Proceedings of the
47th IEEE Conference on Decision and Control
Cancun, Mexico, Dec. 9-11, 2008

A Nonlinear Least Squares Estimation Procedure without Initial Parameter Guesses

Julián Bonilla, *Member, IEEE*, Moritz Diehl, Bart De Moor *Fellow, IEEE* and Jan Van Impe

Abstract— **This paper introduces a convex formulation approach for the initialization of parameter estimation problems (PEP). The proposed method exploits the parameter-affine feature exhibited by some dynamic systems. The method attempts to solve a related convex problem and uses its result as the initial guess for the solution of the original nonconvex PEP. The proposed approach is illustrated through two nonconvex parameter estimation study cases; the harmonic oscillator and the Lorenz attractor.**

I. INTRODUCTION

Obtaining accurate models of dynamic systems has an enormous impact on science and engineering. Models used to predict and control process dynamics are basically characterized by their structure and the parameter values in this structure. Parameter estimation problems refer to the calculation of a set of values in a predefined model structure, linear or nonlinear, such that the outputs of the model for this given set of parameters fit some measurement data. Additionally, constraints on parameter values and model states are usually required e.g. positive concentrations and reaction rates, upper and lower limits in model outputs. Consequently, PEP are recast as optimization problems leading to convex or nonconvex formulations according to the nature of the cost, model and constraints.

Nonconvex PEP are difficult to solve since they might exhibit local solutions and the true parameter or global solution might be hard to find. Fast and efficient techniques based on Newton type methods have been proposed [1]. These methods are derivative-based, and they can easily lock on to a local solution if the problem is not initialized appropriately. The most reliable approches for PEP are based on the constrained Gauss-Newton method, with simultaneous optimization [2]. Nevertheless, all these approaches require a starting point to initialize the algorithms.

Other optimization techniques have been also introduced for these kind of problems, such as non-deterministic global optimization methods based on random search, genetic algorithms and simulated annealing and deterministic approaches based on branch and bound B&B [3]. These techniques usually are computationally more expensive than the derivative-based optimization approach.

This work proposes an initialization method for nonconvex PEP involving a particular class of dynamic models, namely parameter affine systems. The approach leads to a convex problem where initialization is not required and a global solution can easily be obtained by convex optimization methods [4]. Other well-known procedures leading to convex problems have been proposed for parameter affine system, such as Least Squares Prediction Error Methods LS-PEM [5]. These methods are sensitive to noisy data and in order to work well in practice, they need to filter the residual errors.

The approach presented in this paper attempts to find an initial guess of the parameter by solving a related convex formulation. Hereafter, the solution of this convex problem can be used to initialize the original Nonlinear Least Squares Parameter Estimation Problem (NLS-PEP) using simultaneous optimization techniques based on multiple shooting [2] or collocation [7].

The paper is organized as follows: section II introduces the least squares parameter estimation problem for nonlinear systems. Section III shows a well known approach to the PEP involving input affine models. Section IV presents the proposed approach for parameter estimation using a least squares norm and a parameter affine model. Numerical examples comparing the known approaches with the proposed method are developed in section V. Conclusions follow in section VI.

II. THE PARAMETER ESTIMATION PROBLEM

Consider a dynamic system of the form

$$\dot{x}(t) = \Phi(x(t), p) \quad \forall \quad t \in [0, T] \tag{1}$$

where the vector $p \in \mathbb{R}^{n_p}$ and $x(t) \in \mathbb{R}^{n_x}$ denote model parameters and states respectively. In order to estimate the value of the vector p, a set of measurements $y(t) \in \mathbb{R}^{n_y}$ with $n_y \geq n_p$ is collected along the time interval $[0, T]$. The set of measurements $y(t)$ does not necessarily correspond to the model states $x(t)$, however, here it is assumed that the measurement set corresponds to the measurements of the system states, i.e. $y(t) = \bar{x}(t)$. The mismatch between the output of the model (1) and the measurements are usually quantified using a least squares (LS) norm

$$L(x(t), p, t) = \frac{1}{2}\|x(t) - \bar{x}(t)\|_Q^2, \tag{2}$$

J. Bonilla is with the Departments of Chemical Engineering CIT/BioTeC and Electrical Engineering ESAT/SCD, Katholieke Universiteit Leuven, B-3001 Leuven, Belgium julian.bonilla@cit.kuleuven.be

M. Diehl is professor at the Department of Electrical Engineering ESAT/SCD, Katholieke Universiteit Leuven, B-3001 Leuven, Belgium moritz.diehl@esat.kuleuven.be

B. De Moor is full professor at the Department of Electrical Engineering ESAT/SCD, Katholieke Universiteit Leuven, B-3001 Leuven, Belgium bart.demoor@esat.kuleuven.be

J. Van Impe is full professor at the Department of Chemical Engineering CIT/BioTec, Katholieke Universiteit Leuven, B-3001 Leuven, Belgium jan.vanimpe@cit.kuleuven.be

978-1-4244-3124-3/08/$25.00 ©2008 IEEE

where Q weights the mismatch[1]. Following the presented notation, the PEP can be formulated in the form:

$$\min_{x(\cdot),p} \int_0^T L(x(t), p, t)\, dt, \tag{3}$$

subject to

$$\begin{aligned} (x(0) &= x_0), & (4)\\ \dot{x}(t) &= \Phi(x(t), p), \quad t \in [0, T], & (5)\\ x(t) &\in \mathbb{X}, \quad t \in [0, T], & (6)\\ p &\in \mathbb{P}. & (7) \end{aligned}$$

The constraint (4) can be suppressed if the initial condition is also considered a degree of freedom for the optimization problem.

Least squares approach lies in the group of methods which attempt to minimize the errors between data and model outputs with respect to a given norm. It corresponds to maximize the probability of p^* given the set $\bar{x}(t)$ assuming a normal distribution in the measurements errors with a covariance matrix Q^{-1}. Constraints on the parameters and the model states can be introduced into the optimization framework by the sets $\mathbb{P}$ and $\mathbb{X}$ respectively. Consequently, PEP are considered as optimization problems and might lead to nonconvex formulations, particularly when nonlinear models are considered.

III. LS-PREDICTION ERROR METHODS

If the model (1) is affine in the parameters or can be reformulated in such a form, and if the state measurements or estimations are available, the prediction error method (PEM) can be used to estimate the parameter vector p. In the least squares prediction error methods (LS-PEM), an unconstrained optimization problem of the form

$$L_{\text{pem}}(x(t), p, t) = \frac{1}{2}\|\dot{\bar{x}}(t) - \Phi(\bar{x}(t), p, t)\|_Q^2, \tag{8}$$

$$\min_p \int_0^T L_{\text{pem}}(x(t), p, t) dt \tag{9}$$

is solved. The method uses a predictor based on the model (1) and attempts to minimize the difference between measured data and the prediction one-step ahead of this data using past samples. By using this approach along with a parameter-affine model, the problem of estimating p becomes convex leading to the global solution for any initial value. However, the solution can be biased if the process uncertainty and noise are not modelled appropriately by the model used for prediction. In order to do so, the use of an appropriate noise model is advised. For linear systems the use of different noise models is equivalent to filter the error sequence [5]. Thus, the design of a filter for (8) is a neccesary step to avoid biased estimations when PEM are used.

[1]This formulation corresponds to the weighted least squares

IV. A NEW CONVEX FORMULATION

Consider the problem presented in (3)-(7). It is assumed that the optimization task is described by a set of state measurements $\bar{x}(t)$. The measurements need not necessarily be feasible with respect to the model equations (5). However, we assume that $\bar{x}(t) \in \mathbb{X}(t)$, i.e. the data set is feasible with respect to the inequality constraints. We also assume that $\mathbb{X}(t)$ is bounded.

Consider the following extra assumptions:

- The sets $\mathbb{P}$ and $\mathbb{X}$ are convex.
- $\Phi(x(t), p, t)$ is affine in p

Remark 1 : The above assumptions are satisfied if the parameters and the states are constrained by simple bounds (thus, $\mathbb{P} \times \mathbb{X}$ is a hypercube), and if the dynamic model-equations is defined by

$$\Phi(x(t), p) = f(x) + g(x)p. \tag{10}$$

The approach proposed here is inspired by homotopy methods [6]. This kind of methods attempt to solve an optimization problem by first solving a related problem which is connected to the original one by a homotopy path. The nonconvex PEP (3)-(7) is reformulated by introducing a homotopy parameter $\lambda \in (0, 1)$, a new variable $w(t)$ and a norm on this new variable in the cost:

$$\begin{aligned} P(\lambda) : \min_{w(\cdot),x(\cdot),p} \frac{1}{2\lambda} \int_0^T \| x(t) - \bar{x}(t) \|_Q^2\, dt \\ + \frac{1}{2(1-\lambda)} \int_0^T \| w(t) \|_Q^2\, dt \end{aligned} \tag{11}$$

subject to

$$\begin{aligned} x(0) &= x_0 - w(0), & (12)\\ \dot{x}(t) &= f(x(t)) + g(x(t))p - \dot{w}(t),\ t \in [0, T]. & (13) \end{aligned}$$

It is not difficult to observe from (13) that $w(t)$ corresponds to the integral of the modelling errors along the time interval T. The norm on $w(t)$ in (11) is a natural choice due to the nature of the original cost formulation.

The condition $\lambda \rightarrow 1$ leads to the original problem (3)-(7) while $\lambda \rightarrow 0$ force $x(t) \rightarrow \bar{x}(t)$ leading to the convex problem in $w(t)$ and p:

$$P(0) : \min_{w(\cdot),p} \frac{1}{2} \int_0^T \| w(t) \|_Q^2\, dt \tag{14}$$

subject to

$$\begin{aligned} \bar{x}(0) &= x_0 - w(0), & (15)\\ \dot{\bar{x}}(t) &= f(\bar{x}(t)) + g(\bar{x}(t))p - \dot{w}(t),\ t \in [0, T] & (16) \end{aligned}$$

which corresponds to the convex extreme of the homotopy map generated by the parametric PEP (11)-(13). An example of this homotopy map is illustrated in figure 1.

In order to clarify this formulation, the definition

$$\dot{z}_a(t) = f(\bar{x}(t)) + g(\bar{x}(t))p, \tag{17}$$

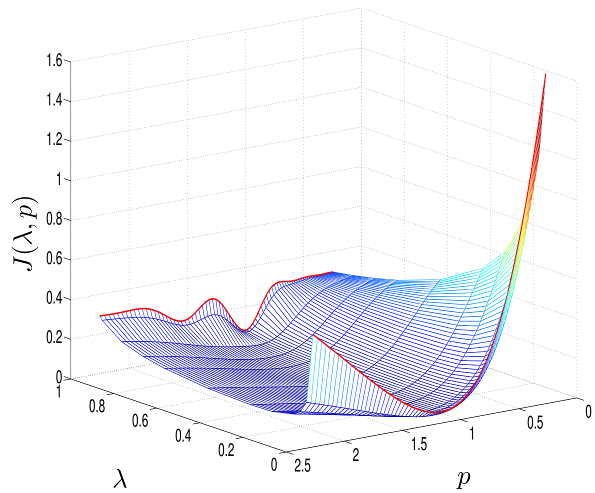

Fig. 1. Homotopy map for the parameter estimation problem of a parameter-affine model. The original PEP is nonconvex ($\lambda \to 1$) but if the measured state is used as the real state, it is possible to achieve convexity in one of the extremes of the map ($\lambda \to 0$).

is introduced in (16), leading to

$$w(t) = z_a(t) - \bar{x}(t). \tag{18}$$

Consequently, the convex PEP (14)-(16) can be recast as

$$\min_{z_a(.),p} \frac{1}{2} \int_0^T \|z_a(t) - \bar{x}(t)\|_Q^2 dt, \tag{19}$$

subject to

$$\begin{aligned}
(z_a(0) &= x_0), & (20)\\
\dot{z}_a(t) &= f(\bar{x}(t)) + g(\bar{x}(t))p, & (21)\\
z_a(t) &\in \mathbb{X}(t) & (22)\\
p &\in \mathbb{P}. & (23)
\end{aligned}$$

The model (21) in terms of $z_a(t)$, can be seen as a linearization of the original function (5) around the measurements $\bar{x}(t)$, where the term involving the variation of $\Phi(x(t), p)$ with respect to $x(t)$ has been neglected. This approach is closely related to prediction error methods (PEM) for parameter estimation. However, the proposed approximation differs from the classic PEM in the filter used to process the residuals. In the PEM the pre-filter is a degree of freedom the user has to define in the pre-processing stage. Here, the use of an arbitrary filter is avoided and it is obtained in a more natural form by simply recasting the problem (3)-(7) as in (19)-(23). Consequently, after setting up the NLS-PEP all the work by the user is done.

Following the presented formulation the NLS-PEP becomes convex enabling the use of more reliable tools and algorithms to solve the problem. This method is referred in the following as *Least Squares Convex Approach* (LS-CA).

V. NUMERICAL EXAMPLES

In order to illustrate the proposed approach to solve a PEP, the following study cases are considered:

A. The Harmonic Oscillator

Regard the dynamic system described by

$$\Phi(x(t), p) = \begin{pmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \end{pmatrix} = \begin{pmatrix} x_2(t) \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ -x_1(t) \end{pmatrix} p. \tag{24}$$

Assume that a set of noisy measurements for the states $x_1(t)$ and $x_2(t)$ is available along a time horizon T. For emulation purposes the data sequence is generated by integrating (24) during 12 seconds using the Runge-Kutta 4^{th} order method with a fixed integration step of $h = 13.3$ ms. The data set used for the PEP is illustrated at the top of Fig. 2. In order to emulate more realistic data, the sequence has been contaminated with colored noise in high and low frequency ranges as depicted at the bottom of Fig. 2.

Given the data set $\bar{x}(t)$ of length N and defining $x(t) = [x_1(t)\ x_2(t)]^T$, the NLS-PEP is discretized with a sampling period h along the horizon T leading to the nonlinear programming problem (NLP)

$$\min_{x(\cdot),p} \frac{1}{2} \sum_{k=0}^{N-1} \| x(k) - \bar{x}(k) \|_Q^2, \tag{25}$$

subject to

$$\begin{aligned}
x(k+1) &= x(k) + \Phi_h(x(k), p),\ k \in [0, N-1], & (26)\\
x(k) &\in \mathbb{X}(k),\ k \in [0, N-1], & (27)\\
p &\in \mathbb{P}. & (28)
\end{aligned}$$

Notice that the constraint related to the initial condition $x(0)$ has been suppressed here since the problem is firstly pre-optimized with respect to $x(.)$. Figure 3 shows the cost

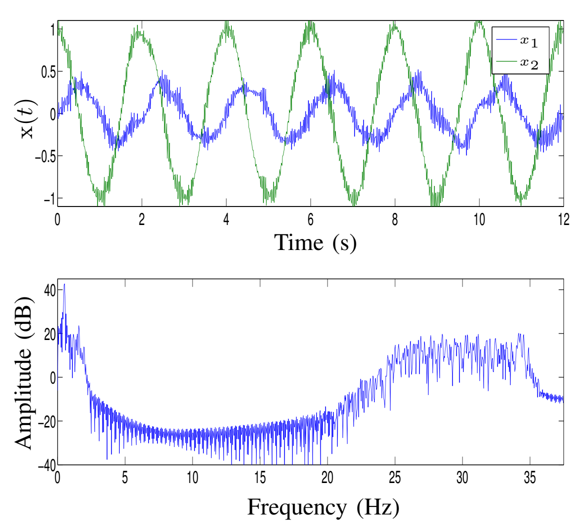

Fig. 2. Noisy measurements (top) and the frequency spectrum for $x_1(t)$ (bottom).

function for this particular data set. The problem exhibits the global solution at $p^* = \pi^2$ and several local minima at $p = 1.53\pi^2,\ 2\pi^2,\ 2.507\pi^2$ among other. Thus, a derivative-based optimization technique could lead to the wrong solution if not initialized close to the global minimum.

The LS-PEM approach can be easily applied to the problem (25)-(28) leading to the optimization problem

$$\min_{p} \frac{1}{2} \sum_{k=1}^{N-1} \|x(k) - \bar{x}(k)\|_Q^2. \tag{29}$$

subject to

$$x(k+1) = \bar{x}(k) + \Phi_h(\bar{x}(k), p),\ k \in [0, N-2], \tag{30}$$
$$x(k) \in \mathbb{X}(k), \quad k \in [0, N-1], \tag{31}$$
$$p \in \mathbb{P}. \tag{32}$$

The cost obtained for the LS-PEM approach is illustrated in Fig. 3. Although the problem is now convex, the minimum of this convex function is biased with respect to the original NLS-PEP, due to the colored noise. Consequently, in order to appropriately apply PEM, the noise model in (26) has to be improved or the prediction errors should be filtered. If the minimum given by this method is used to initialize the original NLS in order to refine the optimum, a local optimization routine would lead to a wrong local solution, since the biased value is closer to a one of the other local minima

On the other hand, the proposed approach (19)-(23) is discretized, leading to

$$\min_{z_a(\cdot), p} \frac{1}{2} \sum_{k=0}^{N-1} \| z_a(k) - \bar{x}(k) \|_Q^2, \tag{33}$$

subject to

$$z_a(k+1) = z_a(k) + \Phi_h(\bar{x}(k), p),\ k \in [0, N-2], \tag{34}$$
$$z_a(k) \in \mathbb{X}(k), \quad k \in [0, N-1], \tag{35}$$
$$p \in \mathbb{P}. \tag{36}$$

The cost function derived from this problem, with the noisy measurements, has been pre-optimized with respect to $z_a(.)$ and illustrated in Fig. 3. The LS-CA approach provides a better characterization of the global minimum for this test. Although the original measurement set contains high and low frequency noise they do not significantly affect the estimation of the global optimum. It is possible to obtain similar results with PEM methods, however, the user must then define a suitable filter in order to decrease the bias in the estimation. This noise model can be obtained by fitting the noise spectrum through filters of the form:

$$\begin{aligned} \dot{w}(t) &= A_n w(t) + K_n e(t) \\ v(t) &= C_n w(t) + D_n e(t) \end{aligned} \tag{37}$$

where $e(t) \in \mathbb{R}^{n_x}$ and $v(t) \in \mathbb{R}^{n_x}$ correspond to white and colored noise respectively.

For this particular example, colored noise in the bands [20-35]Hz and [1-2.5]Hz can be identified by simple inspection of spectrum in figure 2. Thus, the noise spectrum is approximated by designing a second order bandpass filter for each band using the state space representation (37). A more accurate but involved approach is to use spectral factorization to reconstruct the noise model from the noise spectrum [5].

The designed filter is added to the original model (24) and the one-step ahead predictor is formulated using the state space model. Figure 3 shows that the LS-PEM with the noise filter (LS-PEM+NF) provides a less-biased optimum value than without the filter. However, the optimum obtained does not exactly match the real one since an exact description of the noise is necessary to achieve perfect fit of the estimated parameter. Consequently, the design of the noise filter in the predictor is an additional step the user should do in order to get a less biased estimation when using PEM. This step is avoided with the proposed LS-CA approach.

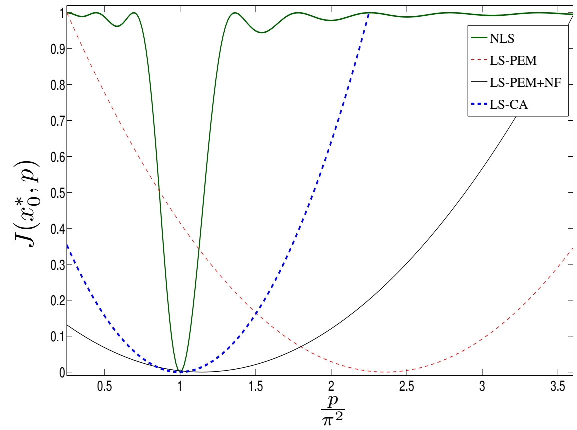

Fig. 3. Cost function for harmonic oscillator PEP using different approaches.

TABLE I
COST VALUES FOR THE ORIGINAL NLS-PEP AND ITS PROPOSED VARIANTS FOR THE HARMONIC OSCILLATOR EXAMPLE

Cost	$\mathbf{J}_{\min}$	$\mathbf{J}_{\max}$	$\frac{\mathbf{p}_n^*}{\pi^2}$	$\mathbf{e}(\mathbf{p}_n^*)$
NLS	7.605e-3	281.83e-3	0.9989	0.0
LS-PEM	20.395e-3	22.42e-3	2.354	135.605
LS-PEM+NF	9.893e-3	13.99e-3	1.145	14.573
LS-CA	8.17e-3	402.45e-3	0.9949	0.402

Results are summarized in table I where the distance to the global minimum is calculated using

$$e(p_n^*) = 100 \times \frac{|p^* - p_n^*|}{p^*}, \tag{38}$$

where p^* corresponds to the global solution of the original NLS-PEP.

B. Lorenz Attractor

As a second study case, consider the attractor introduced by Edward Lorenz in 1963. The dynamics of the Lorenz attractor is governed by a set of differential equations of the

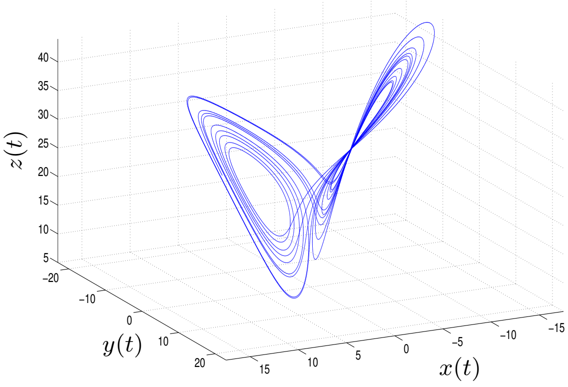

Fig. 4. Chaotic behavior of Lorenz attractor, $\sigma = 10$, $\beta = 8/3$ and $\rho = 28$.

form

$$\begin{aligned} \dot{x}(t) &= \sigma(y-x), \\ \dot{y}(t) &= x(\rho-z)-y, \\ \dot{z}(t) &= xy-\beta z, \end{aligned} \qquad (39)$$

where σ, β and ρ are positive parameters. Typical values to illustrate the chaotic behavior of this set of nonlinear equations are $\sigma = 10$, $\beta = 8/3$ and $\rho = 28$. Note that the system presents the structure (10) with

$$p = \begin{pmatrix} \sigma \\ \rho \\ \beta \end{pmatrix} \qquad (40)$$

The multiple time scale evolution and the chaotic behavior of (39), showed in Fig. (4), is often used to illustrate the internal dynamics of the earth atmosphere [9] and it has been used as study case for parameter estimation methods as well in [10], [11].

In this numerical example, the 4^{th} order Runge-Kutta integration method is used with a fixed step of $h = 0.01$ s. Due to the high sensitivity to initial conditions and parameter changes, the identification of the vector p is a challenging task. In order to illustrate the difficulty of parameter estimation in this example, a simulation is run with initial condition $[x_0, y_0, z_0] = [16, 18, 35]$. Figure 5 illustrates the effect of a small perturbation on z_0 , Δz. Additionally, the effects of a 0.35% change in ρ are simulated. In both cases, the trajectories noticeably diverge from the unperturbed one after some seconds of simulation.

It is assumed that a sequence containing measurements for all states $\bar{x}(t)$ is provided as illustrated in Fig. 6. This sequence is contaminated with colored noise. Thus, it is expected that the LS-PEM without any noise information provides a biased solution.

In order to better visualize the cost functions for the presented approaches, σ and β are considered constant and ρ is estimated. The parameter estimation problem is formulated as in (25)-(28) with the discrete representation of the model presented in (39). The NLS-PEP cost as a function of $\rho \in [10,\ 40]$ is presented in Fig. 7. For this case, no pre-optimization on $x(.)$ has been performed since the initial condition is assumed fixed. Clearly this cost is non-smooth and contains several local minima due to the stochastic behavior of the model for values of $\rho > 24$. Consequently, derivative-based techniques hardly find the right solution when initialized far from the global optimum.

The parameter-affine structure allows for the use of LS-PEM. Following the procedure presented in the harmonic oscillator example, the measurements are used to construct the predictor (26) and the least squares estimation problem is formulated as in (29). The cost to optimize for this problem is illustrated in figure 7. Due to the noise added, and the fact that a noise model is not considered, the global solution provided by LS-PEM methods is biased with respect to the global optimum. Consequently, it is required to define a filter for residuals based on the analysis of the noise spectrum of the collected measurements in figure 6. For this particular example, it is not difficult to recognize noise in the range 100-150Hz. Hence, a simple approach is to propose a filter with a bandwith matching this band. Notice that this is a heuristic approach and not always is easy to isolate the noise spectrum from the measured data. A second order bandpass filter of the form (37) is designed and added to each output of the original model i.e. $v(t), e(t) \in \mathbb{R}^3$ and $w(t) \in \mathbb{R}^6$. The results of the LS-PEM including this filter (LS-PEM+NF) are illustrated in figure 7.

The LS-CA approach is introduced for this parameter estimation problem, by using the Lorenz model (39) and the noisy data set in the optimization problem (33)-(36) with a fixed initial condition. The cost to optimize obtained with this formulation is presented in Fig. 7. Similarly to LS-PEM and LS-PEM+NF, the cost is convex. However, the minimum of the convex cost obtained with LS-CA is quite close to the global solution of the NLS-PEP. Notice that the plots for all the costs in Fig. 7 have been normalized for visualization purposes, nevertheless the minimum and maximum values

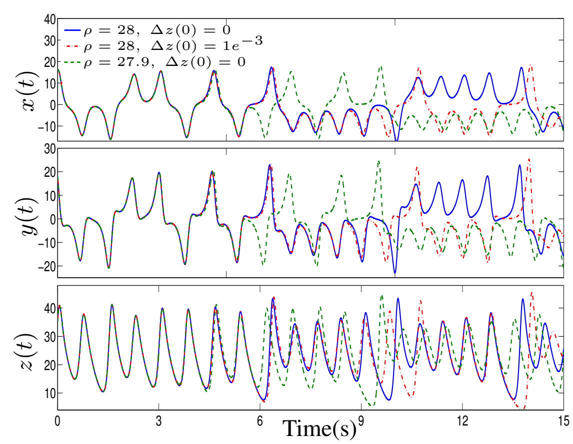

Fig. 5. Response to small perturbations on the initial condition and model parameters ($\sigma = 10$, $\beta = 8/3$)

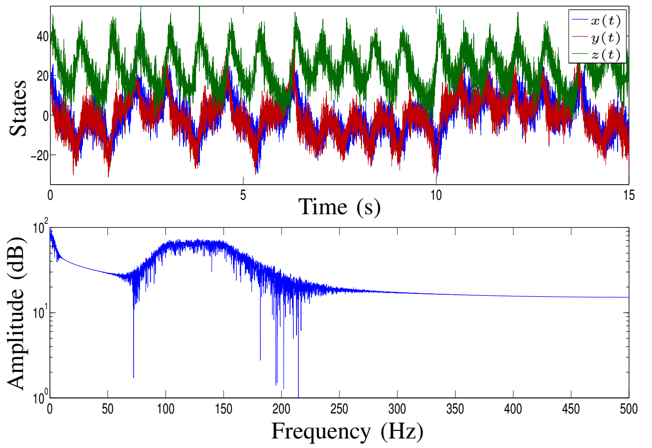

Fig. 6. Data used for the estimation problem, $\sigma = 10$, $\beta = 8/3$ and $\rho = 28$ (top), and the frequency spectrum of x(t) (bottom)

TABLE II

LORENZ ATTRACTOR EXAMPLE PEP. NOISE IN THE RANGE 100-150 HZ

Cost	$\mathbf{J}_{\min}$	$\mathbf{J}_{\max}$	$\rho^*_{\mathbf{n}}$	$\mathbf{e}(\rho^*_{\mathbf{n}})$
NLS	35.063	409.94	28.0	0.0
LS-PEM	20.71	20.75	12.9	53.93
LS-PEM+NF	21.10	21.31	25.8	7.86
LS-CA	58.01	14.69e3	28.6	2.14

for all the costs differ considerably. The same test is run with colored noise in other frequencies, in all cases, the proposed methods showed less bias than its LS-PEM counterpart with and without filter.

Table II and III show the minimum and maximum values for the presented costs over the considered domain of ρ, additionally, the distance of the optimum obtained with the presented methods is calculated using (38) with $p^* = \rho^* = 28$.

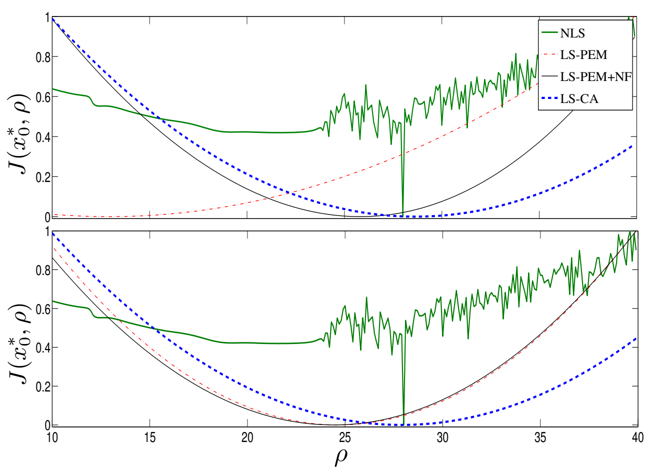

Fig. 7. Costs for different approaches $\sigma = 10$, $\beta = 8/3$. The data has been contaminated with noise in the range 100 to 150 Hz (top) and 400 to 450 Hz (bottom)

TABLE III

LORENZ ATTRACTOR EXAMPLE PEP. NOISE IN THE RANGE 400-450 HZ

Cost	$\mathbf{J}_{\min}$	$\mathbf{J}_{\max}$	$\rho^*_{\mathbf{n}}$	$\mathbf{e}(\rho^*_{\mathbf{n}})$
NLS	60.68	435.56	28	0.0
LS-PEM	227.86	227.86	24.60	12.14
LS-PEM+NF	227.96	227.97	24.40	12.85
LS-CA	70.39	13.56e3	27.90	0.36

VI. CONCLUSIONS

An approach to initialize nonconvex parameter estimation problems involving parameter-affine models has been presented. The approach uses the measurements into the dynamic model states similar to LS-PEM, however no prefiltering of the error residuals is required before the formulation of the optimization problem. The convexity achieved by the method enables the use of efficient convex optimization routines and leads to solutions which demonstrated to be less biased than the one obtained with classical approaches for a set of numerical examples. Consequently, the method is suitable for obtaining a good initial guess of the parameters to be estimated and can be used in combination with simultaneous Gauss-Newton approach were initialization is required.

VII. ACKNOWLEDGMENTS

Research supported by Research Council KUL: CoE EF/05/006 Optimization in Engineering(OPTEC), GOA AM-BioRICS, IOF-SCORES4CHEM, PhD/postdoc and fellow grants; Flemish Government: FWO: PhD/postdoc grants, projects G.0452.04, G.0499.04, G.0211.05, G.0226.06, G.0321.06, G.0302.07, G.0320.08, G.0558.08, G.0557.08, research communities (ICCoS, ANMMM, MLDM); IWT: PhD Grants, McKnow-E, Eureka-Flite; Helmholtz: viCERP; EU: ERNSI,HD-MPC; Contracts: AMINAL; Belgian Federal Science Policy Office: IUAP P6/04 (DYSCO, Dynamical systems, control and optimization, 2007-2011)

REFERENCES

[1] J. Nocedal and S. Wright, *Numerical Optimization*, Springer, NY; 2006
[2] H. Bock and K. Plitt, A Multiple Shooting Algorithm for Direct Solution of Optimal Control Problems. *In Proc. of the 9th IFAC world congress* 1984.
[3] C. Floudas, *Deterministic Global Optimization: theory, methods and applications*, Kluwer Academic Publishers, Dordrecht; 1999.
[4] S. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge University Press, Cambridge; 2006.
[5] L. Ljung, *System Identification, Theory for the Users*, Prentice Hall, NJ; 1999.
[6] L. T. Watson, Theory of globally convergent probability-one homotopies for nonlinear programming. *SIAM J. Optim.*, 11(3):761780, 2000.
[7] L. Biegler, Solution of Dynamic Optimization Problems by Successive Quadratic Programming and Orthogonal Collocation, *Comput. Chem. Engng.* 1984; Vol. 8, pp. 243-248.
[8] L. Ljung, Prediction Error Estimation Methods, *Circuits Systems Signal Processing*, vol. 21, No. I, 2002, pp. 11-21.
[9] T. N. Palmer, Extended Range Atmospheric Prediction and the Lorentz Model, *American Meteorological Society*, vol. 74, No 1, 1993, pp 49-55.
[10] D. J. Lea and M. R. Allen and T. W. N. Haine, Sensitivity analysis of the climate of a chaotic system, *Tellus*, vol. 52A, 2000, pp 523-532.
[11] J.D. Annan and J.C. Hargreaves, Efficient parameter estimation for a highly chaotic system, *Tellus*, vol. 56A, 2004, pp 520-526.