Recursive Estimation of the Preisach Density function for a Smart Actuator

Ram V. Iyer

Department of Mathematics and Statistics, Texas Tech University, Lubbock, TX 79409-1042.

ABSTRACT

The Preisach operator and its variants have been successfully used in the modeling of a physical system with hysteresis. In an application, one has to determine the density function describing the Preisach operator from measurements. In our earlier work, we described a regularization method to obtain an approximation to the density function with limited measurements. In this paper, we describe methods for recursively computing the approximate density function. These methods can be implemented in real-time controllers for smart actuators.

Keywords: Hysteresis, Preisach Operator, Recursive Estimation, Density function Identification, Magnetostriction, Piezoelectric materials, Smart materials, EAPs.

1. INTRODUCTION

The Preisach operator is a mathematical tool that has been used to model the phenomena of hysteresis occurring in ferromagnetism,[1] magnetostriction,[2] shape memory alloys[3] and piezoelectricity[4] for many years. In each phenomenon, the constitutive relation between conjugate variables, for example, magnetic field intensity H and magnetization M in ferromagnetism or electric field intensity E and polarization P in ferroelectricity, is described by a hysteresis operator, usually the Preisach operator or a variant. Part of what is needed to describe the Preisach operator for a particular phenomenon is a non-negative and integrable, density function defined on a subset of $\mathbb{R}^2$ called the Preisach domain. Several researchers starting with Mayergoyz[1] have studied the problem of identifying the density function. A synopsis of the various approaches can be found in our earlier work.[5, 6]

In our earlier work,[6] the identification problem was systematically studied and it was shown to be an ill-posed problem. This was shown using the fact that the Preisach operator with the density function as the independent variable, is a compact operator and, hence, has an unbounded inverse. The unboundedness of the inverse operator indicates a need for regularization, so that one obtains a sequence of bounded approximations. A non-parametric, SVD truncation-based regularization was proposed and numerical results were presented for magnetostrictive actuators and electro-active polymers.[6] Collocation in time and discretization of the input signal levels were used, along with truncation of singular directions corresponding to very small singular values.

It was found that finer discretizations lead to much larger computation times[6](which is to be expected). In this paper, we propose iterative methods with a view to (a) reduce computation times, (b) produce a sequence of approximate density functions that can be used in a real-time controller. The second item is more important from an implementation point of view, and slow changes in the density function can be accommodated by periodic resetting or using short recurrence sequences.

As far as we are aware, adaptive identification of the Preisach density function and its use in the inverse control of hysteresis has only been studied by Tan and Baras.[7] They proposed a recursive least squares algorithm to identify a discretized density function. They accommodated the non-negativity condition for the density function by a simple projection at each iteration. It is well-known that a simple projection into the feasible region does not minimize a convex cost function. Figure 1 shows the level surfaces of a quadratic cost function in two variables - three cases are considered, with the unconstrained minimum in the second, third and fourth quadrants. If the minimum has to satisfy a non-negativity constraint, then it is clear that in each case, the projection of the unconstrained minimum x_p into the first quadrant does not yield the desired solution x_{min}. Furthermore, the implementation of the algorithm requires the inversion of a matrix, which leads to the *persistency of excitation* condition in Tan and Baras. The fulfilment of this condition would require an enormous amount of data for fine mesh selections. Our method in this paper is to numerically obtain the best approximation of the density function

at each iteration, even in the absence of *persistency of excitation* - or in other words, when only limited data is available. This is achieved by applying a primal-dual interior point method after a partial SVD computation. As the SVD computation is independent of the interior point method, convergence follows from standard results for each of these methods.
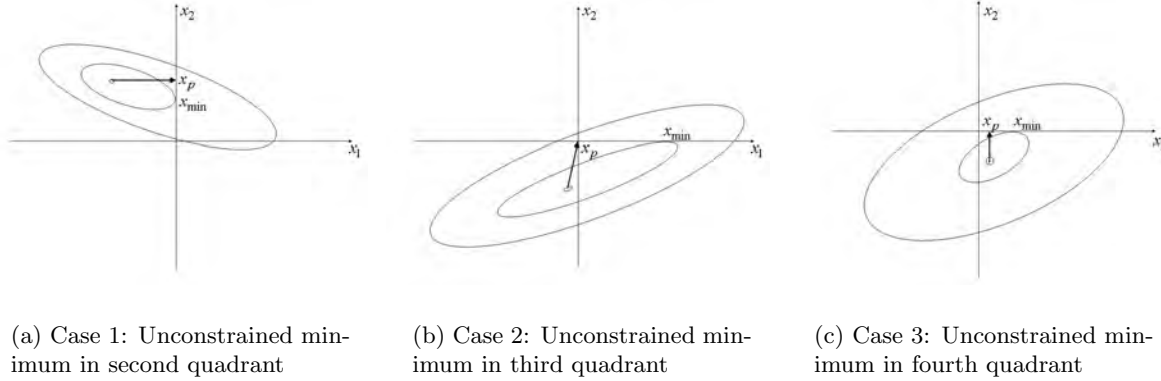

(a) Case 1: Unconstrained minimum in second quadrant

(b) Case 2: Unconstrained minimum in third quadrant

(c) Case 3: Unconstrained minimum in fourth quadrant

Figure 1. Examples of simple projection after unconstrained minimization.

2. PRELIMINARIES

In this section, we briefly recall the Preisach operator and the need for regularization. A detailed treatment on the Preisach operator can be found in.[1, 8, 9] For a pair of thresholds (β, α) with $\beta \leq \alpha$, consider a relay $\mathcal{R}_{\beta,\alpha}[\cdot, \cdot]$ (called a Preisach *hysteron*), as illustrated in Fig. 2. For $u \in C[0, T]$

$$v_{\beta,\alpha}(t) \triangleq \begin{cases} -1 & \text{if } u(t) < \beta \\ 1 & \text{if } u(t) > \alpha \\ v_{\beta,\alpha}(t^-) & \text{if } \beta \leq u(t) \leq \alpha \end{cases},$$

where $v(0^-) = \zeta$ and $t^- \triangleq \lim_{\epsilon > 0, \epsilon \to 0} t - \epsilon$.
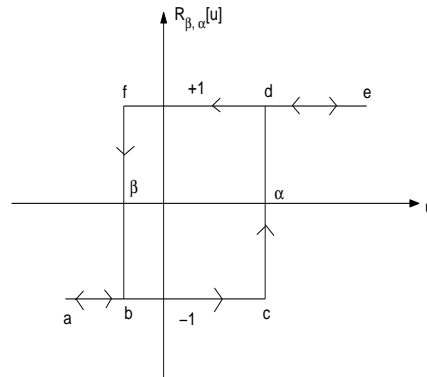

Figure 2. Input-output relationship for the relay.

To construct the Preisach operator from the elementary hysterons when the input is $u(\cdot)$, denote $v_{\beta,\alpha}(\cdot) = R_{\beta,\alpha}[u](\cdot)$. The Preisach operator's input is $u(\cdot)$, and the output is given by[1]:

$$\bar{y}(t) = \iint\limits_{\alpha \geq \beta} \mu(\beta, \alpha) R_{\beta,\alpha}[u](t)\, d\beta d\alpha \tag{1}$$

where $\mu(\cdot, \cdot) \in L^2(K)$ where K is a compact region in the (β, α) plane with $\alpha \geq \beta$ and support$(\mu) = K$. Thus for a given μ, we can define a Preisach operator to be a map $\Gamma_\mu : C_{pm}[0,T] \rightarrow C_{pm}[0,T]$, where $C_{pm}[0,T]$ denotes the space of piecewise monotone continuous functions on $[0,T]$. For compatibility with experimental evidence, we restrict $\mu(\cdot, \cdot)$ to be a non-negative function. During the identification experiments, the fixed input $u \in C_{pm}[0,T]$ usually only affects a portion of the set K in the Preisach plane. Without loss of generality, we can restrict attention to this portion. *In the following, the set K_u is the subset of K affected by the input u.* Define the set of density functions:

$$\mathcal{K}_u \coloneqq \{\mu \in L^2(K_u)\,|\,\mu \geq 0\}. \tag{2}$$

Due to the assumption on K_u, the output at time 0 given by $\Gamma[u](0)$ is the same for all density functions in $\mathcal{K}_u$.

An important class of Preisach operators from the applications point of view, are the *piecewise, strictly increasing* (PSI) Preisach operators.

DEFINITION 2.1. *A Preisach operator is said to be piecewise strictly increasing (PSI) if $(\Gamma_\mu[u](T) - \Gamma_\mu[u](0))(u(T) - u(0)) > 0$ for a monotone input $u \in C[0,T]$ with $u(0) \neq u(T)$. .*

Under the mild condition that the density function is continuous, and greater than zero along the diagonal $\alpha = \beta$, it is easy to show that the Preisach operator is PSI. Clearly, the definition of PSI is for a fixed density function and varying monotone increasing input functions u. However, in the identification problem the input function is fixed while the density function is the independent variable. For a non-constant $u \in C_{pm}[0,T]$, let $0 = T_0 < T_1 < \cdots < T_N = T$ be the standard partition (see Brokate and Sprekels[9]). As u is non-constant, $u(T_{i-1}) \neq u(T_i)$ for $i = 1 \cdots N$. Denote $\Delta_i = [T_{i-1}, T_i]$ for $i = 1 \cdots N$. Define the set of density functions:

$$
\begin{aligned}
\mathcal{K}_{u,PSI} \;\; \coloneqq \;\; & \{0\} \cup \big\{\mu \in L^2(K)\,|\,\mu \geq 0; \text{ and} \\
& (\Gamma[u](t_1) - \Gamma[u](t_2))(u(t_1) - u(t_2)) > 0 \\
& \text{for } t_1, t_2 \in \Delta_i, \;\; t_1 \neq t_2; \;\; i = 1, \cdots, N.\big\}
\end{aligned}
\tag{3}
$$

The idea is that the density function for a PSI Preisach operator (that can be identified) should come from this set. The sets $\mathcal{K}_u$ and $\mathcal{K}_{u,PSI}$ are clearly non-empty with $\mathcal{K}_{u,PSI} \subset \mathcal{K}_u$. To further study their properties, recall that a convex set C is a subset in a vector space, such that if $x_1, x_2 \in C$, then $\theta x_1 + (1-\theta) x_2 \in C$ for $0 \leq \theta \leq 1$ (Luenberger[10]). A *cone with vertex at the origin* is a set C in a vector space with the property that if $x \in C$, then $\theta x \in C$ for all $\theta \geq 0$. A convex cone is defined as a set that is both convex and a cone. Closed convex sets are important from the point of view of numerical methods for constrained optimization problems. The following lemma was proved in Iyer and Shirley.[6]

LEMMA 2.2. *The set $\mathcal{K}_u$ is a closed convex cone with vertex at the origin, while the set $\mathcal{K}_{u,PSI}$ is a convex cone with vertex at the origin that is not closed.*

Now, for all μ in $\mathcal{K}_u$, the initial output $w = \Gamma[u](0)$ is fixed. For a given $u(\cdot) \in C_{pm}[0,T]$, define the operator:

$$
\Phi_u : \quad
\begin{array}{rcl}
\mathcal{K}_u & \rightarrow & L^2[0,T] \\
\mu(\beta, \alpha) & \mapsto & y = \Phi_u\,\mu(\cdot) = \Gamma_\mu[u](\cdot) - w
\end{array}
\tag{4}
$$

The operator Φ_u is a *linear* operator between $L^2(K_u)$ and $L^2[0,T]$ which is not true for the Preisach operator Γ! The key is that Φ_u maps the function $\mu = 0$ to the function $y = 0$. The following important theorem concerning the identification problem for a PSI Preisach operator was proved in Iyer and Shirley.[6]

THEOREM 2.3. *Let u be a non-constant function in $C_{pm}[0,T]$ and $\Phi_u : \mathcal{K}_u \rightarrow L^2[0,T]$ be as defined in (4). Then,*

1. $\{\mu \in \mathcal{K}_u\,|\,\Phi_u\,\mu(t) = 0\} = \{0\}$;

2. $\Phi_u : \mathcal{K}_u \to L^2[0,T]$ *is injective;*

3. $\overline{Range[\Phi_u]} \neq L^2[0,T]$;

4. *Suppose that $y \in Range[\Phi_u]$ and Γ is PSI. Then there exists a unique $\mu \in \mathcal{K}_{u,PSI}$ such that $\Phi_u \, \mu = y$.*

Remarks

We conclude from the above theorem and Lemma 2.2 that:

- Even though $\mu \in \mathcal{K}_{u,PSI}$ for a PSI Preisach operator, we have to carry out the identification over the set $\mathcal{K}_u$, as $\mathcal{K}_{u,PSI}$ is not a closed convex set.

- The set $\mathcal{K}_u$ is a closed convex cone, and as Φ_u is a linear injective operator on this set, it is "well-suited" from a numerical point of view.

- However, as the closure of the range of Φ_u is not all of $L^2[0,T]$, we need to construct a regularization scheme to counter-act noise.

□

The following lemma proved in Iyer and Shirley[6] shows that the density identification problem is an ill-posed one!

LEMMA 2.4. *The operator $\Phi_u : L^2(K_u) \to L^2[0,T]$ is a compact linear operator.*

Due to this lemma, a regularization scheme based on (i) the SVD of a finite dimensional approximation of Φ_u, and (ii) truncation of singular directions corresponding to small singular values was proposed in Iyer and Shirley,[6] and is described below.

Let $0 = t_1 < \cdots < t_n = T$ be a discretization of time, so that we have :

$$\left. \begin{array}{rcl} (\Phi_u \, \mu)(t_j) & = & y(t_j), \\ \int_K R_{\beta,\alpha}[u](t_j)\, \mu(\beta,\alpha)\, d\beta d\alpha & = & y(t_j), \end{array} \right\} \quad j = 1, \cdots, n. \tag{5}$$

where $\mu(\beta,\alpha)$ is unknown, while $u(\cdot)$ and $y(t_j)$ are known. The time instants $t_1, \cdots, t_n$ are known as *collocation* points. Once a set of collocation points have been determined, one can discretize the input values so that a uniform grid is established in the region K_u in the Preisach plane. Then the equation $\Phi_u \, \mu = y$ can be written as a linear equation

linear equation

$$Y = AX + \epsilon, \tag{6}$$

as described in Shirley and Venkataraman.[5] Each element of X (except the last) denotes the area under the density function for a particular grid element. The last element denotes the initial output value (at time 0). To account for noise, we estimate this value also.

We need is a way to solve for X that will best fit the data, but at the same time keep $x_i \geq 0$, since X represents the integral of a density function over a grid element.

We would like to minimize the function

$$f(X) = \frac{1}{2} \|AX - Y\|^2 \tag{7}$$

where

$$A : \mathbb{R}^n \to \mathbb{R}^m, \quad X \in \mathbb{R}^n, \quad Y \in \mathbb{R}^m,$$

with the inequality constraint $g(X) = X \geq 0$. Let $rank(A) = q \leq \min\{m, n\}$. If we perform a singular value decomposition on $A^T A$, then we get:

$$A^T A = V S V^T, \tag{8}$$

where S is an $n \times n$ diagonal matrix with rank $q < n$, and $V^T V = I_{n \times n}$ (see page 54, Bellman[11]). The n singular values of $A^T A$ are the diagonal elements of S and be ordered as $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n \geq 0$. By the remarks following Picard's theorem,[12] we see that "small" singular values should be discarded as they contribute to an amplification of the noise in the solution. We perform the following steps:

1. Pick a tolerance value $\epsilon > 0$, and set those singular values that are less than ϵ to 0. Call the resulting matrix that is obtained from S as S_1.

2. Remove the rows and columns of S_1 that are identically zero, and also remove the corresponding columns of U and V. By this procedure, we obtain a $\bar{q} \times \bar{q}$ diagonal matrix $\hat{S}$, and a $n \times \bar{q}$ matrices $\hat{V}$ that satisfies

$$\hat{V}^T \hat{V} = I_{\bar{q} \times \bar{q}}. \tag{9}$$

The operator norm of the matrix $A^T A - \hat{V} \hat{S} \hat{V}^T$ can be seen to be:

$$
\begin{aligned}
\|A^T A - \hat{V} \hat{S} \hat{V}^T\| &= \|A^T A - V S_1 V^T\| \\
&= \|V (S - S_1) V^T\| \\
&= \|S - S_1\| \\
&< \epsilon.
\end{aligned}
$$

We seek solutions of the type $X = \hat{V} Z$ for the Problem (7). Using Equations (8) and (9), the cost function can be transformed to $Z^T \hat{S} Z - Y^T A \hat{V} Z$. Thus we form the constrained optimization problem:

$$\text{minimize} \quad f(Z) = \tfrac{1}{2} Z^T \hat{S} Z - Y^T A \hat{V} Z, \tag{10}$$

$$\text{subject to} \quad g(Z) = \hat{V} Z \geq 0. \tag{11}$$

Once we have a minimizer Z^* to the above problem, then the desired solution is given $X^* = \hat{V} Z^*$.

For a commercial magnetostrictive actuator made by ETREMA Inc., the relation between the average Magnetic field Intensity H and average Magnetization M along the axis of the rod was modeled by a classical Preisach operator in Shirley and Venkataraman[5] and Iyer and Shirley.[6] Figure 3 reproduces the density functions obtained by the above mentioned procedure for different discretizations (20 Oe, 10 Oe, 6.25Oe) of the input magnetic field. Figure 4 shows a plot of the singular values of the A matrix for each discretization level. The discretization level of 20 Oe leads to a matrix A of full rank, while for the other two discretization levels, the matrix A is rank-deficient. Figure 5 compares the density functions obtained for the discretization level of 10 Oe, using two different truncation levels for the singular values: $\sigma \leq 10^{-3}$ and $\sigma \leq 2$. Unfortunately, using a truncation level of $\sigma \leq 4$ (corresponding to 207 singular values) did not yield a good result, which implies that the problem is only mildly ill-posed.

3. ITERATIVE COMPUTATION OF THE PREISACH DENSITY FUNCTION

In this paper, we consider a iterative version of the procedure in Section 2. The problem now is to minimize the function (at time t_p)

$$f_p(X) = \frac{1}{2} \|A_p X - Y_p\|^2 \tag{12}$$

where

$$A_p : \mathbb{R}^n \to \mathbb{R}^{m_p}, \quad X \in \mathbb{R}^n, \quad Y_p \in \mathbb{R}^{m_p},$$

with the inequality constraint $g(X) = X \geq 0$. In the problem (12), the row dimension of A_p and Y_p increases by one after sampling of the input and output signals is performed. We assume that $rank(A_p) \leq \min \{m_p, n\}$.

The solution procedure outlined in Section 2 is composed of two parts: (a) singular value decomposition of the matrix $A^T A$ and truncation of singular directions corresponding to small singular values; (b) solution of the constrained minimization problem (11). To construct an iterative scheme, we need iterative methods to accomplish both parts.
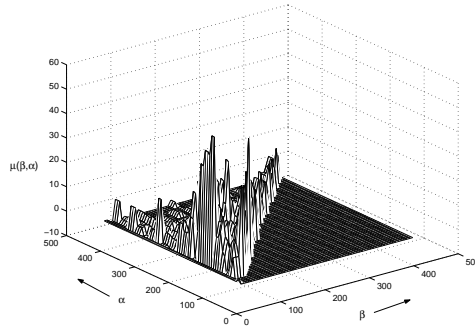
An iterative method for SVD is the double-bracket flow due to Brockett.[13] This method leads to the computation of *all* the singular values of the matrix A_p, which is not necessary due to truncation. In light of
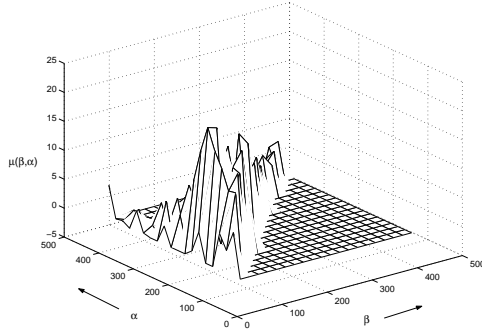
(a) Discretization = 20 Oe
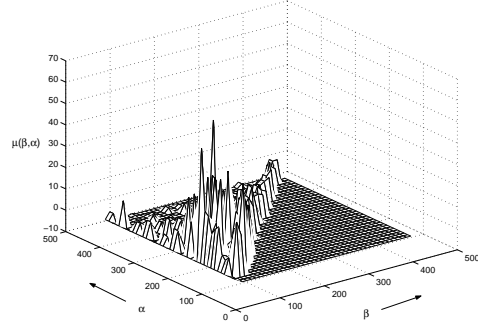
(b) Discretization = 10 Oe

(c) Discretization = 6.25 Oe

Figure 3. Identification experiments for a commercial magnetostrictive actuator with different magnetic field discretizations

the rapid decay in the singular values (both for the full-rank and rank-deficient cases) as seen in Figure 4, we only need to carry out a partial SVD by computing the significant singular values. This can be accomplished in several ways - one of them being the Lanczos iteration method, which seems to require fewer computations for the same level of accuracy than the subspace-iteration method.[14] Other possibilities for sequential SVD can be found in Strobach.[15] Note that ideally, one would like to compute only one iteration of the partial SVD algorithm per time-step.

Once the significant singular values and directions have been computed (resulting in the matrices $\hat{S}_p$ and $\hat{V}_p$), the next task is to minimize the cost function $f(Z)$ in (11). One way to accomplish this is to use a line-search Newton-Raphson algorithm due to Krishnaprasad and Barakat.[16] Another method is to use a Primal-Dual Interior point method.[17,18] Ideally, one would like to apply only one iteration of (either) solution method at time t_p. At time t_p, the Kuhn-Tucker theorem yields the existence of a $\lambda \in \mathbb{R}^n$, such that the pair (Z_p^*, λ_p) satisfy the necessary conditions (see Luenberger[10]):

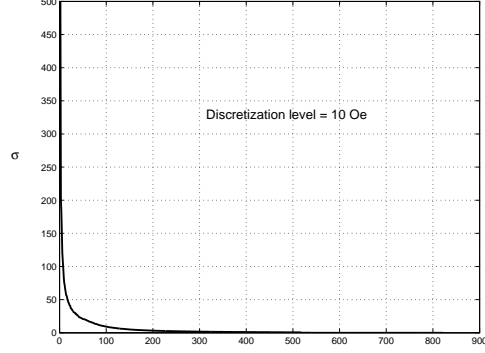$$\hat{S}_p Z_p^* + \hat{V}_p^T \lambda - \hat{V}_p^T A_p^T Y_p = 0 \qquad (13)$$

$$\lambda_p^T \hat{V}_p Z_p^* = 0 \qquad (14)$$
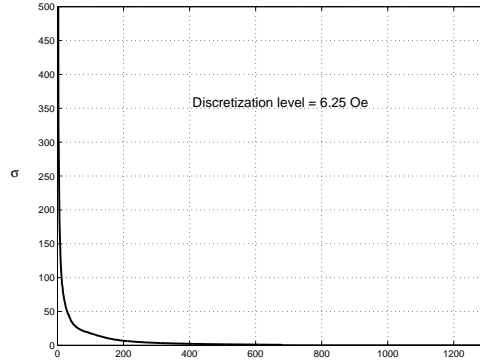
$$\lambda_p \geq 0 \qquad (15)$$

The solution to this set is found using Newton's method. Let $(Z_p^{(i)}, \lambda_p^{(i)})$ be the values at the i-th iteration of
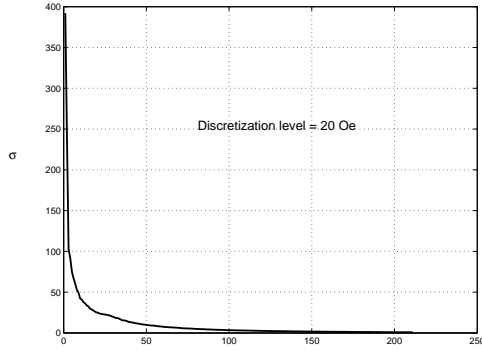
(a) Discretization= 20 Oe

(b) Discretization = 10 Oe

(c) Discretization = 6.25 Oe

Figure 4. Plot of singular values of the A matrix

the Newton's method. Then, the values at the $i + 1$-th step are found using the update direction $(\delta Z, \delta \lambda)_p^{(i)}$) given by the solution to the equation:

$$\begin{bmatrix} \hat{S}_p & \hat{V}_p^T \\ (\lambda_p^{(i)})^T \hat{V}_p & (Z_p^{(i)})^T \hat{V}_p^T \end{bmatrix} = - \begin{bmatrix} \hat{S}_p Z_p^{(i)} + \hat{V}_p^T \lambda_p^{(i)} - \hat{V}_p^T A_p^T Y_p \\ (\lambda_p^{(i)})^T \hat{V}_p Z_p^{(i)} \end{bmatrix}. \tag{16}$$

The update is:

$$(Z, \lambda)_p^{(i+1)} = (Z, \lambda)_p^{(i)} + \alpha_p^{(i)}) (\delta Z, \delta \lambda)_p^{(i)}, \tag{17}$$

where $\alpha_p^{(i)}$ is chosen to ensure that $\lambda_p^{(i+1)} \geq 0$. Obviously, it has to be chosen to be:

$$\alpha_p^{(i)} = \arg \min_{\nu} \lambda_p^{(i)}(k) + \nu \, \delta \, \lambda_p^{(i)}(k) \geq 0 \quad k = 1 \cdots n.$$

In the above equation, $\lambda_p^{(i)}(k)$ denotes the k-th component of $\lambda_p^{(i)}$.

Convergence analysis:

We outline the convergence analysis for our method. Note that the SVD computation is independent of the interior point method, which simplifies the convergence analysis. Furthermore, the rank of the A_p matrix can only increase with p, and that too only in finite steps. As rank(A_p) is monotone increasing and bounded
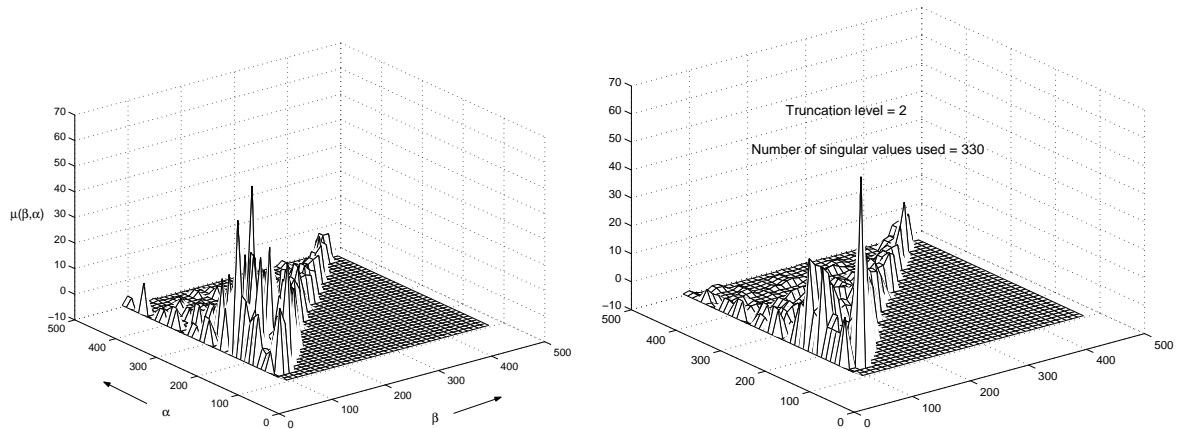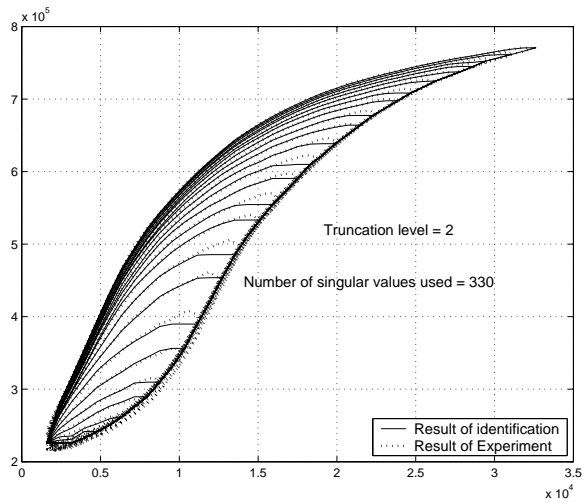
(a) Density function using 515 singular values

(b) Density function using 330 singular values
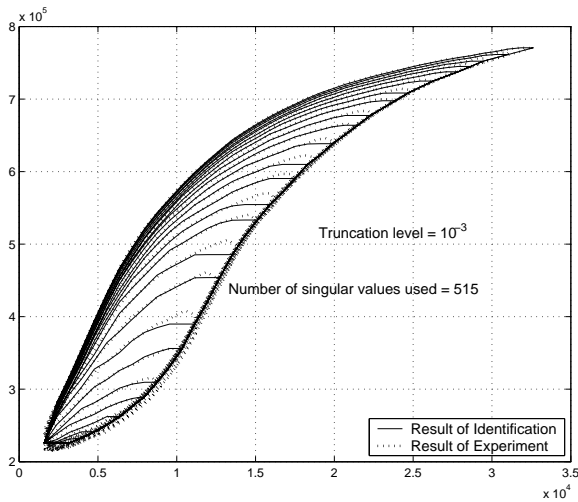
(c) H versus M, using 515 singular values

(d) H versus M, using 330 singular values

Figure 5. Comparison of different SVD truncation schemes for input discretization of 10 Oe.

above by n, the partial SVD computation is convergent by Theorem 4.1 of Vogel and Wade[14] assuming that $A_p^T A_p$ converges to a matrix Σ as $p \to \infty$. That the interior-point method is convergent, is a standard result.[18] Numerical results will be addressed in a future publication.

4. CONCLUSION

In this paper, we examined the problem of recursively determining the Preisach density function given input-output data, in a parameter-free manner. This problem has two features:(a) ill-posedness and (b) a non-negativity constraint on the solution. We address both issues in our recursive implementation. Ill-posedness of the operator equation is addressed by a partial SVD iteration, and the non-negativity constraint is addressed by a Primal-Dual interior point method. As the SVD iteration is independent of the Interior point method computation, convergence of our proposed method follows from standard results for each of these methods.

REFERENCES

1. I. Mayergoyz, *Mathematical Models of Hysteresis*. Springer, 1991.

2. R. Venkataraman, X. Tan, and P. Krishnaprasad, "Control of hysteresis: Theory and experimental results," in *Smart Structures and Materials 2001: Modeling, Signal Processing, and Control in Smart Structures* (V. Rao, ed.), vol. 4326, Mar. 2001.

3. R. Gorbet, D. Wang, and K. Morris, "Preisach model identification of a two-wire SMA actuator," in *Proceedings of IEEE International Conference on Robotics and Automation*, pp. 2161–2167, 1998.

4. W. Galinaitis and R. C. Rogers, "Control of a hysteretic actuator using inverse hysteresis compensation," in *Mathematics and Control in Smart Structures, Smart Structures and Materials 1998* (V. V. Varadhan, ed.), vol. 3323, pp. 267–277, Mar. 1998.

5. M. Shirley and R. Venkataraman, "On the identification of Preisach measures," in *Modeling, Signal Processing and Control in Smart Structures and Materials* (R. Smith, ed.), vol. 5049, pp. 326–336, July 2003.

6. R. V. Iyer and M. E. Shirley, "Hysteresis identification with limited experimental data." submitted to IEEE Trans. Magnetics, preprint available at http://www.math.ttu.edu/ rvenkata, 2003.

7. X. Tan, "Adaptive identification and control of hysteresis in smart materials." preprint, 2003.

8. A. Visintin, *Differential Models of Hysteresis*. Applied Mathematical Sciences, Springer, 1994.

9. M. Brokate and J. Sprekels, *Hysteresis and Phase Transitions*. Applied Mathematical Sciences, Springer Verlag, 1996.

10. D. G. Luenberger, *Optimization by Vector Space Methods*. Wiley, 1968.

11. R. Bellman, *Introduction to Matrix Analysis*. SIAM, Philadelphia, PA 19104., 1995.

12. A. Kirsch, *A Mathematical Introduction to the Theory of Inverse Problems*. Springer Verlag, 1996.

13. U. Helmke and J. B. Moore, *Optimization and Dynamical Systems*. Springer, New York., 1996.

14. C. R. Vogel and J. G. Wade, "Iterative SVD-based methods for ill-posed problems," *SIAM J. SCI. COMPUT.*, vol. 15, no. 3, pp. 736–754, 1994.

15. P. Strobach, "Bi-iteration SVD subspace tracking algorithms," *IEEE Transactions on signal processing*, vol. 45, no. 5, pp. 1222–1240, 1997.

16. P. S. Krishnaprasad and R. Barakat, "A descent approach to a class of inverse problems," *J. Comp. Phys*, vol. 24, pp. 339–347, 1977.

17. S. Wright, *Primal-Dual Interior Point Methods*. SIAM, Philadelphia, PA, 1997.

18. F. A. Potra and S. J. Wright, "Interior-point methods," *Journal of Computational and Applied Mathematics*, vol. 124, pp. 281–302, 2000.