not fully automatic and requires manual intervention (e.g. we need to collect raw sampling data of the P.U.L. parameters from simulations or measurements). Even in Example 2 which involves temperature variations, the parameterization of P.U.L. parameters on temperature is provided directly by [15]; therefore, we may not be able to provide its CPU time. However, it does not affect the efficiency analysis for our PMOR algorithm. The reason is that, whether MOR is performed or not, it is always a necessary step to collect the inputs (in Algorithm 1) to be parameterized if one wants to do variational analysis. Therefore, this issue would not be an efficiency problem for our algorithm.

Table II summarizes the speed-up performance of these three examples. The MOR times of the proposed variational TBR are compared with transient simulation times of the original and reduced models. It is shown that, although the variational MOR takes more (but still comparable) time to construct the ROMs than their original simulation, a very large speed-up is obtained for simulation of ROMs. Because at the early stage of VLSI design when designers need to perform many worst-case simulations in order to obtain optimal system performance, it is very time-consuming to repeatedly simulate the original system corresponding to various parameter settings. Therefore, if we perform variational MOR to obtain parameterized ROMs of much smaller size, the repeated simulations will save a lot of time. Note that the parameterized ROMs are constructed only once and they can obtain about 50–70× speed-up in our experiments; the proposed variational TBR for TDSs is therefore an efficient scheme for variational analysis.

It is noted that we do not provide examples with 10000 states or beyond in this paper. The first reason is the computational cost for MOR process is very high partly because the Cholesky factorization and SVD take a lot of time when the order \( n \) is large. The matrix factorization is one of the bottlenecks of balance truncation-based MOR. Similar to other BT methods, TDS BT can provide higher accuracy but cannot deal with extremely large-scale systems as moment-matching methods do. However, BT methods have recently been made suitable for large-scale systems [47, 48]. Second, high computational cost means more time in MOR process. When \( n \) is too large, the time for MOR process is much higher than the simulation time for the original system. See Table II and example 3, the computational time for MOR is 180 s, which is already four times larger than the simulation time of 40 s. Third, order near 2000 is the highest order one can see in other TDS papers in the literature (please see references [15, 23] for TDS model examples).

| Table II. Speed-up performance and MOR times of the proposed variational TBR (s). |
|----------------------------------|-----------------|-----------------|-----------------|-----------------|
| Example 1 \( \text{Order 688} \) | 14.2 \( \text{12.6 (Order 688)} \) | 0.26 \( \text{0.27 (Order 20)} \) | 48× |
| Example 2 \( \text{Order 688} \) | 24.1 \( \text{12.8 (Order 688)} \) | 0.27 \( \text{0.55 (Order 50)} \) | 47× |
| Example 3 \( \text{Order 1740} \) | 180 \( \text{40 (Order 1740)} \) | 0.55 \( \text{0.55 (Order 50)} \) | 73× |

6. CONCLUSIONS

This paper has presented a parameterized TBR-type MOR scheme for TDSs with design or process parameter variations. To the best of our knowledge, in the literature, there is little work to derive variational ROMs for TDSs via a Gramian-based approach. The experimental results have shown that the proposed algorithm provides good accuracy and efficiency to capture the variational effects of TDSs, making it viable for efficient simulation and TDS design optimization.

ACKNOWLEDGEMENT

This work is supported in part by the Hong Kong Research Grant Council under project 718509E, and in part by the University Research Committee of The University of Hong Kong. This work is also supported by Small Project Funding of HKU from HKU SPACE Research Fund under grant 201007176165.

REFERENCES


