INTERNATIONAL JOURNAL FOR NUMERICAL METHODS IN ENGINEERING
Int. J. Numer. Meth. Engng 2005; **64**:789–818
Published online 23 August 2005 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/nme.1394

Elastic PMLs for resonator anchor loss simulation

David S. Bindel[1] and Sanjay Govindjee[2,*,†]

[1]*Department of Electrical Engineering and Computer Science, University of California Berkeley, Berkeley, CA 94720, U.S.A.*
[2]*Structural Engineering, Mechanics, and Materials, Department of Civil Engineering, University of California Berkeley, Berkeley, CA 94720, U.S.A.*

SUMMARY

Electromechanical resonators and filters, such as quartz, ceramic, and surface-acoustic wave devices, are important signal-processing elements in communication systems. Over the past decade, there has been substantial progress in developing new types of miniaturized electromechanical resonators using microfabrication processes. For these micro-resonators to be viable they must have high and predictable quality factors (Q). Depending on scale and geometry, the energy losses that lower Q may come from material damping, thermoelastic damping, air damping, or radiation of elastic waves from an anchor. Of these factors, anchor losses are the least understood because such losses are due to a complex radiation phenomena in a semi-infinite elastic half-space. Here, we describe how anchor losses can be accurately computed using an absorbing boundary based on a *perfectly matched layer* (PML) which absorbs incoming waves over a wide frequency range for any non-zero angle of incidence. We exploit the interpretation of the PML as a complex-valued change of co-ordinates to illustrate how one can come to a simpler finite element implementation than was given in its original presentations. We also examine the convergence and accuracy of the method, and give guidelines for how to choose the parameters effectively. As an example application, we compute the anchor loss in a micro disk resonator and compare it to experimental data. Our analysis illustrates a surprising mode-mixing phenomenon which can substantially affect the quality of resonance. Copyright © 2005 John Wiley & Sons, Ltd.

KEY WORDS: perfectly matched layer; anchor loss; resonator loss; high Q; semi-infinite half-space

*Correspondence to: S. Govindjee, Structural Engineering, Mechanics, and Materials, Department of Civil Engineering, University of California Berkeley, Berkeley, CA 94720, U.S.A.
†E-mail: govindjee@ce.berkeley.edu

Contract/grant sponsor: National Science Foundation; contract/grant number: ECS-0426660
Contract/grant sponsor: University of California MICRO program
Contract/grant sponsor: Sun Microsystems

Received 23 February 2005
Revised 2 May 2005
Accepted 20 May 2005

1. INTRODUCTION

Modern communication systems rely on high-frequency electromechanical resonators to act as frequency references and filters. Though designers currently use quartz, ceramic, and surface-acoustic wave devices, surface-micromachined microelectromechanical system resonators (MEMS resonators) in development offer an attractive alternative. Because they can be integrated into standard complementary metal-oxide semiconductor (CMOS) technology, MEMS resonators have the potential to use less area and power, and cost less money than existing commercial devices [1]. But to be viable, energy losses in these MEMS resonators must be minimized. The usual measure of this energy loss is the *quality factor* Q of a resonant peak, defined as

$$Q = 2\pi \left(\frac{\text{Stored energy}}{\text{Energy lost per period}} \right) \tag{1}$$

For an ideal linear single degree of freedom oscillator $Q = |\omega|/2\,\mathrm{Im}[\omega]$, where ω is the oscillator's complex-valued eigenvalue [2, p. 158]. Resonators in cell phone filters, for example, require Q values greater than 1000 for good performance, and higher values are preferable [1, 3].

Depending on scale, geometry, and materials, the energy losses that lower Q may come from material damping, air damping, thermoelastic damping, or radiation of elastic waves from an anchor [4]. While losses in low-frequency resonators are dominated by air damping, for which increasingly accurate compact models are available [5, 6], high-frequency disk resonators are known to have similar measured performance in vacuum or air [7]. Many resonators are also vacuum packed in order to reduce air damping effects. Thermoelastic damping is another frequently-cited source of losses at high frequencies [8–12]. In most cases, the damping is estimated by fitting parameters in a model originally developed by Zener [13–15]; unfortunately, this parameter-fitting makes it difficult to tell what should be attributed to thermoelastic effects and what should be attributed to other sources of damping with similar functional form. Though anchor damping is a recognized source of losses [4], there are relatively few MEMS papers (see e.g. References [16–18]) dealing with losses at the anchor.

Although it is not well studied, in several designs for high MHz or GHz frequency resonators, the dominant loss mechanism appears to be radiation of elastic energy through anchors. In these designs, the resonating device is much smaller than the silicon substrate on which it sits, and waves radiating from the anchor are so attenuated by the time they reflect from the sides of the microchip that the reflected waves are negligible. That is, the bulk of the chip can be modelled without loss as a semi-infinite half-space. To simulate the response of a semi-infinite domain, one usually employs boundary dampers, infinite elements, boundary integrals, or exact Dirichlet-to-Neumann (DtN) boundary conditions so that a domain of simulation can be finite and allow for the application of finite element or finite difference methods; see e.g. References [19, Chapter 8], [20–22]. Each of these methods truncates the simulation domain with an artificial boundary at which outgoing waves are absorbed. For an elastic half-space, Green's function is not known in closed form, and so highly accurate global conditions, such as DtN conditions, cannot be used. Instead we model the semi-infinite domain using a *perfectly matched layer* (PML), which can be designed to absorbs waves over a large range of angles of incidence, but which does not require knowledge of Green's function [23].

Basu and Chopra [23] demonstrated the superior performance of their PML method for problems related to earthquake engineering. Here, we examine the utility of a PML for anchor

loss computations in MEMS resonators. We begin with a brief review of PMLs for time-harmonic motion. This is followed by a discussion of the relevant finite element expressions. Our presentation, while similar to that of Reference [23], leads to a simpler implementation. We analyse the effects of discretization on the PML behaviour in the time-harmonic case, and describe how to choose the PML parameters to obtain good accuracy. Because of the large computational scale of resonator problems we also investigate the use of reduced-order models that preserve the complex symmetric structure of the PML equations.

To illustrate the PML technology, we analyse the behaviour of a family of MEMS disk resonators. Several disk resonators have already been built [7, 24, 25], and in the laboratory they have shown quality factors as high as 55 000 at frequencies of up to 1.14 GHz. Through numerical experiments, we explain in detail the mechanism of anchor loss in these devices; the predictions of our model are supported by recent experimental work [25], which we also partially report here.

2. PERFECTLY MATCHED LAYERS

Except for scale, a microresonator atop a silicon chip is much like a structure on the earth's surface during an earthquake. While we are concerned with waves radiating away from a structure and the earthquake engineer is concerned with waves radiating toward a structure, in both cases the substrate is much larger than the structure, and it can be modelled as an elastic half-space (possibly heterogeneous). This infinite-domain approximation occurs in many physical models: acoustic waves radiating from a musical instrument, electromagnetic waves reflecting from aircraft, elastic waves scattering from a crack in a solid, and water waves in an open harbor are only a few additional examples [19, 26, Chapter 8]. The essential characteristic of the infinite-domain solution is that only outgoing waves are allowed. To model infinite-domain problems on a computer, we need finite-size discretizations which enforce this radiation condition.

One way to enforce the radiation condition is to discretize an exact boundary equation satisfied by outgoing waves. For example, outside of a sphere containing any radiators and scatterers, waves can be written as a multipole expansion; in this expansion, the radiation condition just says that certain coefficients corresponding to incoming waves should be zero. A related global condition is the DtN map, which specifies how Dirichlet and Neumann conditions must be related at a surface [26]. These boundary conditions are rigorously derived and highly accurate, but they usually require that the artificial boundary have a particular shape. They are also non-local in space: every boundary unknown is directly related to every other boundary unknown, and consequently the matrix of boundary terms is dense, and expensive to form and to solve. Furthermore, exact boundary conditions may be unavailable for problems in which no analytically tractable Green's function is known, as in our case.

A second approach is to build approximate boundary conditions based on the asymptotic behaviour of outgoing waves. Low-order versions of these approximate conditions are local and inexpensive, but only absorb waves over a small range of angles of incidence [23, 26]. Consequently, a large computational domain may be needed for accurate results. Further, they often have difficulty with surface waves and interface waves. Higher-order versions perform better but require special finite element architectures [27]. Yet another approach is to add a non-physical 'sponge layer' to dissipate waves before they reach the artificial boundary. Waves

passing through the sponge layer are damped on the way to the artificial boundary, and are further damped when they are reflected back, so that most of the signal entering the layer is absorbed. To be effective, though, the layer must be designed so that there is no impedance mismatch to reflect waves back from the interface between the layer and the rest of the domain.

A PML is a refinement of a sponge layer. Bérenger invented the perfectly matched layer for problems in electromagnetic wave propagation [28], and it was later re-interpreted as a complex-valued change of co-ordinates which could be applied to any linear wave equation [29–31]. Not only do these layers rapidly attenuate waves, they also 'perfectly match' the rest of the domain; that is, there are no spurious reflections at the interface due to perfect impedance matching. In Reference [23], a perfectly matched layer for time-harmonic elastodynamics is described which—unlike previous elastodynamic PMLs such as those in References [32–34]—can be implemented with finite elements in a standard displacement framework, with no non-standard global unknowns; see also References [35, 36]. We describe an alternate interpretation of the PML described in Reference [23], and show how our interpretation further simplifies implementation in a finite element code.

2.1. A motivating example

A PML model of an infinite domain problem is composed of two parts: a sub-domain where the actual equation of interest is dealt with explicitly, and a sub-domain that produces the desired effect of a far-field radiation boundary condition. To set terminology and provide insight into the workings of PMLs, we review a simple 1-D example.

2.1.1. 1-D elastic wave. Consider a longitudinal wave propagating in a homogeneous, semi-infinite rod with axial co-ordinate $x \in [0, \infty)$. If waves travel with speed c, the one-dimensional wave equation that describes this system is

$$\frac{\partial^2 u}{\partial x^2} - \frac{1}{c^2}\frac{\partial^2 u}{\partial t^2} = 0 \tag{2}$$

where $u(x,t)$ is the displacement. Time-harmonic solutions $u(x,t) = \hat{u}(x)\mathrm{e}^{\mathrm{i}\omega t}$ are governed by a Helmholtz equation

$$\frac{\mathrm{d}^2 \hat{u}}{\mathrm{d}x^2} + k^2 \hat{u} = 0 \tag{3}$$

where $k = \omega/c$ is the wave number and $\mathrm{i} = \sqrt{-1}$. Solutions to this problem have the form

$$\hat{u} = c_{\text{out}}\mathrm{e}^{-\mathrm{i}kx} + c_{\text{in}}\mathrm{e}^{\mathrm{i}kx} \tag{4}$$

where c_{out} is the magnitude of the outgoing wave travelling from the origin toward infinity, and c_{in} is the magnitude of the incoming wave travelling from infinity toward the origin. In general, we assume there is no source at infinity, so physically meaningful solutions to such problems have $c_{\text{in}} = 0$.

2.1.2. 1-D elastic wave in a perfectly matched medium. We now consider the Helmholtz equation (3) under a change of co-ordinates. Let $\lambda : \mathbb{R} \to \mathbb{C}$ be a continuous function which

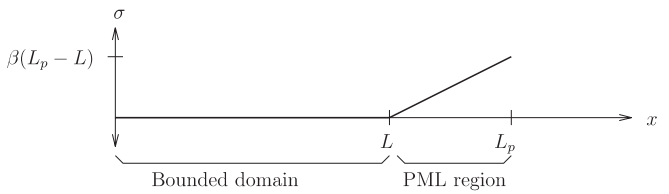

Figure 1. Piecewise linear attenuation function for a plane wave.

is nowhere zero, and define a new co-ordinate

$$\tilde{x} = \int_0^x \lambda(s)\,\mathrm{d}s \tag{5}$$

By definition, $\tilde{x}$ and x are differentially related

$$\frac{\mathrm{d}\tilde{x}}{\mathrm{d}x} = \lambda(x), \quad \frac{\mathrm{d}}{\mathrm{d}\tilde{x}} = \frac{1}{\lambda(x)}\frac{\mathrm{d}}{\mathrm{d}x} \tag{6}$$

Now suppose that the stretched co-ordinate $\tilde{x}$ is used as the independent variable in Equation (3). Then in terms of x, the equation is

$$\frac{1}{\lambda}\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{1}{\lambda}\frac{\mathrm{d}\hat{u}}{\mathrm{d}x}\right) + k^2\hat{u} = 0 \tag{7}$$

where the derivative may be taken in a weak sense, since λ need not be C^1. Equation (7) describes wave propagation in a *perfectly matched medium* (PMM).

Suppose

$$\lambda(s) = 1 - \mathrm{i}\sigma(s)/k \tag{8}$$

then the solutions to the PMM equation (7) are

$$\hat{u} = c_{\mathrm{out}} \exp\left(-\int_0^x \sigma(s)\,\mathrm{d}s\right)\exp(-\mathrm{i}kx) + c_{\mathrm{in}} \exp\left(\int_0^x \sigma(s)\,\mathrm{d}s\right)\exp(\mathrm{i}kx) \tag{9}$$

So as long as $\sigma = 0$, both the incoming and outgoing solutions to the PMM equation (7) agree with the solutions to the original Helmholtz equation (3). Where $\sigma > 0$, the wave decays in the direction of travel. Since the outgoing wave and the incoming wave travel in opposite directions, the outgoing wave amplitude decays with increasing x, while the incoming wave amplitude decays with decreasing x. For example, assume σ is defined to be zero on $[0, L]$ and $\sigma = \beta(s - L)$ on $[L, \infty)$; see Figure 1. Then for $x > L$, the outgoing wave amplitude is $c_{\mathrm{out}} \exp(-\beta(x - L)^2/2)$, and the incoming wave amplitude is $c_{\mathrm{in}} \exp(\beta(x - L)^2/2)$.

Because waves decay so rapidly as they travel through the PMM region, we obtain a good approximation to the infinite-domain problem even if we force $\hat{u}(L_p) = 0$ for some finite $L_p > L$. This generates the concept of a PML; i.e. a PML is a finite PMM attached to a region

with regular wave behaviour. For example, suppose we prescribe $\hat{u}(0) = 1$ and $\hat{u}(L_p) = 0$. For convenience, define $\gamma = \beta(L_p - L)^2$; then the boundary conditions become

$$\begin{bmatrix} \hat{u}(0) \\ \hat{u}(L_p) \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ \mathrm{e}^{-(\gamma/2+\mathrm{i}kL_p)} & \mathrm{e}^{\gamma/2+\mathrm{i}kL_p} \end{bmatrix} \begin{bmatrix} c_{\text{out}} \\ c_{\text{in}} \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \tag{10}$$

and therefore

$$c_{\text{out}} = \frac{1}{1 - \mathrm{e}^{-\gamma-2\mathrm{i}kL_p}} = 1 + O(\mathrm{e}^{-\gamma}), \quad c_{\text{in}} = \frac{-\mathrm{e}^{-\gamma-2\mathrm{i}kL_p}}{1 - \mathrm{e}^{-\gamma-2\mathrm{i}kL_p}} = -O(\mathrm{e}^{-\gamma}) \tag{11}$$

Even for modest γ, the bounded-domain solution is a good approximation to the infinite domain solution. For $\gamma \approx 4.6$, only 1% of the outgoing wave is reflected. Increasing γ decreases the reflection in the continuous case; however, in the discrete equations obtained from finite difference or finite element approximations, we must be careful about how we increase γ. If β is too large, the waves entering the PML will decay rapidly, effectively creating a boundary layer; if the discretization is too coarse to resolve this decay, the numerical solution will be polluted by spurious reflections. We discuss this phenomenon and its implications further in Section 4.

3. MULTI-DIMENSIONAL ELASTICITY

The multi-dimensional equations of motion for a linear time-harmonic elastodynamic medium with no body forces are

$$\omega^2 \rho u + \nabla \cdot \sigma = 0 \tag{12}$$

$$\sigma = \mathsf{C} : \varepsilon \tag{13}$$

$$\varepsilon(u) = \left(\frac{\partial u}{\partial x} \right)^s \tag{14}$$

where u is the displacement field, ε is the infinitesimal strain tensor, σ is the stress tensor, C is the material stiffness tensor, and ρ is the density. A simple isotropic elastic medium admits propagating disturbances moving at two characteristic velocities—compression waves (P waves) and shear waves (S waves). An anisotropic medium admits further characteristic wave speeds, and inhomogeneities and interfaces add yet more wave types. However, as in the one-dimensional case, a complex-valued co-ordinate transformation can be used to attenuate each of these time-harmonic waves in the direction of travel without spurious reflections from artificial interfaces.

3.1. Multi-dimension PMM equations

Though it is possible to introduce the co-ordinate transformation into the local form of the equations [23], it is simpler to first recast the equations in weak form and then transform.

The weak form of the time-harmonic elastodynamic equation is

$$\int_\Omega \varepsilon(w) : \sigma(u)\,\mathrm{d}\Omega - \omega^2 \int_\Omega \rho w \cdot u\,\mathrm{d}\Omega = \int_\Gamma w \cdot t\,\mathrm{d}\Gamma \tag{15}$$

where the domain is Ω, part of the boundary $\Gamma \subset \partial\Omega$ is subject to tractions t, and w is a weight function. As before, suppose $\tilde{x}$ is a transformed co-ordinate such that the Jacobian $\partial\tilde{x}/\partial x = \Lambda$ is continuously defined and everywhere non-singular. Replacing x with $\tilde{x}$ everywhere in (15), we have

$$\int_{\tilde{\Omega}} \tilde{\varepsilon}(w) : \tilde{\sigma}(u)\,\mathrm{d}\tilde{\Omega} - \omega^2 \int_{\tilde{\Omega}} \rho w \cdot u\,\mathrm{d}\tilde{\Omega} = \int_{\tilde{\Gamma}} w \cdot \tilde{\sigma}(u) \cdot \tilde{n}\,\mathrm{d}\tilde{\Gamma} \tag{16}$$

We now map back to the x co-ordinate system:

$$\int_\Omega \tilde{\varepsilon}(w) : \tilde{\sigma}(u)\det(\Lambda)\,\mathrm{d}\Omega - \omega^2 \int_\Omega \rho w \cdot u \det(\Lambda)\,\mathrm{d}\Omega = \int_\Gamma w \cdot \tilde{\sigma}(u) \cdot (\Lambda^{-T} n)\det(\Lambda)\,\mathrm{d}\Gamma \tag{17}$$

In the x co-ordinate system, the transformed strain and stress tensors are

$$\tilde{\varepsilon}(u) = \left(\frac{\partial u}{\partial \tilde{x}}\right)^s = \left(\frac{\partial u}{\partial x}\Lambda^{-1}\right)^s \tag{18}$$

$$\tilde{\sigma}(u) = \mathsf{C} : \tilde{\varepsilon}(u) \tag{19}$$

The local form of (17), which can be derived either from (17) or directly from transforming (12), is

$$\operatorname{trace}\left(\frac{\partial \tilde{\sigma}(u)}{\partial x}\Lambda^{-1}\right) + \omega^2 \rho u = 0 \tag{20}$$

or, in indicial form,

$$\frac{\partial \tilde{\sigma}_{ij}}{\partial x_k}(\Lambda^{-1})_{kj} + \omega^2 \rho u_i = 0 \tag{21}$$

3.2. Anisotropic medium interpretation

We now present a different way to look at the PML equations, in which the original form of the elasticity equations is maintained, but with different material coefficients. This leads to a succinct and intrinsically symmetric implementation for multi-dimensional elastic PMLs.

In indicial form, we write the strain associated with a displacement field u as

$$\varepsilon_{ij}(u) = \frac{1}{2}(\delta_{ip}\delta_{jq} + \delta_{iq}\delta_{jp})\frac{\partial u_p}{\partial x_q} \tag{22}$$

The PML-transformed strain has the same form, except with one of the Kronecker δ functions replaced by Λ^{-1}:

$$\tilde{\varepsilon}_{ij}(u) = \frac{1}{2}(\delta_{ip}\delta_{jr} + \delta_{ir}\delta_{jp})\frac{\partial u_p}{\partial x_q}(\Lambda^{-1})_{qr} \tag{23}$$

$$= \tilde{\mathsf{I}}_{ijpq}\frac{\partial u_p}{\partial x_q} \tag{24}$$

where $\tilde{\mathsf{I}}_{ijpq} := \frac{1}{2}(\delta_{ip}(\Lambda^{-1})_{qj} + (\Lambda^{-1})_{qi}\delta_{jp})$. Now by substitution,

$$\tilde{\varepsilon}_{ij}(w)\mathsf{C}_{ijkl}\tilde{\varepsilon}_{kl}(u) = \frac{\partial w_p}{\partial x_q}\tilde{\mathsf{I}}_{ijpq}\mathsf{C}_{ijkl}\tilde{\mathsf{I}}_{klrs}\frac{\partial u_r}{\partial x_s} \tag{25}$$

$$= \frac{\partial w_p}{\partial x_q}\tilde{\mathsf{C}}_{pqrs}\frac{\partial u_r}{\partial x_s} \tag{26}$$

where we define

$$\tilde{\mathsf{C}}_{pqrs} := \tilde{\mathsf{I}}_{ijpq}\mathsf{C}_{ijkl}\tilde{\mathsf{I}}_{klrs} \tag{27}$$

Note that $\tilde{\mathsf{C}}_{pqrs}$ inherits the major and minor symmetries of C_{ijkl}. That is,

$$\mathsf{C}_{ijkl} = \mathsf{C}_{klij} \implies \tilde{\mathsf{C}}_{pqrs} = \tilde{\mathsf{C}}_{rspq} \tag{28}$$

$$\mathsf{C}_{ijkl} = \mathsf{C}_{jikl} \implies \tilde{\mathsf{C}}_{pqrs} = \tilde{\mathsf{C}}_{qprs} \tag{29}$$

Because of the minor symmetries ($\tilde{\mathsf{C}}_{pqrs} = \tilde{\mathsf{C}}_{qprs}$ and $\tilde{\mathsf{C}}_{pqrs} = \tilde{\mathsf{C}}_{pqsr}$), we can rewrite (26) as

$$\tilde{\varepsilon}_{pq}(w)\mathsf{C}_{pqrs}\tilde{\varepsilon}_{rs}(u) = \frac{\partial w_p}{\partial x_q}\tilde{\mathsf{C}}_{pqrs}\frac{\partial u_r}{\partial x_s} = \varepsilon_{pq}(w)\tilde{\mathsf{C}}_{pqrs}\varepsilon_{rs}(u) \tag{30}$$

If we substitute (30) into the weak form of the PML equation (17), and assume that there is no loading on the transformed part of the boundary, we have

$$\int_\Omega \varepsilon(w) : \tilde{\mathsf{C}}\varepsilon(u)\det(\Lambda)\,\mathrm{d}\Omega - \omega^2\int_\Omega \rho w\cdot u\det(\Lambda)\,\mathrm{d}\Omega = \int_\Gamma w\cdot t\,\mathrm{d}\Gamma \tag{31}$$

Now define

$$\mathsf{C}^{\mathrm{PML}} = \tilde{\mathsf{C}}\det(\Lambda) \tag{32}$$

$$\rho^{\mathrm{PML}} = \rho\det(\Lambda) \tag{33}$$

so that (31) becomes

$$\int_\Omega \varepsilon(w) : \mathsf{C}^{\mathrm{PML}}\varepsilon(u)\,\mathrm{d}\Omega - \omega^2\int_\Omega \rho^{\mathrm{PML}} w\cdot u\,\mathrm{d}\Omega = \int_\Gamma w\cdot t\,\mathrm{d}\Gamma \tag{34}$$

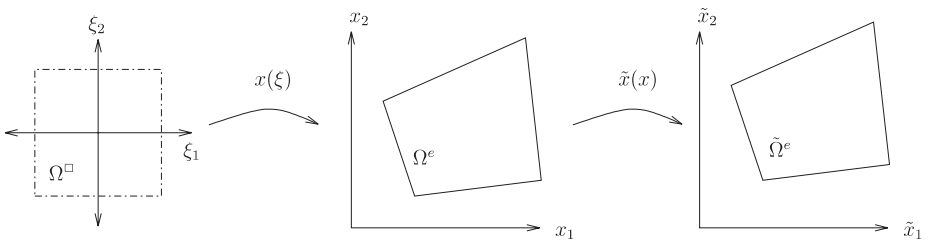

Figure 2. Concatenated isoparametric mapping and PML co-ordinate mapping.

The form of (34) is identical to the form of the standard elasticity equation (15), but with inhomogeneous, anisotropic, complex-valued material properties.

3.3. Finite element implementation

To derive the weak form of the PML equations in the x co-ordinate system, Equation (17), we performed a change of variables in the integrals of Equation (16). Because isoparametric finite elements already use mapped integration, we can combine the change of variables associated with the PML mapping with the change of variables associated with the isoparametric co-ordinate transformation.

Consider the element in Figure 2. Suppose we choose shape functions N_I, so that we have interpolations within elements of the form $u = \sum_I N_I u_I$ and $w = \sum_I N_I w_I$. Then the nodal submatrices for the element stiffness and mass are given by

$$k_{IJ}^e = \int_{\Omega^\square} \tilde{B}_I^{\mathrm{T}} D \tilde{B}_J \tilde{J} \, \mathrm{d}\Omega^\square \tag{35}$$

$$m_{IJ}^e = \left(\int_{\Omega^\square} \rho N_I^{\mathrm{T}} N_J \tilde{J} \, \mathrm{d}\Omega^\square \right) \mathbf{1} \tag{36}$$

where $\mathbf{1}$ is the second-order identity tensor and the nodal matrices $\tilde{B}_I$ come from transforming co-ordinates in the standard B-matrix formulation [37, Chapter 4], D is the standard matrix of material parameters, and $\tilde{J}$ is the Jacobian of the composition of the PML mapping with the isoparametric mapping:

$$\tilde{J} = \det\left(\frac{\partial x}{\partial \xi}\right) \det(\Lambda) \tag{37}$$

In practice, we evaluate the integrals numerically by Gaussian quadrature in the parent domain. Whether the quadrature is done analytically or numerically, the form of the integrands in (35) and (36) guarantees that the mass and stiffness matrices will be complex symmetric.

Remarks

1. This interpretation of the PML in terms of an additional co-ordinate transformation works with plane stress, plane strain, axisymmetric, or three-dimensional problems. In the

axisymmetric case, however, the factor of r that appears in the integrands should *not* be transformed into the PML co-ordinate systems, since that factor of r comes from the Jacobian of the mapping to the (r, z) co-ordinates, and not from the mapping to the $(\tilde{r}, \tilde{z})$ co-ordinates.

2. For many problems, a reasonable choice of co-ordinate transformations is to independently stretch each co-ordinate x_i, so that Λ is a diagonal matrix; i.e. $\Lambda = \text{diag}(\lambda_i)$. If we further choose stretching functions so that Λ can be described by low-order polynomials, then it makes sense to also use isoparametric interpolation to compute the values of the stretching function. That is, given values for λ_i at each node, we compute $\Lambda = \text{diag}(\lambda_i)$ by interpolation at the Gauss points where it is evaluated.
3. By writing the PML equations in this form we can easily institute an economy of programming where every element in a mesh is a 'PML element'. Regular elements are formed using the co-ordinate transformation $\Lambda = \mathbf{1}$ and true PML elements by $\Lambda = \text{diag}(\lambda_i)$. Thus the creation of PML elements only requires a minor modification of the traditional element mapped integration routines.

4. EFFECTS OF DISCRETIZATION AND ANGLE OF INCIDENCE

The analysis of Section 2.1.2 describes the behaviour of PMMs in the one-dimensional continuum setting. In multi-dimensional problems, we also need to consider the important issue of angle of incidence which is intimately tied to the level of discretization. In this section, we show how to extend the low-order finite difference analysis of Reference [38] to the arbitrary order finite element setting. The methodology we employ is similar in concept to that found, for example, in Reference [39] but utilizes block difference equations as a simplifying formalism. As will be shown, this will allow us to nicely separate the spurious reflection coefficient into contributions from the far-end and interface reflections. This will lead to simple heuristics for selecting PML parameters in multi-dimensional time-harmonic problems.

4.1. Model problem: continuous case

Let us consider as a model problem the two-dimensional Helmholtz equation on $[0, L_p] \times (-\infty, \infty)$, where a PML transformation is applied to the x co-ordinate for $x \in [L, L_p]$:

$$\frac{1}{\lambda}\frac{\partial}{\partial x}\left(\frac{1}{\lambda}\frac{\partial u}{\partial x}\right) + \frac{\partial^2 u}{\partial y^2} + k^2 u = 0 \tag{38}$$

$$u(0, y) = \exp(\mathrm{i}k_y y) \tag{39}$$

$$u(L_p, y) = 0 \tag{40}$$

Because of the special nature of the boundary conditions and the infinite extent of the domain in the y-direction, this equation admits plane-wave type solutions of the form $u(x, y) = v(x) \exp(\mathrm{i}k_y y)$ where v satisfies the *one-dimensional* PML equation

$$\frac{1}{\lambda}\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{1}{\lambda}\frac{\mathrm{d}v}{\mathrm{d}x}\right) + k_x^2 v = 0 \tag{41}$$

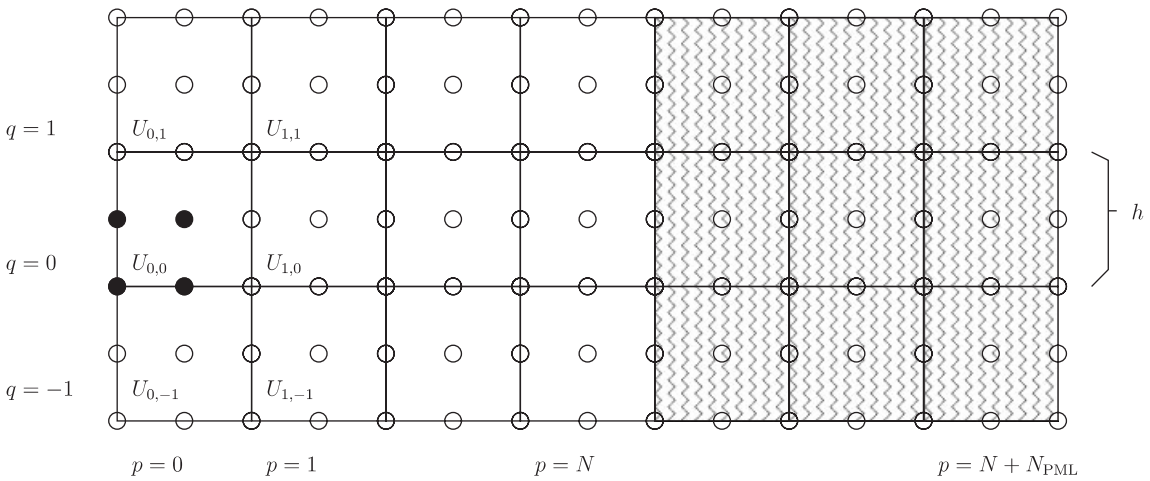

Figure 3. Discrete 2D plane wave test problem.

$$v(0) = 1 \tag{42}$$

$$v(L_p) = 0 \tag{43}$$

We analysed this one-dimensional problem in Section 2.1.1. Recall that in the untransformed part of the domain, the solution $v(x)$ is a linear combination of the free-space left- and right-travelling waves:

$$v(x) = c_{\mathrm{in}} \exp(\mathrm{i}k_x x) + c_{\mathrm{out}} \exp(-\mathrm{i}k_x x) \quad \text{for } x \in [0, L] \tag{44}$$

The continuous reflection coefficient is defined as $r_{\mathrm{continuum}} := |c_{\mathrm{in}}/c_{\mathrm{out}}|$. If $\lambda(x) = 1 - \mathrm{i}\sigma(x)/k$, then the reflection coefficient has the form $\exp(-\gamma k_x/k)$, where γ is a function of the PML length and the choice of parameters and $k = \sqrt{k_x^2 + k_y^2}$. Thus in the continuous case plane waves travelling nearly perpendicular to the PML interface are more strongly absorbed by the PML than are waves travelling at a shallow angle.

4.2. Model problem: discrete case

In the discrete PML case we can perform an analysis similar to the prior section. Starting with a finite element model of our problem, we can employ a *discrete Fourier transform* in the y-direction to obtain a one-dimensional system. This one-dimensional system is finite, so we can solve it numerically. In the ordinary part of the domain where the PML is not in effect, it is then possible to extract the discrete left- and right-travelling parts of the solution; the ratio of the magnitude of these components defines a discrete reflection coefficient.

Our discrete model problem is illustrated in Figure 3. For the purpose of concreteness, but without loss of generality, we will consider a mesh of square biquadratic elements of uniform

size h. Let (x_l, y_l) be the positions of the nodes in $[0, h) \times [0, h)$, i.e.

$$(x_1, y_1) = (0, 0) \qquad (x_2, y_2) = \left(\frac{h}{2}, 0\right)$$

$$(x_3, y_3) = \left(\frac{h}{2}, \frac{h}{2}\right) \quad (x_4, y_4) = \left(0, \frac{h}{2}\right)$$

Every node in the mesh can be written uniquely as $(x_l + ph, y_l + qh)$ for some integers p and q. Let U_{pq}^l be the field value at node $(x_l + ph, y_l + qh)$ and define $U_{pq} = (U_{pq}^1, U_{pq}^2, U_{pq}^3, U_{pq}^4)^{\mathrm{T}} \in \mathbb{C}^4$. Grouping together field values in this way, we can write the discretized Helmholtz equation in terms of a system of difference equations

$$\sum_{p,q} H_{rspq} U_{pq} = 0 \tag{45}$$

where $H_{rspq} \in \mathbb{C}^{4\times 4}$ represents the interaction between U_{pq} and U_{rs}. Note that $H_{rspq} = 0$ for $|r - p| > 1$ or $|s - q| > 1$. Also, because of translational invariance in the y-direction, H_{rspq} really only depends on $s - q$:

$$H_{rspq} = C_{rp,s-q} \tag{46}$$

We now seek solutions of the form $U_{pq} = V_p \exp(\mathrm{i}k_y qh)$. The new variable V satisfies

$$\sum_{p} \hat{H}_{rp} V_p = 0 \tag{47}$$

where

$$\hat{H}_{rp} := \exp(-\mathrm{i}k_y h) C_{rp,-1} + C_{rp,0} + \exp(\mathrm{i}k_y h) C_{rp,1} \tag{48}$$

Equation (47) is a finite-size system of linear equations which may be solved numerically.

To decompose the resulting solution in the regular part of the domain, we note that in the untransformed interior ($0 < p < N$), we have a constant coefficient difference equation

$$B^{\mathrm{T}} V_{p-1} + A V_p + B V_{p+1} = 0 \tag{49}$$

where $A = \hat{H}_{pp}$ and $B = \hat{H}_{pp+1} = \hat{H}_{pp-1}^{\mathrm{T}}$. Therefore for $0 < p < N$, we know V_p is of the form

$$V_p = \sum_m c_m \xi_m^p W_m \tag{50}$$

where (ξ_m, W_m) are solutions to the quadratic eigenvalue problem

$$(B^{\mathrm{T}} + \xi A + \xi^2 B) W = 0 \tag{51}$$

For biquadratic elements, for example, there are eight solutions to the eigenvalue problem (51). Two eigenvalues are zero, and two eigenvalues are infinite; these eigenvalues correspond to modes which only appear at boundaries of the discrete system. Two of the eigenvalues are purely real, and correspond to evanescent waves with a wavelength in the y-direction of $h/2$.

The remaining two eigenvalues are a complex conjugate pair on the unit circle. We call these eigenvalues ξ_1 and ξ_2. They are, respectively, algebraic approximations to $\exp(-ik_x h)$ and $\exp(ik_x h)$ and thus correspond to the right- and left-travelling discrete waves. The reflection coefficient in the discrete case can now be defined as $r_{\text{discrete}} = |c_2/c_1|$. The coefficients c_m are computed by solving (50) using the values for V_1 and V_2 from the numerical solution to (47).

Because ξ_1 is an algebraic function of $k_x h$, it can only approximate the transcendental function $\exp(-ik_x h)$. Thus, there is always *numerical dispersion*; that is, for a fixed h the discrete wave speed depends on the wave number. Similarly, if the mesh size h is not uniform, there will be *numerical reflections* at interfaces where the mesh density—and consequently the discrete wave speed—changes. These spurious effects come from the error implicit in the discretization, and they vanish in the limit as $kh \to 0$. While the mesh size in our model problem is uniform in the original co-ordinate x, the elements are *not* uniformly sized with respect to the transformed co-ordinate $\tilde{x}$; it is unsurprising, then, that there is a mismatch in the discrete wave behaviour at the PML interface that causes reflections.

We therefore view the discrete reflection as the sum of two parts: an interface reflection due to discretization, and a far-end reflection due to the finite termination of the PML. We can easily estimate the magnitude of both effects. To estimate the far-end reflection, we use the continuum reflection coefficient $r_{\text{continuum}}$. To estimate the interface reflection, we use a PML sufficiently long that the supposed continuum reflection should be small, on the same order as the roundoff threshold. We gain a further simplification by the realization that the interface reflection depends only weakly on k_y, assuming that $k_y h$ is not too large; as with the far-end reflection, only the k_x component of the wave vector matters. That is, the reflection coefficient for a two-dimensional finite element mesh of the model problem (38) is nearly the same as the reflection coefficient for a one-dimensional finite element mesh of the model problem (41).

4.3. Example: decomposition concept

We illustrate the decomposition in Figure 4. For a mesh of biquadratic elements, we consider waves with $kh = 2\pi/10$, $L = 10h$, and a five-element PML with the parabolic stretch profile $\lambda(x) = 1 - i0.3h^{-2}(x - L)^2$ for $x > L$. We vary the angle of incidence from 1 to 90° and compute the discrete reflection coefficient r_{discrete} (solid line). At the same time, we compute the continuum reflection coefficient (dotted line) and an estimated interface reflection (dashed line) as outlined in the prior section. In determining the estimated interface reflection, we have chosen a PML length such that the continuum reflection coefficient would be 10^{-15}.

Note that in Figure 4, the continuum reflection coefficient is a strictly monotone function of k_x: the larger k_x is, the smaller $r_{\text{continuum}}$ is. In contrast, the smaller k_x is, the smaller the interface reflection; with more elements per wavelength in the x-direction, the continuum wave behaviour is better resolved. Depending on the exact choice of PML parameters, the interface reflection may not be completely monotone. Nevertheless, in general r_{discrete} is dominated by interface reflection when k_x is large and far-end reflection when k_x is small.

4.4. PML parameter selection heuristics

In computations it is important to be able to set the PML parameters to minimize spurious reflections. The results of the two prior sections motivate the following heuristics.

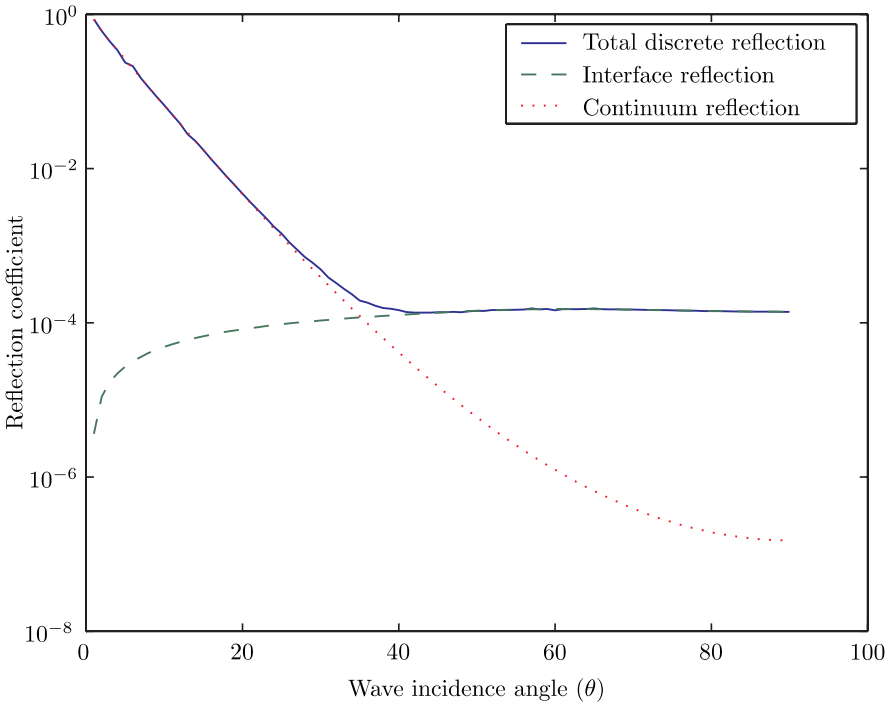

Figure 4. Discrete reflection coefficient r_{discrete} as a function of varying angle. The discrete reflection coefficient (solid line) is closely approximated by the sum of the continuum reflection coefficient $r_{\text{continuum}}$ (dotted line) and the reflection coefficient for one-dimensional waves of wave number k_x directly entering a long PML (dashed line).

First, to be more general, consider PML transformations of the form

$$\lambda(x) = \begin{cases} 1 - \mathrm{i}\beta|x - L|^p, & x > L \\ 1 & x \leqslant L \end{cases} \tag{52}$$

For a reasonably resolved mesh, we will approximate r_{discrete} via a model, r_{model}, where

$$r_{\text{model}}(k_x, \beta, L_p - L, h) := r_{\text{continuum}}(\hat{k}_x, \beta, L_p - L) + r_{\text{interface}}(\hat{k}_x h, \beta h^p) \tag{53}$$

Here, $\hat{k}_x = \mathrm{i}\log(\xi_1)/h$ is the discrete wave speed,

$$r_{\text{continuum}} = \exp\left(-\frac{2\beta}{p+1}(L_p - L)^{p+1}\hat{k}_x\right) \tag{54}$$

is the reflection coefficient in the continuum case, and $r_{\text{interface}}$ is the discrete reflection for a PML with length chosen so that $r_{\text{continuum}}$ is equal to the machine unit roundoff threshold:

$$L_p - L = h\,\mathrm{ceil}\left(h^{-1}\left(-\frac{p+1}{2\beta\hat{k}_x}\log\varepsilon_{\text{machine}}\right)^{1/(p+1)}\right) \tag{55}$$

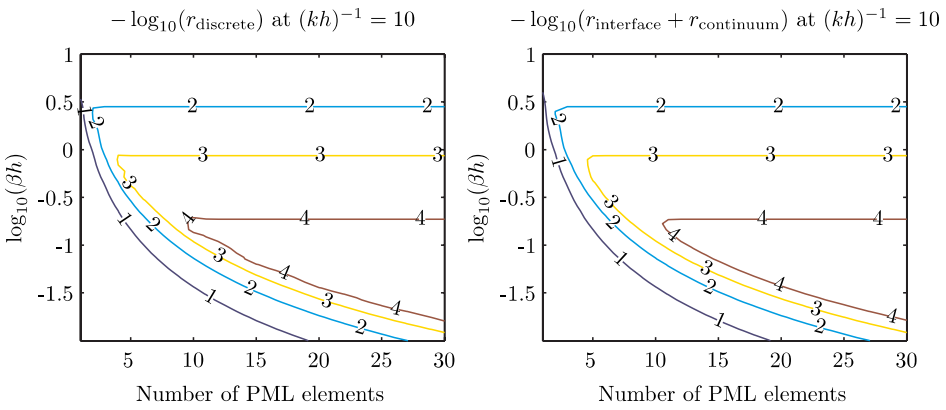

Figure 5. Actual (left) and estimated (right) numerical reflection for varying PML parameters at 10 quadratic elements per wave.

To test this model, we computed r_{discrete} and r_{model} for all combinations of

1. Bilinear, biquadratic, and bicubic elements.
2. Linear and parabolic damping profiles.
3. Incidence angles between 1 and 90°, in steps of one degree.
4. Mesh densities between 8 and 20 elements per wavelength, in steps of one element per wavelength.
5. Continuum reflections of $10^{-2}, 10^{-3}, \ldots, 10^{-12}$.
6. 1–10 elements through the PML.

Over this range of parameters, we found that r_{model} was never more than twice r_{discrete}. In 97% of our test cases, r_{model} was within a factor of two of r_{discrete}. In the cases when r_{model} was substantially larger than r_{discrete}, we typically found that $r_{\text{continuum}}$ and $r_{\text{interface}}$ were comparable and the effects cancelled.

We illustrate the behaviour of the discrete PML with two simple experiments with plane waves launched directly into a PML with a linear damping profile. Figure 5 shows the estimated and actual reflection coefficients for a fixed value of kh as the PML parameter β and the PML length are varied. Notice that when the PML reaches a critical length, the interface reflection begins to dominate, and further lengthening the PML leads to no improvement in the numerical performance. In Figure 6, we show the effects of interface reflections in the PML for linear and quadratic elements. We see from these plots that fine discretizations (either from smaller values of kh or from higher-order elements) and low values of β lead to decreased interface reflections. We also note that quadratic elements perform substantially better than linear elements.

This model for the behaviour of the discrete PML suggests the following simple heuristic for choosing the PML parameters given a polynomial profile and a reflection tolerance r_{tol}. First, consider the maximum wave number in the system $k_{\max}$, and choose β so that the interface reflection for waves of length $k_{\max}$ is $r_{\text{tol}}/2$. Second, choose the PML length so that the continuum reflection is $r_{\text{tol}}/2$ for waves with the minimal propogating wave number $k_{\min}$

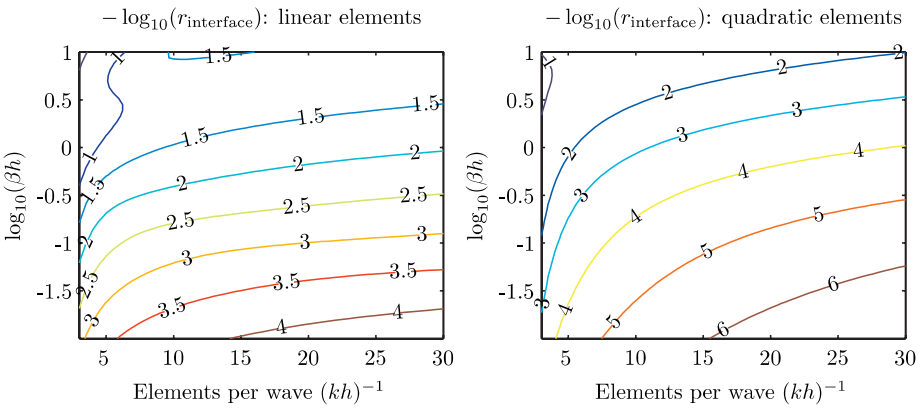

Figure 6. Amount of spurious reflection from a discrete PML interface for varying mesh densities and PML parameters. Both linear (left) and quadratic (right) elements are shown.

entering at the shallowest allowed angle $\theta_{\min}$:

$$L_p - L = h \operatorname{ceil}\left(h^{-1} \left(-\frac{p+1}{2\beta k_{\min} \sin(\theta_{\min})} \log(r_{\text{tol}}/2) \right)^{1/(p+1)} \right) \tag{56}$$

Remark

1. Errors of the type discussed here are sometimes referred to in the context of dispersion analysis—a technique which is very effective for constant-coefficient differential or difference equations. However, the discrete PML equations do not have constant coefficients. While analysing the behaviour of waves propogating through a PML with a constant stretch function can provide useful intuition about the effects of discretization [38, 40], in practice the stretch function is never chosen to be a constant. We therefore prefer to avoid the term *dispersion* when referring to waves propagating through the PML, and instead simply refer to the numerical *reflection* caused by mismatches in the discrete wave behaviour in the PML and the discrete wave behaviour in the untransformed domain.

5. QUALITY FACTORS AND FORCED MOTION COMPUTATIONS

In the MEMS problems of interest to us, we wish to compute quality factors and to compute forced motion responses. The governing equations after spatial discretization of the weak form are given by

$$K_{\text{dyn}}(\omega)u = F \tag{57}$$

where $K_{\text{dyn}}(\omega) := K(\omega) - \omega^2 M(\omega)$. As we noted at the end of Section 3, the PML coordinate transformations suggested in Reference [23] are dependent on the frequency, so that the attenuation through the PML layer will be independent of the forcing frequency. This makes the system matrices dependent upon the drive frequency. Also note that the system matrices are

complex symmetric. The points taken together create a somewhat involved problem. However, a number of basic observations can be used to greatly simplify the situation.

5.1. Quality factors via an eigencomputation

In a properly designed high quality MEMS resonator the drive pattern and frequency are always chosen to excite some resonant mode. Because of this, it suffices for many cases to simply compute the complex-valued eigenvalue of the system closest to the real-valued drive frequency. From this complex-valued eigenvalue, Q is formally defined according to Equation (1) as the ratio of the stored energy to the energy loss per radian, which in terms of a single mode damped oscillator can be expressed as

$$Q = \frac{|\omega|}{2\,\mathrm{Im}(\omega)} \tag{58}$$

where ω is the computed eigenvalue [2].

Because the system matrices depend upon ω, the eigenvalues ω will correspond to solutions of the non-linear eigenvalue problem $\det(K_{\mathrm{dyn}}(\omega)) = 0$. However, when the frequency range of interest is not too wide, the parameters of the co-ordinate transformation may be chosen once to give acceptable attenuation over the desired range, so that the approximate dynamic stiffness for ω near a fixed reference frequency ω_0 is

$$K^0_{\mathrm{dyn}}(\omega) := K(\omega_0) - \omega^2 M(\omega_0) \tag{59}$$

Finding roots such that $\det(K^0_{\mathrm{dyn}}(\omega)) = 0$ is a linear (generalized) eigenvalue problem, which we can approach using standard tools.

To find the damped eigenvalues near some specified (real-valued) reference frequency ω_0, we use a shift-and-invert Arnoldi procedure, described in standard references on numerical linear algebra [41, Chapter 9], [42, Chapter 7]. This procedure computes an orthonormal basis V for the Krylov subspace

$$\mathscr{K}_n(K_{\mathrm{dyn}}(\omega_0), u_0) = \mathrm{span}\{u_0, K^0_{\mathrm{dyn}}(\omega_0)^{-1}u_0, \ldots, K^0_{\mathrm{dyn}}(\omega_0)^{-(n-1)}u_0\} \tag{60}$$

the eigenvalues are then approximated by the eigenvalues of the much smaller problem $V^H K^0_{\mathrm{dyn}}(\omega)V$, where $(\cdot)^H$ indicates complex conjugate transpose. For computing a few isolated eigenvalues near ω_0, the main cost of the shift-and-invert Arnoldi procedure is to compute the factorization needed to apply $K^0_{\mathrm{dyn}}(\omega_0)^{-1}$. In our numerical experiments, we use UMFPACK [43] to factor the shifted matrix, and we use `eigs`, MATLAB's interface to the implicitly-restarted Arnoldi code ARPACK [44], to compute the desired eigenvalues.

5.2. Efficient forced motion computations

Though damped mode eigencomputations are illuminating, they do not give a complete picture of the frequency response behaviour. In practice, we are interested in systems in which there is a single periodically-forced input and a single output determined by some sensed displacement. Suppose F is a time-harmonic load pattern vector, and the output is a linear function of displacement $P^{\mathrm{T}}u$. Then we are really interested in computing the transfer function

$$H(\omega) = P^{\mathrm{T}} K_{\mathrm{dyn}}(\omega)^{-1} F \tag{61}$$

which we approximate for a range of frequencies near ω_0 by

$$H^0(\omega) = P^{\mathrm{T}} K^0_{\mathrm{dyn}}(\omega)^{-1} F \tag{62}$$

Even if $K^0_{\mathrm{dyn}}(\omega)$ is nearly singular at some frequency, the response amplitude $|H^0(\omega)|$ may not peak, since F may be nearly orthogonal to the forcing that drives the mode, or P may be orthogonal to the modal displacement pattern. Therefore, one normally examines $H^0(\omega)$ directly using a Bode plot.

Since the dimension N of $K^0_{\mathrm{dyn}}(\omega)$ will be large, it is expensive to evaluate $H^0(\omega)$ directly. Instead, we construct a *reduced-order model* of dimension $n \ll N$, which we use to approximate $H^0(\omega)$. Krylov-subspace projections are often used to build reduced models of large systems [45, 46]. If we build an orthogonal basis V for a small Krylov subspace using shift-and-invert Arnoldi, then we can approximate $H^0(\omega)$ near the shift ω_0 by

$$\hat{H}^0(\omega) := (V^H P)^H (V^H K^0_{\mathrm{dyn}}(\omega) V)^{-1} (V^H F) \approx H^0(\omega) \tag{63}$$

Though $\hat{H}^0(\omega)$ is often a good approximation to $H^0(\omega)$, the projected system matrix $V^H K^0_{\mathrm{dyn}}(\omega) V$ does not preserve the complex symmetric structure of the original discretization. We can construct a symmetry-preserving reduced-order model by choosing an orthonormal projection basis W such that

$$\mathrm{span}(W) = \mathrm{span}([\mathrm{Re}(V), \mathrm{Im}(V)]) \tag{64}$$

Because the span of W contains the span of V, a reduced model based on W will be at least as accurate as the standard Arnoldi-based reduced model. Also, because W is a real-valued basis, projection onto the space spanned by W corresponds to a Bubnov–Galerkin discretization of the PML equation with shape functions $N_I^{\mathrm{reduced}} = \sum_J W_{IJ} N_J$. While these facts alone might induce us to use W rather than V as a projection basis [47], we expect projection onto W to yield much better accuracy than standard Arnoldi projection, as we now describe.

If (K, M) is a Hermitian pencil and M is positive definite, then the pencil will have an orthonormal eigensystem, so that if v is a column eigenvector, v^H will be a corresponding row eigenvector. In the study of such eigenproblems, the Rayleigh quotient

$$\rho(v) = \frac{v^H K v}{v^H M v} \tag{65}$$

plays a special role. When v is an eigenvector, $\rho(v)$ is a corresponding eigenvalue; and further, since ρ is stationary when and only when the argument is an eigenvector, the Rayleigh quotient produces second-order accurate eigenvalue estimates from eigenvector estimates which are only accurate to first order. For general pencils, a column eigenvector v and the corresponding row eigenvector w^H need have no such simple relationship, and so the Rayleigh quotient only provides first-order accurate eigenvalue estimates. The appropriate generalization of the Rayleigh quotient to the non-Hermitian case is $(w^H K v)/(w^H M v)$, a ratio which again yields second-order accuracy (so long as the degenerate case $w^H M v = 0$ is avoided). When K and M are complex symmetric, we know the left and right eigenvectors are simply (non-conjugated) transposes of each other, and so we re-write the second-order accurate quotient estimate as

$$\theta(v) = \frac{v^{\mathrm{T}} K v}{v^{\mathrm{T}} M v} \tag{66}$$

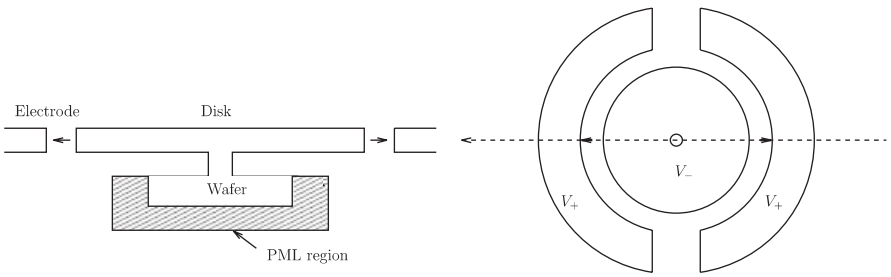

Figure 7. Schematic of a radial disk resonator. An overhead view (right) shows the arrangement of the resonating disk and the electrodes which force it. An idealized cross-section (left) is used in an axisymmetric simulation, where the wafer substrate is treated as semi-infinite using a PML.

This *modified Rayleigh quotient* was used in Reference [48] as part of a Jacobi–Davidson strategy for solving complex symmetric eigenvalue problems from PML discretizations of problems in electromagnetics.

The usefulness of having both left *and* right eigenvectors explains why model reduction methods based on non-symmetric Lanczos iteration often approximate better than Arnoldi methods: a non-symmetric Lanczos iteration simultaneously builds a basis for a right Krylov subspace, which typically contains good approximations for column eigenvectors; and a left Krylov subspace, which typically contains good approximations for row eigenvectors. For complex symmetric matrices, however, left and right subspaces are simply conjugates of each other, and the definition of $\text{span}(W)$ given above is equivalent to

$$\text{span}(W) = \text{span}([V, \text{conj}\, V]) \tag{67}$$

That is, if V is an Arnoldi basis for a right Krylov subspace, then both the right Krylov subspace and the corresponding left Krylov subspace are subspaces of $\text{span}(W)$. As explained in the previous paragraph, then, the position of any eigenvalues (poles) which are estimated by vectors in V will be determined to second-order accuracy by projection onto W, where projection onto V would typically attain only first-order accuracy. For similar reasons, if P and F are proportional to each other, the estimated transfer function obtained from projecting onto W will match H^0 in $2n$ moments, rather than the n moments typical of a standard Arnoldi projection [45].

6. STUDY OF DISK RESONATORS

To exercise and test our proposed methods we examine in detail the response of several radial disk resonators which have been fabricated recently [7, 24, 25]. A schematic of these devices is shown in Figure 7. A thin disk is supported on a post above a substrate. The disk is surrounded by drive electrodes, and the potential difference between the disk and the drive electrodes pulls the disk radially outward at the rim. The disk is driven near the frequency of the first or second axisymmetric, bulk radial, in-plane mode. In Reference [7], the disk is made of polysilicon; Reference [24] describes both polysilicon and polydiamond disks; Reference [25] is concerned

with poly-SiGe disks. The choice of these disk resonators allows us to idealize the systems as axisymmetric.

A major source of energy loss in these devices is the propagation of elastic waves down the supporting post and into the wafer below, where they are largely dissipated. Relative to the size of the resonating device, the wafer is large. We assume that the wafer is effectively infinite in extent, so that none of the waves that radiate into the substrate will be reflected. We use a PML to model the wafer as a semi-infinite half-space.

In the actual devices, there are nitride and oxide films between the device and the wafer. For our simulations we have ignored these geometric features as they typically only have a minor effect on resonator performance. Note that this is a simplification in our model, not an inherent limitation of the PML technology; indeed, one of the attractions of PMLs is the ability to handle layered media and other embedded scatterers.

In the examples to follow, we examine issues of

1. Mesh convergence – h and p.
2. Elucidation of the physical mechanism of anchor loss.
3. Design sensitivity in such resonators.
4. Performance of the model reduction method.

For all examples $\lambda(x)$ is linear in each co-ordinate direction and takes on a value of $1 - \mathrm{i}40$ at the outer edges of the PML.

6.1. Convergence of Q

We first consider the polysilicon disk resonator described in Reference [7], which has a disk 20 μm in diameter and 2 μm thick, supported 0.5 μm above the substrate by a post 2 μm in diameter. When the disk was driven in the second radial mode, the measured Q was 7330 in vacuum and 6100 in air, and the centre frequency was 733 MHz. In our best-resolved simulation, we computed a Q value of 6250 at a centre frequency of 715.6 MHz. Perhaps surprisingly, relatively fine resolution was required to obtain convergence. When the mesh was under-resolved, the Q factor was drastically underestimated, possibly because the small flux reaching the anchor base could not be resolved by the mesh, and was therefore overestimated. We computed the value of Q using both linear and higher-order elements at several mesh densities; see Figure 8. For a given mesh density parameter m, we chose elements so that nodes were as near as possible to $1/m$ μm apart. We computed Q and ω_{centre} by using shift-and-invert Arnoldi with an initial shift of 715 MHz to find the closest complex-valued eigenvalue; most of the time in these computations was spent in the LU factorization of the shifted matrix. The plot shows a clear advantage to p-refinement for this class of problems.

6.2. Observed energy loss mechanism

Figure 9 shows the behaviour of the same disk when driven at 715 MHz, just slightly below the resonant frequency. The amplitudes of both the radial and vertical displacements are shown at the point when the forcing is maximal. Even though the design intent was to excite a pure radial mode, it is clear that the mode contains a bending component. This is a reflection of the fact that pure radial modes are not possible in supported structures. The majority of the displacement occurs in the disk itself, but there is some motion in the post as well. Though the forcing is in the radial direction, the Poisson effect leads to motion in the vertical direction

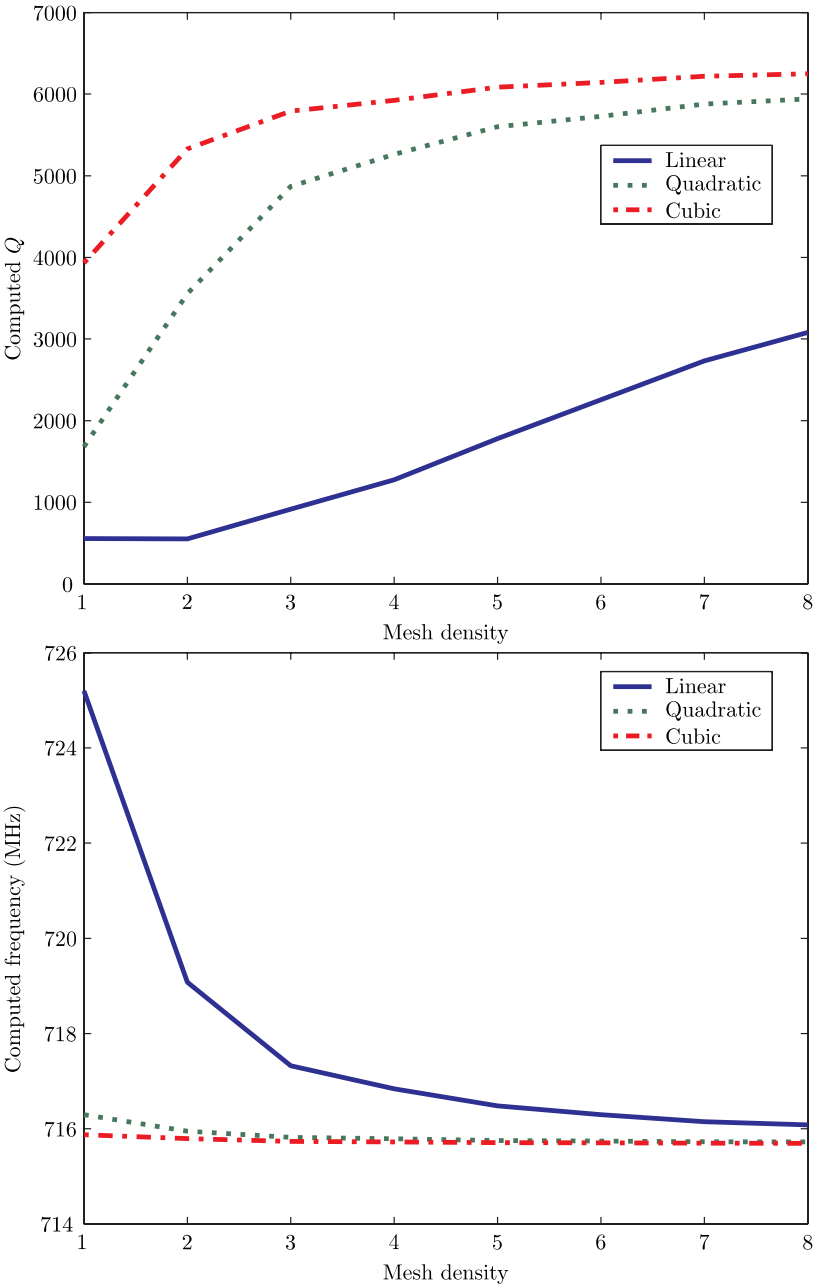

Figure 8. Convergence of Q and ω_{centre} for the second radial mode of the polysilicon disk in Reference [7].

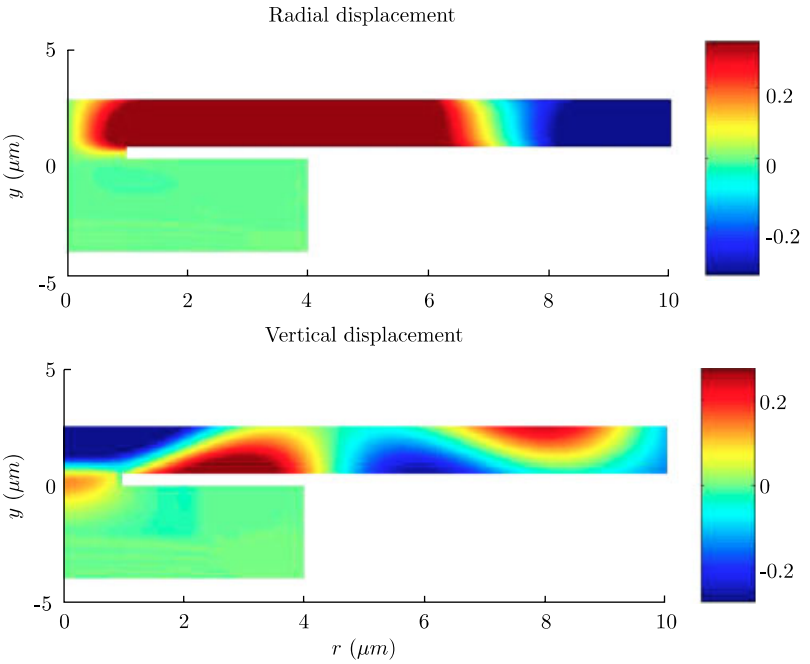

Figure 9. Forced displacement (real part) for the disk resonator model at 715 MHz.

at the centre of the disk. This vertical 'pump' motion results in displacement waves that travel down the post and into the substrate. In an animation, it is possible to see low-amplitude waves radiating away from the post to be absorbed into the perfectly matched layer.

To better understand the behaviour shown in Figure 9, we compute the energy density flux:

$$F(t) = -\operatorname{Re}(\sigma e^{i\omega t})\operatorname{Re}(v e^{i\omega t}) \tag{68}$$

where σ is the stress tensor and v is the velocity vector. Since the energy flux changes over time, we time-average over a single period to obtain the mean energy density flux:

$$\bar{F} = -\tfrac{1}{2}\operatorname{Re}(\sigma v^*) \tag{69}$$

where v^* is the complex conjugate of the velocity. For a standing wave, the displacement and stress are pure real, and $v = iku$ is pure imaginary, so there is no mean energy flux. Therefore, the mean energy flux field tells us something about the departure from the lossless standing-wave behaviour. Figure 10 shows the mean energy flux for a region of the resonator near the edge of the post. The flux vectors in the body of the disk form cycles which carry energy around inside the disk, but do not let it escape into the substrate. Near the post, however, the

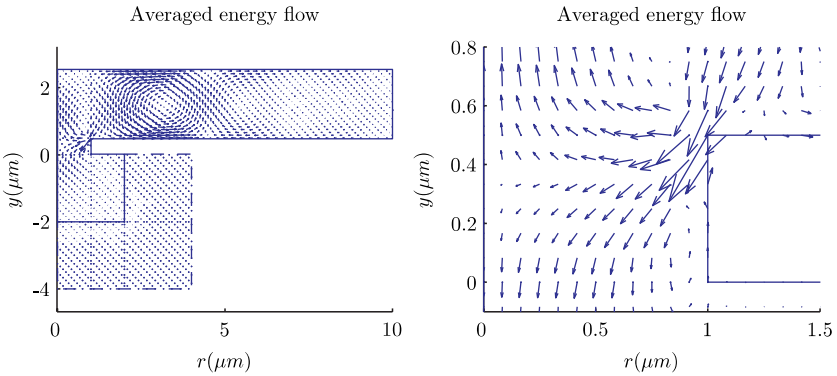

Figure 10. Time-averaged energy flux vector field in the disk resonator driven at 715 MHz. The left plot shows the full field; the right plot shows only the region in the vicinity of the post.

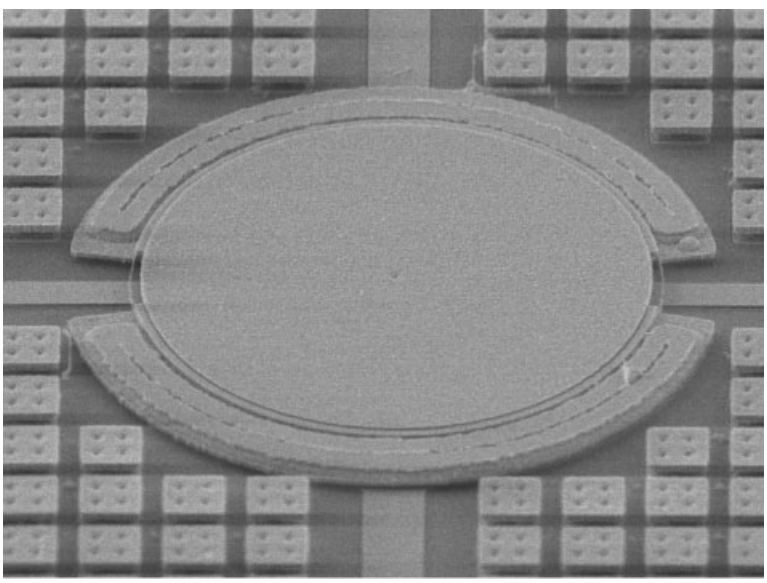

Figure 11. SEM of 41.5 μm radius poly-SiGe disk resonator.

cycle pattern is broken, and the energy flux plot shows a 'spray' of energy that travels down the post and into the substrate.

6.3. Mode mixing and design sensitivity in the disk resonator

To further test our simulation technology on anchor loss, we now consider a series of five poly-$Si_{0.4}Ge_{0.6}$ disk resonators with 41.5 μm radii disks with different thicknesses [25]. Figure 11

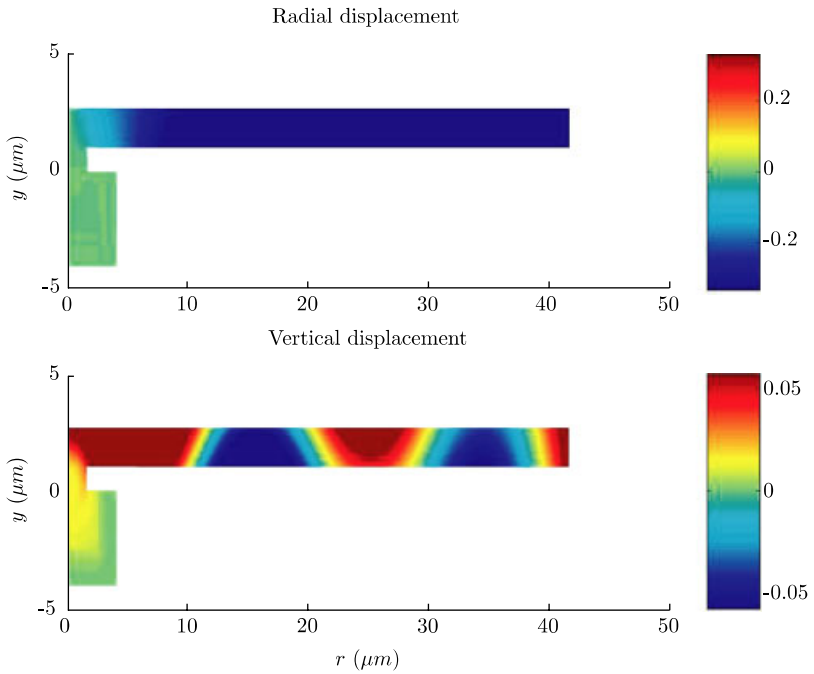

Figure 12. Radial and vertical displacement fields illustrating the mode mixing. Disk radius is 41.5 μm and film thickness is 1.6 μm. The drive frequency is 45 MHz and $Q = 140\,000$. Displacement contours are in units of μm.

shows one such 41.5 μm radius resonator. The disk itself is supported on a conical post with upper radius 1.49 μm, lower radius 1.61 μm, and nominal height 1 μm. The drive is clearly not fully axisymmetric but we model it as such for simplicity. For the material we use a density of 4127 kg/m^3 computed by linear interpolation and assume a Poisson ratio of 0.28; Young's modulus was estimated from an acoustic measurement as 139 GPa.

Figure 12 shows the computed in-phase radial and vertical displacements in one of the disks when it is driven with a radial forcing on its outer edge; the computed Q of the dominant mode is 140 000. The radial motion is coupled to a small bending motion due to the stem as mentioned earlier. Thus, as in the prior example, the dominant mode for this disk is not a pure radial motion. The bending motion of the mode along with the Poisson effect induces a vertical motion in the stem which pumps displacement waves into the substrate, where they carry away the energy of the resonance.

Figure 13 shows measured Q values from the five 41.5 μm disks. The error bars on thickness are indicative of the limits of the SEM geometry measurement method. Also shown in Figure 13 is simulation data using the measured geometry from the resonators. The two curves correspond to Q values for the two eigenvalues nearest the shift (45 MHz). The high Q curve is associated with the radial extension mode we wish to drive; the low Q mode is dominated by a bending motion. The agreement between the measured and computed Q values is good, and the trend with respect to changing disk thickness is captured well.

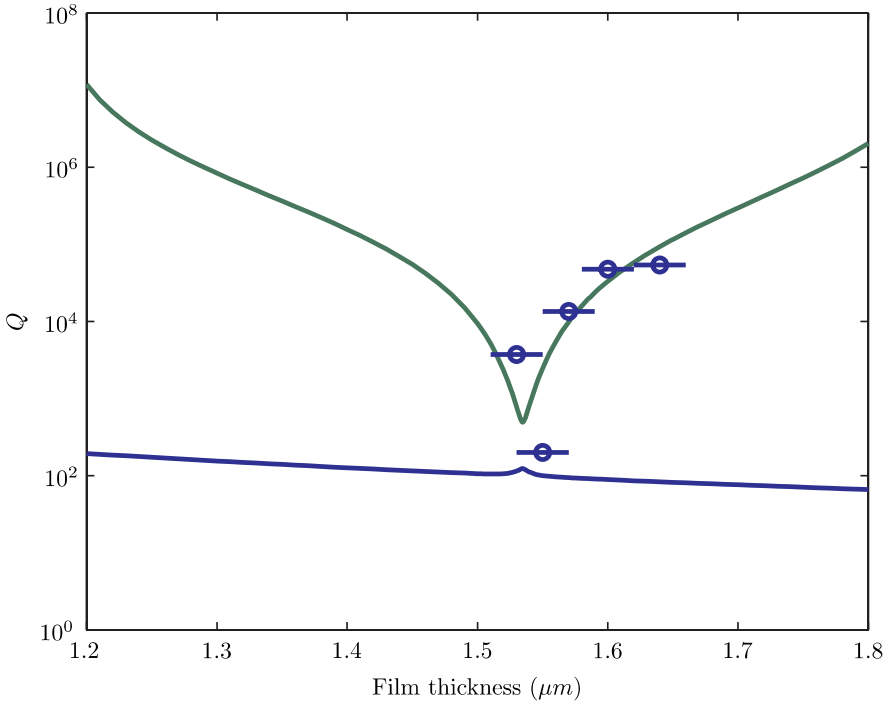

Figure 13. Measured and computed quality factors in 41.5 μm radius disk with varying film thickness. Upper curve indicates Q from eigenvalue closest to the shift. Lower curve indicates Q from next nearest eigenvalue.

The presence of the nearby second mode has a large influence on the high Q mode's quality factor. This is seen in the large swings in the curves of Figure 13 which are computed solely from the system eigenvalues. A good method of visualizing the pole interaction is to examine a root-locus diagram for the two interacting poles (eigenvalues) parameterized by film thickness. Figure 14 was computed for the 41.5 μm radius disks. As the thickness changes the two poles approach each other. The first mode's frequency first moves away from the real axis increasing damping, then back toward the real axis decreasing damping. The speed of the first pole increases the closer it is to the second pole in the complex plane. The dip in the resonator's Q correlates well with the thickness at which the poles are closest.

6.4. Performance of model reduction method

As seen above, the bending-dominated mode significantly affects the resonant peaks associated with the radial dominated modes but not in a way that is immediately obvious. For this reason, Bode plots are helpful in understanding such systems. As an example of our model reduction technology, we start with a shift drawn from a hand estimate for the disk frequency:

$$\omega_{\text{shift}} = 2.405c/R \tag{70}$$

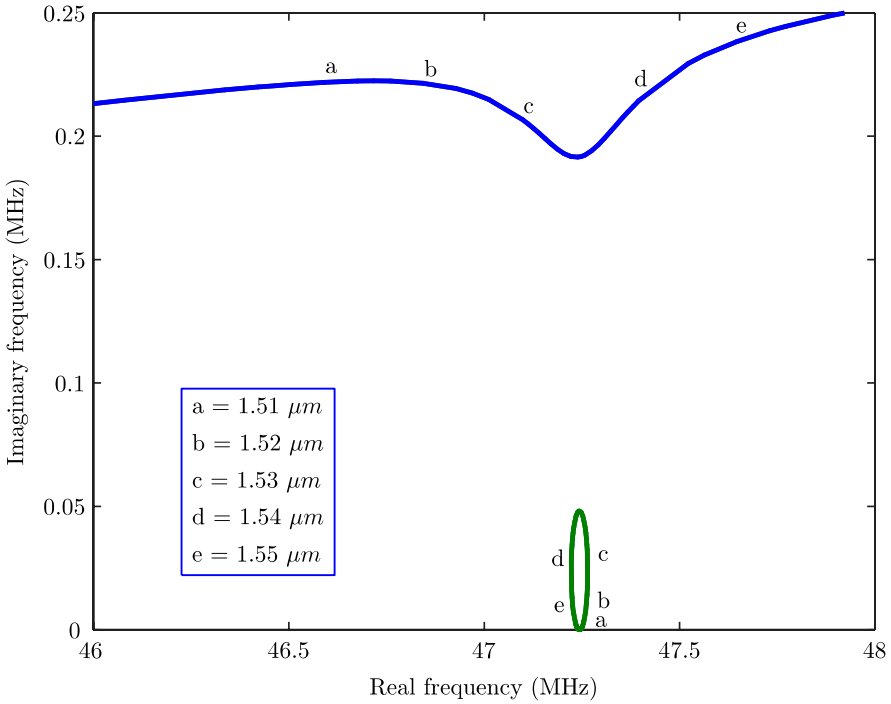

Figure 14. Root-locus plot parameterized by varying film thickness. Lower curve corresponds to the first mode and the upper curve to the next nearest mode.

where R is the disk radius and c is the compression-wave velocity in the disk material. With this shift, we require only two steps of shift-and-invert Arnoldi to resolve the two-dimensional invariant subspace for both relevant eigenvalues. We perform the projections in two ways: first, using a standard Arnoldi projection (with three complex vectors); and second, using a symmetry-preserving projection based on splitting the real and imaginary parts of the Arnoldi vectors (with five real vectors). Both reduced models produce Bode plots which closely match the original 24 265 degree of freedom system (Figure 15). In Figure 16, we compare the accuracy of the two models; clearly, the structure-preserving algorithm is more accurate.

6.5. Conclusions

In this paper, we have developed tools to simulate quality factors in very high frequency MEMS resonators. Anchor losses which dominate such systems have to date been largely ignored in the literature. The primary numerical advances made are as follows:

1. We have described an alternate interpretation of a PML for time-harmonic elasticity which was introduced in Reference [23], and in doing so have shown how to construct a particularly simple finite element implementation for all relevant classes of analysis.

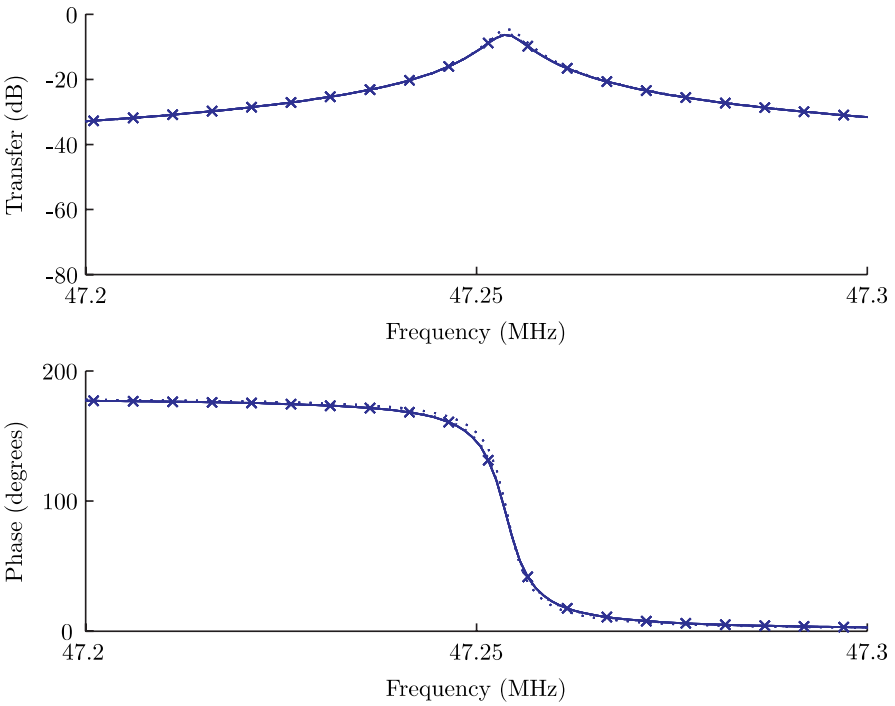

Figure 15. Bode plot of the disk resonator with the full model (×), a structure-preserving reduced model (solid line), and a standard Arnoldi reduced model (dashed line).

2. Through the use of discrete Fourier analysis, we have elucidated the effect of discretization and angles of incidence on the perfect matching property in order to derive heuristics for choosing the PML parameters.
3. We have performed convergence studies which show the care which must be taken to resolve eigenvalues with small imaginary parts compared to their magnitudes, and we have highlighted the advantage of p-refinement in resolving such eigenvalues.
4. By exploiting the complex symmetry inherent in the PML equations, we have described how to improve the accuracy of standard methods for computing free vibrations and for building reduced models for forced frequency-response analysis.

We have demonstrated the utility of our method by analysing the behaviour of a family of disk resonators. Our tools allow us, in detail, to describe the physical mechanism by which these resonators lose energy by radiation of elastic waves from the anchor. Because MEMS fabrication processes are not exact, we have analysed the effect of variations in the thickness of the resonating disks. Our analysis shows that the quality factor of these resonators is highly sensitive to the film thickness, and experimental results confirm this analysis. We explain the variations in the quality factor in terms of interactions between the desired radial extension mode and a parasitic bending mode which resonates at nearly the same frequency. The sensitivity to film thickness was first discovered numerically, and later was verified experimentally.

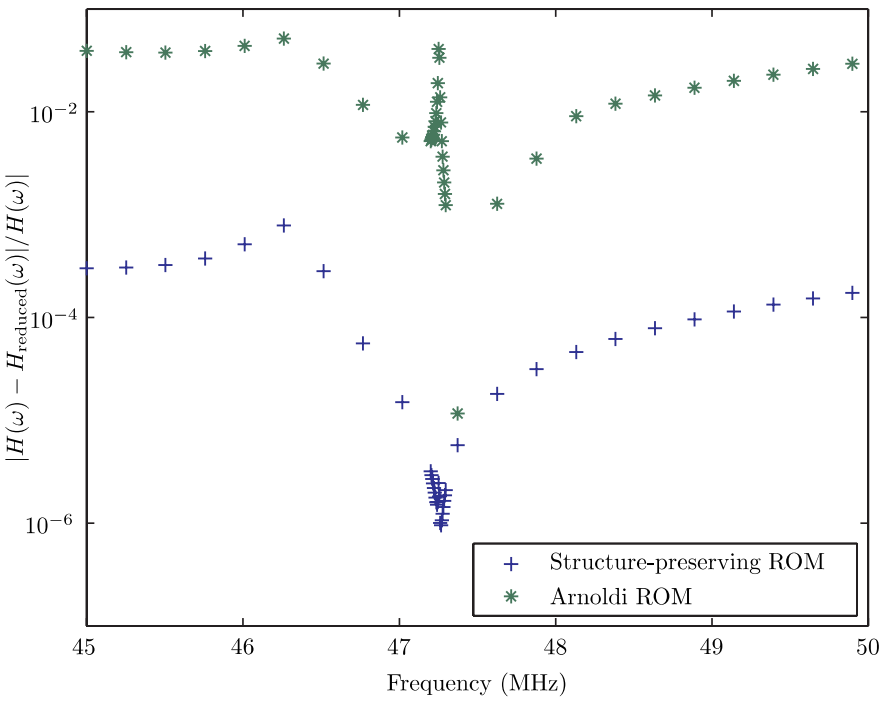

Figure 16. Errors in a structure-preserving reduced model (+) and a standard Arnoldi reduced model (*). Both reduced models are generated from the same number of Krylov vectors.

ACKNOWLEDGEMENTS

The authors wish to thank Prof. Roger Howe, Dr Emmanual Quévy, Dr Jong Woo Shin, Dr Sunil Bhave, Prof. James Demmel, and Dr Ushnish Basu for their interest and assistance in this work.

REFERENCES

1. Nguyen CT-C. Vibrating RF MEMS for low power wireless communications. *Proceedings of the 2001 International MEMS Workshop* (*iMEMS01*), Singapore, July 2001; 21–34.
2. Senturia S. *Microsystem Design*. Kluwer Academic Publishers: Norwell, MA, 2001.
3. Aigner R, Marksteiner S, Elbrecht L, Nessler W. RF-filters in mobile phone applications. *Proceedings of the 12th International Conference on Solid State Sensors*, *Actuators and Microsystems* (*Transducers 03*), Boston, June 2003; 891–894.
4. Candler RN, Li H, Lutz M, Park W-T, Partridge A, Yama G, Kenny TW. Investigation of energy loss mechanisms in micromechanical resonators. *Proceedings of the 12th International Conference on Solid State Sensors*, *Actuators and Microsystems* (*Transducers 03*), Boston, June 2003; 332–335.
5. Ye W, Wang X, Hemmert W, Freeman D, White J. Air damping in laterally oscillating microresonators: a numerical and experimental study. *Journal of Microelectromechanical Systems* 2003; **12**:557–566.
6. Bao M, Yang H, Sun Y, French PJ. Modified Reynolds' equation and analytical analysis of squeeze-film air damping of perforated structures. *Journal of Micromechanics and Microengineering* 2003; **13**:795–800.

7. Wang J, Ren Z, Nguyen CT-C. Self-aligned 1.14 GHz vibrating radial-mode disk resonators. *Proceedings of the 12th International Conference on Solid State Sensors, Actuators and Microsystems* (*Transducers 03*), 2003; 947–950.
8. Duwel A, Gorman J, Weinstein M, Borenstein J, Ward P. Experimental study of thermoelastic damping in MEMS gyros. *Sensors and Actuators A* 2003; **103**:70–75.
9. Houston BH, Photiadis DM, Marcus MH, Bucaro JA, Liu X, Vignola JF. Thermoelastic loss in microscale oscillators. *Applied Physics Letters* 2002; **80**:1300–1302.
10. Abdolvand R, Ho GK, Erbil A, Ayazi F. Thermoelastic damping in trench-refilled polysilicon resonators. *Proceedings of the 12th International Conference on Solid State Sensors, Actuators and Microsystems* (*Transducers 03*), Boston, June 2003; 324–327.
11. Lifshitz R, Roukes ML. Thermoelastic damping in micro- and nanomechanical systems. *Physical Review B* 2000; **61**:5600–5609.
12. Srikar VT, Senturia SD. Thermoelastic damping in fine-grained polysilicon flexural beam resonators. *Journal of Microelectromechanical Systems* 2002; **11**:499–504.
13. Zener C. Internal friction in solids I: theory of internal friction in reeds. *Physical Review* 1937; **52**: 230–235.
14. Zener C. Internal friction in solids II: general theory of thermoelastic internal friction. *Physical Review* 1938; **53**:90–99.
15. Zener C, Otis W, Nuckolls R. Internal friction in solids III: experimental demonstration of thermoelastic internal friction. *Physical Review* 1938; **53**:100–101.
16. Sherman D. An investigation of MEMS anchor design for optimal stiffness and damping. *Technical Report*, University of California, Department of Mechanical Engineering, 1996 (M.S. Report).
17. Park Y-H, Park KC. High-fidelity modeling of MEMS resonators—Part 1: anchor loss mechanisms through substrate. *Journal of Microelectromechanical Systems* 2004; **13**:248–257.
18. Park Y-H, Park KC. High-fidelity modeling of MEMS resonators–part II: coupled beam-substrate dynamics and validation. *Journal of Microelectromechanical Systems* 2004; **13**:248–257.
19. Zienkiewicz OC, Taylor RL. *The Finite Element Method*, *Volume 3*: *Fluid Dynamics* (5th edn). Butterworth-Heinemann: Woburn, MA, 2000.
20. Engquist B, Majda A. Absorbing boundary conditions for the numerical simulation of waves. *Mathematics of Computation* 1977; **31**:629–651.
21. Givoli D. *Numerical Methods for Problems in Infinite Domains*. Elsevier: Amsterdam, 1992.
22. Astley RJ. Infinite elements for wave problems: a review of current formulations and an assessment of accuracy. *International Journal for Numerical Methods in Engineering* 2000; **49**:951–976.
23. Basu U, Chopra A. Perfectly matched layers for time-harmonic elastodynamics of unbounded domains: theory and finite-element implementation. *Computer Methods in Applied Mechanics and Engineering* 2003; **192**:1337–1375.
24. Wang J, Butler JE, Feygelson T, Nguyen CT-C. 1.51-GHz nanocrystalline diamond micromechanical disk resonator with material-mismatched isolating support. *Proceedings of MEMS 2004*, Boston, 2004; 641–644.
25. Bindel DS, Quévy E, Koyama T, Govindjee S, Demmel JW, Howe RT. Anchor loss simulation in resonators. *Proceedings of MEMS 2005*, IEEE: New York, January 2005; 133–136.
26. Turkel E. Introduction to the special issue on absorbing boundary conditions. *Applied Numerical Mathematics* 1998; **27**:327–329.
27. Givoli D. High-order local non-reflecting boundary conditions: a review. *Wave Motion* 2004; **39**:319–326.
28. Bérenger J-P. A perfectly matched layer for the absorption of electromagnetic waves. *Journal of Computational Physics* 1994; **114**:185–200.
29. Collino F, Monk P. The perfectly matched layer in curvilinear coordinates. *SIAM Journal on Scientific Computing* 1998; **19**:2061–2090.
30. Turkel E, Yefet A. Absorbing PML boundary layers for wave-like equations. *Applied Numerical Mathematics* 1998; **27**:533–557.
31. Teixeira F, Chew W. Complex space approach to perfectly matched layers: a review and some new developments. *International Journal of Numerical Modelling* 2000; **13**:441–455.
32. Collino F, Tsogka C. Application of the perfectly matched absorbing layer model to the linear elastodynamic problem in anisotropic heterogeneous media. *Geophysics* 2001; **66**:294–307.
33. Bécache E, Joly P, Tsogka C. Fictitious domains, mixed finite elements, and perfectly matched layers for 2d elastic wave propagation. *Journal of Computational Acoustics* 2001; **9**:1175–1201.

34. Cohen G, Fauqueux S. Mixed spectral finite elements for the linear elasticity system in unbounded domains. *SIAM Journal on Scientific Computing* 2005; **26**(3):864–884.
35. Zheng Y, Huang X. Anisotropic perfectly matched layers for elastic waves in cartesian and curvilinear co-ordinates. *Technical Report*, MIT Earth Resources Laboratory, 2002. Available at http://eaps.mit.edu/erl/research/report1/reports.html
36. Harari I, Albocher U. Studies of FE/PML for exterior problems of time-harmonic elastic waves. *Computer Methods in Applied Mechanics and Engineering*, in press.
37. Zienkiewicz OC, Taylor RL. *The Finite Element Method*, *Volume 1*: *The Basis* (5th edn). Butterworth-Heinemann: Woburn, MA, 2000.
38. Collino F, Monk P. Optimizing the perfectly matched layer. *Computer Methods in Applied Mechanics and Engineering* 1998; **164**:157–171.
39. Mullen R, Belytschko T. Dispersion analysis of finite element semidiscretizations of the two-dimensional wave equation. *International Journal for Numerical Methods in Engineering* 1982; **18**:11–29.
40. Harari I, Slavutin M, Turkel E. Analytical and numerical studies of a finite element PML for the Helmholtz equation. *Journal of Computational Acoustics* 2000; **8**:121–137.
41. Golub GH, Van Loan CF. *Matrix Computations*. John Hopkins University Press: Baltimore, MD, 1989.
42. Demmel JW. *Applied Numerical Linear Algebra*. SIAM: Philadelphia, PA, 1997.
43. Davis TA. Algorithm 832: UMFPACK—an unsymmetric-pattern multifrontal method. *ACM Transactions on Mathematical Software* 2004; **30**:196–199.
44. Lehoucq RB, Soensen DC, Yang C. *ARPACK User's Guide*: *Solution of Large-Scale Eigenvalue Problems with Implicitly Restarted Arnoldi Methods*. SIAM: Philadelphia, PA, 1998.
45. Bai Z. Krylov subspace techniques for reduced-order modeling of large-scale dynamical systems. *Applied Numerical Mathematics* 2002; **43**:9–44.
46. Antoulas AC, Sorensen DC. Approximation of large-scale dynamical systems. *Technical Report TR01-01.pdf*, Department of Computational and Applied Mathematics, Rice University, 2001.
47. Bindel DS, Bai Z, Demmel JW. Model reduction for RF MEMS simulation. *Proceedings of PARA 04*, Lecture Notes in Computer Science. Springer: Berlin, June 2004, in press.
48. Arbenz P, Hochstenbach ME. A Jacobi–Davidson method for solving complex symmetric eigenvalue problems. *SIAM Journal on Scientific Computing* 2004; **25**:1655–1673.

