Construction and Experimental Implementation of a Model-Based Inverse Filter to Attenuate Hysteresis in Ferroelectric Transducers

Andrew G. Hatch and **Ralph C. Smith**
Department of Mathematics
Center for Research in Scientific Computation
North Carolina State University
Raleigh, NC 27695
rsmith@eos.ncsu.edu, aghatch@eos.ncsu.edu

Tathagata De and **Murti V. Salapaka**
Electrical Engineering Department
Iowa State University
Ames, IA 50011
tatha@iastate.edu, murti@iastate.edu

Abstract

Hysteresis and constitutive nonlinearities are inherent properties of ferroelectric transducer materials due to the noncentrosymmetric nature of the compounds. In certain regimes, these effects can be mitigated through restricted input fields, charge- or current-controlled amplifiers, or feedback designs. For general operating conditions, however, these properties must be accommodated in models, transducer designs, and model-based control algorithms to achieve the novel capabilities provided by the compounds. In this paper, we illustrate the construction of inverse filters, based on homogenized energy models, which can be used to approximately linearize the piezoceramic transducer behavior for linear design and control implementation. Attributes of the inverse filters are illustrated through numerical examples and experimental open loop control implementation for an atomic force microscope stage.

1 Introduction

Ferroelectric materials, including the compound lead zirconate titanate (PZT), exhibit novel actuator and sensor capabilities due to the unique electromechanical coupling which they exhibit. This provides them with the capability for providing broadband transduction and nanometer-level set point accuracy. Furthermore, piezoelectric transducers are moderately inexpensive and can be designed to minimally affect the passive dynamics of underlying structures. However, the noncentrosymmetric ion structure that imbues the materials with unique actuator and sensor properties also produces hysteresis and constitutive nonlinearities at all drive levels.

To illustrate, consider the prototypical atomic force microscope (AFM) stage depicted in Figure 1(a) which employs stacked piezoceramic actuators to position the sample in the x and y directions. An additional PZT mechanism provides transverse positioning capabilities. Nested minor loops collected at 0.1 Hz are plotted in Figure 1(b) and data collected at frequencies ranging from 0.279 Hz to 27.9 Hz is plotted in Figure 2 to illustrate the frequency-dependent nature of the hysteresis inherent to field-displacement data.

At low frequencies, the inherent hysteresis can be accommodated through proportional-integral-derivative (PID) or robust control designs [5, 6, 15, 19]. However, at the higher frequencies required for applications ranging from real-time monitoring of biological processes — e.g., protein unfolding — to comprehensive product diagnostics, increasing noise-to-data ratios and diminishing high-pass characteristics of control filters preclude a sole reliance on feedback laws to eliminate hysteresis.

Alternatively, it is illustrated in [12, 13], that use of charge- or current-controlled amplifiers can essentially eliminate hysteresis. However, this mode of operation can be prohibitively expensive when compared with the more commonly employed voltage-controlled amplifiers, and current control is ineffective if maintaining DC offsets as is the case when the x-stage of an AFM is held in a fixed position while a sweep is performed with the y-stage.

This motivates the development of models and model-based control designs which incorporate and accommodate the hysteresis and constitutive nonlinearities. Numerous approaches have been employed to characterize these nonlinear effects including Preisach models [7, 18], domain wall models [25, 26], micromechanical models [4, 10, 11], mesoscopic energy relations [3, 9] and homogenized energy models [23, 30]. We employ the homogenized energy framework due to its energy basis, its capability to quantify a wide range of physical phenomena and operating regimes, its unified nature

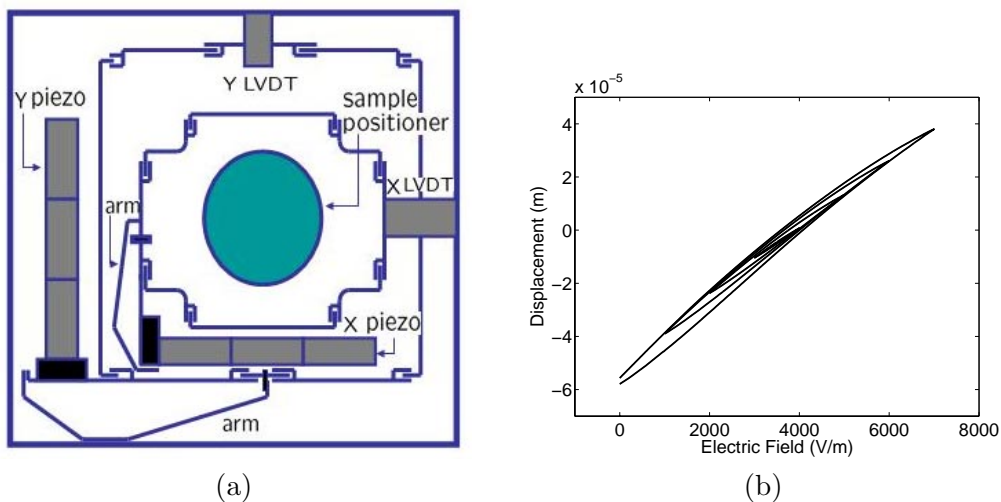(a)   (b)

Figure 1: (a) PZT-based AFM stage, and (b) nested minor loops in data collected at 0.1 Hz.
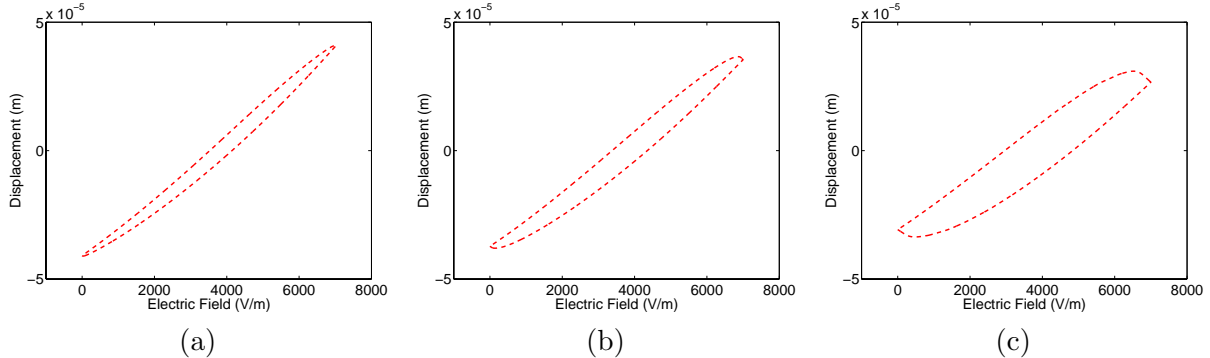
Figure 2: Frequency-dependent field-displacement behavior of a stacked PZT stage: sample rates of (a) 0.279 Hz, (b) 5.58 Hz, and (c) 27.9 Hz.

for characterizing hysteresis in ferroelectric, ferromagnetic and ferroelastic compounds [28, 29], and the potential it provides for real-time implementation. Details regarding the development of this modeling framework and its relation to other characterization techniques can be found in [21, 30].

Model-based control design for piezoceramic transducers operating in highly hysteretic and non-linear regimes can be roughly segregated into two categories: (i) nonlinear control designs, and (ii) linear control designs employing nonlinear inverse filters. Examples of the first technique in the context of optimal control design for smart material transducers are provided in [20, 32]. The second technique is based on the concept of employing either an exact or approximate inverse model to linearize the transducer behavior in the manner depicted in Figure 3. This approach has been employed with a variety of models and control designs — e.g., see [31] for details regarding the development of adaptive control designs utilizing piecewise linear Preisach models and their inverses — and is the technique which we focus on in this paper.

In Section 2 we summarize constitutive relations developed in [21, 30] for ferroelectric materials and provide a highly efficient algorithm for implementing the model when thermal relaxation is negligible. A corresponding inverse polarization-field algorithm is summarized in Section 3 and illustrated through a numerical example. The constitutive model is subsequently employed in Section 4 to develop a lumped model for the stacked actuator employed in the AFM stage shown in Figure 1(a) to illustrate the construction of a macroscopic transducer model. The accuracy of the transducer model is illustrated through a comparison with the frequency-dependent data plotted in Figure 2. In Section 5, an algorithm for the inverse displacement-field relation to linearize the transducer response is developed and, in Section 6, the algorithm is experimentally implemented as an inverse filter for the open loop tracking of a triangular input signal. It is demonstrated that this model-based filter design effectively linearizes the nonlinear and hysteretic transducer dynamics and provides an approximately tenfold increase in accuracy at higher frequencies as compared with
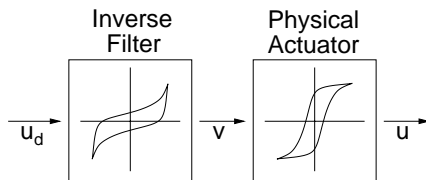

Figure 3: Use of an inverse filter to linearize the response u of a hysteretic actuator to achieve a desired output u_d.

2

the unfiltered case. This significantly improves the accuracy of the transducer and diminishes the sole reliance on feedback laws whose authority decrease as tracking speeds and noise-to-data ratios increase.

2 Constitutive Relations

Constitutive relations quantifying the electromechanical behavior of piezoceramic materials are constructed in two steps. In the first, Helmholtz and Gibbs energy relations at the lattice level are used to characterize the local field-polarization and field-strain behavior of ferroelectric compounds for thermally inactive and active operating regimes. In the second step of the development, material nonhomogeneities and variable effective field effects are incorporated through the assumption that certain material properties are manifestations of underlying distributions rather than constants. This yields low-order macroscopic constitutive relations which are efficient to implement.

2.1 Local Constitutive Relations

Let E, P, ε and σ respectively denote the electric field, polarization, strain and stress. It is illustrated in [30] that an appropriate formulation of the Helmholtz energy for fixed temperatures in the absence of stresses is

$$\psi_P(P) = \begin{cases} \frac{1}{2}\eta(P + P_R)^2 & , \ P \leq -P_I \\ \frac{1}{2}\eta(P - P_R)^2 & , \ P \geq P_I \\ \frac{1}{2}\eta(P_I - P_R)\left(\frac{P^2}{P_I} - P_R\right) & , \ |P| < P_I. \end{cases}$$

As shown in Figure 4, P_I is the positive inflection point which delineates the transition between stable and unstable regions, P_0 denotes the unstable equilibrium, and P_R is the value of P at which the positive local minimum of ψ occurs. The parameter η is the reciprocal of the slope of the E-P relation after switching occurs. This fact can be used to establish an initial parameter value for η when modeling a specific data set.
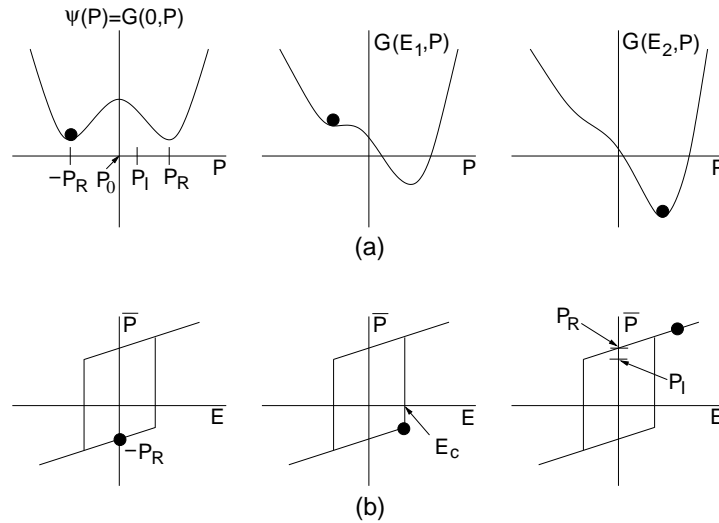

Figure 4: (a) Helmholtz energy ψ and Gibbs energy G for $\sigma = 0$ and increasing fields E. (b) Switch in the local polarization $\overline{P}$ that occurs as E is increased beyond the local coercive field E_c given by (5) in the absence of thermal activation.

The corresponding Gibbs energy relation

$$G_P(E, P) = \psi_P(P) - EP \tag{1}$$

incorporates the electrostatic energy due to the applied field E when $\sigma = 0$.

Elastic effects and electromechanical coupling are incorporated in the Helmholtz energy relation

$$\psi(P, \varepsilon) = \psi_P(P) + \frac{1}{2}Y\varepsilon^2 - a_1\varepsilon P - a_2\varepsilon P^2.$$

The Gibbs energy is then given by

$$G(E, \sigma, P, \varepsilon) = \psi_P(P) + \frac{1}{2}Y\varepsilon^2 - a_1\varepsilon P - a_2\varepsilon P^2 - EP - \sigma\varepsilon \tag{2}$$

where $\sigma\varepsilon$ incorporates the elastic energy. Note that Y denotes the Young's modulus and a_1, a_2 are ferroelastic coupling coefficients associated with linear piezoelectric and quadratic electrostrictive effects.

Polarization Kernel — Negligible Thermal Activation

In the case of negligible thermal activation, the local average polarization kernel $\overline{P}$ is determined from the necessary conditions

$$\frac{\partial G}{\partial P} = 0 \ , \ \frac{\partial^2 G}{\partial P^2} > 0.$$

Applying these conditions to (1) yields the piecewise linear E-P relation

$$\overline{P}(E) = \frac{1}{\eta}E + P_R\delta \tag{3}$$

where $\delta = -1$ for negatively oriented dipoles and $\delta = 1$ for those with positive orientation. To specify δ, and hence $\overline{P}$, more specifically in terms of the initial dipole orientations and previous switches, we employ Preisach notation and take

$$[\overline{P}(E; E_c, \xi)](t) = \begin{cases} [\overline{P}(E; E_c, \xi)](0) \ , \ \tau(t) = \emptyset \\ \frac{E}{\eta} - P_R \qquad , \ \tau(t) \neq \emptyset \text{ and } E(\max \tau(t)) = -E_c \\ \frac{E}{\eta} + P_R \qquad , \ \tau(t) \neq \emptyset \text{ and } E(\max \tau(t)) = E_c. \end{cases} \tag{4}$$

Here

$$[\overline{P}(E; E_c, \xi)](0) = \begin{cases} \frac{E}{\eta} - P_R \ , \ E(0) \leq -E_c \\ \xi \qquad , \ -E_c < E(0) < E_c \\ \frac{E}{\eta} + P_R \ , \ E(0) \geq E_c \end{cases}$$

defines initial kernel values in terms of the parameter $\xi = \frac{E_0}{\eta} \pm P_R$, $\emptyset$ designates the empty set, and the set of switching times is given by

$$\tau(t) = \{t_s \in (0, t] \mid E(t_s) = -E_c \text{ or } E(t_s) = E_c\}.$$

The local coercive field

$$E_c = \eta(P_R - P_I) \tag{5}$$

quantifies the field at which the negative well ceases to exist and hence a dipole switch occurs. To illustrate, the condition $\tau \neq \emptyset$ and $E(\max \tau(t)) = E_c$ designates that switching has occurred and the last switch was at E_c; hence the local polarization is $[\overline{P}(E; E_c, \xi)](t) = \frac{E(t)}{\eta} + P_R$

4

Polarization Kernel — Thermal Activation

If thermal activation is significant, dipoles can achieve the thermal energy required to switch in advance of the minimum Gibbs energy so the relative thermal energy kT/V and Gibbs energy G must be balanced through Boltzmann principles. The probability density for achieving an energy level G is given by

$$\mu(G) = Ce^{-GV/kT} \tag{6}$$

where k is Boltzmann's constant, V is a reference volume and C is a constant that is selected so that when $\mu(G)$ is integrated over all possible dipole orientations, a probability of unity is achieved. If we let 2ϵ be the separation between possible polarization states around P_0, the probabilities of reaching a polarization state having sufficient energy to switch orientations are given by

$$r_{+-} = \frac{\int_{P_0-\epsilon}^{P_0+\epsilon} e^{-G(E,P)V/kT}dP}{\int_{P_0-\epsilon}^{\infty} e^{-G(E,P)V/kT}dP} \quad , \quad r_{-+} = \frac{\int_{P_0-\epsilon}^{P_0+\epsilon} e^{-G(E,P)V/kT}dP}{\int_{-\infty}^{P_0+\epsilon} e^{-G(E,P)V/kT}dP}. \tag{7}$$

The likelihoods of reaching the required energy and thus of the dipoles switching from a positive to a negative orientation and conversely are then

$$p_{+-} = \frac{r_{+-}}{\mathcal{T}(T)} \quad , \quad p_{-+} = \frac{r_{-+}}{\mathcal{T}(T)} \tag{8}$$

where $\mathcal{T}(T)$ is the relaxation time at temperature T. The fractions of dipoles in each orientation evolve according to the ordinary differential equations

$$\frac{dx_+}{dt} = -p_{+-}x_+ + p_{-+}x_-$$

$$\frac{dx_-}{dt} = -p_{-+}x_- + p_{+-}x_+.$$

The expected polarizations due to positively and negatively oriented dipoles are

$$\langle P_+ \rangle = \int_{P_0+\epsilon}^{\infty} P\mu(G)dP \quad , \quad \langle P_+ \rangle = \int_{-\infty}^{P_0-\epsilon} P\mu(G)dP$$

so that evaluation of C yields

$$\langle P_+ \rangle = \frac{\int_{P_0+\epsilon}^{\infty} Pe^{-G(E,P,T)V/kT}dP}{\int_{P_0+\epsilon}^{\infty} e^{-G(E,P,T)V/kT}dP} \quad , \quad \langle P_- \rangle = \frac{\int_{-\infty}^{P_0-\epsilon} Pe^{-G(E,P,T)V/kT}dP}{\int_{-\infty}^{P_0-\epsilon} e^{-G(E,P,T)V/kT}dP}. \tag{9}$$

The local average polarization is subsequently

$$\overline{P} = x_+\langle P_+ \rangle + x_-\langle P_- \rangle. \tag{10}$$

In the manner detailed in [30], the evaluation of the integrals in (7) and (9) can be simplified through approximations employing the inflection points $\pm P_I$ rather than the unstable equilibrium P_0.

2.2 Global Constitutive Relations

For homogeneous compounds with uniform effective fields E_e, the local lattice relations (3), (4) or (10) can be extrapolated throughout the material to provide global constitutive relations. This yields the nearly instantaneous transitions at coercivity that are associated with certain single crystal

compounds — e.g., the hysteresis kernels depicted in Figure 4 provide a reasonable characterization of the single crystal $BaTiO_3$ behavior shown on pages 72-76 of [14] — but provide a poor characterization of the mollified transition behavior of general ferroelectric compounds. To incorporate the effects of material nonhomogeneities, polycrystallinity, and variable effective fields $E_e = E + E_I$, we assume that the interaction field E_I, due to neighboring dipoles and certain electromechanical interactions [1], and local coercive field E_c given by (5) are manifestations of underlying distributions rather than constants. If we designate the associated densities by ν_1 and ν_2, the macroscopic field-polarization behavior is quantified by the relation

$$[P(E)](t) = \int_0^\infty \int_{-\infty}^\infty [\overline{P}(E + E_I; E_c, \xi)](t)\nu_1(E_c)\nu_2(E_I)dE_I dE_c \qquad (11)$$

where the kernel $\overline{P}$ is given by (3), (4) or (10).

As detailed in [23], the densities ν_1 and ν_2 are assumed to satisfy the physical criteria

$$
\begin{array}{ll}
\text{(i)} & \nu_1(x) \ \text{ defined for } \ x > 0, \\
\text{(ii)} & \nu_2(-x) = \nu_2(x), \\
\text{(iii)} & |\nu_1(x)| \le c_1 e^{-a_1 x}, \\
& |\nu_2(x)| \le c_2 e^{-a_2|x|}
\end{array}
\qquad (12)
$$

for positive c_1, a_1, c_2, a_2. The restricted domain in (i) reflects the fact that the coercive field E_c is positive whereas the symmetry enforced in the interaction field through (ii) yields the symmetry observed in low-field Rayleigh loops. Hypothesis (iii) incorporates the physical observation that the coercive and interaction fields decay as a function of distance and guarantees that integration against the piecewise linear kernel yields finite polarization values.

By employing numerical integration routines tailored to the infinite domains or truncated intervals resulting from the decay criteria (12), the integrals in (11) can be approximated to obtain the discretized model

$$[P(E)](t) = \sum_{i=1}^{N_i} \sum_{j=1}^{N_j} [\overline{P}(E + E_{I_j}; E_{c_i}, \xi_j)]\nu_1(E_{c_i})\nu_2(E_{I_j})v_i w_j. \qquad (13)$$

Specific choices for the weights v_i, w_j and abscissas E_{e_j}, E_{c_i} are detailed in [21, 30].

Techniques for identifying the densities ν_1 and ν_2 are illustrated in [23]. For certain applications, reasonable accuracy is provided by *a priori* functions satisfying the physical criteria (12) and having a small number of parameters to be estimated through least squares fits to data — e.g., variances and means in normal and lognormal relations. For more general applications requiring high accuracy for a wide range of operating conditions, the $N_i + N_j$ discretized density values $\nu_1(E_{c_i})$ and $\nu_2(E_{I_j})$ can be estimated through least squares techniques.

To obtain an elastic constitutive relation, the equilibrium condition

$$\frac{\partial G}{\partial \varepsilon} = 0$$

is invoked to obtain

$$\sigma = Y\varepsilon - a_1 P - a_2 P^2. \qquad (14)$$

When $P = 0$, (14) reduces to Hooke's law. To incorporate internal damping, we posit that when $P = 0$, stress is proportional to a linear combination of strain and strain rate (Kelvin–Voigt damping hypothesis). This yields the constitutive relation

$$\sigma = Y\varepsilon - c_D \dot{\varepsilon} - a_1 P - a_2 P^2 \qquad (15)$$

6

where c_D is the Kelvin–Voigt damping parameter. The combination of the field-polarization model (11) or (13) and the electromechanical relation (15) are employed in Section 4 to construct a lumped model for a stacked PZT actuator operating in hysteretic and nonlinear regimes.

2.3 Implementation Algorithm

The efficiency of inverse algorithms used to construct inverse filters is dependent on the efficiency of forward algorithms used to implement the discretized polarization model (13). In this section, we summarize a highly efficient algorithm to evaluate (13) when the kernel $\overline{P}$ is given by (3) or (4) for regimes in which relaxation processes are negligible. Analogous algorithms for the thermally active kernel (10) are reported in [2].

From (4), it is observed that for each field value E_{I_j}, it is necessary to determine whether a transition has occurred relative to the coercive value E_{c_i}. This yields $N_i \times N_j$ conditions to be checked for each input value. While this can be easily accomplished using an `if-then` construct, implementation in this manner diminishes significantly the efficiency of the algorithm. This motivates consideration of an algebraic technique for evaluating the conditional statements.

To retain the history of whether or not effective field values $E_{e_j} = E + E_{I_j}$ have switched due to encounters with coercive field values E_{c_i}, we employ (3) to motivate the matrix formulation

$$\overline{P} = \frac{E}{\eta} + P_R \Delta(E; E_c, E_I)$$

where $\Delta = 1$ if evaluating on the upper branch of the hysteron (hystersis kernel) and $\Delta = -1$ if on the lower branch. For the evaluation of (13), Δ is an $N_i \times N_j$ matrix whose ijth component specifies whether E_{I_j} has reached the coercive value E_{c_i}. To specify Δ, we define the matrices

$$\Delta_{init} = \begin{bmatrix} -1 & \cdots & -1 & 1 & \cdots & 1 \\ \vdots & & \vdots & \vdots & & \vdots \\ -1 & \cdots & -1 & 1 & \cdots & 1 \end{bmatrix}_{N_i \times N_j} , \quad \mathcal{E}_c = \begin{bmatrix} E_{c_1} & \cdots & E_{c_1} \\ \vdots & & \vdots \\ E_{c_{N_i}} & \cdots & E_{c_{N_i}} \end{bmatrix}_{N_i \times N_j}$$

$$\mathcal{E}_k = \begin{bmatrix} E_k + E_{I_1} & \cdots & E_k + E_{I_{N_j}} \\ \vdots & & \vdots \\ E_k + E_{I_1} & \cdots & E_k + E_{I_{N_j}} \end{bmatrix}_{N_i \times N_j}$$

and weight vectors

$$V^T = \left[v_1 \nu_1(E_{c_1}), \cdots , v_{N_i} \nu_1(E_{c_{N_i}}) \right]_{1 \times N_i}$$

$$W^T = \left[w_1 \nu_2(E_{I_1}), \cdots , w_{N_j} \nu_2(E_{I_{N_j}}) \right]_{1 \times N_j}$$

where $E_k = E(t_k)$ is the kth value of the input field. The points E_{c_i} and E_{I_j} in the definitions are determined by the quadrature rule being employed on intervals $[0, E_{c_{max}}]$ and $[E_{I_{min}}, E_{I_{max}}]$ chosen according to the physical decay conditions (12) — i.e., the densities ν_1 and ν_2 are negligible outside these regions.

The polarization $P_k \approx P(E_k)$ is specified by Algorithm 1. In this algorithm, $.*$ indicates componentwise matrix multiplication and sgn denotes the signum function. The first step in the `for`-loop updates Δ by incorporating the status of the previous coercive field switch.

Depending on the methods used for programming, the use of Algorithm 1 rather than utilizing conditional `if-then` constructs can reduce runtimes by factors in excess of 100 so that full multiloop

model simulations run in the order of seconds on a workstation. This level of efficiency is necessary to achieve real-time implementation of control algorithms utilizing the model. Finally, we note that the algorithms reported in [2] for the kernel (10) which incorporates thermal activation are on the order of 2-3 times slower than Algorithm 1. While this produces analogous reductions in the speed of inverse algorithms, the resulting model is still sufficiently efficient to facilitate real-time implementation.

Algorithm 1.

$\Delta_{prev} = \Delta_{init}$

for $k = 1 : N_k$

$\quad \Delta = \text{sgn}(\mathcal{E}_k + \mathcal{E}_c \cdot * \Delta_{prev})$

$\quad \overline{P} = \frac{1}{\eta}\mathcal{E}_k + P_R \Delta$

$\quad P_k = V^T \overline{P} W$

$\quad \Delta_{prev} = \Delta$

end

3   Inverse Relation Between Polarization and Field

The model (11) and discretized model (13) quantify the relation between input fields and the polarization generated in ferroelectric materials. To construct an inverse filter of the type illustrated in Figure 3, it is necessary to quantify the inverse P-E relation. We summarize here an algorithm which accomplishes this when the E-P relation is assumed monotone and illustrate the filtering process depicted in Figure 3 through a numerical example. Extension of the algorithm to accommodate the non-monotone field-displacement behavior shown in Figure 2 is addressed in Section 5.

The first step in the construction of an inverse filter involves the determination of an initial (E_0, P_0) value. This is typically done with $E_0 = 0$ so that $P_0 = \pm P_R$ is the positive or negative remanence value or $P_0 = 0$ for depoled materials. The values of $\pm P_R$ can be computed using the forward model (11) or (13) by taking one step to $\pm E_{max}$ chosen sufficiently large that all dipoles have switched and then stepping back to $E = 0$. For a specified value of P, monotonicity in the E-P relation is exploited and the forward model is subsequently advanced until the prescribed polarization is crossed. Interpolation is then used to specify a final field value corresponding to the prescribed polarization. This process is outlined in Algorithm 2 where specified polarization values are designated by $\{P_k^*\}$ and computed values by $\{P_k\}$ for $k = 1, \cdots N_k$.

Algorithm 2.

for $k = 2 : N_k$

$\quad$ Specify $E_{step} > 0$ as fixed or adaptive

$\quad dP = P_k^* - P_{k-1}^*$

$\quad \Delta E = dP \cdot E_{step}$

$\quad E_{tmp} = E_{k-1}, \ P_{tmp} = P_{k-1}$

$\quad$ while $\text{sgn}(dP) \cdot (P_k^* - P_{tmp}) >= 0$

$\quad\quad E_{tmp} = E_{tmp} + \Delta E$

$\quad\quad P_{tmp}$ given by (13)— e.g., as implemented in Algorithm 1

$\quad$ end

$\quad E_k$ given by linear interpolation

end

The flexibility and robustness provided by the inverse Algorithm 2 are illustrated in Figure 5. The polarization plotted in Figure 5(c) is employed as input to Algorithm 2 to yield the P-E relation plotted in Figure 5(a). At each time step, the resulting field value is then employed as input to the forward Algorithm 1 to yield the E-P curve shown in Figure 5(b). These output polarization values P_{out} are compared with inputs P_{in} in Figure 5(c) and the absolute errors $|P_{in} - P_{out}|$ are plotted in Figure 5(d).

From these results, a number of conclusions can be drawn. (i) We first note that the model and its inverse provide the capability for characterizing a wide range of symmetric and biased minor loop behavior — e.g., see [21, 23, 30]. (ii) The composition of the inverse and model in the manner depicted in Figure 3 can effectively linearize the nonlinear transducer behavior with the numerical accuracy $|P_{in} - P_{out}|$ limited only by the stepsize dP. Whereas the accuracy in a physical system will be diminished due to modeling error, linearization in this manner can significantly improve control authority since less control effort is focused on unmodeled or nonlinear dynamics. This forms the crux of various linear control designs [15, 16]. (iii) Although faster implementation algorithms can be constructed for the inversion process [8], the algorithm described here is highly robust and avoids the potential for losing track of the memory incorporated in the model. Furthermore, the use of adaptive stepsizes ΔE ensures that Algorithm 2 is approximately a factor of two slower than the forward algorithm which is reasonable for physical implementation. (iv) Whereas Algorithm 2 employs the limiting piecewise linear kernel $\overline{P}$ given by (4), analogous algorithms have been developed for the more general kernel (10) which incorporate thermal relaxation and additional dynamic effects [8].
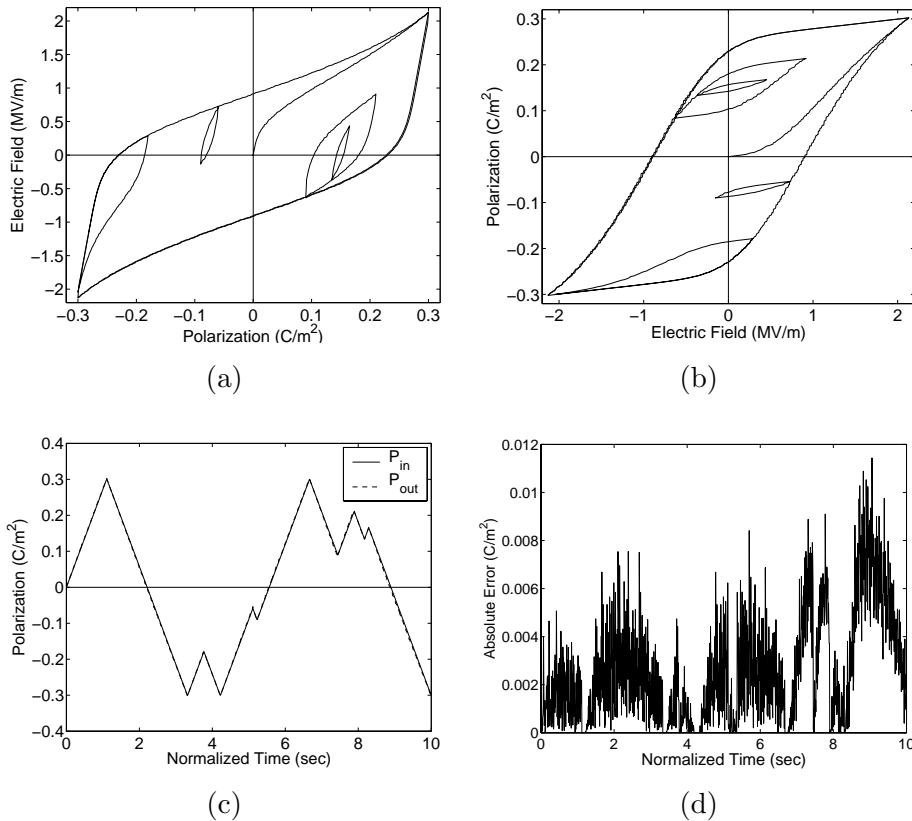

Figure 5: (a) Inverse relation P_{in}-E_{out} given by Algorithm 2. (b) Forward relation E_{out}-P_{out} from Algorithm 1. (c) Comparison between P_{in} and P_{out}. (d) Absolute error $|P_{in} - P_{out}|$ for complete inversion process depicted in Figure 3.

9

4 Lumped Model for the Stacked Actuator

The constitutive relation (15) quantifies the electromechanical behavior of piezoceramic materials operating below the coercive stress σ_c where ferroelastic switching commences. In this section we use this relation to construct a macroscopic model for the stacked PZT rod employed as an AFM stage in the manner depicted in Figure 1(a). While this illustrates macroscopic model development for a specific application, similar principles hold for other transducer designs — e.g., see [21].

We assume that the stacked actuator or rod has length ℓ, cross-sectional area A, density ρ, Young's modulus Y, and Kelvin–Voigt damping parameter c_D. We also assume that the end at $x = 0$ is fixed whereas the end at $x = \ell$ is subjected to inertial, elastic and damping forces associated with the stage mechanisms. Because material properties and forces along the length of the rod are uniform, we consider a lumped model quantifying the displacement $u(t)$ at $x = \ell$. The validity of the lumped ODE model as compared with a distributed PDE model is established in [24]. The geometry, mass m_ℓ, stiffness k_ℓ, and damping mechanisms c_ℓ associated with the end-forces at $x = \ell$ are depicted in Figure 6.

From the assumption of uniform stresses and strains through the length of the rod, it follows that

$$\varepsilon(t) = \frac{u(t)}{\ell}$$

in the stress relation (15). Balancing the forces σA for the rod with those of the stage mechanisms yields the lumped model

$$\rho A \frac{d^2 u}{dt^2}(t) + \frac{c_D A}{\ell} \frac{du}{dt}(t) + \frac{Y A}{\ell} u(t) - A a_1 P(E(t)) - A a_2 P^2(E(t)) = -m_\ell \frac{d^2 u}{dt^2}(t) - c_\ell \frac{du}{dt}(t) - k u(t)$$

or, equivalently,

$$m \frac{d^2 u}{dt^2}(t) + c \frac{du}{dt}(t) + k u(t) = \widetilde{a}_1 P(E(t)) + \widetilde{a}_2 P^2(E(t)) \tag{16}$$

where

$$m = \rho A + m_\ell \quad , \quad c = \frac{c_D A}{\ell} + c_\ell \quad , \quad k = \frac{Y A}{\ell} + k_\ell \quad , \quad \widetilde{a}_1 = A a_1 \quad , \quad \widetilde{a}_2 = A a_2$$

and the initial conditions are $u(0) = u_0$ and $\frac{du}{dt}(0) = u_1$. The polarization P is specified by the model (11) or discretized model (13).

The model can also be written as the first-order system

$$\dot{\vec{u}}(t) = A \vec{u}(t) + \vec{\mathcal{P}}(E(t))$$
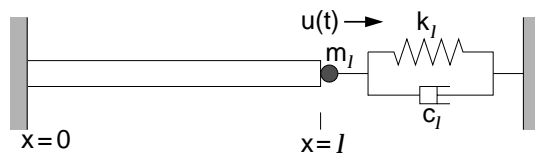$$\vec{u}(0) = \vec{u}_0 \tag{17}$$

Figure 6: Rod geometry used when modeling the stacked PZT actuator employed in the AFM stage depicted in Figure 1(a).

where $\vec{u}(t) = [u(t), \dot{u}(t)]^T$, $\vec{u}(0) = [u_0, u_1]^T$ and

$$A = \begin{bmatrix} 0 & 1 \\ -k/m & -c/m \end{bmatrix} \quad , \quad \vec{\mathcal{P}}(E(t)) = \frac{1}{m}\left[\tilde{a}_1 P(E(t)) + \tilde{a}_2 P^2(E(t))\right] \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

This formulation proves advantageous in the next section when establishing notation used in the construction of the inverse algorithm.

The accuracy of the framework is illustrated in Figure 7 where the lumped model (16) with P specified by (13) is used to characterize the frequency-dependent dynamics of the PZT stacked actuator employed in the AFM stage depicted in Figure 1(a). When constructing the polarization model, the general densities ν_1 and ν_2 were identified using the least squares techniques detailed in [23]. It is noted that the combined model quantifies both the hysteresis and dynamic effects observed as frequencies are increased. Additional details regarding the construction and validation of the stacked actuator model for the AFM stage are provided in [8,24] whereas additional examples demonstrating properties of the model for characterizing hysteresis in various PZT compounds can be found in [21–23,30].
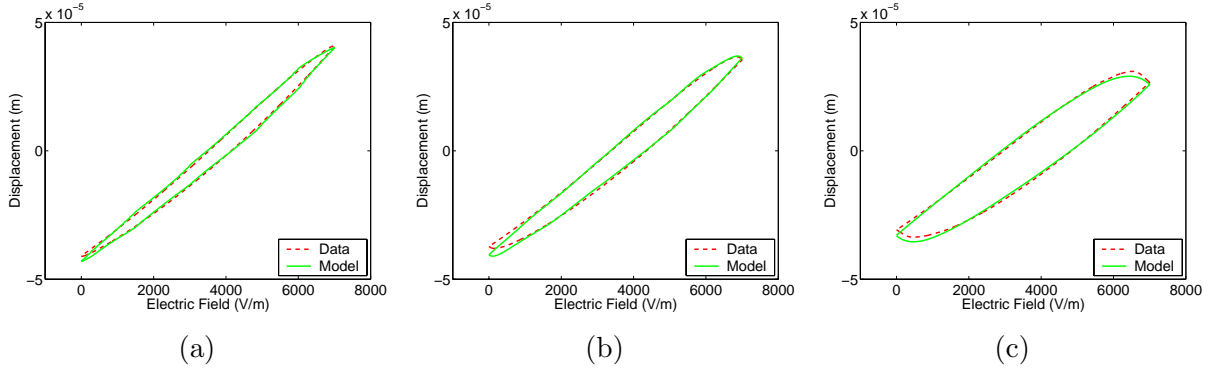

Figure 7: Characterization of AFM field-displacement behavior with sample rates of (a) 0.279 Hz, (b) 5.58 Hz and (c) 27.9 Hz.

5 Inverse Relation Between Displacements and Fields

The inversion algorithm summarized in Section 3 relies on the monotonicity of the E-P relation. As illustrated in Figure 2, this property is not shared by the E-u relation as frequencies are increased so we develop here an extended inversion algorithm which incorporates this non-monotone behavior. The crux of the modification focuses on the accommodation of dynamic effects in the E-u behavior.

To establish notation used when quantifying dynamic effects, we employ modified semigroup notation to define solution values

$$
\begin{aligned}
u_{k+1} &= u(t_{k+1}, t_k, E, \vec{u}_k) = Ce^{A(t_{k+1}-t_k)}\vec{u}_k + C\int_{t_k}^{t_{k+1}} e^{A(t_{k+1}-s)}\vec{\mathcal{P}}(E(s))ds \\
\tilde{u}_{k+1} &= u(t_{k+1}, t_k, E_k, \vec{u}_k) = Ce^{A(t_{k+1}-t_k)}\vec{u}_k + C\int_{t_k}^{t_{k+1}} e^{A(t_{k+1}-s)}\vec{\mathcal{P}}(E_k)ds
\end{aligned}
\tag{18}
$$

where $C = [1, 0]$ and $E_k = E(t_k)$. Hence u_{k+1} is the solution to (16) or (17) with the electromagnetic force applied throughout the time interval $[t_k, t_{k+1}]$ whereas $\tilde{u}_{k+1}$ denotes the displacement of the

rod at $x = \ell$ with the force maintained constant at the kth value throughout the interval $[t_k, t_{k+1}]$ as depicted in Figure 8. The definition of u_k is similar to that of u_{k+1}.

The upper and lower hysteresis curves are denoted by

$$c_L = \left\{ (E, u) \,|\, \frac{dE}{dt} \geq 0 \right\}$$

$$c_U = \left\{ (E, u) \,|\, \frac{dE}{dt} \leq 0 \right\}.$$

Due to dynamic (inertial) effects, it is observed that

$$\widetilde{u}_{k+1} \leq u_{k+1} \text{ on } c_L$$

$$\widetilde{u}_{k+1} \geq u_{k+1} \text{ on } c_U$$

(19)

as depicted in Figure 8(b) and (c). To motivate the relation on c_L, we note that (18) yields

$$u_{k+1} - \widetilde{u}_{k+1} = C \int_{t_k}^{t_{k+1}} e^{A(t_{k+1}-s)} \left[\vec{\mathcal{P}}(E(s)) - \vec{\mathcal{P}}(E_k) \right] ds. \tag{20}$$

We now establish that the right hand side of (20) is nonnegative given that the monotonicity of the E-P relation implies that $\vec{\mathcal{P}}(E(s)) - \vec{\mathcal{P}}(E_k) \geq 0$ for $s \geq t_k$. We first note that the Cayley–Hamilton theorem dictates that $e^{A(t_{k+1}-s)} = \alpha_0 I + \alpha_1 A$. Furthermore, it follows from the definitions of C and $\vec{\mathcal{P}}(E(t))$ that only the $(1,2)$ entry of $e^{A(t_{k+1}-s)}$ contributes to the right hand side of (20). Moreover, it follows from the definition of A that the $(1,2)$ entry of $e^{A(t_{k+1}-s)}$ is simply α_1.

To determine α_1, we note that there are two possibilities for the eigenvalues of A: (i) both are real, distinct, and negative, and (ii) they are a conjugate pair with negative real part. We consider the first case where the eigenvalues satisfy $\lambda_2 < \lambda_1 < 0$. It follows from the Cayley–Hamilton theorem that

$$e^{\lambda_1(t_{k+1}-s)} = \alpha_0 + \alpha_1 \lambda_1$$

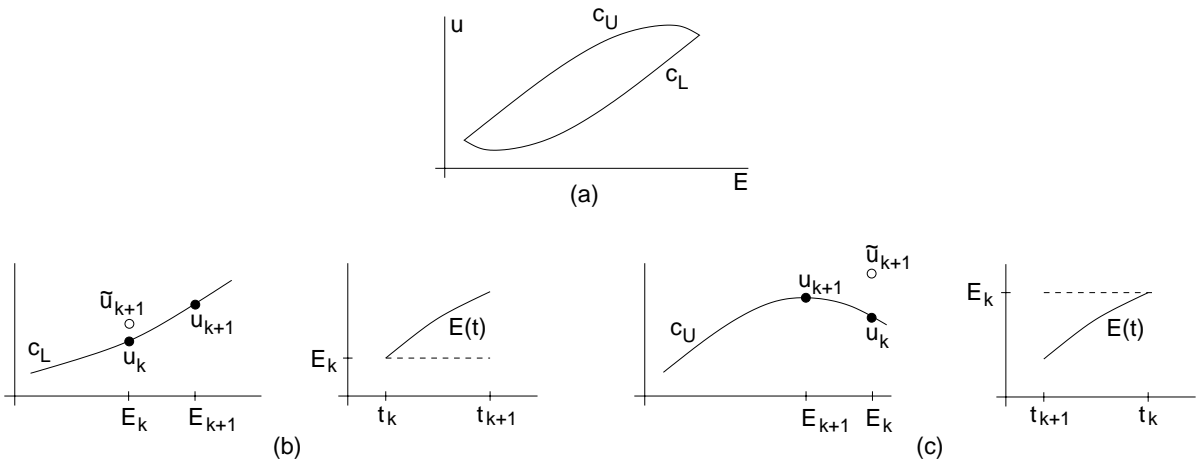$$e^{\lambda_2(t_{k+1}-s)} = \alpha_0 + \alpha_1 \lambda_2$$

Figure 8: Non-monotone behavior of the field-displacement relation measured at 27.9 Hz as shown in Figure 2(c). Solution values $\widetilde{u}_{k+1}$ and u_{k+1} respectively due to field inputs E_k and $E(t)$ on the (b) lower loop c_L, and (c) upper loop c_U of the curve.

so that

$$\alpha_1 = \frac{e^{\lambda_1(t_{k+1}-s)} - e^{\lambda_2(t_{k+1}-s)}}{\lambda_1 - \lambda_2} > 0.$$

Similar analysis holds for the second case. Thus the integrand $Ce^{A(t_{k+1}-s)}[\vec{\mathcal{P}}(E(s)) - \vec{\mathcal{P}}(E_k)]$ of (20) is nonnegative and consequently so is the integral. As a result, $u_{k+1} - \widetilde{u}_{k+1} \geq 0$ or equivalently $\widetilde{u}_{k+1} \leq u_{k+1}$. The argument for c_U is analogous.

In the inverse algorithm used to specify the u-E relation given data values $\{u_k^*\}$, $k = 1, \cdots, N_k$, the inequalities (19) are applied to either the exact or discretized solution of (16) to determine the appropriate sign of ΔE when implementing the polarization component (13) of the model. The notation u_{tmp} and $\widetilde{u}_{tmp}$ designate either the exact or approximate solutions to (13) having the interpretation specified in (18). The resulting inversion process is outlined in Algorithm 3. Attributes of the algorithm are illustrated in the next section in the context of experimental open loop control implementation.

Algorithm 3.

> for $k = 2 : N_k$
>> Specify $E_{step} > 0$ as fixed or adaptive
>> Specify Δt
>> $\Delta E = \text{sgn}(u_{k+1}^* - \widetilde{u}_{k+1}) \cdot E_{step}$
>> $E_{tmp} = E_{k-1}$, $P_{tmp} = P_{k-1}$, $t_{tmp} = t_{k-1}$
>> while $\text{sgn}(u_{k+1}^* - \widetilde{u}_{k+1}) \cdot \Delta E >= 0$
>>> $t_{tmp} = t_{tmp} + \Delta t$
>>> compute u_{tmp} as true or approximate solution to (16)
>>> $E_{tmp} = E_{tmp} + \Delta E$
>>> P_{tmp} given by (13)— e.g., as implemented in Algorithm 1
>> end
>> E_k given by linear interpolation
> end

6 Open Loop Control Implementation

To illustrate the effect of filters employing the inverse model developed in Section 5 on open loop tracking performance, we summarize experiments conducted at 0.279 Hz and 27.9 Hz. The trajectory to be tracked consisted of triangle waves having amplitudes of 40.56 μm and 27.04 μm as shown in Figures 9–11.

To specify the model, parameters in the polarization model (13) and lumped rod model (16) were estimated through a least squares fit to field-displacement data collected at 0.279 Hz, 5.58 Hz and 27.9 Hz using the techniques detailed in [23]. This yielded model fits similar to those shown in Figure 7. The model with these parameters was then used to construct an inverse filter using Algorithm 3 of Section 5. In a series of experiments, the specified trajectories were input to the filter and the resulting field was applied to the AFM stage. To provide a metric for comparison, a second input field for each case was determined through a linear scaling of the field-displacement relation. This linear filter accommodated the scaling difference between inputs and outputs but neglected inherent hysteresis and constitutive nonlinearities.
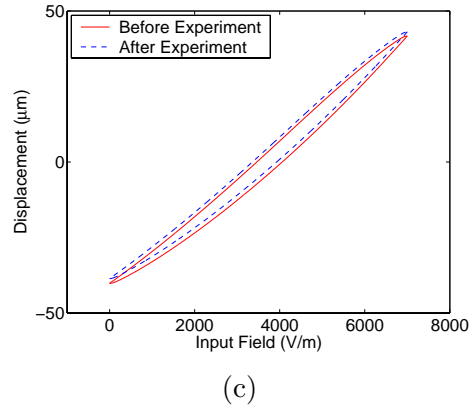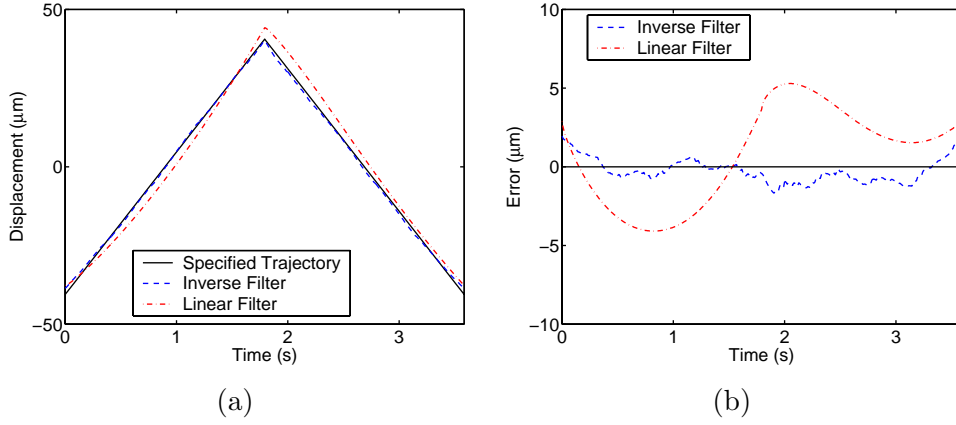
Figure 9: Tracking performance utilizing the model-based inverse filter of Section 5 and a linear filter at 0.279 Hz and amplitude of 40.56 μm. (a) Specified trajectory and tracking provided by the inverse and linear filters, and (b) errors obtained with the two filters). (c) Field-displacement data measured before and after the open loop control experiment.
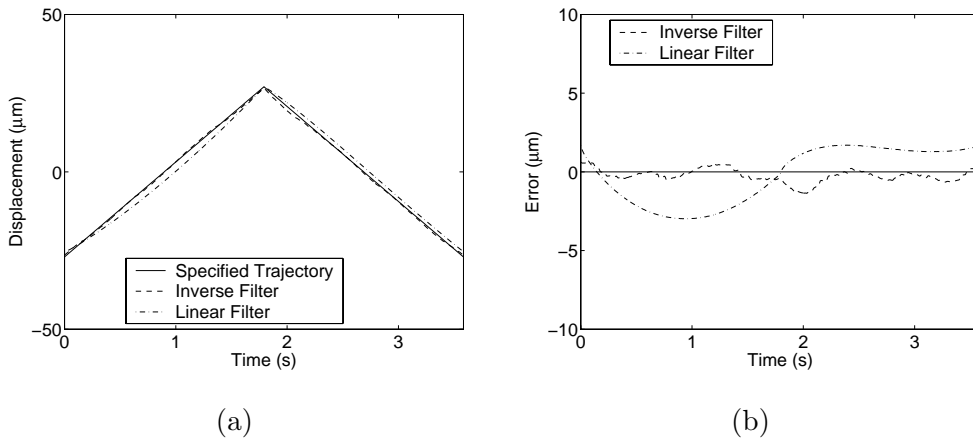

Figure 10: Tracking performance utilizing the model-based inverse filter of Section 5 and a linear filter at 0.279 Hz and amplitude of 27.04 μm. (a) Specified trajectory and tracking provided by the inverse and linear filters, and (b) errors obtained with the two filters).
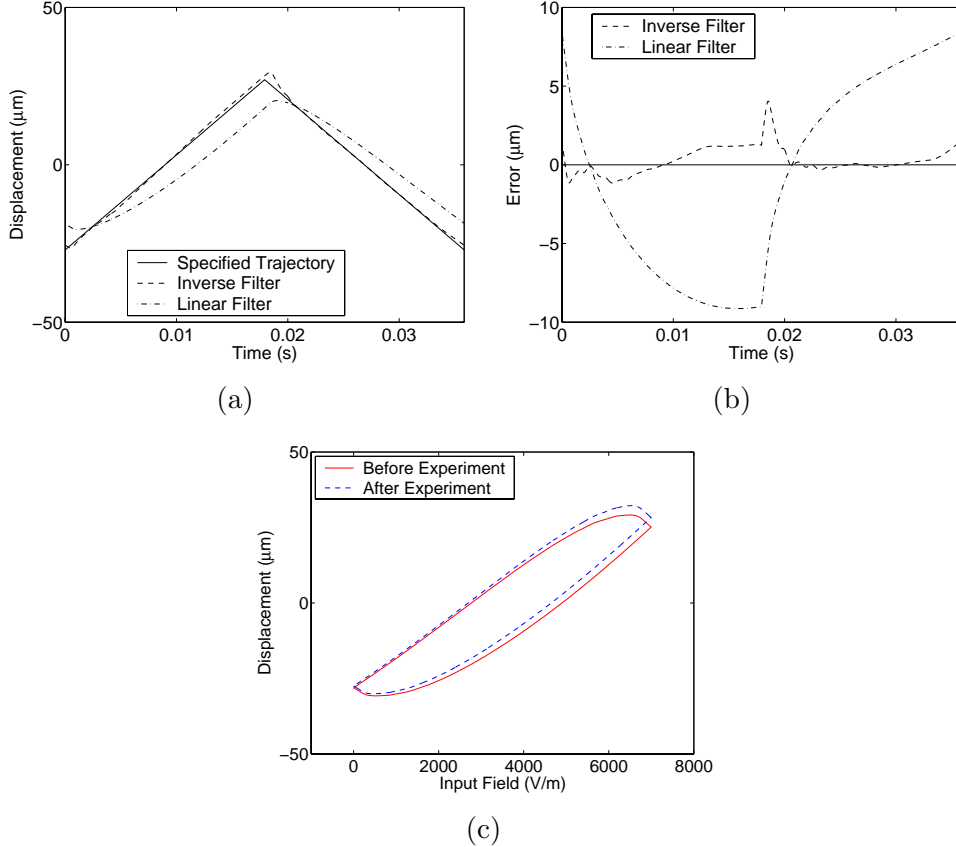
14

Figure 11: Tracking performance utilizing the model-based inverse filter of Section 5 and a linear filter at 27.9 Hz and amplitude of 27.04 μm. (a) Specified trajectory and tracking provided by the inverse and linear filters, and (b) errors obtained with the two filters). (c) Field-displacement data measured before and after the open loop control experiment.

The specified and achieved trajectories, errors, and selected hysteresis plots for the combinations (i) 0.279 Hz, 40.56 μm amplitude, (ii) 0.279 Hz, 27.04 μm amplitude, and (iii) 27.9 Hz, 27.04 μm amplitude are plotted in Figures 9–11. At 0.279 Hz, the filtered design provides only marginally improved accuracy due to the low degree of hysteresis. For the more hysteretic response at 27.9 Hz, however, the inverse filter provides a significant increase in accuracy and yields errors that are approximately a factor of 10 smaller than the linearly scaled case. This illustrates the advantage of incorporating the frequency-dependent model inverse in the control design. We note that the same model parameters were employed in each case thus illustrating the capability of the model and model-based inverse to compensate for the frequency-dependent hysteresis.

The primary source of errors in the filtered design is variability between experiments as illustrated by the variation in the hysteresis plots measured at the two frequencies before and after the open loop control experiments. This is hypothesized to be due to variations in the true applied voltage and illustrates one reason feedback is necessary in final control designs.

7 Concluding Remarks

This paper addresses the development, implementation and experimental validation of a model-based inverse filter to accommodate hysteresis and constitutive nonlinear inherent to the field-polarization

and field-displacement behavior of ferroelectric materials. The nonlinear hysteresis effects are quantified using a previously developed framework consisting of energy relations at the lattice level in combination with stochastic homogenization techniques to provide low-order macroscopic constitutive relations. The development of a lumped transducer model based on these constitutive relations is illustrated in the context of a PZT stage for an AFM. The inverse displacement-field model exploits monotonicity in the E-P relation, the efficiency of forward E-P algorithms, and dynamic properties of the transducer model.

To illustrate attributes of the inverse displacement-field algorithm, it was employed as a filter in open loop tracking experiments with an AFM stage. These experiments illustrate that at low frequencies, where hysteresis and constitutive nonlinearities are minimal, incorporation of the inverse filter provides only marginal improvement in tracking accuracy as compared with a linear filter. However, at higher frequencies where hysteresis becomes significant, the inverse filter yields an approximately tenfold improvement in accuracy compared with the linear filter thus maintaining tracking accuracy even though the transducer is operating in highly hysteretic and nonlinear regimes. Linearization of the electromechanical behavior in this manner reduces the degree to which feedback mechanisms must expend energy linearizing the transducer response and increases control authority for stabilization or tracking.

We note that a significant advantage of the energy-based model is the fact that it provides a unified framework for characterizing hysteresis and constitutive nonlinearities in ferroelectric, ferromagnetic and ferroelastic (e.g., SMA) compounds [28, 29]. One facet of present investigations focuses on the extension and implementation of the inverse filtering techniques for the latter two classes of compounds.

Present investigations are also focused on the development of robust feedback control designs which employ the inverse filters to linearize transducer dynamics. Initial investigations focused on the numerical implementation of these energy-based inverse filters for magnetic transducers will be reported in [17] and the experimental implementation of feedback designs exploiting the filters are under present investigation.

Acknowledgements

References

Note: Center for Research in Scientific Computation Technical Reports can be accessed at the web site `http://www.ncsu.edu/crsc/reports.html`.

[1] J.C. Anderson, *Dielectrics*, Reinhold Publishing Corporation, New York, 1964.

[2] T. Braun and R.C. Smith, "Implementation techniques for a ferromagnetic hysteresis model," Smart Structures and Materials 2005, Proceedings of the SPIE, to appear.

[3] W. Chen and C.S. Lynch, "A model for simulating polarization switching and AF-F phase changes in ferroelectric ceramics," *Journal of Intelligent Material Systems and Structures*, 9, pp. 427–431, 1998.

[4] W. Chen and C.S. Lynch, "A micro-electro-mechanical model for polarization switching of ferroelectric materials," *Acta Materialia*, 46(15), pp. 5303–5311, 1998.

[5] D. Croft, G. Shed and S. Devasia, "Creep, hysteresis, and vibration compensation for piezoactuators: Atomic force microscopy application," *Journal of Dynamic Systems, Measurement, and Control*, 23, pp. 35–43, 2001.

[6] A. Daniele, S. Salapaka, M.V. Salapaka and M. Dahleh, "Piezoelectric scanners for atomic force microscopes: Design of lateral sensors, identification and control," Proceedings of the America Control Conference, San Diego, CA, pp. 253–257, 1999.

[7] P. Ge and M. Jouaneh, "Modeling hysteresis in piezoceramic actuators," *Precision Engineering*, 17, pp. 211–221, 1995.

[8] A.G. Hatch, *Model Development and Control Design for Atomic Force Microscopy*, PhD Dissertation, North Carolina State University, Raleigh, NC, 2004.

[9] L. Huang and H.F. Tiersten, "An analytic description of slow hysteresis in polarized ferroelectric ceramic actuators," *Journal of Intelligent Material Systems and Structures*, 9, pp. 417–426, 1998.

[10] C.M. Landis, "Non-linear constitutive modeling of ferroelectrics," *Current Opinion in Solid State and Materials Science*, 8, pp. 59–69, 2004.

[11] W. Lu, D.-N. Fang and K-C. Hwang, "Nonlinear electric-mechanical behavior and micromechanics modelling of ferroelectric domain evolution," *Acta Materialia*, 47(10), pp. 2913–2926, 1999.

[12] J.A. Main, E. Garcia and D.V. Newton, "Precision position control of piezoelectric actuators using charge feedback," *Journal of Guidance, Control, and Dynamics.* 18(5), pp. 1068–73, 1995.

[13] J.A. Main, D. Newton, L. Massengil and E. Garcia, "Efficient power amplifiers for piezoelectric applications," *Smart Materials and Structures*, 5(6), pp. 766–775, 1996.

[14] A.J. Moulson and J.M. Herbert, *Electroceramics: Materials, Properties, Applications*, Chapman and Hall, New York, 1990.

[15] J.M. Nealis and R.C. Smith, "$\mathcal{H}_\infty$ Control Design for a Magnetostrictive Transducer," Proc. 42nd IEEE Conf. Dec. and Control, Maui, HA, pp. 1801–1806, 2003.

[16] J.M. Nealis and R.C. Smith, "Model-Based Robust Control Design for Magnetostrictive Transducers Operating in Hysteretic and Nonlinear Regimes," CRSC Technical Report CRSC-TR03-25; *IEEE Transactions on Control Systems Technology*, submitted.

[17] W. Oates and R.C. Smith, "Robust control design for nonlinear smart systems," Smart Structures and Materials 2005, Proceedings of the SPIE, to appear.

[18] G. Robert, D. Damjanovic and N. Setter, "Preisach modeling of piezoelectric nonlinearity in ferroelectric ceramics," *Journal of Applied Physics*, 89(9), pp. 5067–5074, 2001.

[19] S. Salapaka, A. Sebastian, J.P. Cleveland and M.V. Salapaka, "High bandwidth nano-positioner: A robust control approach," *Review of Scientific Instruments*, 73(9), pp. 3232–3241, 2002.

[20] R.C. Smith, "A nonlinear optimal control method for magnetostrictive actuators," *Journal of Intelligent Material Systems and Structures*, 9(6), pp. 468–486, 1998.

[21] R.C. Smith, *Smart Material Systems: Model Development*, SIAM, Philadelphia, PA, 2005.

[22] R.C. Smith and A. Hatch, "Parameter estimation techniques for nonlinear hysteresis models," Smart Structures and Materials 2004, Proceedings of the SPIE, Volume 5383, pp. 155–163, 2004.

[23] R.C. Smith, A. Hatch, B. Mukherjee and S. Liu, "A homogenized energy model for hysteresis in ferroelectric materials: General density formulation," CRSC Technical Report CRSC-TR04-23; *Journal of Intelligent Material Systems and Structures*, to appear.

[24] R.C. Smith, A.G. Hatch, M. Salapaka, T. De and J. Raye, "Model development for atomic force microscope stage mechanisms," preprint.

[25] R.C. Smith and C.L. Hom, "Domain wall theory for ferroelectric hysteresis," *Journal of Intelligent Material Systems and Structures*, 10(3), pp. 195–213, 1999.

[26] R.C. Smith and Z. Ounaies, "A domain wall model for hysteresis in piezoelectric materials," *Journal of Intelligent Material Systems and Structures*, 11(1), pp. 62–79, 2000.

[27] R.C. Smith and M. Salapaka, "Model development for the positioning mechanisms in an atomic force microscope," *International Series of Numerical Mathematics*, Vol 143, pp. 249–269, 2002.

[28] R.C. Smith, S. Seelecke, M.J. Dapino and Z. Ounaies, "A unified model for hysteresis in ferroic materials," Smart Structures and Materials 2003, Proceedings of the SPIE, Volume 5049, pp. 88–99, 2003.

[29] R.C. Smith, S. Seelecke, M.J. Dapino and Z. Ounaies, "A unified framework for modeling hysteresis in ferroic materials," *Journal of the Mechanics and Physics of Solids,* submitted.

[30] R.C. Smith, S. Seelecke, Z. Ounaies and J. Smith, "A free energy model for hysteresis in ferroelectric materials," *Journal of Intelligent Material Systems and Structures*, 14(11), pp. 719–739, 2003.

[31] G. Tao and P. V. Kokotović, *Adaptive Control of Systems with Actuator and Sensor Nonlinearities*, John Wiley and Sons, New Jersey, 1996.

[32] J. Zhong, S. Seelecke, R.C. Smith and C. Büskens, "Optimal control of piezoceramic actuators," Smart Structures and Materials 2003, Proceedings of the SPIE, Volume 5049, pp. 264–274, 2003.