An iterative method for simulation of large scale modular systems using reduced order models

Muruhan Rathinam [1] and Linda R. Petzold
Computer Science Department
University of California
Santa Barbara, CA 93106
muruhan@engineering.ucsb.edu, petzold@engineering.ucsb.edu
Fax: 805-893-5435, Ph: 805-893-5728

Abstract

We describe a new iterative method for simulation of large scale modular systems using reduced order models that preserve the interconnection structure. Our technique essentially involves simulating in turn each subsystem connected to model reduced versions of the rest of the subsystems. The data from this simulation is then used to update the reduced model for that particular subsystem. We illustrate the method using a power grid example modelled by nonlinear swing equations.

1 Introduction

Modeling and simulation are some of the important steps in designing control methodologies for complex dynamical systems. Especially large scale systems such as the power grid of the United States present challenges in modeling and simulation. One failed power line can cause a cascade of failures which can be hard to predict without accurate modeling and simulation. On such systems conventional methods of direct numerical integration of the entire system may not be practical without massive computing resources.

An example from outside the control community is the simulation of VLSI circuits, where the need for efficient and accurate simulation has resulted in the development of new techniques that use an iterative approach that takes advantage of the modular structure of circuits. This technique is known as *waveform relaxation* (WR) in circuit simulation community. See Lelarasmee *et al* [8], Miekkala and Nevanlinna [10], and Miekkala [9] for instance. Other advantages of such modular approach are that they help exploit parallel computation, exploit multirate nature of some problems, and the potential to use different numerical techniques for different subsystems.

[1]Research supported in part by EPRI WO-8333-06 and NSF/KDI ATM-9873133.

Another way to deal with models that are too complex is via model reduction techniques. Several model reduction techniques have been studied by researchers in various fields. *Balanced truncation* has been studied by the control community [17, 7], *proper orthogonal decomposition* (POD) has been applied in the study of turbulence [5], cascading failures in power grids [13], and control of compressors [3]) etc., *selective modal analysis* has been developed by researchers in the electrical power field [14]), and *singular perturbation* techniques have been applied to modular control systems [16]), to name a few. In the case of large scale modular systems it is also desirable that the reduced order model maintains the modular structure (i.e. "respects" the modular structure). See [13] and [16] for some ideas.

In this paper we present an iterative approach to simulation that takes advantage of the modular nature of a large scale system and uses reduced order models that "respect" the modular structure. Our approach involves simulation of each subsystem in turn while it is connected to reduced order models of the rest of the subsystems. The simulation results are then used to update the reduced order model for that particular subsystem. If the reduced order models are small enough, then the combination of an unreduced subsystem with the rest of the reduced subsystems results in a system small enough not to pose any computational difficulties. In principle any model reduction method that uses data from trajectories could be used in this iteration. In this paper we use POD (also known as *Karhunen-Loève decomposition*) for the model reduction. We shall also provide an error analysis of the computed solution of a reduced model obtained from POD.

Our method may be compared with WR which involves simulating each subsystem in turn with the couplings coming from other systems treated as inputs. At the end of each iteration these couplings are updated. In our method, during each simulation of a subsystem the couplings are accounted for in a more direct, albeit approximate manner. This contrast makes our method different and hence could be expected to be a comple-

mentary method to the WR technique.

2 Model reduction using proper orthogonal decomposition

The model reduction technique described here essentially consists of first finding a subspace in the full phase space of a given dynamical system and then constructing an approximating dynamical system in that subspace. The original dynamical system may be nonlinear and in that case the resulting lower dimensional model will also typically be nonlinear.

2.1 Proper orthogonal decomposition

Given a set of data points $x^{(\alpha)} \in \mathbb{R}^n$, POD seeks a subspace $S \subset \mathbb{R}^n$ so that the total square distance

$$D = \sum_{\alpha=1}^{N} \|x^{(\alpha)} - P_S x^{(\alpha)}\|^2$$

(where P_S is the orthogonal projection onto the subspace S) is minimized. The norm considered is the 2-norm. (Thus we assume that the phase space comes equipped with a notion of inner-product.) The solution to this problem requires the construction of the *correlation matrix* defined by

$$R = \sum_{\alpha=1}^{N} x^{(\alpha)} (x^{(\alpha)})^T.$$

Note that R is symmetric positive semidefinite. Let $\lambda_1 \geq \lambda_2 \cdots \geq \lambda_N \geq 0$ be the ordered eigenvalues of R. Then the minimum value of D over all $k(\leq n)$ dimensional subspaces S is given by $\sum_{j=k+1}^{n} \lambda_j$ ([5]). In addition the S that minimizes D is the invariant subspace corresponding to the eigenvalues $\lambda_1, \ldots, \lambda_k$.

Remark 1 We assumed that the data $x^{(\alpha)}$ were discrete set of points. If we have a continuous trajectory $x(t)$ as data then the summations need to be replaced by integrals.

Often it may be best to find an affine subspace as opposed to a linear subspace. This requires first to find the mean value of the data points

$$\bar{x} = \frac{1}{N} \sum_{\alpha=1}^{N} x^{(\alpha)}$$

and then construct the *covariance matrix* $\bar{R}$ given by

$$\bar{R} = \sum_{\alpha=1}^{N} (x^{(\alpha)} - \bar{x})(x^{(\alpha)} - \bar{x})^T.$$

Let S_0 be the invariant subspace of the largest k eigenvalues of $\bar{R}$. Then the best approximating affine subspace S passes through $\bar{x}$ and is obtained by shifting S_0 by $\bar{x}$. Algebraically the projection onto the subspace S is given by

$$z = P(x - \bar{x}) \tag{1}$$

where $z \in \mathbb{R}^k$ are coordinates in the subspace S, $x \in \mathbb{R}^n$ are coordinates in the original coordinate system in $\mathbb{R}^n$, and the matrix P of the projection consists of row vectors ϕ_i^T $(i = 1, \ldots, k)$ where ϕ_i are the unit eigenvectors corresponding to the largest k eigenvalues of $\bar{R}$. Note that given any point $p \in S$ with coordinates $z \in \mathbb{R}^k$ the coordinates $x \in \mathbb{R}^n$ of the same point in the original coordinate system are given by

$$x = P^T z + \bar{x}.$$

For more information see Holmes *et al* [5], Moore [11], Lall *et al* [7], Glavaski *et al* [3] and references therein.

2.2 Galerkin projection

The procedure for constructing a vector-field on the approximating subspace is known as Galerkin projection and has been widely used in reducing PDEs to ODEs, see [5] for more details. Suppose the original dynamical system in $\mathbb{R}^n$ is given by a vector-field f,

$$\dot{x} = f(x, t).$$

Let $S \subset \mathbb{R}^n$ be the best k dimensional approximating affine subspace with projection given by (1). A vector-field f_a in the subspace S is constructed by the following rule: for any point $p \in S$ compute the vector-field $f(p, t)$ and take the projection $Pf(p, t)$ onto the subspace S to be the value of $f_a(p, t)$. If z are the subspace coordinates of p then $f_a(z, t) = Pf(P^T z + \bar{x}, t)$. Thus we obtain the following reduced model,

$$\dot{z} = f_a(z, t) = Pf(P^T z + \bar{x}, t). \tag{2}$$

If we are solving an initial value problem with $x(0) = x_0$, then in the reduced model one has the initial condition $z(0) = z_0$ where

$$z_0 = P(x_0 - \bar{x}).$$

The approximating solution $\hat{x}(t)$ in the original coordinates in $\mathbb{R}^n$ is given by

$$\hat{x}(t) = P^T z(t) + \bar{x}.$$

From the above it is easy to see that the approximating solution $\hat{x}(t)$ is the solution to the following initial value problem:

$$\dot{\hat{x}} = \mathcal{P} f(\hat{x}, t); \quad \hat{x}(0) = \hat{x}_0 = \mathcal{P}(x_0 - \bar{x}) + \bar{x} \tag{3}$$

where $\mathcal{P} = P^T P \in \mathbb{R}^{n \times n}$ is the matrix of the projection expressed in the original coordinate system in $\mathbb{R}^n$. Also note that $\hat{x}_0$ is just the projection of x_0 onto the affine subspace S.

2.3 Error analysis of POD method

Consider solving the initial value problem : $\dot{x} = f(x,t)$, $x(0) = x_0$, using a reduced order model in interval $[0,T]$. Then in effect we are solving the initial value problem (3). We shall derive an estimate for the error $e(t) = \hat{x}(t) - x(t)$. Let us denote the component of $e(t)$ orthogonal to the subspace S by $e_o(t)$ and the component parallel to S by $e_i(t)$. Thus $e_o(t)$ and $e_i(t)$ are orthogonal vectors. Hence by definition $\mathcal{P}e_o(t) = 0$ and $\mathcal{P}e_i(t) = e_i(t)$. It is important to observe that $e_o(t)$ comes from the first part of the method, i.e. the subspace approximation. It is the error between $x(t)$ and its projection onto the subspace S. If one is considering a data compression problem then $e_o(t) = e(t)$. But since we form a reduced order model by projecting the vector-field onto S, we make further approximations resulting in the additional error $e_i(t)$.

Remark 2 Note that for any function $g : [0,T] \to \mathbb{R}^n$, $\|g(t)\|$ is a norm in $\mathbb{R}^n$ which shall be the 2-norm throughout this paper. The function norm will be denoted by $\|g\|$ and we shall explicitly state which function norm is being used.

We can derive an error estimate for $e_i(t)$ in terms of $e_o(t)$. Differentiating $e_o(t) + e_i(t) = \hat{x}(t) - x(t)$ and substituting the ODEs for $\hat{x}$ and x we get

$$\dot{e}_o + \dot{e}_i = \mathcal{P}f(\hat{x},t) - f(x,t).$$

Multiplying on the left by $\mathcal{P}$ and using $\mathcal{P}^2 = \mathcal{P}$ we obtain the initial value problem for $e_i(t)$:

$$\dot{e}_i = \mathcal{P}(f(x(t) + e_o(t) + e_i, t) - f(x(t),t)); \quad e_i(0) = 0. \tag{4}$$

Note that $e_i(0) = 0$ since the starting point $\hat{x}_0$ is the projection of x_0 onto S. We treat $x(t)$ and $e_o(t)$ like forcing terms. Let γ_1 and γ_2 be Lipschitz constants for the vector-field $f(x,t)$ (in the x-variable) in a neighbourhood of $x(t)$, respectively in the directions parallel to S and orthogonal to S. In other words suppose there exist γ_1, γ_2 such that for all x_1, x_2 in a neighbourhood of $x(t)$ for $t \in [0,T]$ that is large enough to contain $\hat{x}(t)$ we have

$$\|f(x_1,t) - f(x_2,t)\| \le \gamma_1\|v\| + \gamma_2\|w\|,$$

where $v = \mathcal{P}(x_1 - x_2)$ and $w = x_1 - x_2 - v$. (Note that γ_1 and γ_2 are essentially sensitivities of $f(x,t)$ with respect to x near $x(t)$ in the directions parallel to and orthogonal to S respectively.)

We obtain the following inequality from (4) using the fact that $\|\mathcal{P}\| = 1$ (in the induced 2-norm).

$$\|\dot{e}_i\| \le \gamma_1\|e_i\| + \gamma_2\|e_o\|.$$

Noting $e_i(0) = 0$, and applying some basic results from differential inequalities (see Hale [4], pages 30-32) we get

$$\|e_i(t)\| \le \gamma_2 \int_0^t e^{\gamma_1(t-\tau)}\|e_o(\tau)\|d\tau.$$

Applying the Cauchy-Schwarz inequality on the right side, we get

$$\|e_i(t)\| \le \frac{\gamma_2}{\sqrt{2\gamma_1}}\sqrt{e^{2\gamma_1 t} - 1}\sqrt{\int_0^t \|e_o(\tau)\|^2 d\tau}. \tag{5}$$

Now let the root-mean-square value of $e_o(t)$ in $[0,T]$ be ϵ, i.e. let $\|e_o\|_2 = \epsilon$. Then we obtain the following bounds for e_i. The maximum value in $[0,T]$ of e_i is bounded by,

$$\|e_i\|_\infty \le \epsilon\frac{\gamma_2}{\sqrt{2\gamma_1}}\sqrt{e^{2\gamma_1 T} - 1}. \tag{6}$$

Also bounding $\sqrt{\int_0^t \|e_o(\tau)\|^2 d\tau}$ by ϵ and integrating (5) we obtain an upper bound for $\|e_i\|_2$ which can be combined with $\|e_o\|_2$ to get a bound for total mean-square error $\|e\|_2$,

$$\|e\|_2 \le \epsilon\sqrt{1 + \frac{\gamma_2^2}{4\gamma_1^2}(e^{2\gamma_1 T} - 1 - 2\gamma_1 T)}. \tag{7}$$

Remark 3 The value of ϵ depends only on the true solution $x(t)$, and the pair $(\mathcal{P}, \bar{x})$ which determine the reduced order model, but not on f. If $\mathcal{P}$ and $\bar{x}$ were computed from the true solution $x(t)$ in the interval $[0,T]$ (this is somewhat an ideal situation), then $\epsilon = \sqrt{\sum_{j=k+1}^n \lambda_j}$. But if the reduced model was computed from some other trajectories as often is the case in applications of model reduction methods, then ϵ will depend on how close $x(t)$ was to the trajectories used as data in addition to the quantity $\sum_{j=k+1}^n \lambda_j$ (typically ϵ will be larger than $\sum_{j=k+1}^n \lambda_j$).

2.4 Modular model reduction

In this paper *modular system* shall mean any system expressed in the form

$$\dot{x}_i = f_i(x_1, \ldots, x_m, t), \quad i = 1, \ldots, m. \tag{8}$$

Note that any system $\dot{x} = f(x,t)$ can be written in this form. All that is involved is a partitioning of the states $x = (x_1, \ldots, x_m)$ where $x_i \in \mathbb{R}^{n_i}$ are vectors. This partitioning may arise naturally from the physical interpretation of the system as in the power grid example presented later on, or may be introduced according to some optimal criteria for the simulation problem at hand. In this paper we are considering situations where the overall system is very large and hence we modularize the system by breaking it into manageable smaller parts.

Given such a partitioning one may be interested in model reduction methods that "respect" the partitioning.

In POD method this could be accomplished by forming separate covariance matrices for each subsystem states x_i,

$$\bar{R}_i = \sum_{\alpha=1}^{N} (x_i^{(\alpha)} - \bar{x}_i)(x_i^{(\alpha)} - \bar{x}_i)^T$$

and computing separate projections $P_i \in \mathbb{R}^{k_i \times n_i}$ that operate within the state space of a subsystem. Thus the reduced model will be

$$\dot{z}_i = P_i f_i(P_1^T z_1 + \bar{x}_1, \ldots, P_m^T z_m + \bar{x}_m, t), \quad i = 1, \ldots, m.$$

3 Iterative method of simulation using reduced order models

In this section we shall describe our iterative approach for simulating a large scale modular system of the form (8). Consider the modular system (8) with initial conditions $x_i(0) = x_{i,0}$ and suppose we are interested in a simulation interval $[0, T]$. Our method is described as follows.

Start with some initial reduced model for each subsystem: $(P_i^{(0)}, \bar{x}_i^{(0)})$ for $i = 1, \ldots, m$. One way to generate these is to simulate each subsystem in isolation setting the states of the other subsystems to some constant values, for instance the initial conditions. In other words, simulate the following equations:

$$\dot{x}_i = f_i(x_{1,0}, x_{2,0}, \ldots, x_{i-1,0}, x_i, x_{i+1,0}, \ldots, x_{m,0}, t), \quad i = 1, \ldots, m,$$

with initial conditions $x_i = x_{i,0}$. The resulting solutions $x_i(t)$ are then used to compute the covariance matrices $\bar{R}_i$.

At the jth step in the iteration we have the reduced models from the previous step $(P_i^{(j-1)}, \bar{x}_i^{(j-1)})$. We also have the trajectories $x_i^{(j-1)}(t), t \in [0, T]$ which were used in constructing these reduced models. Now for $i = 1, \ldots, m$ connect the unreduced subsystem i with the reduced versions of all other subsystems and simulate the resulting system which is given by

$$\begin{aligned} \dot{x}_i &= f_i(X, t), \\ \dot{z}_l &= P_l f_l(X, t), \quad l = 1, \ldots, i-1, i+1, \ldots, m, \end{aligned} \tag{9}$$

where X is the following list of vector arguments.:

$$\begin{aligned} X = (P_1^{(j-1)})^T z_1 + \bar{x}_1^{(j-1)}, \ldots, (P_{i-1}^{(j-1)})^T z_{i-1} + \bar{x}_{i-1}^{(j-1)}, \\ x_i, (P_{i+1}^{(j-1)})^T z_{i+1} + \bar{x}_{i+1}^{(j-1)}, \ldots, (P_m^{(j-1)})^T z_m + \bar{x}_m^{(j-1)}. \end{aligned}$$

We use the resulting trajectory for the ith subsystem $x_i^{(j)}(t)$ to compute an updated reduced order model $(P_i^{(j)}, \bar{x}_i^{(j)})$ for the ith subsystem. We use as convergence criterion

$$\sup_{t \in [0,T]} \{\|x_i^{(j)}(t) - x_i^{(j-1)}(t)\|\} \leq tol, \quad i = 1, \ldots, m,$$

where tol is some specified tolerance.

Remark 4 Note that the solutions $z_l^{(j)}(t)$ obtained from the solution of (9) at the jth step are not directly used to evaluate anything. Only the trajectory $x_i^{(j)}(t)$ is used. Thus model reduction only affects the coupling dynamics but not the dynamics internal to subsystems.

For any technique involving reduced order models, accuracy is an important issue. Since reduced order models are computed from the trajectories obtained from the given initial value problem, our situation is reasonably close to the circumstances under which the accuracy of the POD method could be expected to be as good as possible as (see Remark 3).

We have observed from various examples that our method generally converges. We have also found examples where it fails to converge, but on those occasions breaking up the time interval $[0, T]$ into smaller ones $[t_i, t_{i+1}], \quad i = 0, \ldots M-1$, where $t_0 = 0$ and $t_M = T$ and running the algorithm successively in each interval achieves convergence. But of course there is an optimal length beyond which making the intervals smaller results in higher computational effort. In addition we have theoretical results on convergence and accuracy of this method ([15]) which are beyond the scope of this paper.

In some examples our method has a clear advantage over WR. For instance if subsystem i does not have direct coupling with subsystem j but is coupled only via a sequence of some other subsystems, then the effect of j on i will take a few iterations to be accounted for in WR, resulting in an initial overhead which could be large when there are many subsystems involved, while in our method it is immediate. In general we expect our method to be complementary to WR.

4 Example: Power Grid

We applied our technique to simulate the transient behaviour of a power grid model (*swing equations*) taken from the paper [13]. Swing dynamics may interact with protection mechanisms and lead to cascading failures. See [13] for a study of cascading failures, as well as [1], [6], and [12] for more details.

We used a power grid consisting of 36 nodes arranged in a 6 by 6 square grid. Each node is either a generator or

a load. We assumed that the loads were all induction motors. The swing equations involve the variables δ_i (the generator or motor rotor angle), where indices $i = 1, \ldots, N$. The equations are given by

$$M_i\ddot{\delta}_i + D_i\dot{\delta}_i = P_{mi} - P_{gi}, \quad i = 1, \ldots, N, \tag{10}$$

where M_i and D_i are inertia and damping terms for the generator or motor, P_{mi} is the mechanical power and P_{gi} is the electrical power. It is assumed that the voltage magnitudes at the nodes are maintained fixed by regulators. The electrical power P_{gi} is given by

$$\begin{aligned} P_{gi} &= Re(V_i^* I_i) = Re(V_i^* \sum_{j=1}^{N} Y_{ij} V_j) \\ &= -\sum_{j=1}^{N} |V_i||V_j| b_{ij} \sin(\delta_i(t) - \delta_j(t)), \quad i = 1, \ldots, N \end{aligned} \tag{11}$$

where $V_i = |V_i| e^{j\delta_i}$, and $Y = G + jB$ is the admittance matrix for the network. We assumed that the lines were lossless ($G = 0$). The b_{ij} are the terms of the *susceptance* matrix B.

The grid we chose is shown in Figure 1. The generators are marked by circles, and all other nodes are motors. We chose all the voltage magnitudes to be the same (constant) value. The mechanical powers P_{mi} were chosen to be -0.2 for motors and 1.6 for generators so that the total sums to zero. Given these parameters the swing equations have a nontrivial (not all δ_i are zero) steady state solution. We picked a random initial condition. A plot of the "reference solution" (obtained by direct integration of the entire system) for δ_i for some of the nodes is shown in Figure 2. In order to apply our iterative method we modularized the system so that the square grid of 6 by 6 nodes was split into 4 subsystems each consisting of a square subgrid of 3 by 3 nodes. Breaking the interval $[0, 10]$ into smaller intervals of length 0.5 each achieved convergence. Initial reduced models were obtained by simulating each subsystem in isolation (all other subsystem states were assumed to remain zero) for the time interval under consideration. The reduced order models of all subsystems were chosen to be of dimension 3 (the full dimension of each subsystem is 18). The solution obtained is shown in Figure 3 for the same nodes as in Figure 2. The solutions of the reference method and ours are visually indistinguishable and so we plotted them on separate figures. In fact the maximum pointwise error $\|\hat{x}(t) - x(t)\|$ (includes all states) was computed to be less than 0.005.

An important point during swing dynamic transients is that large deviations of δ_i and $\dot{\delta}_i$ can trigger the protection mechanisms that can shut down a line or a generator and this may cascade into a large failure (see [13]). Since our scheme predicts the solutions accurately, we expect it to predict the first failure (location and time) accurately. However beyond the first failure, for accurate prediction we have to restart our iteration for a time interval beginning at the failure since failure changes the system parameters discontinuously. Numerical simulations done by Cao and Petzold [2] revealed that the reduced model formed from trajectories obtained before a failure was not accurate after the failure and could not be used to predict further failures. If we include failures in our model, our iterative method may need to be modified. During our iterative simulations if a subsystem indicates failure, then the time interval of simulation may be shortened so that the reduced models remain more accurate. We have not yet numerically investigated this type of scenario, but it is an important next step.

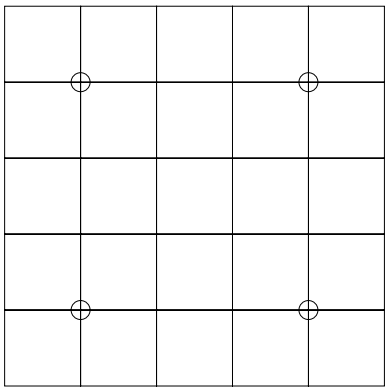

Figure 1: Power grid

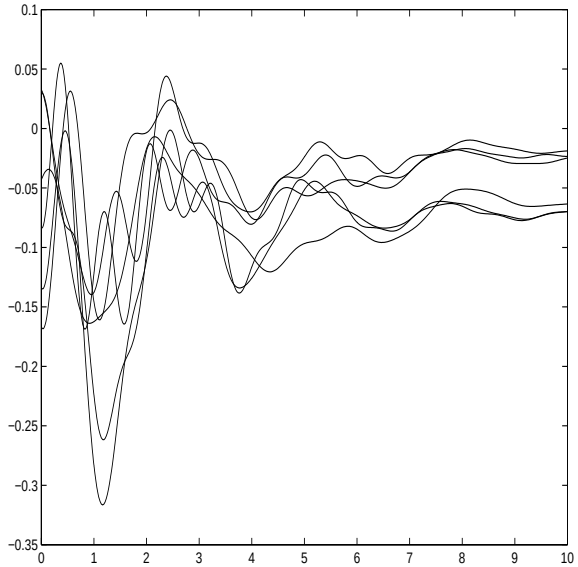

Figure 2: δ_i vs time: Reference method

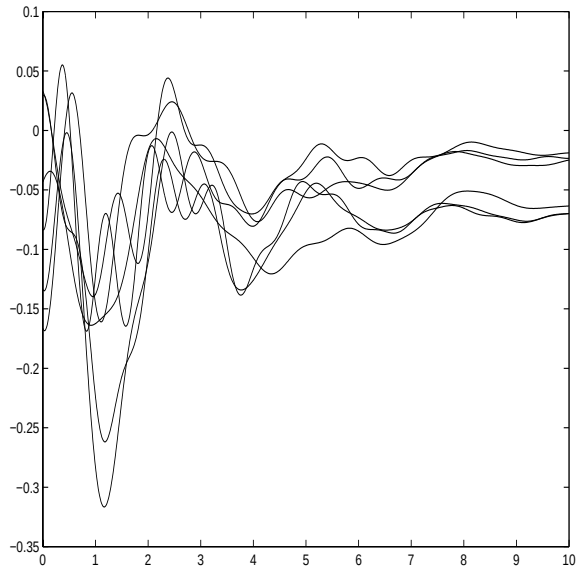

Figure 3: δ_i vs time: Iterative scheme

5 Conclusions and future work

We presented an iterative approach to simulation of large scale interconnected systems that uses reduced order models of subsystems which are also refined during the iterations. We used the POD method for model reduction and also provided an error analysis of POD obtaining error bounds for a computed solution of an initial value problem.

We presented a power grid example where our method converged well and gave very accurate simulation results. Future work will include failure models, and as outlined in Section 4 we may need to modify the iterations to make sure that accuracy is maintained.

Acknowledgement We would like to thank John C. Doyle for suggesting the general notion of iterating reduced order models of subsystems. We would also like to thank Pablo Parrilo for providing us with the software for the power grid generation and his useful comments.

References

[1] A. Bergen. *Power Systems Analysis.* Prentice-Hall, 1986.

[2] Y. Cao and L.R. Petzold. A note on model reduction for analysis of cascading failures in power systems. In preparation.

[3] S. Glavaski, J.E. Marsden, and R.M. Murray. Model reduction, centering, and the Karhunen-Loève expansion. In *Proceedings of the IEEE Control and Decision Conference*, Tampa, Florida, 1998.

[4] J. Hale. *Ordinary Differential Equations.* McGraw-Hill, New York, 1969.

[5] P. Holmes, J.L. Lumley, and G. Berkooz. *Turbulence, Coherent Structures, Dynamical Systems and Symmetry.* Cambridge university press, 1996.

[6] P. Kundur. *Power System Stability and Control.* McGraw-Hill, 1994.

[7] S. Lall, J.E. Marsden, and S. Glavaski. Empirical model reduction of controlled nonlinear systems. In *Proceedings of the IFAC world congress*, Beijing, 1999.

[8] E. Lelarasmee, A. Ruehli, and A. Sangiovanni-Vincentelli. The waveform relaxation method for time-domain analysis of large scale integrated circuits. *IEEE Transactions on computer-aided design of integrated circuits and systems*, CAD-1(3):131–145, 1982.

[9] U. Miekkala. Remarks on waveform relaxation method with overlapping splittings. *Journal of Computational and Applied Mathematics*, 88:349–361, 1997.

[10] U. Miekkala and O. Nevanlinna. Quasinilpotency of the operators in Picard-Lindelöf iteration. *Numerical Functional Analysis and Optimization*, 13(1 & 2):203–221, 1992.

[11] B.C. Moore. Principal component analysis in linear systems: Controllability, observability, and model reduction. *IEEE Transcations on Automatic Control*, AC-26(1):17–31, 1981.

[12] M.A. Pai. *Power System Stability.* North-Holland Publishing Company, 1981.

[13] P. Parrilo, F. Paganini, G.C. Verghese, B.C. Lesieutre, and J.E. Marsden. Model reduction for analysis of cascading failures in power systems. In *Proceedings of the American Control Conference*, San Diego, 1999.

[14] I.J. Pérez-Arriaga, G.C. Verghese, F.L. Pagola, J.L. Sancha, and F.C. Schweppe. Developments in selective modal analysis of small-signal stability in electric power systems. *Automatica*, 26(2):215–231, 1990.

[15] M. Rathinam and L.R. Petzold. Dynamic iteration using reduced order models: A method for simulation of large scale modular systems. In preparation.

[16] A.F. Vaz and E.J Davison. Modular model reduction for interconnected systems. *Automatica*, 26(2):251–261, 1990.

[17] K. Zhou and J.C. Doyle. *Essentials of Robust Control.* Prentice Hall, 1998.