Automatic Generation of Compact Electro-Thermal Models for Semiconductor Devices

Tamara Bechtold, Evgenii B. Rudnyi and Jan G. Korvink

IMTEK, University of Freiburg, Georges-Köhler-Allee 103, 79110 Freiburg, Germany

SUMMARY

A high power dissipation density in today's miniature electronic/mechanical systems makes on-chip thermal management very important. In order to achieve quick to evaluate, yet accurate electro-thermal models, needed for the thermal management of microsystems, a model order reduction is necessary. In this paper, we present an automatic, Krylov-subspace-based order reduction of a electro-thermal model, which we illustrate by a novel type of micropropulsion device. Numerical simulation results of the full finite element model and the reduced order model, that describes the transient electro-thermal behavior, are presented. A comparison between Krylov-subspace-based order reduction, order reduction using control theoretical approaches and commercially available reduced order modeling has been performed. A Single-Input-Single-Output setup for the Arnoldi reduction algorithm was proved to be sufficient to accurately represent the complete time-dependent temperature distribution of the device.

key words: compact electro-thermal model, order reduction, Krylov-subspace, Arnoldi process

1. Introduction

The modeling of electro-thermal processes becomes increasingly important during semiconductor device development. For example, with the decreasing size and growing complexity of micro-electronic and micro-electro-mechanic (MEMS) systems, the power dissipation of integrated circuits has become a critical concern. Due to the very high clock frequencies, the power consumption of a modern microprocessor can exceed 50W. The thermal influence upon the device caused by each transistor's self-heating and the thermal interaction with tightly placed neighboring devices cannot be neglected, because excessive temperatures may cause the malfunction of the device or even destroy it. Therefore, it is necessary to develop a electro-thermal model which computes the dependence between power dissipation and temperature distribution over the device. Moreover, such heat transfer analysis needs to be done quickly in response to every design alteration. The model must also provide good accuracy in order to return precise temperature values.

The solution of thermal issues in micro-electronic and MEMS devices is usually performed by extending the electrical model with the thermal part [1],[2], which computes the precise temperature distribution of the microsystem by using either finite element (FE), finite difference or some other spatial discretization method. Such an accurate computation is time-consuming, for the resulting ordinary differential equation (ODE) systems can easily exceed the order of 100 000. Even with the increasing speed of modern computers it is not possible to perform repeated simulations in an acceptable amount of time without simplification or model reduction.

Conventionally, the reduction of electro-thermal models for micro-electronic and MEMS devices is performed through a lumped-element decomposition of the model followed by parameter fitting [3],[4],[5]. Such a non-automatic approach requires the designer to choose the corect reduced model structure without strict guidelines, and to perform a time-consuming parametrization including one or more simulations of the full-scale model. Moreover, sufficient accuracy is provided only for a limited parametric domain.

In order to achieve both efficiency and accuracy in the thermal management of microsystems, we propose an automatic order reduction approach. It is suitable for linear electro-thermal models, and is based on an Arnoldi algorithm [6]. Section 2 contains an overview of the different mathematical order reduction methods. Control theoretical approaches [7], Krylov-subspace based approaches [8] and a commercially available reduced order modeling [9] are discussed. Section 3 describes our MEMS actuator test case: a solid fuel microthruster. The numerical simulation results of the full finite element model and the different reduced order models of the device that describe its transient electro-thermal behavior are presented and discussed.

2. Model Order Reduction

The temperature distribution within a microsystem, T, is the solution of the 3-dimensional heat transfer equation:

$$\nabla \bullet (\kappa \nabla T) + Q - \rho C_p \frac{\partial T}{\partial t} = 0$$
⁽¹⁾

where κ is the thermal conductivity, C_p is the specific heat capacity, ρ is the mass density, T is the temperature distribution and Q is the heat generation rate. Since the temperature changes within the microdevice are expected to be relatively small, the assumption that κ and C_p are temperature independent is often acceptable. Becouse the Joule heating is a dominant heating mechanism, the heat

generation can be replaced through:

$$Q = \frac{j^2}{\sigma} \tag{2}$$

where \boldsymbol{j} is the spatially varying electric current density vector and $\boldsymbol{\sigma}$ is the specific electric conductivity. Assuming that the heat generation Q is uniformly distributed within the heating area, the spatial discretization of the governing Eqs (1) and (2) leads to linear ODE system of the form:

$$\begin{bmatrix} C \end{bmatrix} \dot{\mathbf{T}} + \begin{bmatrix} K \end{bmatrix} \mathbf{T} = \mathbf{F} I(t)^2 R$$

$$y = E^T \cdot T$$
(3)

where $[K], [C] \in \mathbb{R}^{n \times n}$ are the global heat conductivity and heat capacity matrix, $T(t), F, E \in \mathbb{R}^{n}$ are the temperature (state), the load and the output vector respectively and *n* is the dimension of the system. The electric current I(t) through the heater with the total electric resistance *R* is the input to the system. The introduction of the output vector (matrix) *E* allows the user to obtain the temperature values only in predetermined points.

The above number of equations n is, employing an accurate discretization, usually too large for an efficient electro-thermal simulation at the system level. Hence, in order to use behavioral simulators such as SABER or ELDO, a reduced order model of the form:

$$\begin{bmatrix} C_r \end{bmatrix} \dot{\boldsymbol{T}}_r + \begin{bmatrix} K_r \end{bmatrix} \boldsymbol{T}_r = \boldsymbol{F}_r \boldsymbol{I}(t)^2 \boldsymbol{R}$$

$$y_r = E_r^T \cdot \boldsymbol{T}_r$$
(4)

with the dimension $r \ll n$ has to be generated.

Eqs (3) and (4), as written above, represent a Single-Input-Single-Output (SISO) system. The present work considers a special case where E is the identity matrix, that is, y = T, which we call a Single-Input-*Complete*-Output (SICO) system.

2.1 Model Order reduction via Balanced Truncation Approximation

Control theory already has a number of well established tools for the automatic model reduction of stable linear systems [7]. Each linear dynamic system (3) of order *n* has *n* so-called Hankel singular values σ_i which can be computed by solving the Lyapunov equations:

$$AP + PA^{T} = -bb^{T}$$

$$A^{T}Q + QA = -EE^{T}$$
(5)

were $A = -[K]^{-1}[C]$, and $b = -[K]^{-1}F$, for the controllability grammian P and the observability grammian Q. Hereby P is to be seen as the connection between the input function and the state (temperature) vector, and Q as the connection between the state vector and the system output. They determine which states can be controlled and which ones can be observed. The Hankel singular values of the original dynamic system (3) are equal to the square root of the eigenvalues of the product of P and Q:

$$\sigma_i = \sqrt{\lambda_i (P \cdot Q)}, i = 1, ..., n$$
(6)

Once these values are known, there are a number of model reduction methods with guaranteed error bounds for the difference between the transfer function of an original n-dimensional system and its reduced r-dimensional system, as follows:

$$\|G - G\|_{\infty} \le 2(\sigma_{r+1} + \dots + \sigma_n) \tag{7}$$

provided that the Hankel singular values have been sorted in descending order.

The basic idea behind order reduction using a balanced truncation approximation (BTA) after Moore [10] is to transform the state vector in such a way that

$$P = Q = diag\{\sigma_1, ..., \sigma_n\}$$
(8)

so that all n-r states which are simultaneously weakly controllable and weakly observable can be excluded. Practically, this excludes those equations from the system (3) which correspond to the last (and hence smallest) n-r Hankel singular values. The guaranteed error bounds for the difference between the transfer function of original high-dimensional and reduced low-dimensional systems, Eq (7), means that model reduction based on these methods can be made fully automatic. A user just sets an error bound and then the algorithm finds the smallest possible dimension of the reduced system, which satisfies that bound. Alternatively, a user specifies the dimension of the reduced system and the algorithm estimates the error bound for the reduced system.

However, the time required to solve the Lyapunov equations (5), as well as to perform a singular value decomposition, grows as the cubic power of the number of equations $O(n^3)$. Hence, if the system order increases twice, the time required to solve a new problem will increase about eight times. Hence, for computational reasons, these order reduction methods are limited to relatively small systems, typically with state vector dimensions in the range of 1000 to 2000.

2.2 Model Order Reduction via Arnoldi

By performing model reduction on Eq (3), the goal is to effectively describe the behavior of the vector T in time, through some low-dimensional subspace as:

$$T = [V] \cdot T_r + \mathbf{\epsilon}, \quad T_r \in \mathbb{R}^r, \quad r \ll n, \quad \mathbf{\epsilon} \approx 0$$
(9)

Eq (9) states that, with the exception of a small error described by vector $\mathbf{\varepsilon} \in \mathbb{R}^n$, the possible movement of the *n*-dimensional vector \mathbf{T} belongs, for all times, to a *r*-dimensional subspace, and is determined by an $n \times r$ transformation matrix *V*. The matrix $V \in \mathbb{R}^{r \times n}$ is composed from *r n*-dimensional vectors that form a basis for the reduced subspace, and the *r*-dimensional vector \mathbf{T}_r represents a new low order set of coordinates for the given basis. When the subspace is found, Eq (3) is projected onto it, and this projection process produces a reduced order system (4).

It has been shown that, in the case of large-scale systems, very good candidates for the required low-order subspace in Eq (9) are Krylov subspaces [8].

The basic idea behind the Krylov-subspace-based Arnoldi algorithm is to write down the transfer function of (3) in the frequency domain using a Taylor series in the Laplace variable $s_0 = 0$:

$$\{G(s)\} = -\sum_{i=0}^{\infty} \{m\}_{i} s^{i}$$
(10)

where $\{m\}_i = E^T (-C^{-1}K)^{-(i+1)}C^{-1}F$ is called the *i*-th moment, and then to find a much lower order system (4) whose transfer function $\{G_r(s)\}$ has the same moments as $\{G(s)\}$ up to the degree *r*. The moments are not computed explicitly. Instead, a Krylov subspace of dimension *r*:

$$K_{r}\{A, \boldsymbol{b}\} = span\{\boldsymbol{b}, A\boldsymbol{b}, ..., A^{r-1}\boldsymbol{b}\}$$
with $A = -[K]^{-1}[C]$, and $\boldsymbol{b} = -[K]^{-1}\boldsymbol{F}$
(11)

is used, and through the computation of an orthogonal basis for this subspace, the matrices $[C_r]$ and $[K_r]$ and the load vector F_r of the reduced system are computed without taking into account either input function or the output vector (matrix) *E*. All the inputs and outputs of the Arnoldi algorithm are shown in Fig. 1.

The property of the Krylov subspace (11) is such that the first r moments of $\{G_r(s)\}$ and $\{G(s)\}$ match, as required. Since the time required for the Arnoli-based reduction grows approximately as $O(2r^2 \cdot n + k \cdot n)$, where k depends on the sparsity of the system matrices [11], there is no limitation for

applying Arnoldi algorithm to the linear systems with the high dimensional state vector.

2.3 Model Order reduction via Guyan

The large dimension of (3) could be reduced by the elimination of internal nodes i. e., those which do not connect to external circuitry. For steady-state problems ([C] = 0) it is possible to decompose the linear system into terminal and internal equations, by splitting the matrix K into four blocks:

$$\begin{bmatrix} K_{ee} & K_{ei} \\ K_{ie} & K_{ii} \end{bmatrix} \begin{bmatrix} T_e \\ T_i \end{bmatrix} = \begin{bmatrix} F_e \\ F_i \end{bmatrix} I^2(\infty)R$$
(12)

with the index sets e and i ranging over all external and internal nodes respectively. It is now possible to eliminate the equations for the non-terminal nodes by means of linear algebra operations (e.g., the Schur complement [12]), to get the heat conductivity matrix and the load vector of the reduced system:

$$[K_r] = [K_{ee}] - [K_{ei}][K_{ii}]^{-1}[K_{ie}]$$

$$\{F_r\} = \{F_e\} - [K_{ei}][K_{ii}]^{-1}\{F_i\}$$
(13)

1

The commercial finite element solver ANSYS [13] offers the possibility of reduced order modeling also for the transient problems of the form (3). The computation of $[K_r]$ and $\{F_r\}$ is done as in (13) and the reduced heat capacity matrix is given through:

$$[C_r] = [C_{ee}] - [K_{ei}][K_{ii}]^{-1}[C_{ie}](-[C_{ei}][K_{ii}]^{-1}[K_{ie}] + [K_{ei}][K_{ii}]^{-1}[C_{ii}][K_{ii}]^{-1}[K_{ie}])$$
(14)

which is analog to the computation of the reduced mass matrix for structural dynamics (without damping effects), as proposed by Guyan [14]. It is now possible to expand the terminal degree of freedom (DOF) values to gain the complete temperature distribution of the device using again Eq (12) for the steady-state:

$$\{T_i\} = [K_{ii}]^{-1}\{F_i\} - [K_{ii}]^{-1}[K_{ie}]\{T_e\}$$
(15)

3. Results

3.1 MEMS Test Case

A new class of high energy MEMS actuators integrates solid fuel with three bonded silicon micromachined wafers [15],[16]. It delivers either an impulse-bit thrust or pressure waves within a sub millimeter volume of silicon, by producing a high amount of energy from the ignitable substance contained within the microsystem. The microthruster fuel is ignited by passing an electric current through a polysilicon resistor embedded in a dielectric membrane, as shown in Fig. 2.

After the ignition phase, sustained combustion takes place and forms a high-pressure, high-temperature gas mixture. Under the pressure of the gas the membrane ruptures, and an impulse is imparted to the carrier frame as the gas escapes from the tank. The present work considers the initial heating phase of the fuel, right up to the onset of ignition. Although the functioning goal of the microthruster device is to reach the critical (ignition) temperature rather than to avoid it through power dissipation, the mathematics behind the modeling, Eq (3) is exactly the same as for the cases of semiconductor devices described in [1]-[3]. All three order reduction methods described in the section 2 were applied to the thruster model.

3.2 Numerical Simulation Results

A numerical simulation result of the full finite element model is shown in the Fig. 3.

Fig. 4 shows the decay of the Hankel singular values for the microthruster model, computed by the BTA method. The steep decay of the Hankel singular values shows that the finite element basis obtained after mesh generation is redundant. In this case, a low-dimensional model can accurately capture the behavior of the original high-dimensional FE model. We believe that this holds for most electro-thermal models described through equation system (3).

For the specified error bound, Eq (7), of 0.1 a BTA algorithm has estimated the size of reduced system to be of the order 7. An equation system (3) containing 1071 ODEs was reduced to 7 ODEs using each of the three presented algorithms (Fig. 5). Master degrees of freedom (external nodes) needed for the Guyan algorithm were chosen automatically by ANSYS5.7.

A maximal relative error by Arnoldi-based reduction doesn't exceed 4% during the initial phase, whereas by an optimal BTA-based reduction this error amounts 0.3% within the steady-state phase (Fig. 6).

The Guyan-based reduction produces maximal relative error of more than 64% (not shown in the Fig. 6). This large error is mostly due to the transient heating phase, and vanishes within the steady-state response according to Eqs (12) and (13) as is to be expected. The approximation error for the reduced heat capacity matrix (14) decreases as the order r of the reduced system grows (Fig. 7), but maximal relative error between the full-scale solution and the reduced solution of the order 200 still amounts to 6%.

Fig. 8 shows the corresponding SISO transfer functions of the full-scale and both optimal and nonoptimal reduced systems of order 7. A further important result is that, for the electro-thermal model described through equation system (3), the simple SISO setup for the Arnoldi algorithm was sufficient to approximate not only a single output response but also the transient thermal response in all the finite element nodes of the device. Fig. 9 shows a mean relative difference for all the nodes computed as:

$$MSRD(t) = \sqrt{\frac{\sum_{i=1}^{n} \left(\frac{\boldsymbol{T}_{i}(t) - \boldsymbol{T}_{i}(t)}{\boldsymbol{T}_{i}(t)}\right)^{2}}{n}}$$
(16)

between the full-scale and the reduced order models. Herby $\hat{T}_i(t)$ is the temperature of the i-th node over time, $\hat{T}_i(t)$ is the i-th component of the vector $[V] \cdot T_r(t)$ and n = 1071.

Already for the reduced system of order 20 a maximal mean relative error (difference) amounts to only 0.14%. Hence it was possible, after the simulation of the reduced model, to recover the solution for all the 1071 nodes by using Eq (9). In this case, the Arnoldi reduction algorithm can be viewed as a projection from the full space to the reduced Krylov space (11), with an identity output matrix corresponding to the SICO system description.

A software package has been developed which forms a netlist suitable for the behavioral simulator SABER (Fig. 10) from the three-dimensional geometry and governing Eqs (1) and (2) of the model.

The structure of the SABER input file is shown in Fig. 11.

The template microthruster.sin is implemented in the form of ODEs, Eq (4). The SABER plot shows the temperature development in node 1 (Fig. 2) for the case when the pre-heating was performed, needed to improve the subsequent sustained combustion. An approximation that the heat power $I(t)^2 R$ is uniformly distributed over the meander resistor, allows the back coupling of the meander's resistance. This has the advantage that the monitoring of the temperature through a change of resistance is possible, and certain design changes (such as the change of the meander's resistivity) are still possible after the model reduction phase.

4. Discussion

All three algorithms presented offer the possibility for the automatic order reduction of ODE systems. The BTA method additionally offers reduced system order estimation for a given error margin. However, due to the computational cost of $O(n^3)$ for the BTA-method it is impractical for large-scale systems.

The Guyan method doesn't provide sufficient accuracy for the reduction of electro-thermal systems. This is due to the fact that reduced order modeling after Guyan makes an attempt to generalize Eq (12) for a steady-state response to the transient problem using a coordinate transformation of the form:

$$\begin{cases} T_e \\ T_i \end{cases} = \begin{bmatrix} I \\ -K_{ii}^{-1} K_{ie} \end{bmatrix} \{ T_e \}$$
 (17)

This leads to exact matrix condensation for the heat conductivity matrix, but an approximated condensation for the heat capacity matrix.

Arnoldi-based reduction starts with moment matching for the transient problem as it is, and amounts to a coordinate transformation of the form (9), where $[V] \in R^{nxr}$ is gained directly as an output of the Arnoldi algorithm. It has been shown that for the electro-thermal model described by equation system (3), by a reduction from n = 1071 to $r \ge 20$ results in $\varepsilon \approx 0$. This means, that a Single-Input-Single-Output setup for the Arnoldi reduction algorithm produces an excellent approximation for the linear Single-Input-Complete-Output electro-thermal system, which is a significant result. The computational cost for Arnoldi-based order reduction, being less than $O(n^2)$, makes it suitable for reducing models with more than 10 000 degrees of freedom. A further big advantage of the Arnoldi algorithm is its iterative nature, i.e., only consecutive matrix vector multiplications are needed in each iterative step of the Arnoldi orthogonalisation procedure. In this way it is possible to exploit the sparse form of the system matrices [K] and [C], and to create fast application-specific implementations for the required $A \cdot b$ (see Eq (11)) product. Model reduction is automatic and based on the original system matrices only. In this way, a timeconsuming simulation of the full-scale model is circumvented with no loss of spatial distribution information. As already explained, the Arnoldi-based model reduction is not limited to this particular device, but can be applied to a wide spectrum of important electro-thermal MEMS devices, as well as for applications in bonding, IC temperature control, packaging and so on.

Acknowledgments

This work is partially funded by the EU through the project MICROPYROS (IST-1999-29047), partially by the DFG project MST-Compact (KO-1883/6) and partially by an operating grant of the University of Freiburg.

References

- H. Kawashima and R. Dang, "Non-Isothermal Device Simulation of gate Switching and Drain Breakdown Characteristics of Si MOSFET in Transient State", Proc. IEIECE, vol. E82-C, N 6, pp. 894-899, 1999.
- [2] K. Sonoda, M. Tanizawa, K. Ishikawa, N. Kotani and T. Nishimura, "Circuit-Level Electrothermal Simulation of Electrostatic Discharge in Integrated Circuits", Proc. IEIECE, vol. E83-C, N 8, pp. 1317-1323, 2000.
- [3] W. Nakayama, "Thermal Issues in Microsystems Packaging", Proc. IEEE Transactions on Advanced Packaging, vol. 23, N 4, pp. 602-607, 2000.
- [4] P. Schwarz, P.Schneider, "Model Library and Tool Support for MEMS Simulation", Proc. SPIE, vol. 4407, pp. 1-14, 2001.
- [5] V. d'Alessandro, N. Rinaldi, "A critical review of thermal models for electro-thermal simulation", Solid-State Electron., vol. 46, N 4, pp. 487-496, 2002.
- [6] E. B. Rudnyi, J. G. Korvink, "Review: Automatic Model Reduction for Transient Simulation of MEMS-based Devices", Sensors Update, vol. 11, 2002.
- [7] A. C. Antoulas, "Approximation of linear dynamical systems", Wiley Encyclopedia of Electrical and Electronics Engineering, (Ed.: J.G. Webster), 11, pp. 403-422, 1999.
- [8] L. M. Silveira, M. Kamon, I. Elfadel, J. White, "A Coordinate-transformed Arnoldi Algorithm for Generating Guaranteed Stable Reduced-Order Models of RLC Circuits", Comp. Methods. Appl. Mech. Eng., vol. 169, pp. 377-389, 1999.
- [9] D. Ostergaard, M. Gyimesi, "Finite Element Based Reduced Order Modeling of Micro Electro Mechanical Systems (MEMS)", Proc. 3rd International Conf. on Modeling and simulation of Microsys., San Diego, California, pp. 684-687, March 2000.
- [10] B. C. Moore, "Principal Component Analysis in Linear Systems: Controlability, Observability, and Model Reduction", IEEE Trans. on Autom. Control AC, vol. 26, pp. 17-32, 1981.
- [11] Y. Saad, "Iterative Methods for Sparse Linear Systems", International Thomson Publishing Inc., 1996.
- [12] C. Meier, M. Emmenegger, S. Taschini, H. Baltes, J. G. Korvink, "Equivalent Circuit Model of Resistive IC Sensors Derived with the Box Integration Method", IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, vol. 18, N 7, 1999.
- [13] ANSYS Release 5.6 User Manuals-Theory Manual, 1999.
- [14] R. J. Guyan, "Reduction of Stiffness and Mass Matrices", AIAA Journal, vol. 3, N 2, p. 138, 1965.
- [15] C. Rossi, "Micropropulsion for Space", Sensors Update, vol. 10, pp. 257-292, 2002.
- [16] E. B. Rudnyi, T. Bechtold, J. G. Korvink, C. Rossi, "Solid Propellant Microthruster: Theory of Operation and Modelling Strategy", Nanotech 2002 - At the Edge of Revolution, September 9-12, 2002, Houston, USA, AIAA Paper 2002-5755.

Figure Captions

Fig. 1 Model Reduction by the Arnoldi process.

Fig. 2 Microthruster Structure.

Fig. 3 Temperature distribution within the igniting wafer after 0.3 s of heating with 80 mW power; $T_{ref} = 273K$.

Fig. 4 Decay of the Hankel singular values for the order 1071 microthruster model.

Fig. 5 Solution of the full system (order 1071) and of the 7th order reduced systems for a single node (node 1 in Fig. 2).

Fig. 6 Relative error corresponding to the plots in Fig. 5 during the initial 0.15 s.

Fig. 7 Solution of the full system (dashed, order 1071) and of the different order reduced systems for a single node (node 1 in Fig. 2), using Guyan reduction.

Fig. 8 SISO transfer function of the full-scale (order 1071) and of the reduced 7th order system corresponding to node 1 in Fig. 2.

Fig. 9 Mean square relative difference (MSRD) for all the nodes during the initial 0.05 s, for an Arnoldibased reduction from order 1071 to 20, 15, 10, 7 and 5.

Fig. 10 Automatic model order reduction software block diagram.

Fig. 11 Structure of the SABER model.



Fig. 1 Model Reduction by the Arnoldi process.



Fig. 2 Microthruster Structure.



Fig. 3 Temperature distribution within the igniting wafer after 0.3 s of heating with 80 mW power; . $T_{ref} = 273K$.



Fig. 4 Decay of the Hankel singular values for the order 1071 microthruster model.



Fig. 5 Solution of the full system (order 1071) and of the 7th order reduced systems for a single node (node 1 in Fig. 2).



Fig. 6 Relative error corresponding to the plots in Fig. 5 during the initial 0.15 s.



Fig. 7 Solution of the full system (dashed, order 1071) and of the different order reduced systems for a single node (node 1 in Fig. 2), using Guyan reduction.



Fig. 8 SISO transfer function of the full-scale (order 1071) and of the reduced 7th order system corresponding to node 1 in Fig. 2.



Fig. 9 Mean square relative difference (MSRD) for all the nodes during the initial 0.05 s, for an Arnoldibased reduction from order 1071 to 20, 15, 10, 7 and 5.



Fig. 10 Automatic model order reduction software block diagram.



Fig. 11 Structure of the SABER model.