

# Generating Efficient Dynamical Models for Microelectromechanical Systems from a Few Finite-Element Simulation Runs

Elmer S. Hung and Stephen D. Senturia, *Fellow, IEEE*

**Abstract**—In this paper, we demonstrate how efficient low-order dynamical models for micromechanical devices can be constructed using data from a few runs of fully meshed but slow numerical models such as those created by the finite-element method (FEM). These reduced-order macromodels are generated by extracting global basis functions from the fully meshed model runs in order to parameterize solutions with far fewer degrees of freedom. The macromodels may be used for subsequent simulations of the time-dependent behavior of nonlinear devices in order to rapidly explore the design space of the device. As an example, the method is used to capture the behavior of a pressure sensor based on the pull-in time of an electrostatically actuated microbeam, including the effects of squeeze-film damping due to ambient air under the beam. Results show that the reduced-order model decreases simulation time by at least a factor of 37 with less than 2% error. More complicated simulation problems show significantly higher speedup factors. The simulations also show good agreement with experimental data. [399]

**Index Terms**—Karhunen-Loève decomposition, macromodels, microelectromechanical simulation, principal component analysis, reduced-order models, squeeze-film damping.

## I. INTRODUCTION

THE development of increasingly complex microelectromechanical systems (MEMS) demands sophisticated simulation techniques for design and optimization [1]. MEMS devices typically involve multiple coupled energy domains and media that can be modeled using partial differential equations (PDE's). Often the functionality of these devices can only be captured with time-dependent nonlinear PDE's.

Traditional finite-element methods (FEM's) can be used for explicit dynamical simulations of PDE's, but time-dependent FEM's are usually computationally very intensive, making them difficult to use when a large number of simulations are needed, especially if multiple devices are involved in a system.

Manuscript received November 10, 1998. This work was supported by the Defense Advanced Research Project Agency under contract J-FBI-95-215. Subject Editor, W. N. Sharpe, Jr.

E. S. Hung was with the Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, MA 02139 USA. He is now with the Xerox Palo Alto Research Center, Palo Alto, CA 94304 USA (e-mail: elmer@alum.mit.edu).

S. D. Senturia is with the Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, MA 02139 USA.

Publisher Item Identifier S 1057-7157(99)06543-9.

Therefore, a major current goal of modeling and simulation research is to develop efficient methods of creating accurate reduced-order dynamical models that capture the same information contained in the original PDE's (or, equivalently, in a fully meshed dynamic FEM simulation), but in a form that can be used for fast dynamical simulations in the context of a circuit- or system-level simulation environment. These reduced-order models are often referred to as *macromodels*.

Previous MEMS macromodeling efforts have investigated lumped-parameter techniques [2], [3]; however, it is often difficult to construct accurate lumped-element models for continuous systems, especially when arbitrary geometries are involved. It may also be necessary to separately model effects from dissipative and energy-conserving behavior. Another approach uses a linear analysis to generate normal modes which are used as the basis functions for the model. These modes have been used, for example, to provide a reduced-order set of generalized coordinates for nonlinear capacitance-based simulations of electrostatically actuated microstructures [4], [5]. However, linear modes may not adequately capture all the features of nonlinear behavior [6]. Also, static modal functions are sometimes not readily available.

In this paper, we describe a procedure in which a few finite-element or finite-difference simulations are used to create a reduced-order macromodel that permits fast simulation of MEMS devices while capturing most of the accuracy and flexibility of the full model. Although this approach requires an initial overhead cost to run FEM simulations, once the macromodel is generated, it may be used to efficiently carry out any further simulations involving the device. This strategy could be used to greatly speed up multiple simulations of a single device for exploration of a design space, or could be used as input to system-level simulators for designing systems with many coupled devices.

FEM's rely on highly localized interpolation functions (or *basis functions*) to approximate the solution to PDE's. These local basis functions are generated by meshing the domain of interest and parameterizing the desired solution locally on each mesh element. This parameterized solution converts a continuous (PDE) problem to a coupled system of ordinary differential equations (ODE's) that can be integrated in time. The resulting ODE system usually has many degrees of freedom (perhaps several variables per mesh element). If a fine mesh is required, the problem size grows rapidly, with a

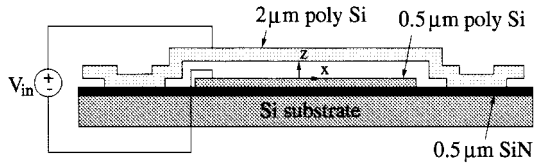


Fig. 1. Fixed-fixed beam pull-in experiment setup [14]. For analysis purposes, we assume that the  $x$  and  $y$  axes are oriented parallel to the length and width of the beam, respectively, and  $z$  is directed up, perpendicular to the substrate.

corresponding rapid growth in computational cost for explicit dynamic simulation.

In contrast to the FEM approach, global basis functions can often be used to capture the solution with fewer degrees of freedom. For example, Fourier decompositions [7] or modal functions [4] may be used. However, it is usually difficult to determine, *a priori*, an optimum set of such functions, particularly when irregular geometries are involved.

This work seeks a middle path, focusing on obtaining macromodels based on global basis functions generation from an approach that is mathematically equivalent to Karhunen-Loève analysis of a small but representative ensemble of dynamic FEM runs. Previous work has shown that Karhunen-Loève decompositions can be used to generate such basis functions for nonlinear dynamical system simulations for turbulence problems in fluid mechanics [8], [9], fluid-structure interactions in aerodynamical systems [10], [11], and in chemical engineering systems [12].

Here we use singular value decomposition (SVD) to generate global basis functions from explicit FEM results and generate the macromodel by the use of Galerkin method with the original governing PDE's. We discuss how basis function macromodeling techniques can contribute to MEMS simulation technology and use the method to simulate a problem involving squeeze-film (air) damping of a microbeam. This approach was first presented in [13].

## II. CASE STUDY EXAMPLE

In order to illustrate the macromodel technique, we first introduce a case study problem that will be referred to throughout the discussion. The device (Fig. 1) consists of a deformable elastic beam microstructure that is electrostatically pulled in by an applied voltage waveform. The time it takes for the beam to pull in is highly sensitive to the air pressure under the beam. In fact, this structure has been proposed for use as a pressure sensor [14].

Simulating the time-dependent dynamics of the device involves a nonlinear squeeze-film damping problem with mechanical, electrostatic, and fluidic components. Efficient simulations of squeeze-film damping problems are an active topic of research [15], [16] and are important for a variety of applications in order to control moving structures and to determine how fast microstructures can be moved in air.

The pull-in time pressure sensor device can be modeled by coupling a 1-D elastic beam equation with electrostatic force [(1)] and the 2-D compressible isothermal squeeze-film

Reynold's equation [17] for air damping with slip flow [(1)]<sup>1</sup>:

$$EI \frac{\partial^4 z}{\partial x^4} - S \frac{\partial^2 z}{\partial x^2} = F_{\text{elec}} + F_{\text{air}} - \rho \frac{\partial^2 z}{\partial t^2} \quad (1)$$

$$\nabla \cdot ((1 + 6K)z^3 p \nabla p) = 12\eta \frac{\partial(pz)}{\partial t} \quad (2)$$

where  $F_{\text{elec}} = -\epsilon_0 w V^2 / (2z^2)$  is the electrostatic force,  $F_{\text{air}} = \int_0^w p dy$  is the mechanical load from the squeezed air film,  $z(x, t)$  is the height of the beam above the substrate, and  $p(x, y, t)$  is the air pressure under the beam. Other parameters include Knudsen's number  $K(x, t) = \lambda/z$ , the mean-free path of air  $\lambda = 0.064 \mu\text{m}$ , elastic modulus  $E = 149 \text{ GPa}$ , moment of inertia  $I = wh^3/12$ , beam width  $w = 40 \mu\text{m}$ , length  $\ell = 610 \mu\text{m}$ , thickness  $h = 2.2 \mu\text{m}$ , undeflected gap  $z_0 = 2.3 \mu\text{m}$ , stress  $[S/(hw)] = -3.7 \text{ MPa}$ , density  $[\rho/(hw)] = 2330 \text{ kg/m}^3$ , and air viscosity  $\eta = 1.82 \times 10^{-5} \text{ kg/(m} \cdot \text{s)}$ . The beam is modeled with fixed supports and the fluid system is assumed to be open (ambient pressure) along the sides of the beam and closed (no flow) at the ends of the beam. All parameters are measured or are from the literature, except for stress which is matched to the measured pull-in voltage  $V_{\text{PI}} = 8.76 \text{ V}$  as in [18].

Explicit large-amplitude dynamical simulations of these coupled nonlinear equations using FEM's have been previously reported [19], with good agreement between model and experiment.

## III. BASIS FUNCTION STRATEGY

Let us now consider the simulation strategy. We would like to simulate a dynamical system described by a PDE of the form

$$L(u) = f \quad (3)$$

where  $L$  is a differential operator (possibly nonlinear), and  $u$  is a vector of state variables. For simplicity, we assume that the state solution  $u(x, t)$  and forcing term  $f(x, t)$  are functions of time  $t$  and a spatial variable  $x$ . Given  $f$  and  $L$ , we would like to determine  $u$ .

We begin by defining a space  $H$  of square integrable functions with dot product  $(v, w) = \int v^T(x)w(x)dx$  and norm  $|v| = (v, v)^{1/2}$ . We would like to approximate the desired PDE solution  $u(x, t)$  in a separable form as a series expansion of time varying coefficients  $\alpha_i(t)$  and spatially varying basis functions  $a_i(x)$

$$\hat{u}(x, t) = \sum_{i=1}^N \alpha_i(t) a_i(x) \quad (4)$$

where  $\hat{u}(x, t)$  is the approximation for  $u(x, t)$ . The choice of orthogonal basis functions  $a_i(x)$  will be addressed in Section IV below.

<sup>1</sup>The compressible isothermal Reynold's equation [(2)] can be derived from the Navier-Stokes, continuity, and ideal-gas equations by assuming 1) isothermal conditions, 2) a pressure that is constant with  $z$  (across the gap), 3) negligible inertial effects, and 4) negligible fluid velocity in the  $z$  direction (perpendicular to the substrate). Because the dimensions of interest are beyond the limit where atmospheric air can be modeled accurately as a continuous fluid, slip flow boundary conditions are used to model the device, parameterized by Knudsen's number  $K$ . See [17] for a derivation.

Assuming for the moment that the basis functions are known, the Galerkin method can then be used to specify the equations of motion for the coefficients  $\alpha_i(t)$ . This method requires that the PDE residual  $L(\hat{u}) - f$  be orthogonal to each of the basis functions in  $H$ . In other words, the Galerkin condition requires that

$$(a_i, L(\hat{u}) - f) = \int \alpha_i^T (L(\hat{u}) - f) dx = 0 \quad (5)$$

for all  $i \in \{1, \dots, N\}$ .

Now let us apply these ideas to the pull-in time pressure sensor. In this case, the state solution  $u(x, t)$  consists of the time histories of both pressure  $p(x, y, t)$  and displacement  $z(x, t)$ . Because of the formulation of the governing PDE's, we choose to write separate approximations for pressure  $\hat{p}(x, y, t)$  and displacement  $\hat{z}(x, t)$  as follows:

$$\hat{p}(x, y, t) = p_a + \sum_{i=1}^N \alpha_i(t) a_i(x, y) \quad (6)$$

$$\hat{z}(x, t) = z_0 + \sum_{i=1}^M \beta_i(t) b_i(x) \quad (7)$$

where  $p_a$  is the ambient pressure,  $z_0$  is the undeflected gap,  $N$  and  $M$  are the number of pressure and displacement basis functions we would like to use, and  $a_i(x, y)$  and  $b_i(x)$  are scalar basis functions for pressure and displacement, respectively.

Using the Galerkin method, the PDE's in (1) and (2) can be reduced to coupled matrix ODE's in terms of the basis function approximations above. A derivation of these equations appears in Appendix I. The Euler beam equation (1) results in the following ODE:

$$M\ddot{\beta} + K\beta + f = 0 \quad (8)$$

where

$$\begin{aligned} M_{ij} &= \int_L \rho b_i b_j dx \\ K_{ij} &= \int_L \left( EI \frac{\partial^2 b_i}{\partial x^2} \frac{\partial^2 b_j}{\partial x^2} + S \frac{\partial b_i}{\partial x} \frac{\partial b_j}{\partial x} \right) dx \\ f_i &= \int_L b_i (F_{\text{elec}} + F_{\text{air}}) dx \end{aligned}$$

where  $L$  indicates integration along the length of the beam. The Reynolds equation (2) results in the following ODE:

$$A\dot{\alpha} + B\alpha + c = 0 \quad (9)$$

where

$$\begin{aligned} A_{ij} &= \int_{\Omega} 12\eta a_i a_j \hat{z} dx dy \quad c_i = \int_{\Omega} 12\eta p_a a_i \frac{\partial \hat{z}}{\partial t} dx dy \\ B_{ij} &= \int_{\Omega} \left\{ (1 + 6K) \hat{z}^3 \hat{p} \left( \frac{\partial a_i}{\partial x} \frac{\partial a_j}{\partial x} + \frac{\partial a_i}{\partial y} \frac{\partial a_j}{\partial y} \right) \right. \\ &\quad \left. + 12\eta a_i a_j \frac{\partial \hat{z}}{\partial t} \right\} dx dy \end{aligned}$$

where  $\Omega$  indicates integration along the beam area. Equations (8) and (9) constitute the macromodel formulation. The equations are integrated numerically to simulate dynamic behavior.

Note that these equations are the same general form as those used for a standard FEM simulation, but because of the model-order reduction with global basis functions, constitute a much smaller computational problem.

The main idea behind the macromodel is that the number of ODE's needed to simulate the system has been reduced from perhaps many thousands in the case of the full FEM simulation, to just a few basis function coordinates. Thus the macromodel simulation can be very efficient computationally compared to the FEM model. Note also that many of the terms in the matrix ODE's above can be precomputed once the basis functions are known [e.g.,  $M_{ij}$  in (8)]. The terms that cannot be precomputed correspond to the nonlinear terms in the original PDE (clearly, for linear PDE's, the resulting ODE's would also be linear). An example of this would be  $c_i$  in (9). Since  $\hat{z}$  depends on time,  $c_i$  must be recomputed at every time step during the ODE numerical integration. Note that the computation of  $c_i$  requires a spatial integration in the original coordinate system,  $(x, y)$ .

As one might expect, much of the time spent in simulating the macromodel equations typically involves evaluation of the nonlinear terms that cannot be precomputed. However, Gabbay and Senturia [5] show that one can sometimes improve performance by precomputing nonlinear terms in the Galerkin formulation. Specifically, in [5] the nonlinear electrostatic force term  $F_{\text{elec}}$  from (8) is sampled in basis function coordinates and a rational function approximation for  $F_{\text{elec}}$  is computed in terms of the basis function coordinates. This makes the computation of  $F_{\text{elec}}$  very efficient during the numerical integration stage; however, it requires initial effort to generate the rational function approximation.

Note also that instead of using independent basis functions for pressure and displacement, it is possible to analyze basis functions for a combined state vector  $u_c(x, y, t) = [z(x, t) \ p(x, y, t)]^T$ . This may produce benefits in reducing the number of basis functions needed to characterize the dynamics of the system because it captures coupling between displacement and pressure. In this case, however, independent displacement and pressure basis functions make the Galerkin derivation simpler and also makes sense given the physics of the problem.

#### IV. GENERATING THE BASIS FUNCTIONS

We now describe how the basis functions are chosen. First the system dynamics are simulated using a slow but accurate technique such as finite elements or finite differences. An ensemble of runs may be used to suitably characterize the operating range of the device.

The spatial distributions of each state variable  $u(x, t)$  are then sampled at a series of  $N_s$  different times during these simulations, and the sampled distributions are stored as a series of vectors,  $\{u_i\}$ , where each  $u_i$  corresponds to a particular "snapshot" in time.

For example, for the microbeam problem, the pull-in dynamics of the beam are simulated using a finite-difference method for an ensemble of step voltages. To determine pressure basis functions, we take a series of snapshots,  $\{u_i\}$ , of

the pressure distributions at various times where the entries in each  $u_i$  vector correspond to pressures at a different node of the finite-difference mesh.

Now suppose we would like to pick  $N$  orthogonal basis functions,  $\{a_1, \dots, a_N\}$ , in order to represent the observed state distributions as closely as possible. One way to do this is to attempt to minimize the following quantity:

$$\sum_{i=1}^{N_s} |u_i - \text{proj}(u_i, \text{span}\{a_1, \dots, a_N\})|^2 \quad (10)$$

where  $\text{proj}(v, W)$  is the projection of the vector  $v$  onto subspace  $W$ . In other words, we minimize a least squares measure of the “error” distances between the observed states and the basis function representation of those states.

It turns out that this can be accomplished quite simply by taking the SVD of the matrix  $U$ , whose columns are  $u_i$ . The SVD of  $U$  is given as follows:

$$U = V\Sigma W^T \quad (11)$$

where  $\Sigma = \text{diag}(\sigma_1, \dots, \sigma_N)$  is a diagonal matrix,  $V$  and  $W$  are orthonormal matrices,  $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_N \geq 0$ , and  $N < N_s$ . In Appendix II we show that the proper orthogonal basis functions  $\{a_i\}_{i=1}^N$  minimizing (10) can be chosen by setting  $a_i = v_i$  for  $i = \{1, 2, \dots, N\}$  where  $v_i$  are the columns of  $V$ .

SVD decompositions have previously been used for generating basis functions for image compression and analysis [20]. The method also turns out to be mathematically equivalent to the Karhunen–Loève approach described in [8]–[12] for generating reduced order dynamical models (see [21] for a derivation).

Note that the columns of  $V$  are the eigenvectors of  $UU^T$  and the columns of  $W$  are the eigenvectors of  $U^TU$  [22]. It turns out that the standard formulation of the Karhunen–Loève approach is equivalent to explicitly computing the eigenvalues and eigenvectors of  $U^TU$  along with substitution back into a matrix multiplication in order to find the basis functions. However, there are algorithms for computing the SVD of  $U$  (e.g., [22]) which are numerically more robust than computing the  $U^TU$  product directly. Thus, we suggest that computation of the SVD of  $U$  is a preferred approach, in practice, for generating the basis functions.

## V. RESULTS

### A. Macromodel Accuracy

We now present simulation results for the pull-in time pressure sensor example. The dynamics of the beam are first simulated using a finite-difference analysis. A quarter of the beam is initially meshed using a  $20 \times 10$  node 2-D grid. The state at each node consists of three quantities:  $z$ ,  $\partial z/\partial t$ , and  $p$ . Since  $z$  and  $\partial z/\partial t$  are simulated in 1-D, this results in  $20 \times 10 + 2 \times 20 = 240$  coupled nonlinear ODE’s which must be integrated in time.

Basis functions are generated for pressure and displacement based on runs of the finite-difference code for an ensemble

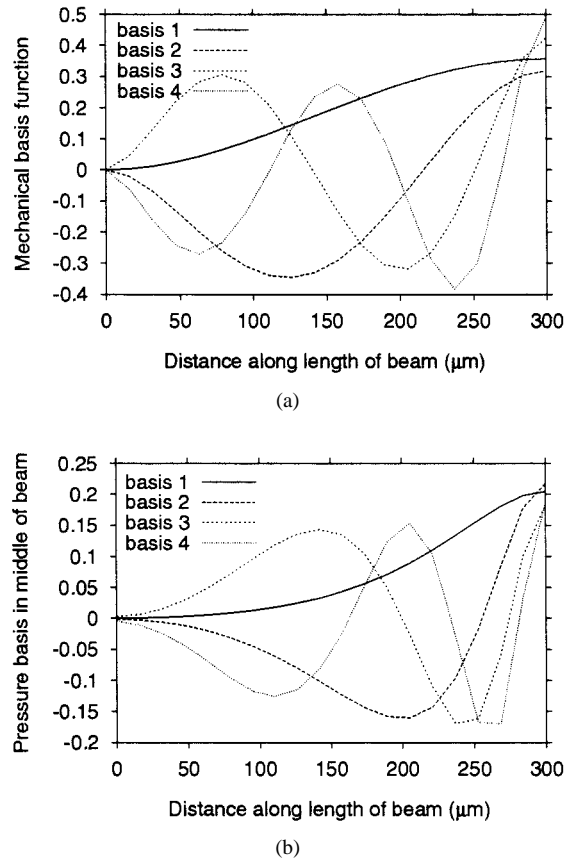


Fig. 2. Basis functions for (a) displacement  $z(x, t)$  and (b) pressure  $p(x, y, t)$  (holding  $y$  constant along the center of the beam). Note that only half of the beam is being displayed.

of four different step voltages:  $V_{in} = 9$  V, 10 V, 12 V, 16 V. One hundred samples of pressure and displacement are taken during these four runs at fixed time intervals. These samples are used to generate the basis functions. The resulting basis functions for displacement and pressure are shown in Fig. 2. As one might expect, the basis functions for displacement look similar to the linear modes for the system. However, the pressure basis functions show a higher spatial frequency (in  $x$  and  $y$ ) toward the middle of the beam, reflective of the fact that pressure is much higher toward the center of the beam as the beam pulls down and the gap decreases.

We then generate macromodels using the basis functions extracted from the ensemble of finite-difference simulations. Fig. 3 shows a comparison of the maximum beam deflection versus time for both the macromodel and the full finite-difference simulation for a 10-V step voltage input. The macromodel matches the finite-difference simulations extremely well, especially for  $M \geq 2$  and  $N \geq 2$  (as few as six ODE’s) where  $M$  and  $N$  are the number of mechanical and pressure basis functions used, respectively. In general, the deflection error is less than 1% of the undeflected gap size ( $0.02 \mu\text{m}$ ).

The accuracy of the macromodel is especially apparent in comparison to the performance of lumped-element methods. For example, linear damping is often used (e.g., [14]) to simulate the first-order effects of squeeze-film damping in microstructures. Fig. 4 shows results from a linear damping

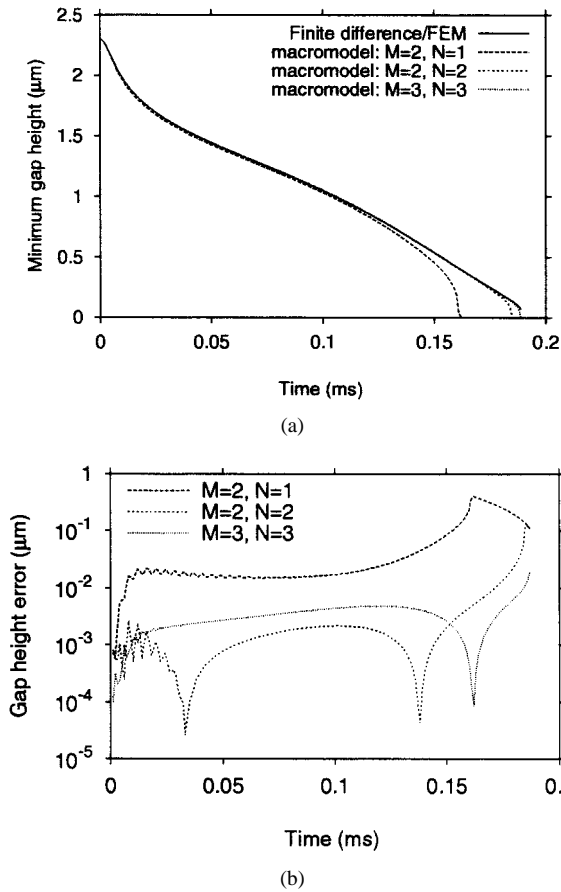


Fig. 3. (a) Comparison of beam deflection versus time simulation results for a 10-V step input.  $M$  and  $N$  are the number of mechanical and pressure basis functions used in the macromodel, respectively. Macromodel results are almost indistinguishable from finite-element simulation for  $M \geq 2$  and  $N \geq 2$ . By contrast, the linear damping simulation shows large errors. (b) Macromodel deflection error versus time for the simulation depicted in (a). Error is measured with respect to the finite-difference simulation.

simulation of the pull-in time pressure sensor. In this simulation, the damping force  $F_d(x, t)$  at each point on the beam is proportional to the velocity of the beam at that point

$$F_d(x, t) = -c \frac{\partial z}{\partial t}$$

where  $c$  is a damping constant. In this case, the damping constant was tuned to match the pull-in time from the finite-difference simulation. Fig. 4(a) shows displacement versus time results for a 10-V step input (compare with Fig. 3), and Fig. 4(b) shows results for a 14-V 10-kHz sine wave input (compare with Fig. 6). We see that the linear damping simulation shows large errors compared with the macromodel, illustrating the type of problems that can result from using linear models to simulate large-amplitude nonlinear squeeze-film damping effects.

The results of the macromodel also correspond well to experimental data from [19] (Fig. 5). A series of step input voltages are applied, and the pull-in time is measured and plotted for both experiment and simulation. Note that the finite-difference results are indistinguishable from the macromodel for these series of step input simulations. The relatively small errors between experiment and simulation may be due to

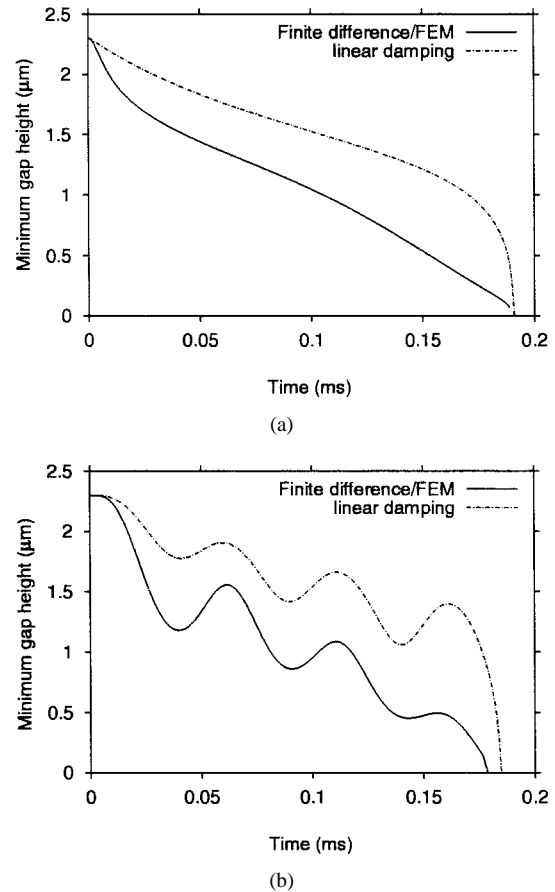


Fig. 4. Beam deflection versus time results for a linear damping simulation compared to the finite-difference model. (a) Results for a 10-V step input. (b) Results for a 14-V 10-kHz sine wave voltage input. The simulations show that first-order lumped element techniques can result in large errors compared to the macromodel results (Figs. 3 and 6).

unmodeled effects such as compliant supports and strain-stiffening of the beam. Note that the order-of-magnitude increase of pull-in time in air and the change in critical pull-in voltage due to inertial effects are well modeled by both simulations. This demonstrates that these types of simulations can achieve reasonable accuracy for modeling transient squeeze-film damping effects.

### B. Changing Device Parameters and Input

One question is, how flexible is the basis-function technique? If the parameters or inputs to a device change, is it necessary to rerun FEM simulations to generate new basis functions for the macromodel? This is important because a primary motivation for the use of the macromodeling technique for MEMS devices is that a single macromodel may be used to run many simulations, exploring the design space of a device or system without having to recompute FEM simulations.

To test this, the same macromodel described above, generated using basis functions from the step voltage input FEM runs, was used to simulate the response to a 14-V 10-kHz sine wave voltage input. The results are compared to a full finite-difference solution for the 14-V 10-kHz sine wave stimulus (Fig. 6). Again, macromodel results for  $M \geq 2$  and  $N \geq 2$  match the finite-element run extremely well, while the linear

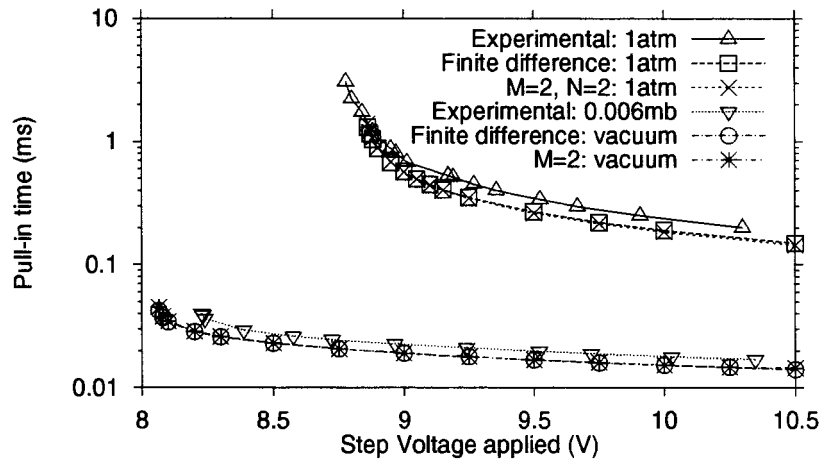
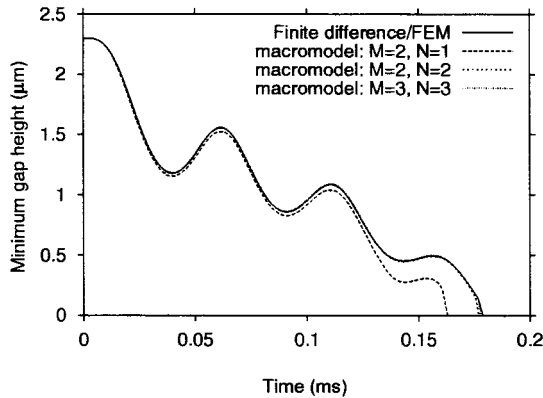
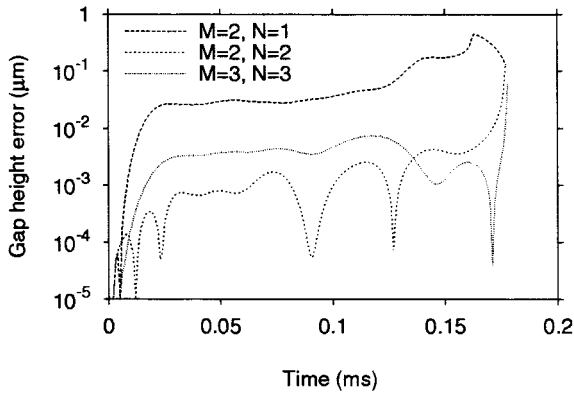


Fig. 5. Comparison of pull-in time macromodel and finite-difference simulation results with experimental measurement in 1 atm pressure and in vacuum.



(a)



(b)

Fig. 6. (a) Comparison of beam deflection versus time simulation results for a 14-V 10-kHz sine wave voltage input.  $M$  and  $N$  are the number of mechanical and pressure basis functions used in the macromodel, respectively. (b) Macromodel deflection error versus time for the simulation depicted in (a). Error is measured with respect to the finite-difference simulation.

damping model performs poorly. This shows that without changing basis functions, the macromodel is quite flexible at handling inputs to the device that are substantially different from the step inputs the macromodel was trained on.

Another test is to change the physical parameters of the device and see how well the macromodel works. Fig. 7 shows

that the basis functions also work for scaled geometry. The figure shows a comparison between finite-difference transient simulations and the macromodel for a beam with scaled dimensions ( $L = 400 \mu\text{m}$ ,  $w = 30 \mu\text{m}$ , and zero residual stress for the new beam, as opposed to  $L = 610 \mu\text{m}$ ,  $w = 40 \mu\text{m}$ , and stress =  $-3.7 \text{ MPa}$  for the original beam). The macromodel used for the simulation of the new beam is generated using linearly scaled versions of the basis functions from the original 610- $\mu\text{m}$ -long beam. The pull-in voltage for the new beam is near 22 V, and the simulation shown in Fig. 7 is for a 25-V step input. Macromodel results for  $M \geq 2$  and  $N \geq 2$  still match the finite-difference run well, despite the change in parameters. While we have not studied the accuracy limits on scaling of dimensions over a large workspace, we have demonstrated that the macromodel for the pull-in time pressure sensor is quite accurate for moderate device parameter changes.

### C. Macromodel Efficiency

Now we address the computational efficiency of the macromodel. Table I shows explicitly the speedups and relative errors we obtain from using the macromodel technique compared with the full finite-difference model. Note that for  $M = 2$  and  $N = 2$ , the macromodel achieves a factor of 37 speedup in computation time over the finite-difference model while suffering pull-in time errors of less than 2%.

The macromodel efficiency results are even more impressive for larger simulations. The finite-difference simulation was initially run using a  $20 \times 10$  node grid because this is the minimum grid size at which reasonably accurate results can be obtained. In more complicated systems, however, much larger grid sizes will likely be necessary. Thus, it is important to know how the performance of the macromodel scales with increasing grid size. Fig. 8 shows how much the macromodel speeds up pull-in simulations for different grid sizes. As the grid size increases, both the finite-difference and macromodel simulations take longer to compute. Nonlinear terms in the macromodel involve spatial integrations/sums that take longer to compute. However, the speed up factor for the macromodel

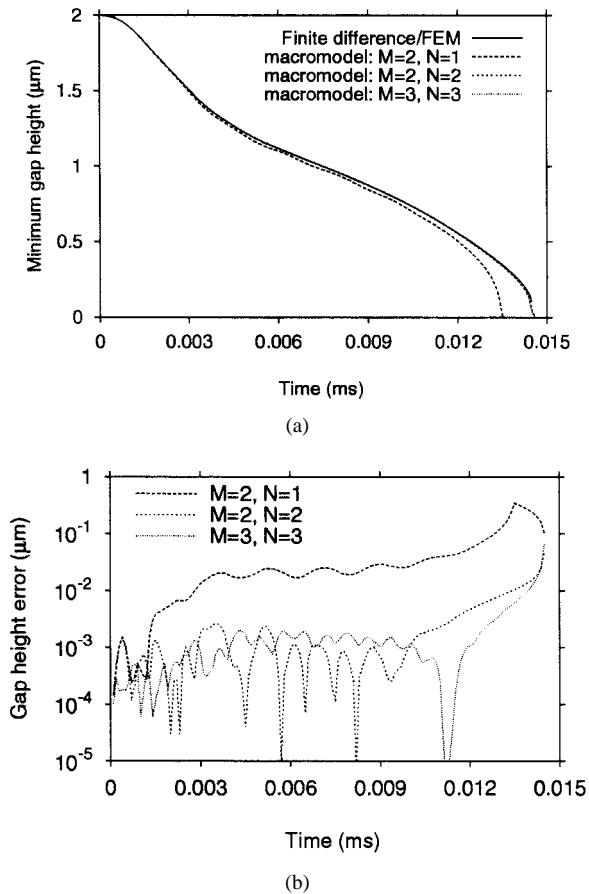


Fig. 7. (a) Comparison of beam deflection versus time simulation results for a 25-V step input for a beam with  $L = 400 \mu\text{m}$ ,  $w = 30 \mu\text{m}$ , and no residual stress. Macromodel is generated using scaled basis functions from finite-difference simulations of a beam with  $L = 610 \mu\text{m}$ ,  $w = 40 \mu\text{m}$ , and stress =  $-3.7 \text{ MPa}$ .  $M$  and  $N$  are the number of mechanical and pressure basis functions used in the macromodel, respectively. (b) Macromodel deflection error versus time for the simulation depicted in (a). Error is measured with respect to the finite-difference simulation.

TABLE I

PERFORMANCE COMPARISON BETWEEN MACROMODELS AND  $20 \times 10$  NODE FINITE-DIFFERENCE SIMULATION ON CALCULATING PULL-IN TIMES FOR THE ENSEMBLE OF STEP VOLTAGES  $V_{in} = 9 \text{ V}, 10 \text{ V}, 12 \text{ V}, 16 \text{ V}$ . ERROR AND SPEED UP FACTORS ARE CALCULATED WITH RESPECT TO THE FULL FINITE-DIFFERENCE MODEL.  $M$  AND  $N$  ARE THE NUMBER OF MECHANICAL AND PRESSURE BASIS FUNCTIONS USED IN THE MACROMODEL, RESPECTIVELY. SIMULATIONS WERE CARRIED OUT ON A SPARC ULTRA 1 MODEL 170

method	number of ODEs	average error	run time (min:sec)	speed up factor
finite difference	240	0.0	31:04	1.0
macromodel M=2 N=1	5	14%	0:18	105
macromodel M=2 N=2	6	1.7%	0:51	37
macromodel M=3 N=3	9	1.0%	3:08	10
macromodel M=4 N=4	12	<1%	7:19	4.2

compared with the finite-difference model increases with larger grid sizes. Thus, the macromodel technique should become even more helpful for more complicated systems where larger mesh sizes are needed.

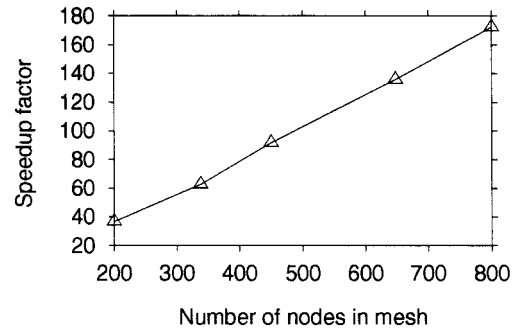


Fig. 8. The effect of changing the grid size of the finite-difference model. Speed up factor is the time required to simulate the ensemble of step voltages  $V_{in} = 9 \text{ V}, 10 \text{ V}, 12 \text{ V}, 16 \text{ V}$  for the finite-difference model divided by the time required to perform the corresponding macromodel simulation using two mechanical and two pressure basis functions.

## VI. DISCUSSION AND CONCLUSIONS

In this paper, we describe a method to speed up the dynamical simulation of micromechanical devices by using a reduced-order macromodel generated from finite-element or finite-difference simulations of the device. This method is demonstrated on a pull-in time pressure sensor device where the macromodel speeds up simulations by a factor of 37 over a finite-difference technique with less than 2% error.

The method is based on generating global orthogonal basis functions from time samples of the state variables from the FEM runs. Using the Galerkin method, these basis functions are combined with the governing PDE of the system to create the macromodel.

For the pull-in time pressure sensor device, it is shown that a single set of basis functions can be used to generate macromodels which accurately simulate the dynamics of the device over a range of different input waveforms and device parameter and geometry changes.

There can be complicating issues when basis function techniques are used to simulate general nonlinear dynamical systems. One example is a system exhibiting the intermittent chaotic behavior studied in [23]. Problems result when the sampled state-space trajectories from which the basis functions are generated do not visit all of the relevant parts of state space often enough. This could occur, for example, if there are multiple attractors, if there are rare intermittent fluctuations in the system, or if there are nearby bifurcations in parameter space which result in trajectories that explore whole new regions of state space.

It is important, in general, to be sure that the sample states used to generate the basis functions are a representative collection of the qualitative behavior encountered by the system. With MEMS devices, this does not appear to be a difficult condition to satisfy as long as one is reasonably careful when setting up the simulations used to determine the basis functions. By design, most MEMS systems are relatively well behaved, and the design spaces for most systems do not include problematic types of bifurcations.

With the caveats described above, the use of basis function macromodeling techniques for MEMS systems seems to be quite promising for rapidly simulating the dynamics and

exploring the design space of MEMS devices. Presently, the overhead for generating the initial FEM simulation is quite high. However, as coupled-domain simulation tools improve, making the initial FEM simulations more accessible, these reduced-order macromodeling techniques will become increasingly attractive for speeding up the design and optimization process, especially as the complexity of the MEMS components increases.

## APPENDIX I

### MACROMODEL FOR PULL-IN TIME PRESSURE SENSOR

In this appendix, we show how to derive the equations of motion for the macromodel simulation of the pull-in time pressure sensor. The original model for the system involves the coupled elastic beam equation with the compressible isothermal Reynold's equation for the air damping [(1) and (2)]. We derive the macromodel form of the Reynolds equation. Note that this derivation follows an approach commonly used for FEM models (see, e.g., [24]).

The macromodel uses basis function expansions as approximations for pressure,  $\hat{p}(x, y, t)$ , and displacement,  $\hat{z}(x, t)$ , given in (6) and (7) where  $\{a_i(x, y)\}$  are the basis functions for pressure. Substituting (6) into (2), Reynold's equation can be rewritten as

$$\frac{\partial}{\partial x} \left( (1 + 6K) \hat{z}^3 \hat{p} \frac{\partial \hat{p}}{\partial x} \right) + \frac{\partial}{\partial y} \left( (1 + 6K) \hat{z}^3 \hat{p} \frac{\partial \hat{p}}{\partial y} \right) = 12\eta \left( \hat{z} \frac{\partial \hat{p}}{\partial t} + \hat{p} \frac{\partial \hat{z}}{\partial t} \right).$$

Galerkin's condition (5) requires that

$$Z_i = \int_{\Omega} a_i \left[ \frac{\partial}{\partial x} \left( (1 + 6K) \hat{z}^3 \hat{p} \frac{\partial \hat{p}}{\partial x} \right) + \frac{\partial}{\partial y} \left( (1 + 6K) \hat{z}^3 \hat{p} \frac{\partial \hat{p}}{\partial y} \right) - 12\eta \left( \hat{z} \frac{\partial \hat{p}}{\partial t} + \hat{p} \frac{\partial \hat{z}}{\partial t} \right) \right] dx dy = 0 \quad (12)$$

for all  $i \in \{1, 2, \dots, N\}$  where  $\Omega$  represents the beam area. Let

$$F_x = (1 + 6K) \hat{z}^3 \hat{p} \frac{\partial \hat{p}}{\partial x} \quad \text{and} \quad F_y = (1 + 6K) \hat{z}^3 \hat{p} \frac{\partial \hat{p}}{\partial y}. \quad (13)$$

Substituting (13) into (12), and since  $a_i \frac{\partial F_x}{\partial x} = \frac{\partial}{\partial x} (a_i F_x) - \frac{\partial a_i}{\partial x} F_x$ , we have

$$Z_i = \int_{\Omega} \left[ \frac{\partial}{\partial x} (a_i F_x) - \frac{\partial a_i}{\partial x} F_x + \frac{\partial}{\partial y} (a_i F_y) - \frac{\partial a_i}{\partial y} F_y - 12\eta a_i \left( \hat{z} \frac{\partial \hat{p}}{\partial t} + \hat{p} \frac{\partial \hat{z}}{\partial t} \right) \right] dx dy.$$

From the component form of the Divergence Theorem, we know  $\int_{\Omega} \frac{\partial}{\partial x} (a_i F_x) dx dy = \oint_C (a_i F_x n_x) \cdot ds$ , where  $C$  represents the boundary of the beam area. Thus,

$$Z_i = \oint_C (a_i F_x n_x + a_i F_y n_y) \cdot ds - \int_{\Omega} \left[ \frac{\partial a_i}{\partial x} F_x + \frac{\partial a_i}{\partial y} F_y + 12\eta a_i \left( \hat{z} \frac{\partial \hat{p}}{\partial t} + \hat{p} \frac{\partial \hat{z}}{\partial t} \right) \right] dx dy. \quad (14)$$

Because of the symmetry of the problem, both contour integrals are zero. Substituting (13) into (14) we have that

$$Z_i = - \int_{\Omega} \left[ \frac{\partial a_i}{\partial x} (1 + 6K) \hat{z}^3 \hat{p} \frac{\partial \hat{p}}{\partial x} + \frac{\partial a_i}{\partial y} (1 + 6K) \hat{z}^3 \hat{p} \frac{\partial \hat{p}}{\partial y} + 12\eta a_i \left( \hat{z} \frac{\partial \hat{p}}{\partial t} + \hat{p} \frac{\partial \hat{z}}{\partial t} \right) \right] dx dy. \quad (15)$$

Substituting the basis function approximation for pressure (6) into (15), we have

$$Z_i = - \int_{\Omega} \left[ \frac{\partial a_i}{\partial x} (1 + 6K) \hat{z}^3 \hat{p} \left( \sum_{j=1}^N \alpha_j(t) \frac{\partial a_j}{\partial x} \right) + \frac{\partial a_i}{\partial y} (1 + 6K) \hat{z}^3 \hat{p} \left( \sum_{j=1}^N \alpha_j(t) \frac{\partial a_j}{\partial y} \right) + 12\eta a_i \left( \hat{z} \left( \sum_{j=1}^N \frac{\partial \alpha_j}{\partial t}(t) a_j \right) + \left( p_a + \sum_{i=1}^N \alpha_i(t) a_i(x, y) \right) \frac{\partial \hat{z}}{\partial t} \right) \right] dx dy. \quad (16)$$

Note that since Reynolds equation is nonlinear, not all instances of  $\hat{p}$  above are substituted for using the basis function expansion. The substitution choices are made in order to produce a macromodel that can be computed efficiently.

Since we require that  $Z_i = 0$  for all  $i \in \{1, 2, \dots, N\}$ , collecting terms from (16) and rewriting the equation in matrix form results in the following ODE:

$$A\dot{\alpha} + B\alpha + c = 0$$

where

$$A_{ij} = \int_{\Omega} 12\eta a_i a_j \hat{z} dx dy; \quad c_i = \int_{\Omega} 12\eta p_a a_i \frac{\partial \hat{z}}{\partial t} dx dy;$$

$$B_{ij} = \int_{\Omega} \left\{ (1 + 6K) \hat{z}^3 \hat{p} \left( \frac{\partial a_i}{\partial x} \frac{\partial a_j}{\partial x} + \frac{\partial a_i}{\partial y} \frac{\partial a_j}{\partial y} \right) + 12\eta a_i a_j \frac{\partial \hat{z}}{\partial t} \right\} dx dy.$$

This is the macromodel formulation which is integrated numerically.

A similar derivation can be carried through with the Euler beam equation (1) resulting in the matrix ODE given in (8).

## APPENDIX II

### THE SVD METHOD FOR CHOOSING BASIS FUNCTIONS

In this appendix, we show that basis functions that are optimal in a least-squares sense can be generated using an SVD of a matrix containing data from time snapshots of the relevant state variable.

Let  $u(x, t)$  be the state variable in the PDE of interest. Suppose that the spatial distributions of  $u(x, t)$  are sampled at a series of  $N_s$  different times  $\{t_1, t_2, \dots, t_{N_s}\}$  during these simulations, and the sampled distributions are stored as a series of vectors  $\{u_i\}$  where each  $u_i$  corresponds to a particular state

“snapshot” in time, and each entry of  $u_i$  corresponds to the state at a local node or mesh element.

For some  $N < r$ , we would like to choose basis functions,  $\{a_1, \dots, a_N\}$ , to minimize  $E_N$ , a sum of the squares measure of the error distances between the state variables and the basis function representations for those variables

$$E_N(a_1, \dots, a_N) = \sum_{i=1}^{N_s} |u_i - \text{proj}(u_i, \text{span}\{a_1, \dots, a_N\})|^2. \quad (17)$$

Below, we show that the optimal  $\{a_1, \dots, a_N\}$  can be chosen by taking the SVD of the matrix  $U$ , whose columns are the sampled state vectors  $u_i$ .

First, recall some results from linear algebra (see [22] for proofs).

*Theorem:* Given a real matrix  $U \in \mathbf{R}^{m \times n}$ , there exists an SVD of  $U$

$$U = V\Sigma W^T \quad (18)$$

into a diagonal matrix  $\Sigma = \text{diag}(\sigma_1, \dots, \sigma_r)$  and orthonormal matrices  $V$  and  $W$  where  $r = \min(m, n)$  and  $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r \geq 0$ .

Let  $v_1, \dots, v_m$  be the columns of  $V$  and  $w_1, \dots, w_n$  be the columns of  $W$ . It turns out that the  $\sigma_i$ 's are the square roots of the eigenvalues of  $UU^T$  and  $U^T U$ , while the  $v_i$ 's and  $w_i$ 's are the corresponding eigenvectors of  $UU^T$  and  $U^T U$ , respectively. Note also that we can write

$$U = \sum_{j=1}^r v_j \sigma_j w_j^T. \quad (19)$$

*Definition:* We define the Frobenius norm [22]  $\|U\|_F$  of matrix  $U$  as follows:

$$\|U_{ij}\|_F = \left( \sum_i \sum_j |u_{ij}|^2 \right)^{\frac{1}{2}}.$$

It is not difficult to show that the Frobenius norm of  $U$  is directly related to the SVD factorization in (18)

$$\|U_{ij}\|_F^2 = \sum_{i=1}^r \sigma_i^2. \quad (20)$$

*Claim:*  $E_N$  is minimized when  $a_i = v_i$  for  $i \in \{1, 2, \dots, N\}$  [where the  $v_i$ 's are the columns of  $V$  in (18)].

*Proof:* The proof contains two parts. First we show that  $E_N(v_1, \dots, v_N) = \sum_{i=N+1}^{N_s} \sigma_i^2$ . Then we show that  $E_N(a_1, \dots, a_N) \geq \sum_{i=N+1}^{N_s} \sigma_i^2$  for any orthonormal set of vectors  $\{a_1, \dots, a_N\}$ .

Given any orthonormal,  $\{a_1, \dots, a_N\}$ , since

$$\text{proj}(u_i, \text{span}\{a_1, \dots, a_N\}) = \sum_{j=1}^N a_j a_j^T u_i$$

we can use the Frobenius norm to rewrite  $E_N$  from (17) in matrix form

$$E_N(a_1, \dots, a_N) = \left\| U - \sum_{j=1}^N a_j a_j^T U \right\|_F^2 \quad (21)$$

where  $U$  is the matrix whose columns are  $u_i$ .

Substituting  $v_i$  for  $a_i$ , and using (21), (19), and (20), we have

$$\begin{aligned} E_N(v_1, \dots, v_N) &= \left\| U - \sum_{j=1}^N (v_j v_j^T U) \right\|_F^2 \\ &= \left\| \sum_{j=1}^r v_j \sigma_j w_j^T - \sum_{j=1}^N v_j v_j^T V \Sigma W^T \right\|_F^2 \\ &= \left\| \sum_{j=1}^r v_j \sigma_j w_j^T - \sum_{j=1}^N v_j \sigma_j w_j^T \right\|_F^2 \\ &= \sum_{i=N+1}^r \sigma_i^2. \end{aligned}$$

This completes the first part of the proof.

Now let  $Y = U - \sum_{j=1}^N a_j a_j^T U$  for some orthonormal set of bases,  $\{a_1, \dots, a_N\}$ . From (21) and (20),  $E_N$  must be the sum of the squares of the eigenvalues of  $Y^T Y$ .

But  $Y^T Y = U^T U + B$  for some matrix  $B$ , where  $\text{rank}(B) \leq N$  since  $\text{rank}(\sum_{j=1}^N a_j a_j^T U) \leq N$ .

Clearly, at least  $r - N$  eigenvalues of  $Y^T Y$  must be the same as eigenvalues of  $U^T U$ . Thus,

$$E_N(a_1, \dots, a_N) \geq \sum_{i=N+1}^r \sigma_i^2.$$

This proves the claim.  $\square$

#### ACKNOWLEDGMENT

The authors wish to thank J. Young, V. Rabinovich, R. Gupta, J. Mehner, M. Grétilat, L. Gabbay, and J. Phillips for useful discussions. They also thank the anonymous reviewers for helpful comments.

#### REFERENCES

- [1] S. D. Senturia, “CAD challenges for microsensors, microactuators, and microsystems,” *Proc. IEEE*, vol. 86, pp. 1611–1626, Aug. 1998.
- [2] H. Tilmans, “Equivalent circuit representation of electromechanical transducers: I. lumped-parameter systems,” *IEEE J. Electromech. Syst.*, vol. 6, no. 1, pp. 157–176, Mar. 1996.
- [3] T. Veijola, H. Kuisma, J. Lahdenperä, and T. Ryhänen, “Equivalent-circuit model of the squeezed gas film in a silicon accelerometer,” *Sens. Actuators*, vol. A48, pp. 239–248, 1995.

- [4] G. K. Ananthasuresh, R. K. Gupta, and S. D. Senturia, "An approach to macromodeling MEMS for nonlinear dynamic simulation," presented at ASME Int. Mechanical Engineering Congr. Exposition, Atlanta, GA, 1996.
- [5] L. D. Gabbay and S. D. Senturia, "Automatic generation of dynamic macro-models using quasistatic simulations in combination with modal analysis," in *Proc. Solid State Sensor and Actuator Workshop*, Hilton Head, SC, 1998, pp. 107–200.
- [6] L. D. Gabbay, "Computer aided macromodeling for MEMS, Ph.D. dissertation, Mass. Inst. Technol., Cambridge, MA, 1998.
- [7] H. S. Cheng and C. H. T. Pan, "Stability analysis of gas-lubricated, self-acting, plain, cylindrical, journal bearings of finite length, using Galerkin's method," *J. Basic Eng.*, pp. 185–192, Mar. 1965.
- [8] L. Sirovich, "Turbulence and the dynamics of coherent structures part I: Coherent structures," *Q. Appl. Math.*, vol. 45, pp. 561–571, 1987.
- [9] G. Berkooz, P. Holmes, and J. L. Lumley, "The proper orthogonal decomposition of turbulent flows," *Ann. Rev. Fluid Mech.*, vol. 25, pp. 539–575, 1993.
- [10] E. H. Dowell, K. C. Hall, and M. C. Romanowski, "Eigenmode analysis in unsteady aerodynamics: Reduced order models," *Appl. Mech. Rev.*, vol. 50, no. 6, pp. 371–386, 1997.
- [11] T. Kim, "Frequency-domain Karhunen-Loève method and its application to linear dynamic systems," *AIAA J.*, vol. 36, no. 11, pp. 2117–2123, 1998.
- [12] H. M. Park and D. H. Cho, "The use of the Karhunen-Loève decomposition for the modeling of distributed parameter systems," *Chem. Eng. Sci.*, vol. 51, no. 1, pp. 81–98, 1996.
- [13] E. S. Hung, Y.-J. Yang, and S. D. Senturia, "Low-order models for fast dynamical simulation of MEMS microstructures," in *Transducers'97*, pp. 1101–1104.
- [14] R. K. Gupta and S. D. Senturia, "Pull-in time dynamics as a measure of absolute pressure," in *Proc. MEMS 1997*, pp. 290–294.
- [15] C.-L. Chen and J. J. Yao, "Damping control of MEMS devices using structural design approach," in *Proc. IEEE Solid-State Sensor and Actuator Workshop*, Hilton Head, SC, 1992, pp. 72–75.
- [16] Y.-J. Yang, M.-A. Grétilat, and S. D. Senturia, "Effect of air damping on the dynamics of nonuniform deformations of microstructures," in *Transducers'97*, pp. 1093–1096.
- [17] B. J. Hamrock, *Fundamentals of Fluid Film Lubrication*. New York: McGraw-Hill, 1994.
- [18] P. M. Osterberg and S. D. Senturia, "M-TEST: A test chip for MEMS material property measurement using electrostatically actuated test structures," *IEEE J. Microelectromech. Syst.*, vol. 6, no. 2, pp. 107–118, 1997.
- [19] R. K. Gupta *et al.*, "Pull-in dynamics of electrostatically-actuated microstructures," *Semiconductor Research Corporation (SRC) TECHCON 1996*.
- [20] I. T. Jolliffe, *Principal Component Analysis*. New York: Springer-Verlag, 1986.
- [21] E. S. Hung, "Positioning, control, and dynamics of electrostatic actuators for use in optical systems," Ph.D. dissertation, Mass. Inst. Technol., Cambridge, MA, 1998.
- [22] G. H. Golub and C. F. Van Loan, *Matrix Computations*, 3rd ed. Baltimore: Johns Hopkins Press, 1996.
- [23] M. D. Graham and I. G. Kevrekidis, "Alternative approaches to Karhunen-Loève decomposition for model reduction and data analysis," *Comput. Chem. Eng.*, vol. 20, no. 5, pp. 495–506, 1996.
- [24] J. N. Reddy, *An Introduction to the Finite Element Method*. New York: McGraw-Hill, 1984.



**Elmer S. Hung** received the S.B., S.M., and Ph.D. degrees in electrical engineering and computer science from the Massachusetts Institute of Technology, Cambridge, in 1991, 1994, and 1998 respectively. His Ph.D. dissertation included work on the design of the Polychromator: a MEMS-based optical spectroscopy system for remote gas sensing.

He is currently a Member of Research Staff at the Xerox Palo Alto Research Center, Palo Alto, CA, where his research interests include work on distributed sensor systems for diagnostic information processing.



**Stephen D. Senturia** (M77-SM'91-F'93) received the B.A. degree in physics, *summa cum laude*, in 1961 from Harvard University, Cambridge, MA, and the Ph.D. degree in physics in 1966 from the Massachusetts Institute of Technology, Cambridge.

He is presently the Barton L. Weller Professor of Electrical Engineering at the Massachusetts Institute of Technology, having joined the department immediately after completing his education. His current principal research activities are the development of

CAD systems for design and simulation of microelectromechanical devices ("MEMS") and the use of microfabricated structures both for microsensor and microactuator applications, and for materials research. In 1982, he founded Micromet Instruments, Inc., and served on its Board of Directors until 1992.

Dr. Senturia is a founding Associate Editor of the IEEE/ASME JOURNAL OF MICROELECTROMECHANICAL SYSTEMS (since 1992), which he now serves as Senior Editor. He was the Solid-State Sensors Associate Editor of the IEEE TRANSACTIONS ON ELECTRON DEVICES (1985–1996) and has been an active participant in the organization and planning of major conferences on solid-state sensors and actuators since 1983. He is co-winner of an IR-100 Award for his work on Automated Reclamation of Urban Solid Waste, and of the 1988 Arthur K. Doolittle Prize of the Division of Polymeric Materials Science and Engineering of the American Chemical Society. Honor and professional societies include Phi Beta Kappa and Sigma Xi. He is a Trustee of the Transducer Research Foundation.