Microfluid Nanofluid (2014) 16:369–380
DOI 10.1007/s10404-013-1210-0

RESEARCH PAPER

A reduced-order model for whole-chip thermal analysis of microfluidic lab-on-a-chip systems

Yi Wang · Hongjun Song · Kapil Pant

Received: 30 November 2012 / Accepted: 20 May 2013 / Published online: 13 June 2013

Abstract This paper presents a Krylov subspace projection-based reduced-order model (ROM) for whole microfluidic chip thermal analysis, including conjugate heat transfer. Two key steps in the reduced-order modeling procedure are described in detail: (1) the acquisition of a 3D full-scale computational model in the state-space form to capture the dynamic thermal behavior of the entire microfluidic chip; and (2) the model order reduction using the block Arnoldi algorithm to markedly lower the dimension of the full-scale model. Case studies using practically relevant thermal microfluidic chip are undertaken to establish the capability and to evaluate the computational performance of the reduced-order modeling technique. The ROM is compared against the full-scale model and exhibits good agreement in spatiotemporal thermal profiles (<0.5 % relative error in pertinent time scales) and over three-orders-of-magnitude acceleration in computational speed. The salient model reusability and real-time simulation capability render it amenable for operational optimization and in-line thermal control and management of microfluidic systems and devices.

1 Introduction

Microfluidic lab-on-a-chip (LoC) systems hold great promise for a variety of applications in biology, medicine, and chemistry (Reyes et al. 2002; Aurouz et al. 2002) and offers numerous advantages such as speedup in analysis time, savings in reagent and sample volumes, salient throughput, and capability for high levels of integration and automation. In many applications, thermal environments need to be precisely controlled for proper execution of biochemical and cellular assays. For example, in thermal cycling (denaturing, annealing, and extension)-based polymerase chain reaction (PCR), robust spatial or temporal control within ±1 °C is normally entailed for consistent amplification performance (Lee et al. 2007). Temperature gradients are often harnessed for high-throughput kinetics analysis and network dynamics studies of biological systems, among other applications. In particular, for microfluidic systems constructed from polymer materials, adequate heat dissipation strategies and operating envelopes are desired to sustain biocompatible environments.

High-fidelity (3D) numerical simulation capable of coupled spatiotemporal thermal analysis is still the mainstream tool for the design of complex microfluidic devices. Erickson et al. (2003) applied 3D "whole-chip" finite element model to examine Joule heating and heat transfer in a poly(dimethylsiloxane) (PDMS) microchip. They observed dramatic temperature gradients at the intersections in the chip as well as the dependence of the buffer temperature on the applied potential, and proposed simple guidelines for improving chip design. Sikanen et al. (2008) carried out multi-physics and multi-scale simulation of coupled electrostatics, fluid dynamics, and heat transfer in an electrokinetic separation microchip using a 3D computational model. Due to the large mesh sizes (268,000 hexahedral elements in the volume mesh and 21,500 elements in the background electrolyte channel including the reservoirs, in total), a simulation time of 30 h was required for a single analysis.

In contrast to the other microfluidic phenomena (e.g., fluidics, electrokinetics, sample transport, and reaction),

Electronic supplementary material The online version of this article (doi:10.1007/s10404-013-1210-0) contains supplementary material, which is available to authorized users.

Y. Wang (✉) · H. Song · K. Pant
CFD Research Corporation, Huntsville, AL 35805, USA
e-mail: yxw@cfdrc.com

thermal analysis in microfluidics poses formidable challenges to the chip design as heat propagates in the entire chip domain rather than only in the fluidic channels, leading to more substantial model sizes and slower thermal diffusion process. Both factors contribute to the prohibitively computational cost associated with whole-chip thermal analysis, rendering it ill-suited for concept evaluation, operational optimization, as well as closed-loop thermal control and management, where rapid (near-real-time) iterative simulation capability is strongly desired for fast turnaround times. This is further exacerbated by slow computational convergence in the multi-timescale system and the use of composite materials with distinctly different thermal properties.

To overcome these limitations, compact analytical models and heuristic models (for instance, based on artificial neural networks (ANN)) were developed for rapid parametric analysis and design evaluation. Gui and Ren (2008) developed analytical models using the Green's function for evaluating Joule heating effects on the temperature distribution in an electrokinetically driven microfluidic PCR chip. The model showed good agreement with the 3D numerical model and was utilized for parametric studies of chip temperature control. Hao (2004) presented an analytical linear model of the steady-state temperature distribution in a 2D multi-layered structure heated by multi-heaters for realizing arbitrary temperature distributions. The analysis was extended to compute the input electricity power for accurate temperature control. A common limitation of the analytical models is that they are normally cast in a 2D conductive domain. Despite their notable speedup and physical insight enabled by the closed-form solution, they are less useful for thermal analysis involving complex, irregular microfluidic structures (in particular at the intersection, junction, and chip periphery) involving convective heat transfer, and hence are ill-suited to capture spatio-temporal thermal distribution at the chip level. On the heuristic model side, Lee et al. (2007) developed an ANN model using a hybridized neurogenetic optimization methodology. The model trained using finite element analysis data could be exploited to predict the steady-state temperature distribution in specific regions of the continuous-flow PCR microchannels. The need for a large amount of training data and the limited capability for transient analysis render ANN model inefficient in design optimization and thermal management. Therefore, there is a strong need for a general thermal modeling methodology amenable to high-fidelity whole-chip analysis with ultrafast simulation speed for the targeted applications.

In this context, this paper presents a generalized model order reduction methodology for thermal analysis of whole microfluidic chips containing arbitrary geometries. The reduced-order model (ROM) is automatically derived by projecting a full-scale 3D computational model onto a low-dimensional Krylov subspace that is constructed by the mathematically formal block Arnoldi algorithm. As a result, a low-dimensional ROM whose size is at least 2–3 orders of magnitude less than the original full-scale model can be obtained. The ROM is then coupled with ODE solvers to compute the transient temperature profile in the entire computational domain. The ROM is verified against high-fidelity 3D computational fluid dynamics (CFD) analysis tool for microfluidic thermal devices, including PCR and thermal gradient generator. As the developed ROM is capable of simultaneously handling multiple independent inputs, it is expected to find widespread use in operational optimization and real-time thermal management.

The facets that clearly distinguish the present work from the existing research in terms of electric thermal modeling in MEMS (Yang and Shen 2005; Rudnyi and Korvink 2002), analytical thermal models, and 3D full-scale model analysis include: (1) The first attempt to develop ROMs for whole-chip thermal analysis of microfluidic lab-on-a-chip systems incorporating convective heat transfer. (2) Our ROMs bridge the gap between the analytical model and high-fidelity physical model, and saliently combine several benefits of both in terms of computational accuracy and speed. Different from the analytical model, our ROM is capable of capturing the complex thermal scenarios, such as chips with irregular shapes and thermal transitioning regions, and allows for comprehensive spatiotemporal inspection of the thermal transport at the 3D chip level. Relative to the full model analysis, in general, 2–3-orders-of-magnitude speedup can be achieved without appreciably compromising the simulation accuracy. Although our linear ROM assumes temperature-independent material properties (e.g., fluid viscosity and thermal conductivity), it is able to resolve the dominant thermal behavior with adequate accuracy. The judicious combination of these merits make ROM-based simulation methodology well suited for initial concept evaluation and screening, parametric analysis, operational optimization, as well as development of reliable thermal management strategies. (3) In addition to the model order reduction algorithms, the present paper also provides a systematic view of the ROM methodology and the procedure for acquiring full-scale models in the state-space form to facilitate implementation. It should be pointed out that while the present work focuses on linear heat transfer, the methodology can be readily extended to the nonlinear domain for enhanced model accuracy, but at the cost of additional ROM generation and simulation times.

The paper is organized as follows: The approach and procedure of the model order reduction is introduced in Sect. 2. Specifically, the description of the full-scale model acquisition and the block Arnoldi algorithm to compute the projection subspace is presented in Sects. 2.1 and 2.2,

respectively. The ROM is then verified against full-scale model data in case studies of various microfluidic devices (Sect. 3). The paper concludes with a summary in Sect. 4.

2 ROM methodology for whole-chip thermal analysis

In this section, we will first present the systematic organization of the ROM simulation methodology. Then, the key elements of full-scale model acquisition and model order reduction for generating reduced thermal models will be described. As shown in Fig. 1, the ROM-based thermal analysis includes three key steps:

1. *Acquisition of full-scale thermal models*: The procedure starts with acquisition of the 3D full-scale thermal model, which encompasses (1) meshing of the microfluidic chip geometry into discrete control volumes; (2) computation of the thermal links (both conductive, D, and convective, F, links), heat sources, Q, and thermal capacitance, C, based on appropriate differencing schemes (e.g., finite volume or finite element); and (3) assembly of the model information into the state-space form to obtain a dynamic system along with appropriately selected system inputs, u. The first two steps are routinely performed by almost all high-fidelity numerical analysis tools.
2. *Generation and computation of ROMs*: The assembled full-scale thermal model is then reduced using the mathematically rigorous model order reduction algorithm. Specifically, a low-dimensional Krylov subspace ($X_q \in \Re^{M\times q}$) is constructed using the block Arnoldi algorithm, onto which the full-scale model is projected via $T = X_q T_r$ (i.e., the "Model Order Reduction" block in Fig. 1) leading to a low-dimensional thermal ROM, viz. $T_r \in \Re^q$ with $q \ll M$, where q is the dimension of the ROM; M is the number of the discrete control volumes used to discretize the spatial terms of the thermal transport equation and is the dimension of the full-scale model. The ROM is computed using ordinary differential equation (ODE) solvers for transient integration of the temperature in the reduced computational domain. Due to its significantly reduced dimension relative to the full model (i.e., $q \ll M$), the ROM can be computed at very fast speed. The results of the generated ROM can then be compared against the full-scale thermal model for verification.
3. *Use of ROMs for microfluidic analysis*: In many instances (in particular for linear cases), the validated ROM can be reused in subsequent iterative simulation sweeping the operating parameter space (e.g., input power and heat dissipation conditions) to identify the best parameter combination for design optimization and in-line thermal management.

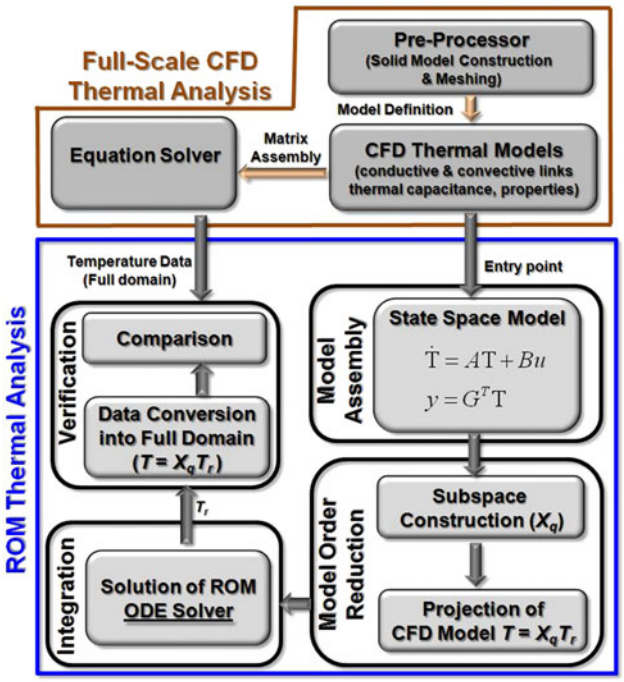

Fig. 1 Schematic of the ROM methodology for whole-chip thermal analysis

It should be noted that ROM generation is a one-time cost given fixed model geometry. Due to its mathematically rigorous nature, the ROM bears excellent extrapolation ability as long as the linear assumption remains valid in the entire parameter space. This is in distinct contrast to the heuristics-based approach (e.g., ANN) whose applicability heavily depends on the range of the training data set.

2.1 Acquisition of the full-scale model

In this section, we will present the equations and assumptions in the whole-chip thermal analysis and the procedure to obtain the full-scale thermal models for model order reduction. The conjugate heat transfer in microfluidic chips involves fluid flow and heat transfer. The former only occurs in the fluidic channel, while the latter takes place across the entire chip. Both are linked via a two-way coupling, that is, the fluidic flow contributes to the convective heat transfer; and the temperature affects the fluid viscosity and flow velocity in the channel. Thus, the governing equation of the heat transfer in the chip can be written as

$$\rho C_p \frac{\partial T}{\partial t} = k\nabla \cdot (\nabla T) - \rho C_p \mathbf{v} \cdot \nabla T + Q \qquad (1)$$

where $\mathbf{v}$ is the velocity vector of fluid flow in the microchannel; T is the temperature in the microchip; ρ, k, and C_p

are, respectively, the density, thermal conductivity, and specific heat of the materials in the chip; Q is the heat source or sink terms (e.g., laser heating and Joule heating). Note that by removing the convection term Eq. (1) is still valid for the solid part of the chip.

Note that proper orthogonal decomposition (POD)-based ROMs have been presented in prior efforts for both pressure-driven (Yang and Kuo 2008) and electrokinetic flow (Qiao and Aluru 2003). Therefore, in the present paper, we primarily focus on the ROMs of heat transfer assuming the background flow is pre-calculated by full models or other reduced models. Accordingly, following assumptions are made to facilitate the analysis:

1. Similar to previous studies (Lee et al. 2007; Gui and Ren 2008; Hao 2004), the fluid viscosity and flow velocity in the microchannel are assumed temperature independent and calculated at a constant temperature. Therefore, our model neglects the impact of temperature variation on fluid properties and captures the dominant, first-order convective effect on heat transfer.
2. In contrast to the thermal diffusion process propagating in the entire chip, the flow equilibrates rather fast in the microchannel and hence is treated steady state.
3. Density, specific heat, and thermal conductivity of the solid chip materials are assumed constant, which is reasonable in the operating temperature range of biologically relevant microchips and similar to previous studies (Erickson et al. 2003; Gui and Ren 2008; Hao 2004).

The spatial differentials (i.e., R.H.S.) of the thermal governing Eq. (1) can be discretized by means of various numerical schemes, including finite volume method (FVM), finite difference method (FDM), or finite element method (FEM). In the current work, the FVM discretization of the convective term by an upwind scheme is used, yielding

$$\begin{aligned}\rho C_p V_{\mathrm{p}} \frac{\partial T_P}{\partial t} &= -(F_e + F_s + D_e + D_s + D_w + D_n)T_{\mathrm{P}} \\ &\quad + (F_w + D_w)T_{\mathrm{W}} + (F_n + D_n)T_{\mathrm{N}} + D_e T_{\mathrm{E}} \\ &\quad + D_s T_{\mathrm{S}} + Q_{\mathrm{P}} V_{\mathrm{P}} = -(F_w + F_n + D_e + D_s \\ &\quad + D_w + D_n)T_{\mathrm{P}} + (F_w + D_w)T_{\mathrm{W}} + (F_n + D_n)T_{\mathrm{N}} \\ &\quad + D_e T_{\mathrm{E}} + D_s T_{\mathrm{S}} + Q_{\mathrm{P}} V_{\mathrm{P}} = -a_p T_{\mathrm{P}} + a_w T_{\mathrm{W}} \\ &\quad + a_n T_{\mathrm{N}} + a_e T_{\mathrm{E}} + a_s T_{\mathrm{S}} + Q_{\mathrm{P}} V_{\mathrm{P}} \end{aligned} \quad (2)$$

where P is the present cell of interest; E, S, W, and N are the center of the cells surrounding cell P; e, s, w, n in the subscript are the interfaces linking the cell P with its adjacent cells (see Fig. 2); T_{P} is the temperature at the cell center; V_{p} is the volume of the cell P; $F_i = (\rho v C_{\mathrm{p}} A)_i$ is convective links at the cell interface; A is the area of the interface; and subscript i is the index of the interface "e", "s", "w", and "n". $D_i = (kA/\delta)_i$ are the conductive links at the cell interface, where δ is the distance between the cell P and its neighboring cell. $Q_{\mathrm{p}} V_{\mathrm{p}}$is the heat generation in the present cell, and t is the time. In derivation of Eq. (2), the mass continuity of flow is used to obtain $F_e + F_s = F_w + F_n$. $a_e = D_e$, $a_s = D_s$, $a_w = F_w + D_w$, $a_n = F_n + D_n$, and $a_{\mathrm{P}} = a_e + a_s + a_w + a_n$. For the solid part, the convection terms (F_i) will vanish. It should be pointed out that in contrast to previous ROM-based electrical thermal analysis (Yang and Shen 2005; Bechtold et al. 2005), the present effort captures the convective contributions arising from the microfluidic flow.

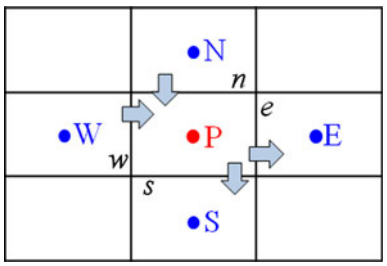

Fig. 2 Finite volume method-based representation of full-scale thermal model in a 2D computational domain

Equation (2) is applied to all the computational cells in the entire domain including those at the boundaries, yielding a large system of ODEs in the state-space form governing the temperature evolution at each cell center, i.e.,

$$\frac{\partial \mathbf{T}}{\partial t} = \mathbf{AT} + \mathbf{Bu} \qquad \mathbf{y} = \mathbf{G}^T \mathbf{T} \quad (3)$$

Here, $\mathbf{T}(t) \in \Re^M$ is a vector storing the temperature at all the discrete control volumes (i.e., cells in FVM), and M is the total number of the discrete control volumes in the domain as defined above. $\mathbf{u}$ is a time-dependent input vector and can be the prescribed temperature or heat flux at chip boundaries, the ambient temperature for the heat dissipation, and volumetric heat generation inside the chip. $\mathbf{A} \in \Re^{M \times M}$ is the thermal exchange matrix containing all the conductive and convective links described in Eq. (2) that are normalized by the thermal capacitance of each cell, $\mathbf{B} \in \Re^{M \times k}$ is the scatter matrix assigning input $\mathbf{u}$ into each control volume. $\mathbf{y}$ is an output vector storing the quantities of interest (e.g., the average temperature or gradients in subdomains) that is gathered by the output matrix $\mathbf{G}$ from $\mathbf{T}$.

2.2 Model order reduction

Since the dimension of Eq. (3) is normally too large for efficient numerical integration, the model order reduction technique is used to reduce the dimension ($q \ll M$) by projecting the original model onto a subspace $\mathbf{X}_q \in \Re^{M \times q}$

(i.e., $\mathbf{T} = \mathbf{X}_q\mathbf{T}_r$) while retaining the same number of inputs and outputs, which is given by

$$\frac{\partial \boldsymbol{T}}{\partial t} = \boldsymbol{A}\boldsymbol{T} + \boldsymbol{B}\boldsymbol{u} \overset{T=X_qT_r}{\leftrightarrows} \frac{\partial \boldsymbol{T}_r}{\partial t} = \boldsymbol{A}_r\boldsymbol{T}_r + \boldsymbol{B}_r\boldsymbol{u}$$
$$\boldsymbol{y} = \boldsymbol{G}^T\boldsymbol{T} \overset{T=X_qT_r}{\leftrightarrows} \boldsymbol{y} = \boldsymbol{G}_r^T\boldsymbol{T}_r \tag{4}$$

where $\mathbf{T}_r \in \Re^q$ is the temperature in the reduced domain, $\mathbf{A}_r = \mathbf{X}_q^{\mathrm{T}}\mathbf{A}\mathbf{X}_q$, $\mathbf{B}_r = \mathbf{X}_q^{\mathrm{T}}\mathbf{B}$, $\mathbf{G}_r = \mathbf{X}_q^{\mathrm{T}}\mathbf{G}$. As the dimensions of $\mathbf{A}_r$, $\mathbf{B}_r$, and $\mathbf{G}_r$ are significantly lower than those of the full system, the ROM governing $\mathbf{T}_r$ can be computed rapidly with significantly lower hardware requirements. The $\mathbf{T}_r$ in the reduced domain can be either converted back to the original full space using projection $\mathbf{T} = \mathbf{X}_q\mathbf{T}_r$ or used to attain the output of interest via $\mathbf{y} = \mathbf{G}_r^T\mathbf{T}_r$ in Eq. (4).

The key step of model order reduction is to identify an appropriate, low-dimensional projection subspace $\mathbf{X}_q$. A good candidate for reducing large-scale linear systems such as Eq. (4) is the Krylov subspace method (Rudnyi and Korvink 2002; Antoulas 2005) to match the moments of the transfer functions of the original full system in the frequency domain. A numerically stable procedure for locating the Krylov subspace $K_q(\mathbf{A},\mathbf{B}) = \mathrm{colspan}\{\mathbf{B}, \mathbf{AB}, \ldots, \mathbf{A}^{k-1}\mathbf{B}\}$ is the Arnoldi process (Yang and Shen 2005) or block Arnoldi for multi-inputs (Odabasioglu et al. 1998) (see online supplemental material for more details).

3 ROM analysis and discussion

In this section, the model order reduction technique will be applied to thermal analysis of practical LoC devices, including a thermal gradient microchip and a continuous-flow PCR microchip. Comparison of ROMs against high-fidelity CFD analysis will be presented to demonstrate its salient computational accuracy and efficiency.

The high-fidelity CFD analysis and model acquisition were performed with the commercial finite volume based multi-physics simulation package CFD-ACE+ (ESI-CFD, Inc.). The computational domain was meshed using a hexahedral volume-based grid within the preprocessor (CFD-GEOM) of the CFD-ACE+ package. As discussed above, the package was used to solve the Naviér–Stokes equations to provide the background, steady-state flow in the microchannel. The definitive information of the full-scale 3D thermal model for the whole microchip (i.e., $\mathbf{A}$ and $\mathbf{B}$ matrices and $\mathbf{u}$ vector in Eq. (3)) was exported from CFD-ACE+ and assembled into the state-space form, which was then reduced and computed in our ROM simulation tool coded in Matlab (MathWorks, Inc.) ODE45, the standard ODE solver based on an explicit, variable time step Runge–Kutta method, was employed as the main integration engine for ROM. In order to characterize the error between the ROM and full-scale CFD model, a performance index was defined

$$e(t) = \frac{\|\boldsymbol{X}_q\boldsymbol{T}_r - \boldsymbol{T}\|}{\|\boldsymbol{T}\|} \tag{5}$$

where $\mathbf{T}_r$ and $\mathbf{T}$ are, respectively, the solution from the ROM and full-scale CFD analysis. As $\mathbf{T}_r$ from the reduced model and $\mathbf{T}$ from the full model have different dimensions, the projection matrix $\mathbf{X}_q$ was used for converting the former back into the full domain.

3.1 Microfluidic thermal gradient chip

The first case study is focused on the thermal analysis of a microfluidic temperature gradient generator. Such a device is widely utilized for parallel and combinational measurements of bioassays, molecular screening, material synthesis, and catalyst optimization (Mao et al. 2002). As shown in Fig. 3, an array of microfluidic channels sits between a hot-source channel on the left and a cold-sink channel on the right. Relative to the microchannels, both the source and sink channels, sleeved with an internal thin brass layer, have larger sizes, in which hot and cold water is continuously supplied to maintain the temperature. Thus, an approximately linear temperature gradient forms along the transverse direction of the chip, and each microchannel experiences a different temperature. Our ROM study focuses on the region involving the thermal gradient generation (enclosed in the dashed rectangular box in Fig. 3) that comprises four domains, the lower substrate made from glass, the microfluidic channel containing buffer solution, the brass layer in the sink/source channels, and the upper substrate made from PDMS. Their properties are summarized in Table 1 (Sikanen et al. 2008). The flow velocity of 0.1 m/s and the temperature of $T_H = 80$ °C and $T_L = 8$ °C, respectively, were specified at the inlet of the source (hot) and sink (cold) channel. As all the microfluidic channels share a common inlet (see Fig. 3), a boundary

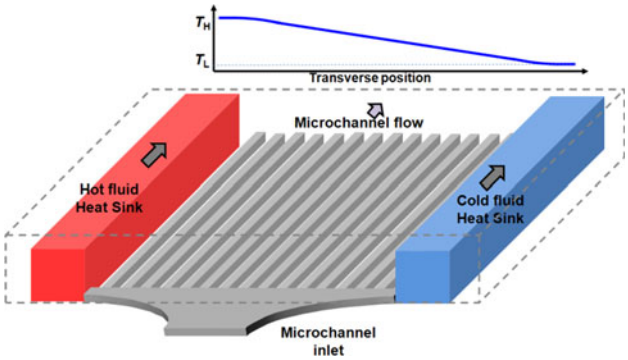

Fig. 3 Schematic of microfluidic thermal gradient generator

Table 1 Thermal and fluidic properties of the chip materials and buffer solution for high-fidelity CFD analysis (Sikanen et al. 2008)

	Thermal conductivity k (W/m K)	Specific heat C_p (J/kg K)	Density ρ (kg/m^3)	Viscosity μ (kg/s)
PDMS	0.18	1,100	1,030	N/A
Glass	1.4	835	2,230	N/A
Buffer	$0.61 + 1.2 \times 10^{-3}(T\text{-}298)$	4,191	1,024	$1.788 \times 10^{-3}\exp\{-1.704 - 5.306\cdot(273/T) + 7.003\cdot(273/T)^2\}$
Brass	109	380	8,500	N/A

Table 2 Benchmark high-fidelity CFD-ACE+ simulation using various parameters

CFD simulation no.	Type	Temperature-dependent liquid property	Pressure at the inlets of the microchannel array (Pa)
1	Linear	No (constant)	4,000 (or 400)
2	Nonlinear/coupled	Yes	4,000
3	Nonlinear/coupled	Yes	400

condition of the constant pressure was specified at their inlets. Ambient pressure of 1 atm was set at all the outlets of the microchannels and source/sink channels. Other thermal boundary conditions were specified as follows based on literature values: A heat transfer coefficient ($h = 10$ W m^{-2}K^{-1}) at the top and bottom surface and zero flux conditions along the side walls due to their small thermal contributions (Erickson et al. 2003).

Three high-fidelity CFD-ACE+ simulations were conducted to compare against the ROM simulation, which are summarized in Table 2. We can see that the effects of temperature dependence of the liquid properties can be captured by comparing simulation case No. 1 and 2, and the influence of the flow velocity in the microchannels on heat transfer will be manifested by the differences between case No. 2 and 3. Accordingly, in case No. 1 the fluid flow was pre-computed, followed by transient thermal analysis using constant material properties. In contrast, flow and heat transfer were solved in a coupled manner in case No. 2 and 3, that is, fluid flow needs to be updated during each iteration due to the temperature dependence of the fluid viscosity. Our ROM simulation, however, used constant property parameters evaluated at 50 °C, which are the same as the case No. 1 in Table 2.

Figure 4 illustrates the contour plot of the temperature in the simulation domain at various times obtained from linear CFD analysis (top row, i.e., No. 1 case in Table 2) and from ROM simulation with 24 dimensions (bottom row), i.e., $q = 24$ in Eq. (4). It reveals that the temperature gradients form in both the transverse and the streamwise direction, and the thermal fronts of the heat source and sink gradually propagate toward the center. Given the pressure head of 4,000 Pa and constant liquid properties, the average flow velocity in all the microchannels is the same and has a value of 2 cm/s. Figure 5 depicts the transverse temperature profile extracted at the end of the microchannel array at various times. The plateaus in the curves represent the temperature values in the heat source and sink channel, between which are the time-varying temperature gradients experienced by the microchannel array. Similar to the 3D contour plot, it clearly shows the transverse migration of the thermal fronts toward the central position starting from both sides, accompanied by the gradual temperature variations in the source and sink. It takes about 250 s for the thermal diffusion in the entire domain to approach steady state yielding a linear profile in the central region. Figure 6 illustrates the error of the ROM relative to the CFD results for all simulation, in which either the linear CFD or the coupled CFD was used as the benchmark. The ROM matches the linear CFD very well except for the initial times, $t \leq 1$ s (see the bold curve in Fig. 6b and the area highlighted in the circle in Fig. 4a as opposed to its CFD counterpart). The relative error continuously decreases with the simulation time from 1 to 10^{-3} %, which can be attributed to the fact that the truncation error due to the moment-matching method using the Krylov subspace becomes less important at the large time scale (i.e., low-frequency regime).

In order to examine the effect of the temperature dependence of the fluid property, coupled nonlinear CFD analysis (i.e., case No. 2 in Table 2) was also undertaken. Its results are compared against ROM using the constant property values in Fig. 7a. Small errors ($\sim$0.1 %, the dotted curve in Fig. 6) were observed at all times, in particular in the gradient-forming region between the source and sink channels where the temperature dependence of the thermal conductivity and viscosity of the buffer solution is the most salient. The large variation in temperature from 8 to 80 °C causes appreciable difference in viscosity and velocity between the coldest and hottest microchannel, contributing to the error in thermal analysis via convection.

To confirm this finding, a fully coupled CFD analysis involving slower flow velocity (i.e., case No. 3 in Table 2) was also carried out and is presented in Fig. 7b. We can see that low flow velocity featuring weaker convective heat

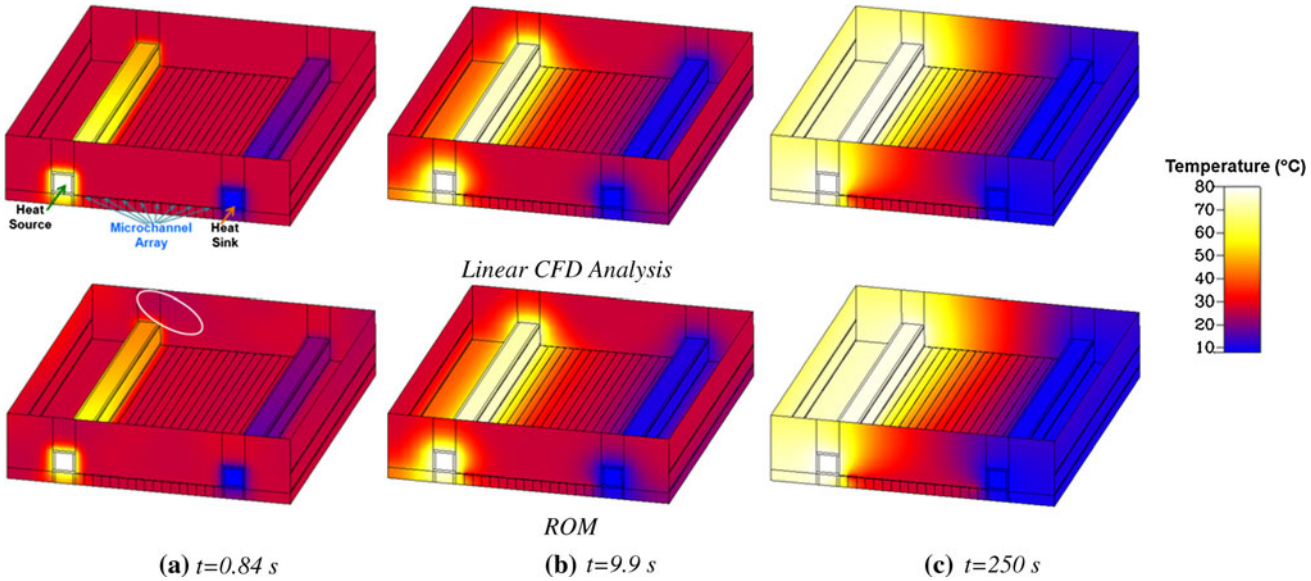

Fig. 4 Contour plot of the temperature in the simulation domain obtained from the linear CFD analysis (*top row*) and the ROM (*bottom row*) at various times **a** 0.84 s, **b** 9.9 s, and **c** 250 s

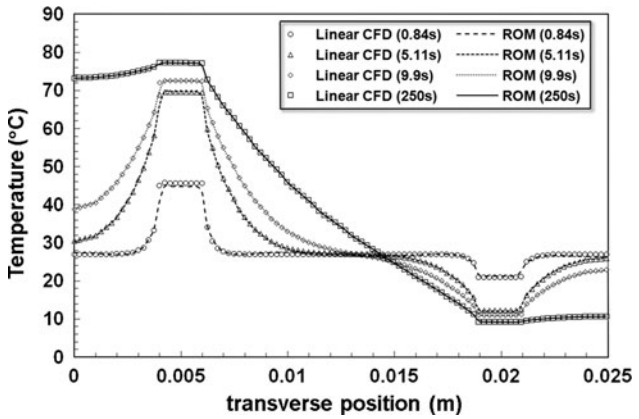

Fig. 5 Comparison between the linear CFD analysis and ROM. Transverse temperature profile extracted at the end of the microfluidic channels with pressure head $\Delta P = 4{,}000$ Pa at various times

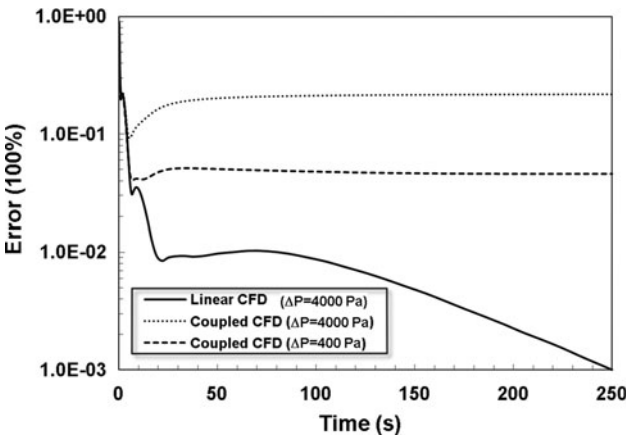

Fig. 6 Relative error of ROM relative to the CFD results for all simulation, where linear and coupled CFD are, respectively, used as the benchmark

transfer results in even smaller error and better agreement between coupled CFD and ROM (i.e., dashed curve in Fig. 6). Another interesting observation is the enhanced linearity in temperature profile between the source and sink channels at the low flow velocity relative to that in Fig. 7a as a result of the conduction-dominated heat transfer. Such an insight is very valuable in terms of guiding practical chip design, e.g., precise configuration of the flow velocity in the microchannel and selection of the sampling/observation spots, which otherwise is unavailable by analytical models (Gui and Ren 2008; Hao 2004). It should be pointed out that in microfluidic systems, the Reynolds number Re is low (e.g., maximum $Re \sim 5$ in our case). Therefore, convection is not the dominant transport mechanism, which is the primary reason why our ROM is capable of capturing the leading spatiotemporal thermal behavior in the whole chip with relatively good accuracy.

A summary of the compute times for various ROM and CFD analysis is summarized in Table 3. The full CFD analysis and ROM simulation were undertaken on a 64-bit Windows 7-based PC with 3.6 GHz Intel® Core™ i7-3820 CPU and 32 GB RAM. In the entry of the simulation time in Table 3, two values are reported: the time for generating the ROM and for computing the ROM, respectively. We can see that it took 922.2, 3,705, and 3,919 s, respectively, for the three high-fidelity CFD analysis. The use of the temperature-dependent properties in the coupled CFD analysis (case No. 2 and 3) requires the update of the flow field at each time step, resulting in longer computational

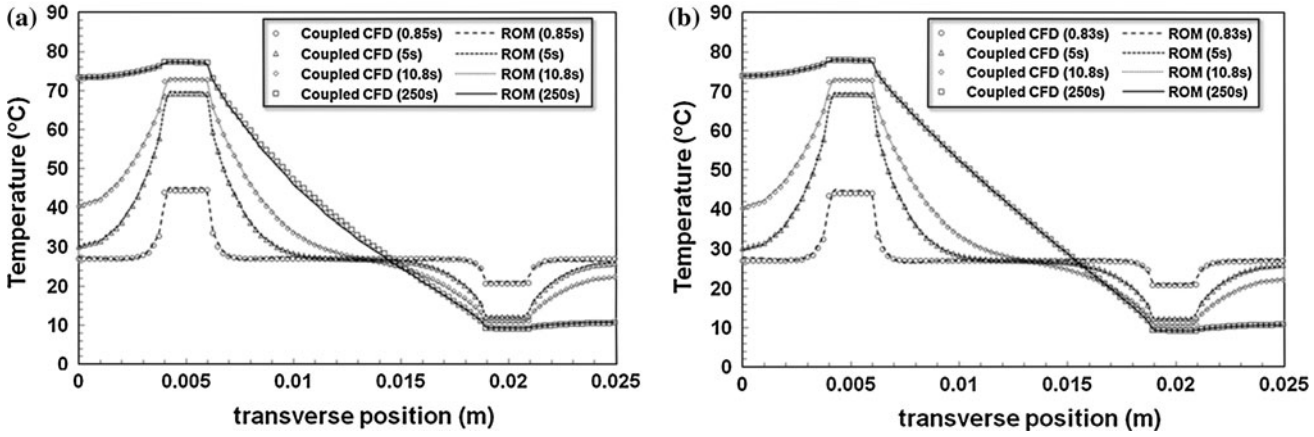

Fig. 7 Comparison between the coupled nonlinear CFD analysis and ROM in terms of the transverse temperature profile extracted at the outlet of the microfluidic channels at various times. **a** Pressure head $\Delta P = 4{,}000$ Pa and **b** pressure head $\Delta P = 400$ Pa

Table 3 Simulation times for CFD and ROM analysis of thermal gradient chip

	ROM (q = 24)	Linear CFD (case No. 1)	Coupled CFD (case No. 2)	Coupled CFD (case No. 3)
Simulation time (generation/computation)	15.04/0.378 s	922.2	3,705 s	3,919 s
Speedup factor (with/without ROM generation)	NA	60/2,400	240/9,800	260/10,000

time in contrast to the linear CFD case. In contrast, it took 15 and <0.4 s to respectively generate and compute our ROM with 24 dimensions. It should be noted that in many occasions, ROM generation is a one-time cost and the generated ROM can be reused for various heating/cooling conditions (e.g., different source/sink inlet temperature and environmental temperature). Accordingly, discounting its generation time, our ROM achieves more than three-orders-of-magnitude acceleration in the simulation speed relative to the full-scale CFD analysis, rendering possible the real-time, in-line thermal control and management.

3.2 Continuous-flow microfluidic PCR chip

The second case study consisted of a continuous-flow, serpentine-shaped microfluidic PCR chip. As shown in Fig. 8, the chip consists of three domains: the lower glass substrate, fluidic domain of buffer solution in microchannels, and the upper PDMS substrate. The fluidic domain consists of multiple thermal cycling channels connected serially. Each individual cycle comprises a denaturing channel (maintained at 95 °C), an annealing channel (at 60 °C), and an extension channel (at 72 °C). The different operating temperatures are generated by aligning three independent heating elements across the channel's flow path below the glass substrate. The sample solution containing the DNA template, primers, and PCR mix enters the device through a single inlet and undergoes thermal cycling and PCR amplification as it traverses through the microchannels. To investigate the thermal cycling behavior at the chip periphery, the first two and a half cycles were considered. Therefore, the planes interfacing the downstream cycling channels were treated as symmetric BCs (see Fig. 8a) in our computational model to simulate the real-world device. Constant temperature values of 95, 72, and 60 °C were, respectively, specified at the denaturing, extension, and annealing zones of the glass substrate that are in direct contact with the heating elements. The flow rate at the microchannel inlet was set at 4.8 or 48 μl/min, respectively, yielding 2 mm/s and 2 cm/s inlet velocity. The material properties and other simulation parameters (e.g., thermal boundary conditions) are the same as in the previous case study. Likewise, both linear and coupled nonlinear CFD analyses were carried out. The CFD model with a fully structured computational domain consisting of ~150,000 hexahedral volumes was constructed.

Figure 8 illustrates the contour plot of the temperature in the chip at various times obtained from coupled CFD analysis (top row) and ROM simulation with 24 dimensions (bottom row) for an inlet flow rate of 4.8 μl/min. It takes about 50 s for the thermal diffusion to stabilize throughout the device and reach the desired PCR temperatures in each individual zone. The agreement between CFD and ROM is excellent except at one of the chip corners in the initial period of the fast time scale (e.g., the region in the circle at $t = 1.04$ s in Fig. 8a). In this case

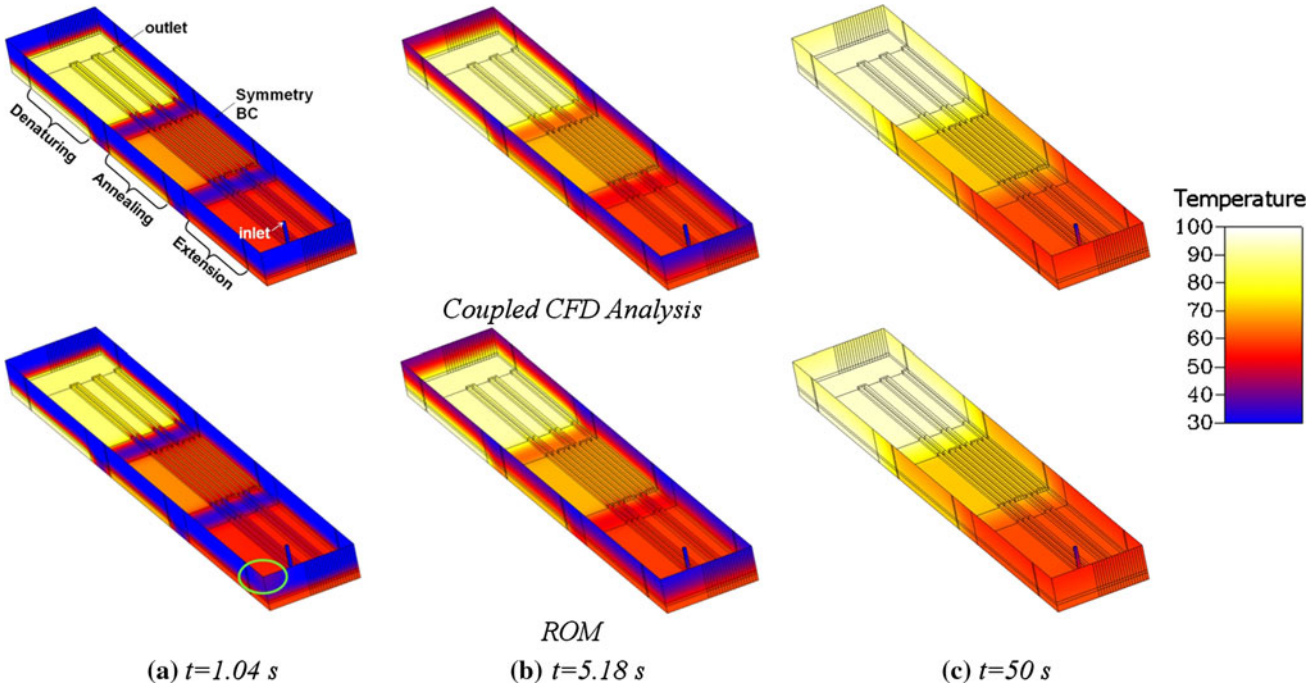

Fig. 8 Contour plot of the temperature in the microfluidic PCR chip obtained from the coupled CFD analysis (*top row*) and ROM (*bottom row*) at various times **a** 1.04 s, **b** 5.18 s, and **c** 50 s

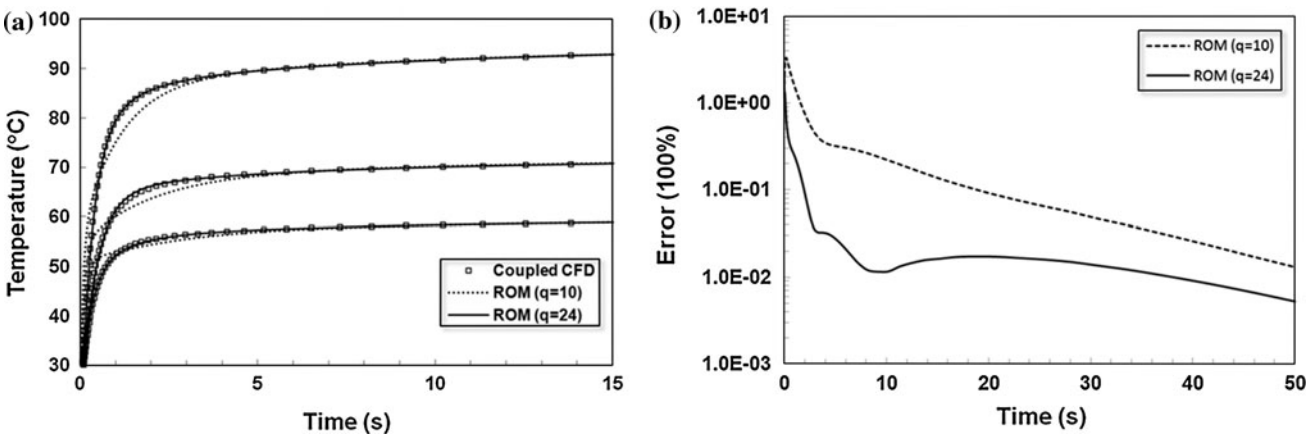

Fig. 9 Comparison between coupled CFD analysis and ROM with different dimensions: $q = 10$ and $q = 24$. **a** Transient temperature profile in the three PCR zones; and **b** relative error of ROM to the CFD analysis

study, we also examined the effect of the ROM dimension on the computational accuracy. Figure 9 portrays the transient temperature profiles extracted at the middle points of the denaturing, annealing, and extension region in the second thermal cycling channel that were obtained from coupled CFD analysis and ROM simulation with $q = 10$ or $q = 24$. For the sake of clarity, only thermal profiles at the first 15 s are presented, which show that all the simulation results match very well at the large time scales because of the moment-matching nature of the Krylov subspace method. The ROM with larger dimensions (e.g., $q = 24$) matching the higher-order moments is able to capture the fast time-scale behavior more accurately at all three temperature zones (e.g., 1–5 s in Fig. 9). Figure 9b quantitatively depicts the error of the ROM relative to the coupled CFD analysis. Both exhibit excellent agreement at all practically relevant time scales with the relative error less than 1 %.

Figure 10 illustrates the steady-state temperature profile along the flow streamlines originating from the geometrical center of the inlet plane of the microchannel for two flow rates: 4.8 μl/min and 48 μl/min. Excellent agreement is

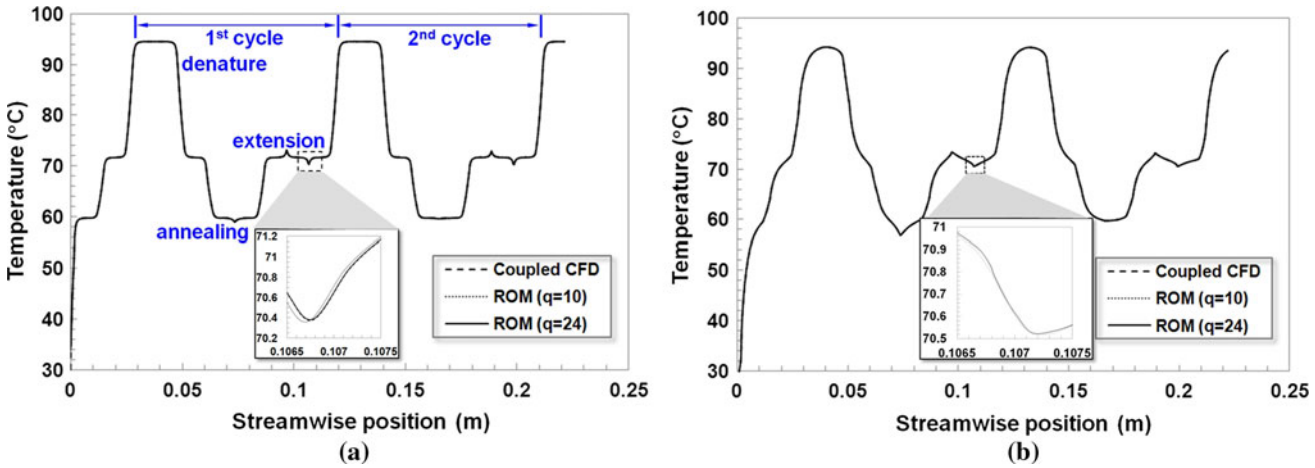

Fig. 10 Steady-state temperature profile along the flow streamlines **a** flow rate 4.8 μl/min **b** flow rate 48 μl/min

Table 4 Simulation times for CFD and ROM analysis for continuous-flow microfluidic PCR chip

	ROM ($q = 10$)	ROM ($q = 24$)	Coupled CFD	Linear CFD
Simulation time (generation/computation)				
$F = 4.8$ μl/min	18.8/0.30 s	21.22/0.34 s	4,217 s	1,215
$F = 48$ μl/min	18.6/0.31 s	21.4/0.38 s	8,627 s	1,097
Speedup factor	3,500–28,000	2,900–23,000	N/A	N/A

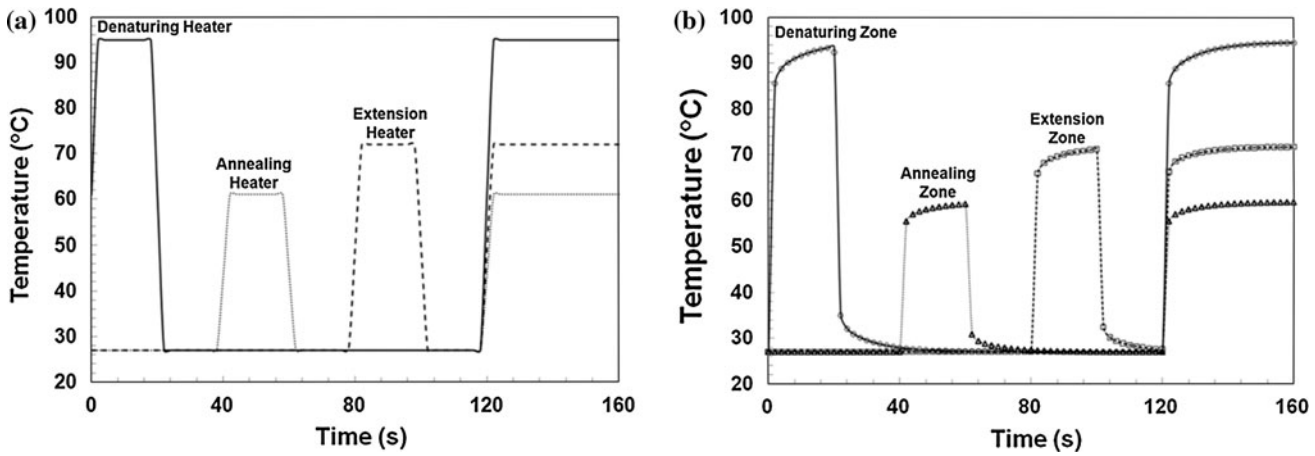

Fig. 11 Application of the ROM for real-time simulation. **a** Temporal inputs at the three heaters and **b** thermal response at the midpoint of the second thermal cycle in the three PCR zones

observed between the coupled CFD analysis and ROM results with all curves virtually indistinguishable. It also shows that the PCR solution indeed undergoes periodic thermal cycles, and the thermal profiles are susceptible to the flow rate. At low flow rate of 4.8 μl/min, the thermal equilibrium is reached in the major portion of the denaturing, annealing, and extension channel. It is interesting to note that even small temperature jumps in the annealing and the extension zone (the inset in Fig. 10) caused by the thermal interference from the channel inlet and the transitional region are clearly resolved, which again is difficult to capture using analytical models. At the higher flow rate, the channel length is not sufficient to establish thermal equilibrium as indicated by the absence of the temperature plateau in all the microchannels. It should be pointed out that in contrast to the previous case study, the ROM matches the coupled CFD analysis very well in this case, and the error due to the temperature dependence of the fluid

viscosity is negligible. This is because in this case, a constant flow rate (rather than a constant pressure head) was applied at the inlet of the microchannel, leading to virtually temperature-independent convection. In addition, linear CFD analysis employing the temperature-independent properties was also undertaken and matched the ROM very well (data are not presented).

The computational times for the coupled CFD analysis and ROM are compared in Table 4. It took 18–21 and $<$0.4 s to, respectively, generate and compute ROM. Likewise, given ROM reusability, our ROM enables more than three-orders-of-magnitude speedup relative to full-scale CFD analysis, rendering possible operational optimization and real-time thermal control at the whole chip level that may involve intensive, in-line simulation and prediction.

To demonstrate its reusability, the ROM generated above was then exploited in a real-time simulation involving arbitrary temporal inputs at the boundaries to predict the dynamic thermal behavior of the whole chip. Specifically, the temperatures of the three heaters in direct contact with the three PCR zones on the glass substrate are specified time-varying as depicted in Fig. 11a. We can see that the heaters are activated individually in the first 120 s and then all are turned on. Figure 11b illustrates the transient temperature response sampled at the midpoints of the denaturing, annealing, and extension region in the second thermal cycling channel obtained respectively by ROM and linear CFD analysis, exhibiting excellent agreement. Most saliently, the computational time of ROM is only 1.6 s, which is negligible to the physical time scale (160 s). It confirms the reusability and real-time simulation capability of our ROM for in-line thermal control and management.

4 Conclusion

This paper presented a mathematically rigorous Krylov subspace projection method to generate reduced-order models for whole microfluidic chip thermal analysis involving conjugate heat transfer. The detailed procedure of acquiring 3D full-scale model to capture the dynamic thermal behavior of the whole microfluidic chip and casting it into a state-space form amenable for model order reduction was elucidated. The block Arnoldi-based algorithm was then utilized to significantly reduce the model dimensions to expedite the simulation. The computational performance of ROM was demonstrated via several case studies using practically pertinent thermal microfluidic chips. The ROM was verified by comparing against the high-fidelity CFD model, which exhibited good agreement in spatiotemporal thermal profiles ($<$0.5 % relative error in practically relevant time scales) and over three-orders-of-magnitude acceleration in computational speed. The salient model reusability and real-time simulation capability enable its widespread applications in operational optimization and in-line thermal control and management of microfluidic devices.

The salient capability to capture spatiotemporal thermal details that are otherwise unavailable by analytical models and the order-of-magnitude speedup over the full model analysis verified the utility of ROMs for initial concept screening, operational optimization, and real-time thermal control and management. The errors arising from the linear assumption embedded in our ROMs are in general minor or even negligible for the temperature range and device materials in the present consideration due to the nature of low Reynolds flow in microfluidic systems.

ROM generation is a one-time cost. For large and reusable models, the overhead of model generation becomes less important or even negligible. It should be noted that the present ROM is built on fixed model geometry and properties and hence has limited applicability to geometric and material optimization (unless the full-scale transient analysis is so overwhelming that the ROM generation for a few uses is acceptable). Therefore, to incorporate the capability for parametric analysis, more sophisticated parameterized model order reduction techniques need to be exploited, which will be the focus of our future work.

Acknowledgments This research is sponsored by NIH/NHGRI under Grant Number 5R44HG004290-03.

References

Antoulas AC (2005) An overview of approximation methods for large-scale dynamical systems. Annu Rev Control 29(2):181–190

Aurouz PA, Lossifidis D, Reyes DR, Manz A (2002) Micro total analysis systems. 2. Analytical standard operations and applications. Anal Chem 74(12):2637–2652

Bechtold T, Rudnyi EB, Korvink JG, Graf M, Hierlemann A (2005) Connecting heat transfer macromodels for array MEMS structures. J Micromech Microeng 15:1205

Erickson D, Sinton D, Li D (2003) Joule heating and heat transfer in poly (dimethylsiloxane) microfluidic systems. Lab Chip 3(3): 141–149

Gui L, Ren CL (2008) Analytical and numerical study of joule heating effects on electrokinetically pumped continuous flow PCR chips. Langmuir 24(6):2938–2946

Hao Q (2004) Analytical heat-transfer modeling of multilayered microdevices. J Micromech Microeng 14:914–926

Lee HW, Arunasalam P, Laratta WP, Seetharamu KN, Azid IA (2007) Neuro-genetic optimization of temperature control for a continuous flow polymerase chain reaction microdevice. J Biomech Eng 129:540–547

Mao H, Yang T, Cremer PS (2002) A microfluidic device with a linear temperature gradient for parallel and combinatorial measurements. J Am Chem Soc 124(16):4432–4435

Odabasioglu A, Celik M, Pileggi LT (1998) PRIMA: Passive reduced-order interconnect macromodeling algorithm. IEEE Trans Comput Aided Des Integr Circuits Syst 17(8):645–654

Qiao R, Aluru NR (2003) Transient analysis of electro-osmotic transport by a reduced-order modelling approach. Int J Numer Meth Eng 56(7):1023–1050

Reyes DR, Lossifidis D, Auroux P-A, Manz A (2002) Micro total analysis systems. 1. Introduction, theory, and technology. Anal Chem 74:2623–2636

Rudnyi EB, Korvink JG (2002) Review: automatic model reduction for transient simulation of MEMS-based devices. Sensors Update 11(1):3–33

Sikanen T, Zwinger T, Tuomikoski S, Franssila S, Lehtiniemi R, Fager CM, Kotiaho T, Pursula A (2008) Temperature modeling and measurement of an electrokinetic separation chip. Microfluid Nanofluid 5(4):479–491

Yang YJJ, Kuo CW (2008) Generating scalable and modular macromodels for microchannels using the Galerkin-based technique. IEEE Trans Comput Aided Design Integr Circuits Syst 27(9):1545–1554

Yang YJ, Shen KY (2005) Nonlinear heat-transfer macromodeling for MEMS thermal devices. J Micromech Microeng 15(2):408–418