Available online at www.sciencedirect.com

ScienceDirect

IFAC PapersOnLine 53-1 (2020) 602–607

Parameter Identification in Nonlinear Systems Using PD Controllers as Penalty Functions

R. Manikantan * **C. P. Vyasarayani** ** **M. Manjuprasad** *

* *National Aerospace Laboratories (NAL), Wind Tunnel Road, Kodihalli, Bengaluru, Karnataka, India, 560037.*
** *Department of Mechanical and Aerospace Engineering, Indian Institute of Technology Hyderabad, Kandi, Sangareddy, Telangana, India, 502285 (e-mail: vcprakash@iith.ac.in).*

Abstract: The identification of parameters in nonlinear systems using a partial set of experimental measurements is considered in this paper. The estimation of these parameters introduces an optimization problem. For parameter estimation, the use of gradient-based optimizers often converges to a local minimum rather than the global optimum. To overcome the local convergence of the parameters, a PD controller algorithm is implemented for estimation. The addition of a morphing parameter with a proportional-derivative controller (PD) to the system equation transforms the objective function into convex, and the optimization is performed using a gradient-based optimizer. To illustrate the nonlinear parameter estimation using the present approach, a numerical example of Van der Pol-Duffing oscillator is presented. A comparative analysis is then carried out with global optimization methods, such as genetic algorithm (GA) and particle swarm optimization (PSO) techniques. The numerical results confirm that the PD controller algorithm is superior in terms of computational effort and convergence efficiency.

Keywords: Nonlinear systems, Van der Pol-Duffing oscillator, Parameter estimation, PD controller.

1. INTRODUCTION

System simulation requires a validated mathematical model of the physical system, whose parameters is to be estimated from the experimental measurements. The identification of parameters is a difficult task, especially when nonlinearities are involved in the mathematical model. Further, it is assumed that only a partial set of experimental measurements are available. The problem of identifying the parameters in a mathematical model given a partial set of experimental data is a matter of interest to many in the field of biological and physical sciences (Gershenfeld (1999)). The parameter estimation is generally considered as an optimization problem, wherein the difference between experimental and model responses of the physical system is minimized (Frewin (2013)).

The method of least squares is one of the fundamental techniques used in the parameter estimation (Ljung and Söderström (1983)) process. Classical optimization techniques, such as Gauss–Newton (Rao (2009)), steepest descent, and Levenberg–Marquardt are gradient-based algorithms that can be employed for estimation, provided the objective function is convex. In other words, these methods mostly identify the global optimum, only when unimodal objective functions are considered. The main limitation of the gradient-based optimization methods is that they often converge to a local minimum instead of global minimum unless the chosen initial guess is near to global minimum. In many mathematical models, the parameters appear nonlinearly in the system equation, and therefore, the objective function becomes multimodal. The stochastic optimization techniques such as genetic algorithm (GA), simulated annealing (SA), and particle swarm optimization (PSO) (Goldberg (2002); Romeijn and Smith (1994)) can be used to obtain the global optimum. However, if a large number of parameters are to be identified in the system, then these methods become computationally expensive and time-consuming. The lack of gradient-based algorithms for the estimation of parameters in the multimodal objective function is a bottleneck in the parameter identification process.

The homotopy methods are one such candidate for obtaining the global optimum (Watson and Haftka (1988)). Vyasarayani et al. (2012b) introduced the homotopy optimization for parameter identification in nonlinear systems. This method was successfully implemented for the estimation of the parameter in multibody systems, wherein, the system dynamics is governed by the differential-algebraic equations (DAEs) (Vyasarayani et al. (2012a)). The current work is a direct extension of the homotopy works of Vyasarayani et al. (2012a). In the homotopy method, a high gain observer and a morphing parameter are introduced into the objective function implicitly. The original multimodal objective function is then transformed into a simple convex function. This convex function is slowly morphed back to the original objective function by re-

10.1016/j.ifacol.2020.06.101

ducing the value of morphing parameter in a series of iterative steps. The optimization is carried out for each value of the morphing parameter using a gradient-based optimizer, considering the previous value of the minimizer as the initial guess for the current steps. This process is continued until the morphing parameter becomes zero, and in the process, the original objective function is regained, and in consequence, global parameters are obtained. To demonstrate the efficacy of the present method, a comparative analysis is performed with the stochastic optimizers, such as GA and PSO.

In the present study, a numerical example of Van der Pol–Duffing (VDPD) oscillator (Quaranta et al. (2010)) is presented to illustrate the parameter estimation process in nonlinear systems. This paper is organized as follows: The introduction about PD controller algorithm is discussed in Sec. 2. In Sec. 3, the parameter identification of VDPD oscillators is presented, followed by comparison with stochastic optimizers is discussed. Finally, the conclusion is given in Sec. 4.

2. PARAMETER ESTIMATION USING PD CONTROLLER ALGORITHM

The differential equation governing the dynamics of the experimental system is assumed to be of the form:

$$\begin{aligned}\dot{\mathbf{x}}_{1e}(t) &= \mathbf{x}_{2e}(t)\\ \dot{\mathbf{x}}_{2e}(t) &= \mathbf{f}(\mathbf{x}_{1e}(t),\ \mathbf{x}_{2e}(t),\ \mathbf{p},\ t)\end{aligned} \tag{1}$$

where, $\mathbf{x}_{1e}(t) = [x_{1e}(t),\ x_{2e}(t), ...,\ x_{ne}(t)]^{\mathrm{T}}$, contains the time series of the n states, and $\mathbf{p} = [p_1,\ p_2, ...,\ p_k]^{\mathrm{T}}$ are the parameters in the experimental system. It is assumed that only one state x_{1e} is measured over a time period $t \in [0,\ t_f]$, and other states in the experimental system are unobserved. The mathematical model representing the experimental system is written in a first-order state-space form as follows:

$$\begin{aligned}\dot{\mathbf{x}}_1(t) &= \mathbf{x_2}(t)\\ \dot{\mathbf{x}}_2(t) &= \mathbf{f}(\mathbf{x}_1(t),\ \mathbf{x}_2(t),\ \hat{\mathbf{p}},\ t)\end{aligned} \tag{2}$$

where, $\mathbf{x}(t) = [x_1(\hat{\mathbf{p}}, t), x_2(\hat{\mathbf{p}}, t), ..., x_m(\hat{\mathbf{p}}, t)]^{\mathrm{T}}$. Our objective is to identify the parameters $\hat{\mathbf{p}} = [\hat{p}_1,\ \hat{p}_2, ...,\ \hat{p}_k]^{\mathrm{T}}$ in the mathematical model (Eq. (2)), such that the error obtained between the predicted response $x_1(t)$ and the experimental measurement $x_{1e}(t)$ is minimized. To identify the parameters, the following objective function is minimized:

$$J(\hat{\mathbf{p}}) = \frac{1}{2}\int_0^{t_f} (x_{1e}(t) - x_1(\hat{\mathbf{p}},\ t))^2 dt \tag{3}$$

The above objective function may contain many local minima, and the minimization may lead to one of them. To avoid the premature convergence to the local minimum, a high gain proportional-derivative (PD) controller with a morphing parameter (λ) is introduced to the mathematical model (Eq. (2)) explicitly as follows:

$$\dot{\mathbf{x}}_1(t) = \mathbf{x_2}(t) + \lambda\mathbf{k}_d(x_{1e}(t) - x_1(t)) \tag{4a}$$

$$\dot{\mathbf{x}}_2(t) = \mathbf{f}(\mathbf{x}_1(t),\ \mathbf{x}_2(t),\ \hat{\mathbf{p}},\ t) + \lambda\mathbf{k}_p(x_{1e}(t) - x_1(t)) \tag{4b}$$

where, $x_{1e}(t) - x_1(t)$ in Eq. (4a) and (4b) represents the derivative and proportional controller respectively. The corresponding gain values are $\mathbf{k}_d$ and $\mathbf{k}_p$. The following transformed objective function is used for estimation:

$$J(\hat{\mathbf{p}},\ \lambda) = \frac{1}{2}\int_0^{t_f} (x_{1e}(t) - x_1(\hat{\mathbf{p}},\ \lambda,\ t))^2 dt \tag{5}$$

The optimization process is initiated by providing an arbitrary initial guess for parameters, and λ is set to 1. As $\lambda = 1$, the objective function becomes convex. The minimization of the objective function $J(\hat{\mathbf{p}}, 1)$ is carried out using Eq. (4) as the mathematical model. Any gradient-based optimizer can be used and the global minimum ${}^1\hat{\mathbf{p}}^*$ is obtained. Subsequently, we reduce λ by a small amount $\delta\lambda$, and minimize $J(\hat{\mathbf{p}}, 1 - \delta\lambda)$ using ${}^1\hat{\mathbf{p}}^*$ as the initial guess. Once we determine the optimum ${}^{1-\delta\lambda}\hat{\mathbf{p}}^*$, the value of λ is reduced further to $1 - 2\delta\lambda$, and proceed the next iteration using ${}^{1-\delta\lambda}\hat{\mathbf{p}}^*$ as initial guess. This procedure is continued until λ becomes zero; by then, the PD controller vanishes from the mathematical model. In the last iteration where $\lambda = 0$, global optimum $\hat{\mathbf{p}}^* = \mathbf{p}$ is achieved.

This method has a very close resemblance to the penalty function method. In the penalty function approach, the constraints are added to the objective function with a penalty parameter r, and it is transformed into an unconstrained optimization problem. The optimization is performed by choosing an appropriate value for r, such that the minimizer satisfies all the constraints. If all the constraints are not satisfied, the objective function is penalized by a factor of $10\,r$, and the next iteration is initiated using the previous value of the minimizer as the initial guess for the current step. This process is continued until the minimizer satisfies all the constraints. The analogy of penalty parameter r to the homotopy parameter λ and penalty function to the PD controller can be easily make out while comparing these two methods. One major drawback of the penalty function method is that it only achieves near-optimal solutions, whereas the present method attains the global optimum exactly. A numerical example of Van der Pol-Duffing (VDPD) oscillator is presented in the next section to demonstrate the parameter estimation using the PD controller algorithm.

3. VAN DER POL-DUFFING OSCILLATORS

Van der Pol-Duffing (VDPD) oscillators are one of the classical nonlinear systems embedded with rich dynamical properties, and it appears in various fields of science and engineering (Zhao and Yang (2015)). The classical duffing oscillator augmented with linear damping is characterized in these systems.

The experimental system governing the externally exited VDPD oscillator (Quaranta et al. (2010)) in first-order state-space form is given as follows:

$$\begin{aligned}\dot{x}_{1e}(t) &= x_{2e}(t)\\ \dot{x}_{2e}(t) &= \mu(1 - x_{1e}^2(t))x_{2e}(t) - \alpha x_{1e}(t) - \beta x_{1e}^3(t) + f(t)\end{aligned} \tag{6}$$

where, $f(t) = g\cos(wt)$ is the input excitation, and $g = 0.50$, $\omega = 0.79$ represent the amplitude and frequency of the exited force respectively, and t is the time variable. The behavior of the VDPD oscillator is attributed to three important parameters: α, β and μ. The parameter α controls the stiffness, β represents the amount of nonlinearity in the system, and μ is the damping coefficient ($\mu > 0$).

Based on the value of these parameters, VDPD oscillators are classified into three main physical states:

- single well ($\alpha > 0$, $\beta > 0$);
- double well ($\alpha < 0$, $\beta > 0$);
- double-hump ($\alpha > 0$, $\beta < 0$).

The system parameters must be estimated from the experimental measurements, as discussed in the next section.

3.1 Parameter estimation in VDPD oscillators

The problem of parameter estimation is formulated as a multi-dimensional optimization problem, where the arguments corresponding to the minimum value of the objective function represent the optimal parameters. In the estimation of VDPD system parameters, it is assumed that both initial conditions and external excitation $f(t)$ are known a priori. With this assumption, we minimize the objective function $J(\hat{\mathbf{p}})$ over a constrained real bounded domain from where the parameters are to be identified.

$$\min(J(\hat{\mathbf{p}})) \quad s.t. \quad \mathbf{p}^l \leq \hat{\mathbf{p}} \leq \mathbf{p}^u$$

where, $\hat{\mathbf{p}} = [\hat{\alpha},\, \hat{\beta},\, \hat{\mu}]$ represent the parameter vector within the search space of lower $\mathbf{p}^l$ and upper $\mathbf{p}^u$ bounds. Since the stability of VDPD oscillator lies only in a certain range of parameter values, the estimation process is formulated as a constrained optimization problem. The experimental parameters $\mathbf{p} = [\alpha,\, \beta,\, \mu]$ for identification in VDPD oscillators are given in Table 1.

We begin the identification process in a single well VDPD oscillator, where the system parameters are $[\alpha,\, \beta,\, \mu] = [0.5,\, 0.5,\, 0.1]$. The initial conditions for the experimental system and the mathematical model are assumed to be $[x_{1e}(0),\, x_{2e}(0)] = [x_1(0),\, x_2(0)] = [1,\, 0]$, and generate a 20 seconds experimental data using Eq. (6) as shown in Fig. 1(A)(a). Note that only the state x_{1e} is measured and another state x_{2e} is unobserved. To identify the global parameters, we minimize the following objective function:

$$J\left(\hat{\mathbf{p}}\right) = \frac{1}{2}\int_0^{t_f} \left(x_{1e}(t) - x_1(\hat{\mathbf{p}},\, t)\right)^2 dt \qquad (7)$$

We now visualize the shape of the objective function $J\left(\hat{\mathbf{p}}\right)$ in the parameter domain. A graph depicting the variation of $J(\beta)$ for constant values of α is generated in Fig. 1(A)(b), by keeping $\mu = 0.1$ constant. Fig. 1(A)(b) shows the presence of peaks and valleys, and it is evident that the existence of multiple local minima in the objective function. A basin of attraction is presented in Fig. 1(A)(c) to distinguish between the local and global minima in the objective function. In Fig. 1(A)(c), the region is shown in green color indicates the initial guess points that are converging to global minimum ($*$), while the remaining points in the parameter domain converge to multiple local minima as depicted in red color. The global parameters are identified only if the chosen initial guesses fall in the region specified in green color, which is difficult to obtain when gradient-based optimizers are used for estimation.

To address the issue of multiple local minima in the objective function (Eq. (7)), we introduce a high gain proportional-derivative (PD) controller with a morphing parameter (λ) to the objective function (Eq. (7)) implicitly, and to the mathematical model (Eq. (6)) explicitly. Thus, we have the following objective function for estimation:

$$J\left(\hat{\mathbf{p}},\, \lambda\right) = \frac{1}{2}\int_0^{t_f} \left(x_{1e}(t) - x_1(\hat{\mathbf{p}},\, \lambda,\, t)\right)^2 dt \qquad (8)$$

and the corresponding mathematical model is used:

$$\dot{x}_1(t) = x_2(t) + \lambda\, k_d(x_{1e}(t) - x_1(t)) \qquad (9a)$$

$$\dot{x}_2(t) = \hat{\mu}(1 - x_1^2(t))x_2(t) - \hat{\alpha}x_1(t) - \hat{\beta}x_1^3(t) + f(t) + \lambda\, k_p(x_{1e}(t) - x_1(t)) \qquad (9b)$$

where, $x_{1e}(t) - x_1(t)$ represents the PD controller, and $[k_d,\, k_p] = 10$ is the gain value selected such that synchronization (Abarbanel et al. (2009)) occurs between experimental and estimated responses of the oscillator when initial parameter guesses are chosen.

The estimation process begins by providing initial guess for each parameters $[\alpha^0,\, \beta^0,\, \mu^0] = [7.00,\, 8.00,\, 0.01]$, which are far away from the actual values, and proceed the optimization by minimizing Eq. (8) with $\lambda = 1$. As $\lambda = 1$, the objective function becomes convex, and the minimum of $J\left(\hat{\mathbf{p}},\, 1\right)$ is determined. The value of λ is now reduced by 0.2, and the next iteration is initiated using the previous value of minimizer as the initial guess to the current step. This procedure is continued till λ becomes zero, and in the process, global parameters are identified (Table 2). The convergence of parameter estimates during the optimization process is shown in Fig. 2(A)(a).

Similarly, the parameters in double well and double-hump VDPD oscillators are estimated, and the simulation results are given in Figs. 2(B)(a) and 2(C)(a). The parameter estimation in the double-hump oscillator is the most difficult one, as the system response is highly unstable, as shown in Fig. 1(C)(a). So, the time window of $t \in [0,\, 2.5]$ seconds is selected for the estimation process. For numerical simulation, we used the *fmincon* built-in optimization routine in MATLAB, along with a fourth-order *Runge-Kutta* method for solving the ODEs. Sequential quadratic programming (SQP) algorithm in the *fmincon* routine is used as an optimizer with a first-order optimality tolerance of 10^{-6} is implemented as the termination criteria. The sensitivity information of the parameter estimates is obtained using finite difference methods. For basin of attraction, we used the *MultiStart* routine with parallel processing option in MATLAB to generate the multiple local minima.

3.2 Comparison with stochastic optimizers

Stochastic optimizers are the global searching techniques adopted universally for minimizing the multimodal objective functions. The search technique employed in these optimizers is based on the principles of evolution and survival of the fittest strategy. We have used the genetic algorithm (GA) and particle swarm optimization (PSO) techniques for the estimation of the parameter in VDPD oscillators. A comparison study is then performed between the PD controller algorithm and the stochastic optimizers.

Firstly, we perform the estimation process using GA and PSO algorithms and identify the optimal parameters. The evolution of parameter estimates during the optimization process is shown in Figs. 2(b) and 2(c) for all the three

Table 1. System parameters for identification in VDPD oscillators.

Parameter	Single well			Double well			Double-hump		
	Experimental Value ($\mathbf{p}$)	Lower Bound ($\mathbf{p}^l$)	Upper Bound ($\mathbf{p}^u$)	Experimental Value ($\mathbf{p}$)	Lower Bound ($\mathbf{p}^l$)	Upper Bound ($\mathbf{p}^u$)	Experimental Value ($\mathbf{p}$)	Lower Bound ($\mathbf{p}^l$)	Upper Bound ($\mathbf{p}^u$)
α	0.5	0.1	10	-0.5	-5.0	5.0	1.5	1.0	5.0
β	0.5	0.1	10	0.5	0.1	10	-0.5	-0.9	-0.1
μ	0.1	0.01	2	0.1	0.01	2	0.1	0.01	1.0

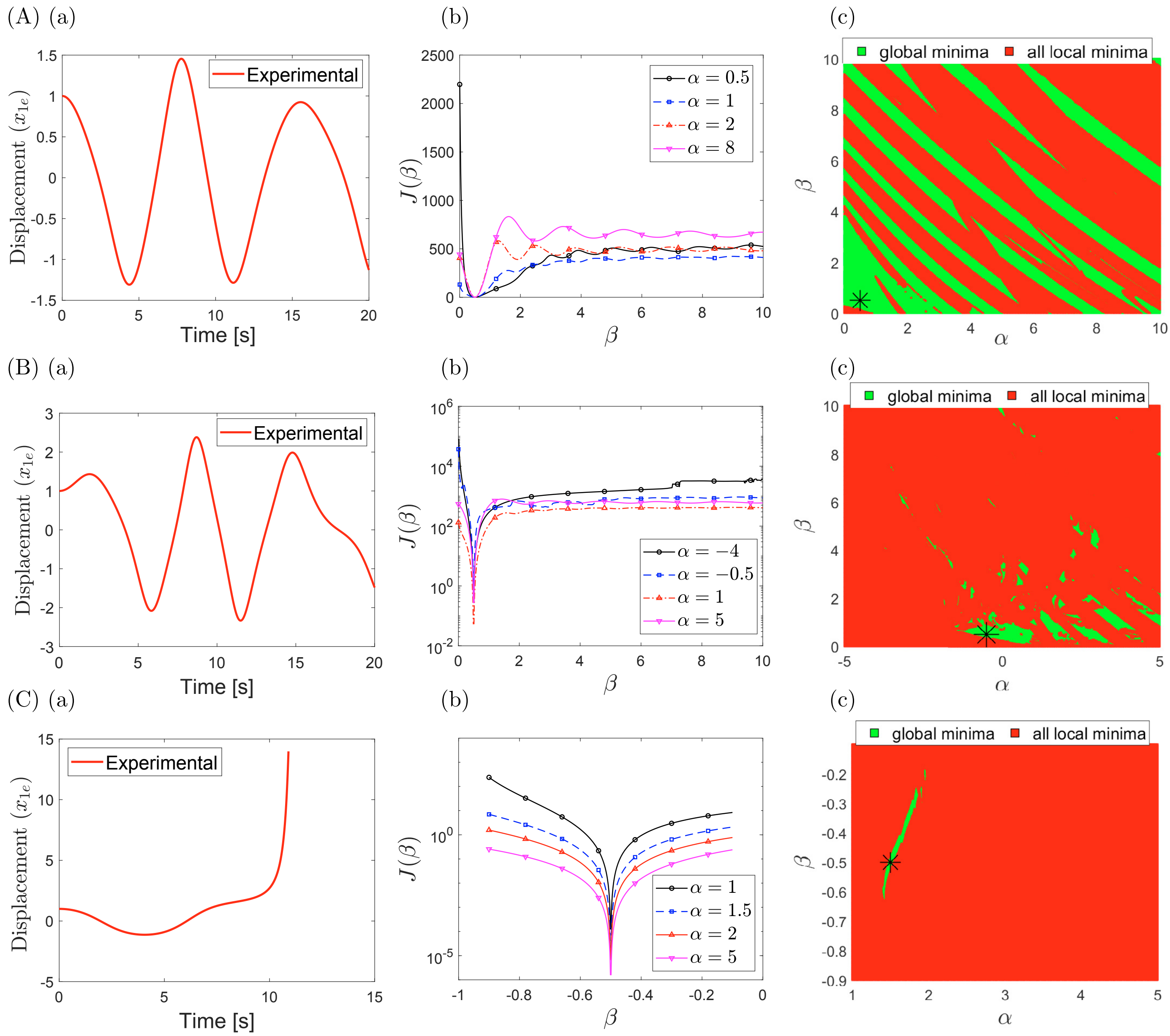

Fig. 1. (a) Experimental data (x_{1e}) for estimation, (b) objective function $J(\beta)$ and (c) basin of attraction for (A) single well, (B) double well and (C) double-hump VDPD oscillators keeping $\mu = 0.1$ constant.

Table 2. Comparison of estimated parameters and relative percent errors in VDPD oscillators.

Algorithm	Single well			Double well			Double-hump		
	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\gamma}$	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\gamma}$	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\gamma}$
PD controller	0.5(0.00%)	0.5(0.00%)	0.10(0.00%)	0.50(0.00%)	0.50(0.00%)	0.10(0.00%)	1.49(0.01%)	-0.49(0.02%)	0.10(0%)
PSO	0.5(0.00%)	0.5(0.00%)	0.10(0.00%)	0.50(0.00%)	0.50(0.00%)	0.10(0.00%)	1.58(5.69%)	-0.59(19.32%)	0.06(36.7%)
GA	0.5(0.00%)	0.5(0.00%)	0.10(0.00%)	0.50(0.02%)	0.50(0.00%)	0.10(0.00%)	1.56(3.87%)	-0.57(13.2%)	0.08(24.2%)

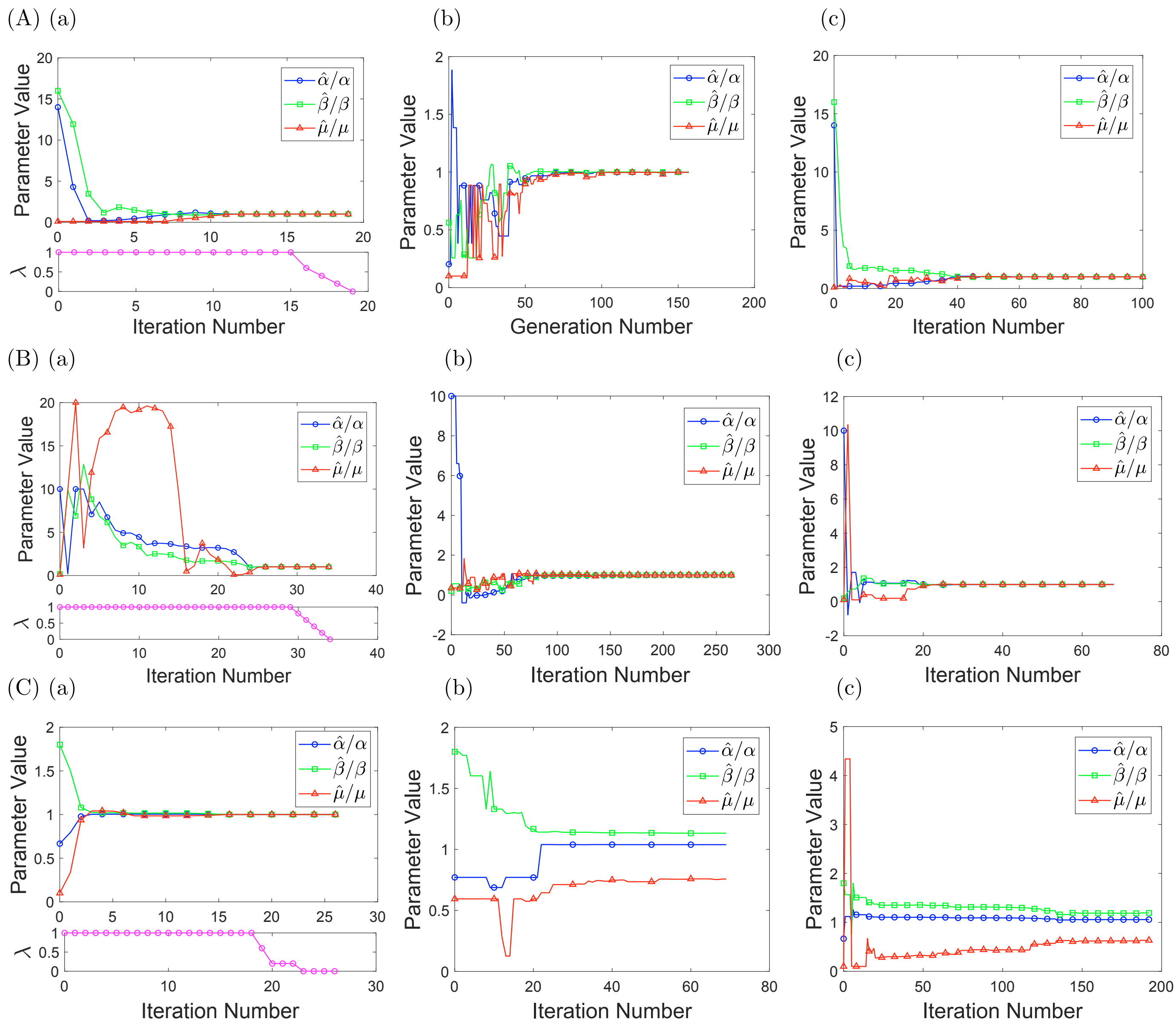

Fig. 2. Evolution of parameter estimates (normalized relative to the experimental values) using (a) PD controller algorithm, (b) genetic algorithm and (c) particle swarm optimizer for (A) single well (B) double well and (C) double-hump VDPD oscillators.

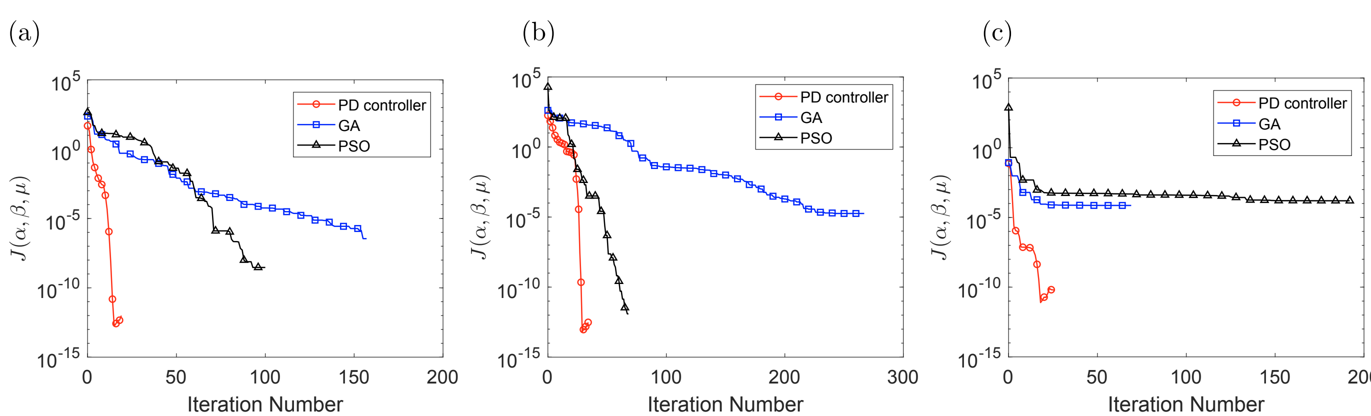

Fig. 3. Error convergence of $J(\alpha, \beta, \mu)$ in (a) single well (b) double well and (c) double-hump VDPD oscillators using PD controller, GA and PSO algorithms.

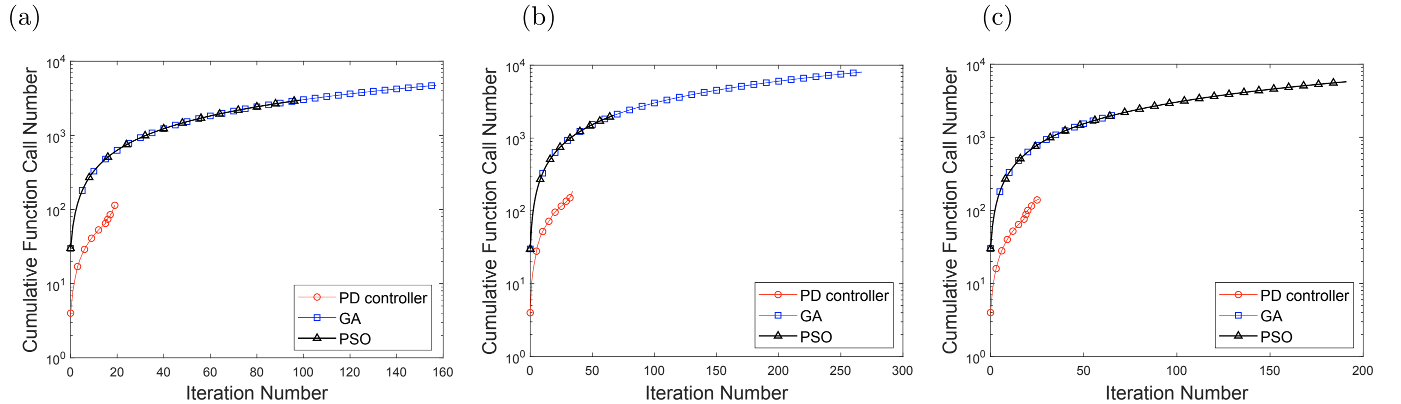

Fig. 4. Computational effort for parameter estimation in (a) single well (b) double well and (c) double-hump VDPD oscillators using PD controller, GA and PSO algorithms.

Table 3. Comparison of computational effort of PD controller, GA and PSO algorithms in VDPD oscillators.

	Single well			Double well			Double-hump		
	PD controller	GA	PSO	PD controller	GA	PSO	PD controller	GA	PSO
No. Iter	19	157	99	34	266	68	26	69	192
No. Fun. Eval	114	4740	3030	186	8040	2070	186	2100	5790
Error (J)	9.89×10^{-13}	3.51×10^{-7}	2.93×10^{-9}	3.05×10^{-13}	1.79×10^{-5}	1.19×10^{-12}	5.92×10^{-11}	7.24×10^{-5}	1.56×10^{-4}

VDPD oscillators. We used the *ga* and *pso* built-in MATLAB optimization routines with a population size of 30 in each case. The population size is $10 \times n$, where n is the number of variables in the objective function. We then compare these results with the PD controller algorithm for error convergence and computational effort as depicted in Figs. 3 and 4. From Figs. 3 and 4, it is clear that the PD controller algorithm converges faster than GA and PSO with fewer iterations and function evaluations. As gradient-based optimizer is used in the PD controller algorithm, convergence is faster. Contrarily, the stochastic optimizers are working on the principle of evolution, and hence they converge slowly and requires more iterations and function evaluations than the PD controller algorithm. The performance of the three algorithms based on the convergence properties is summarized in Table 3.

4. CONCLUSIONS

In this paper, we have estimated the global parameters of the VDPD oscillators using the PD controller algorithm. Since gradient-based optimizers are used for identification, global parameters are identified with a small number of iterations and function evaluations. Comparison with stochastic optimizers demonstrated the efficacy of the present algorithm for error convergence and computational effort.

REFERENCES

Abarbanel, H.D.I., Creveling, D.R., Farsian, R., and Tsimring, M. (2009). Dynamical state and parameter estimation. *SIAM Journal on Applied Dynamical Systems*, Vol.8(4), 1341–1381.

Frewin, C. (2013). Literature review and implementation of parameter identification methods for multibody systems governed by diferential algebraic equations. University of Applied Sciences Upper Austria, Wels, Austria.

Gershenfeld (1999). *The nature of mathematical modeling*. Cambridge University Press, Cambridge.

Goldberg, D.E. (2002). *Genetic algorithm in search, optimization and machine learning*. Adison-Wesley, Boston.

Ljung, L. and Söderström, T. (1983). *Theory and practise of resursive identitifcation*. MIT press, Cambridge.

Quaranta, G., Monti, G., and Marano, G.C. (2010). Parameters identification of Van der Pol-Duffing oscillators via particle swarm optimization and differential evolution. *Mechanical Systems and Signal Processing*, 24, 2076–2095.

Rao, S.S. (2009). *Optimization theory and applications*. John Wiley and Sons, New York.

Romeijn, H.E. and Smith, R.L. (1994). Simulated annealing for constrained global optimization. *J.Global.Optim*, 5(2), 101–126.

Vyasarayani, C.P., Uchida, T., and McPee, J. (2012a). Nonlinear parameter identification in multibody systems using homotopy continuation. *Journal of Computational and Nonlinear Dynamics*, 7, 011012:1–8.

Vyasarayani, C.P., Uchida, T., and McPee, J. (2012b). Single-shooting homotopy method for parameter identification in dynamical systems. *Physical Review E*, 85, 036201.

Watson, L.T. and Haftka, R.T. (1988). Modern homotopy methods in optimization. *Elsevier Science Publishers B.V. (North-Holland)*.

Zhao, Z. and Yang, S. (2015). Application of Van der Pol-Duffing oscillator in weak signal detection. *Computers and electrical Engineering*, 41, 1–8.