Zonotope-based Nonlinear Model Order Reduction for Fast Performance Bound Analysis of Analog Circuits with Multiple-interval-valued Parameter Variations

Yang Song, Sai Manoj P. D. and Hao Yu
School of Electrical and Electronic Engineering
Nanyang Technological University, Singapore 639798
haoyu@ntu.edu.sg

Abstract—It is challenging to efficiently evaluate performance bound of high-precision analog circuits with multiple parameter variations at nano-scale. In this paper, a nonlinear model order reduction is proposed to deploy zonotope-based model for multiple-interval-valued parameter variations. As such, one can have a zonotope-based reachability analysis to generate a set of trajectories with performance bound defined. By further constructing local parameterized subspaces to approximate a number of zonotopes along the set of trajectories, one can perform nonlinear model order reduction to generate the performance bound under parameter variations. As shown by numerical experiments, the zonotope-based nonlinear macromodeling by order of 19 achieves up to 500× speedup when compared to Monte Carlo simulations of the original model; and up to 50% smaller error when compared to previous parameterized nonlinear macromodeling under the same order.

I. INTRODUCTION

At nano-scale, nonlinearity of analog circuits and amount of parasitic have become more prominent than ever. For example, I/O circuits with interconnects are strongly susceptible to parameter variations with performance bound [1], [2], [3], [4], [5], [6], [7] of skew to be determined. The calculation of performance bound is expensive with consideration of multi-dimensional parameter variations, because it involves both large-scale parasitics and nonlinear I/O buffers. Model order reduction (MOR) [8], [9], [10] can reduce complexity of state space by subspace-based approximation. The generation of subspace can be obtained by Krylov iteration or truncation-balanced realization. However, the effectiveness of subspace generation is limited when considering the strong nonlinearity and parameter variations for analog circuits such as I/Os.

The nonlinearity can be addressed during MOR [11], [12], [13], [14] by constructing a series of local subspaces obtained at different operation points of one trajectory, which can be further aggregated to provide one global subspace for approximation. The challenge is to further consider parameter variations since the number of local expanded subspaces by moments can grow substantially [11] when considering all types of parameter variations. In [15], parameter variations are included during MOR with the use of interval-values. Correlation-decoupled parameter variations in a statistical range are modeled by interval values. For example, \( \Delta x_1, \Delta x_2 \) model uncertainties in two state variables of state vector \( x \) with \( c \) as interval center, the neighboring points for variation regions can be modeled as:

\[
x = c + [-1, 1] \Delta x_1 + [-1, 1] \Delta x_2 \text{ in one dimension line. However, there is no method developed to deal with multi-dimensional interval-value problem. What is more, it is unknown how to perform a nonlinear MOR with multi-dimensional interval-values that model parameter variations of analog circuits.}

Zonotope based over-approximation have been widely deployed in reachability analysis for verification of system dynamics by exploring potential trajectories of operating points in state space [16], [17]. It can conveniently provide predicted boundary of multiple trajectories under uncertain interval parameters by one-time computation, in contrast to simulate different trajectories one by one. The zonotope based reachability analysis has been deployed for a number of hard analog circuit verifications [18]. What is more, implicit integration is further developed in [19] for a fast numerical estimation of zonotope with consideration of nonlinearity.

In this work, zonotope based over-approximation has been applied to describe multi-dimensional interval-values for parameter variations, which are included into interval state matrices. By running reachability analysis with interval state matrices, a zonotope-expanded state space can be constructed to enclose a set of trajectories, each used to describe one type of combination of parameter variations. By further constructing local Krylov subspaces in term of interval state matrices at a number of zonotopes along the set of trajectories, one can construct global subspaces to approximate the original nonlinear analog circuits with consideration of multiple parameter variations. Note that in the previous parameterized nonlinear macromodeling [11], [12], [13], [14], a weighting sphere is introduced to achieve an approximated subspace based on the neighboring points. The approximation is a weighted combination in which the weights are calculated by Kernel distance. It is not well defined how to build weights in presence of multiple parameter variations. In contrast, the parameterized subspace is well covered by the zonotope in the proposed method, which results in higher accuracy as shown by experiments. Experimental results have shown that the proposed method achieves up to 500× speedup when compared to Monte Carlo simulation with 1000 samples by 1% error; and up to 50% smaller error when compared to previous parameterized nonlinear macromodeling [11], [12], [13], [14] under the same order.

The rest of this paper is organized as follows. Section II reviews the nonlinear MOR. Section III introduces basic principles of zonotope and reachability analysis. Further, the zonotope-based nonlinear model order reduction considering multiple parameter variations is proposed. The proposed method is validated by experiments in Section IV. Conclusions are drawn in Section V.

II. NONLINEAR MODEL ORDER REDUCTION

Model order reduction (MOR) reduces complexity by building subspace to approximate the original full state space. Unlike well-developed linear MOR [8], [9], one problem of nonlinear MOR for
The dynamics of a nonlinear system can be described by differential algebraic equation (DAE) as
\[
\frac{d}{dt} q(x(t)) + f(x(t)) + Bu(t) = 0
\] (1)
in which \( x(t) \in \mathbb{R}^n \) is the state variable vector and \( u(t) \) is the input vector.

\[ V^T \frac{d}{dt} q(Vz) + f(Vz) + Bu = 0. \] (2)

To derive the projection matrix \( V \), (1) needs to be linearized first as below
\[
C \frac{dx(t)}{dt} + Gx(t) = Bu(t) = b
\] (3)
Here, \( x^* \) is the operating point at which the linearization is performed, \( C \) is a non singular linearized capacitance matrix, \( G \) is linearized conductance matrix, and the right hand side vector \( b \) contains both the input vector \( u(t) \) and the linearization residue of \( f(x) \).

At the \( j \)th local operating points \( x_j \), the Krylov subspace with an order of \( p \) can be constructed by
\[
K(\lambda_j A_j, r_j, p) = \text{colsp}(r_j, A_j r_j, A_j^2 r_j, ..., A_j^{p-1} r_j)
\] (4)
where \( A_j = -G_j^{-1} C_j \) and \( r_j = -G_j^{-1} b_j \). By orthogonalizing base vectors for Krylov subspace, one can obtain the local subspace for projection matrix \( V_j \). To approximate a nonlinear system, one needs to assemble those \( V_j \) together for an aggregated global subspace.

In [14] the subspace of the nonlinear system is constructed by two types of manifolds in the state space. As shown in Fig.1, by simulating (DC or TRAN) the full system with the training inputs, a series of sample points are first generated and scattered in the state space as a 1D manifold, called the DC-manifold. Second, the nonlinear system is linearized and further reduced at each sampled points. The spanned subspace is the so-called AC-manifold.

### III. Zonotope-based Nonlinear Model Order Reduction

#### A. Zonotope for Multiple-interval-valued Parameter Variation

Zonotope-based reachability analysis [16], [17] can efficiently determine a reachable region that one dynamic system can evolve with uncertain states as shown in Fig.2. As such, if we model the multiple-interval-valued parameter variations to the state \( x \) by a zonotope or a polytope, one can perform reachability analysis to obtain the performance bound of analog circuits. In this paper, we further show how nonlinear model order reduction can be introduced during the zonotope-based reachability analysis.
enclosing polytope. One simple and symmetrical type of polytope, called \textit{zonotope} [16] is defined as follows.
\[
Z = \{ x \in \mathbb{R}^{n \times 1} : x = c + \sum_{i=1}^{q} [-1, 1] g^{(i)} \}
\]
where \( c \in \mathbb{R}^{n \times 1} \) is the zonotope center; \( q \) represents number of zonotope generators; and \( g^{(i)} \in \mathbb{R}^{n \times 1} \) is a zonotope generator. A zonotope spans a polytope in the space state to include multiple-interval-valued state variables.

The interval-valued state variable can come from the uncertainty of interval-valued parameter variation, which contains a constant nominal value and a variation range. Such uncertain parameters with variations can be modeled as zonotopes in the parameter space, which is the first step of the proposed zonotope-based nonlinear model order reduction. This is achieved as follows. A state matrix with uncertain entries can be described in the form of matrix zonotopes by
\[
M = \{ M \in \mathbb{R}^{n \times n} : M = M^{(0)} + \sum_{i=1}^{q} [-1, 1] M^{(i)} \}
\]
which can be conveniently transformed into an interval matrix by \([M^{(0)} = \sum [M^{(i)}], M^{(0)} + \sum [M^{(i)}]]\). Similar to zonotopes, the matrix \( M^{(0)} \) is called the \textit{center matrix} and the matrix \( M^{(i)} \) is called the \textit{generator matrix}. Addition and multiplication rules for zonotopes and matrix zonotopes are similarly defined in [17].

\( M^{(0)} \) can be the linearized state matrix for the nominal operating point; and \( M^{(i)} \) contains variation ranges due to multiple parameter variations. Uncorrelated parameter variations after decoupling are filled in different generator matrices. As a result, the interval values of multiple parameter variations can be considered in the zonotope matrix for circuit simulation, including the use in nonlinear model order reduction.

In the following, zonotope-based nonlinear model order reduction with multiple parameter variations is discussed first. Performance bound analysis for the reduced system will be illustrated at last.

\[ \text{B. Nonlinear Model Order Reduction with Zonotopes} \]

To account for the impact by parameter variations, the linearized system equation (3) needs to be formulated by zonotopes as below
\[
\mathcal{C}' \frac{dX(t)}{dt} + \mathcal{G}X(t) = Bu(t)
\]
where \( X \) is the state variable zonotope, and \( \mathcal{G} \) and \( \mathcal{C} \) are the zonotope matrices of \( G \) and \( C \) respectively as discussed in (9).

The construction of zonotope conductance matrix \( \mathcal{G} \) is straightforward for linear devices. For example, the zonotope conductance of a multiple-interval-valued resistor is expressed as
\[
g = g_o + \sum_{i=1}^{q} [-1, 1] \Delta g^{(i)},
\]
which includes the nominal value and the range of parameter variations. \( \mathcal{G} \) is built up by assembling interval conductance stamps.

As for nonlinear devices such as MOSFETs by BSIM models, the construction of state matrix needs one more step. Suppose that one transistor width \( W \) has parameter variation, the variation of its transconductance satisfies
\[
\Delta g_m = \frac{\partial g_m}{\partial W} \Delta W.
\]
Here, \( \frac{\partial g_m}{\partial W} \) needs to be computed at one nominal operating point. Other interval conductances including \( g_{ds} \) and \( g_{mb} \) can be derived in the same fashion. Inclusion of parameter variations can be performed below
\[
\Delta G = \begin{pmatrix}
\vdots \\
\frac{\partial g_m}{\partial W} & -\frac{\partial g_m}{\partial W} \\
-\frac{\partial g_m}{\partial W} & \frac{\partial g_m}{\partial W}
\end{pmatrix} \Delta W.
\]

After the above uncertain system equation is formulated with parameter variations modeled by zonotope matrix with multiple-interval-values by
\[
\mathcal{G} \in [\mathcal{G}^{(0)} - \sum_i |\mathcal{G}^{(i)}|, \mathcal{G}^{(0)} + \sum_i |\mathcal{G}^{(i)}|].
\]

One can build the local subspace by the Krylov method in the next step. In addition, note that the correlated zonotope generators can be merged into the same zonotope generator matrix. To consider multiple-interval-valued parameter variations in zonotopes, the base vectors in the Krylov subspace are formulated as zonotopes too by
\[
K_r(A, R, p) = \cos p(R, A^t R, A^t R^2, ..., A^t R^{p-1})
\]
where \( A = -[\mathcal{G}]^{-1} \mathcal{C} \) and \( R = -[\mathcal{G}]^{-1} \mathcal{B} \). As such, the zonotope vectors in (13) form the parameterized Krylov subspace.

What is more, QR decomposition is necessary to orthogonalize the base vectors, but the interval-valued \( A^t R \) cannot be handled directly by the conventional QR decomposition routine. In this paper, we orthogonalize the interval-valued matrices based on the center matrix as below
\[
Q^{(i)} = M^{(i)}((Q^{(0)})^T M^{(0)})^{-1}
\]
where the center matrix \( M^{(0)} \) is orthogonalized to \( Q^{(0)} \), and the zonotope generator matrix \( M^{(i)} \) is orthogonalized to \( Q^{(i)} \).

As such, the orthogonalized parameterized subspace can be obtained by
\[
\mathcal{V} = \{ V \in \mathbb{R}^{n \times n} : V = Q^{(0)} + \sum_{i=1}^{q} [-1, 1] Q^{(i)} \}.
\]

The construction of parameterized subspace is performed at each sample point of DC-manifold, and is stored to produce the compact macromodel in the following to provide a fast performance bound analysis.

\[ \text{C. Parameterized Macromodel for Performance Bound} \]

As a result, one can obtain the performance bound efficiently based on the order-reduced macromodel. The state variable \( z \) in (7) is now replaced by the zonotope \( Z \) with a nominal center \( z^{(0)} \) and a series of generators \( z^{(i)} \) caused by multiple parameter variations. The zonotope-based DAE is shown below based on (7) and (15) in the reduced state space
\[
V^T CV \frac{dZ}{dt} + F(Z) + V^T Bu = 0.
\]

With obtained parameterized subspaces at the \( j \)th sample point, the projection of \( f(z) \) from the original space to the AC-manifold in (6) becomes
\[
F(Z) = V_j^T f(x_j) + V_j^T G_j v_j (Z - z_j).
\]
Here, the multiplication of three zonotope matrices are used to evaluate the interval function \( F(Z) \). Note that higher-order of variation
products are discarded since their contribution is quite small compared with first order ones:

$$V_j^T G_j V_j = (V_j + \Delta V_j)^T (G_j + \Delta G_j)(V_j + \Delta V_j)$$

$$\approx V_j^T G_j V_j + \Delta V_j^T G_j V_j + V_j^T \Delta G_j V_j + V_j^T G_j \Delta V_j$$

(18)

where the variations $\Delta G_j$, $\Delta V_j$ refer to the sum of generators in $G_j$ and $V_j$. Note that the sample point $x_j$ in (17) is selected based on the Euclidean distance from the operating point. The center of zonotope $Z$ of state vector is used to calculate the Euclidean distance.

To solve the zonotope-based DAE in (16), implicit Euler method is applied with discretized time-step $h$ at $k$th time-step by

$$V^T C V Z_k - Z_{k-1} + F(Z_k) + V^T B u = 0.$$  

(19)

By substituting $F(Z_k)$ with (17), one can obtain

$$\left( \frac{V^T C V}{h} + V^T G V \right) Z_k = \frac{V^T C V Z_{k-1}}{h} - V^T f(x_j) + V^T G V z_j - V^T B u.$$  

(20)

Here $V$ has same dimension for all intervals. Performance bound is obtained by transforming the final state set/zonotope into an interval as in (12).

Note that the full multiplication between a zonotope matrix and a zonotope leads to an increased number of generators. The Minkowski summation rule [17] is used to merge zonotopes while preserving new generators created during multiplication

$$Z_k = A^{-1} \left( \frac{V^T C V Z_{k-1}}{h} \oplus -V^T f(x_j) \oplus V^T G V z_j \oplus -V^T B u \right),$$  

(21)

where $A = \frac{V^T C V}{h} + V^T G V$.

Moreover, there is no real inverse of zonotope matrix $A = (A^{(0)}, \ldots, A^{(i)}, \ldots)$ in (21) but by the two-step below. The first step is the approximated expansion of $A^{-1}$ by

$$A^{-1} = ((A^{(0)})^{-1}, \ldots, (A^{(0)})^{-1}A^{(i)}(A^{(0)})^{-1}, \ldots).$$  

(22)

The second step is to calculate $(A^{(0)})^{-1}$ by LU decomposition.

$$(A^{(0)})^{-1} = U^{-1} L^{-1} P^T I$$  

(23)

where $I$ is the identity matrix and $P$ is the permutation matrix. This approach enables a cost-effective numerical implementation of zonotope-based circuit analysis similar to a SPICE-like simulator.

IV. EXPERIMENTAL RESULTS

The proposed zonotope-based nonlinear model order reduction (ZMOR) is implemented in MATLAB on the basis of a SPICE-like simulator. As such, one can have efficient performance bound analysis with multiple-interval-valued parameter variations by the reduced macromodel. Manipulations of zonotopes are performed by a MATLAB toolbox named Multi-Parametric Toolbox (MPT) [21].

Experiment data is collected on a desktop with Intel Core i5 3.2GHz processor and 8GB memory. 40hm is used as the technological node for the test circuits: RC-interconnect with inverter-buffer chain; and transmission-line (T-line) with current-mode-logic (CML) buffers.

A. Inverter-buffer Chain with RC-interconnect

The first test-case is a CMOS inverter-buffer chain containing four stages as shown in Fig.3. Each resistor has an independent variation of 10%; and transistors share a local variation of 5% on their widths. The inverter chain is simulated for 40ns with a square-wave signal input active between 5ns and 25ns. The reduced macromodel of the inverter-buffer chain with variations is generated by the proposed ZMOR. For comparison, the nonlinear MOR [14] is also deployed. The performance bound evaluation by reachability analysis is undertaken afterwards for both models. To verify the accuracy of reduced macromodels and the Monte Carlo of the full model are performed considering all parameter variations. The performance bound is defined as the spatial variation of voltage-waveform difference (skew) in the state space, where the horizontal axis $v_1$ represents the voltage on the gate of the third inverter and the vertical axis $v_2$ is for the voltage on the gate of the fourth inverter.

Performance bound analysis of the reduced macromodels and the Monte Carlo of the full model is shown in Fig.4 and Fig.5. Both models are reduced to eighth order ($p = 8$). In Fig.4, reachability analysis of the reduced macromodel by ZMOR is shown by white blocks with pink envelopes. Each of the blocks is a zonotope for the potential operating points at one time step. The blue curves are generated by the Monte Carlo. One can observe that the trajectories move towards the bottom-right corner under the input $V_{in} = 1$ starting from the top-left corner. After the input signal flips to zero, the trajectories move back towards the initial state. What is more, one can observe that the reduced macromodel with variations manages to fit the trajectories of the full model by the Monte Carlo. There is small error that happens when over-approximation increases as $\Delta V$ reaches about 0.6$V$. As shown in Fig.4, more than 90% of the operating points at the final instant by the Monte Carlo settle within the final set of reachability analysis based on the proposed model. In contrast, the result of the reduced macromodel by [14] is shown in Fig.5. One can observe that due to the incapability of handling parameter variations, only 40% of operating points of the Monte Carlo at the final time-instant are contained in the final zonotope, which is nearly 50% lower than what is achieved by the proposed model.

Furthermore, the macromodel is obtained with process variation with 15% independent variations of resistors and 10% local variations on transistor widths. Performance summary for different conditions is listed in Table I in which the Monte Carlo is performed with 1000 samples, variation gives the variation of local and global parameters and the proposed ZMOR refers to the zonotope-based MOR, min and max stand for the lower and upper bounds of the voltage-waveform difference at the final state. One can observe that the difference of the boundaries compared with the Monte Carlo can be limited within 1% for $p = 8$. Up to 500x of speedup can be achieved by the proposed method.

B. CML Buffer with Transmission Line

Next, we consider a CMOS current-mode-logic (CML) buffer with transmission line as shown in Fig.6. The reduced macromodel
Performance bound analysis by the reduced macromodels and

with variations is generated by the proposed method with comparison of [14]. The accuracy of the reduced model is validated by comparing with the Monte Carlo of the full model considering all parameter variations. Each resistor in the transmission line has an independent variation of 10%. The amount of variations in transistor widths are set as 10%. The number of channels can be increased. The performance bound is defined as the spatial variation of voltage-waveform difference (skew) of the CML buffer at different channels.

Performance bound analysis by the reduced macromodels and

by the Monte Carlo of the full model is shown in Fig.7 and Fig.8. Both the models are reduced to seventh order \((p = 7)\). In Fig.7, reachability analysis by the ZMOR with consideration of interval-valued parameter variations is shown by white blocks with pink envelope. Each zonotope is a polytope (zoomed-in) that can be clearly observed at the point of over-approximation in the beginning and in the close-up of the final state. The blue curves are generated by the Monte Carlo with parameter variations. Over-approximation is observed at the initial points of the trajectory. The proposed method can include the trajectories of the full model within the zonotopes most of the time as shown in Fig.7. In contrast, as shown in Fig.8, only 40% of the operating points are included when using the macromodel by [14].

Further, we show performance comparison by increasing the number of channels. Performance summary for different number of channels is listed in Table II in which the Monte Carlo is performed with 1000 samples, ZMOR refers to the reachability-based model order reduction, \( \text{min} \) and \( \text{max} \) stand for the lower and upper bounds of the voltage-waveform difference at the final state. \( \text{Channel} \) represents the number of channels. As shown, the order of reduced model varies with the number of channels. The difference compared to the Monte Carlo can be limited by increasing the order at the cost of extra runtime. For 8-channel of order 16, up to 400 \( \times \) of speedup can be achieved by the proposed method. The difference compared to the Monte Carlo is nearly 10%. One can observe that increase in speedup due to reduction in order but at the cost of accuracy.

Lastly, Fig.9 shows the transient waveforms by ZMOR and the full model. Transient waveforms obtained at two outputs of buffers by ZMOR of tenth orders \((p = 10)\) and full order Monte Carlo are shown in pink and blue color curves respectively. Transient waveforms of the full model are nearly overlapped with the ZMOR of order 10 \((p = 10)\)
can generate a set of trajectories with formed performance bound, with less than 1% error. There is a slight difference between the two curves shown in the close-up. The performance bounds with different orders are shown in equal column of Table I and II.

V. CONCLUSIONS

In this paper, the zonotope-based nonlinear model order reduction is developed for nonlinear analog circuits such as I/Os with multiple parameter variations. By zonotope-based reachability analysis, one can generate a set of trajectories with formed performance bound, and each trajectory is associated with one combination of parameter variations. The parameterized subspaces for model order reduction can be thereby constructed by a number zonotopes along the set of trajectories with well-defined weights. The developed method can be conveniently applied to evaluate the performance bound for large-scale nonlinear analog circuits with multiple parameter variations. As shown by numerical experiments, the zonotope-based nonlinear macromodeling by order of 19 achieves up to 500× speedup when compared to Monte Carlo simulations of the original model; and up to 50% smaller error when compared to previous parameterized nonlinear macromodeling under the same order.

REFERENCES