Missing point estimation in models described by proper orthogonal decomposition

Patricia Astrid   Siep Weiland   Karen Willcox   Ton Backx

Abstract— **The method of proper orthogonal decomposition (POD) has been proven to be very useful for constructing low dimensional models of large scale systems. However, despite the model order reduction, low-order models derived from truncations of POD bases remain computationally intensive for the simulation of large scale linear time-varying (LTV) and nonlinear models. The main bottleneck lies in the requirement to have full spatial information from the original model to construct the reduced-order models. In this paper, we propose criteria to select a suitable subset of the original spatial coordinate system using information from the snapshot matrix and the POD basis functions. We show that the states of the POD-based reduced order model can be estimated much more efficiently by conducting projections on these selected states. The method is applied to a representative industrial model of a glass feeder.**

I. INTRODUCTION

The vast development of computing resources has enabled the simulation of complex physical processes, such as systems whose behavior is governed by coupled mass, momentum, and energy balances. Such systems are often described by nonlinear partial differential equations (PDE's). Prominent examples of large scale PDE-based models include computational fluid dynamics (CFD) models. Such models are widely applied in diverse engineering fields including the chemical, aerospace, mechanical, and seismographical domains.

CFD models typically include variables in both spatial and temporal coordinates. The spatial domains of the governing PDE's are discretized typically into 10^3 to 10^6 grid cells to attain a required level of accuracy. Due to this discretization, CFD models tend to be of high order and complex. In addition, these models require considerable computational effort, so that implementation of fast predictions and on-line model-based control is often infeasible. Reduced order modeling is therefore an essential tool for model-based control design of such systems.

Proper orthogonal decomposition (POD), also known as Karhunen-Loève expansions or principal component analysis, have been applied as a model reduction technique for large-scale models. The method is data based, in that a suitable orthonormal basis is determined from observed or

P. Astrid, S. Weiland and A.C.P.M. Backx are with the Department of Electrical Engineering, Control Systems Group, Eindhoven University of Technology, P.O. Box 513, 5600 MB Eindhoven, The Netherlands `p.astrid@tue.nl`, `s.weiland@tue.nl`, `a.c.p.m.backx@tue.nl`

K. Willcox is with the Department of Aeronautics and Astronautics, Aerospace Computational Design Laboratory, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, U.S.A. `kwillcox@mit.edu`

experimental data, so as to capture and order, in a well defined sense, relevant information from the spatial dynamics of the system. The reduced order model is obtained by a Galerkin projection of the system dynamics on the first few, and most relevant basis functions.

In this paper, we mainly consider linear time-varying (LTV) systems obtained from a discretization of the PDE-based models. Suppose the general full order model has K states, where the state vector $\mathbf{z}(k) \in \mathbb{R}^K$, representing the spatial information at time k, evolves according to:

$$\begin{cases} \mathbf{A}(k)\mathbf{z}(k+1) = \mathbf{A}_0(k)\mathbf{z}(k) + \mathbf{B}\mathbf{u}(k) \\ \mathbf{y}(k) = \mathbf{C}\mathbf{z}(k). \end{cases} \quad (1)$$

Here, the matrices $\mathbf{A}(\cdot)$, $\mathbf{A}_0(\cdot)$, $\mathbf{B}(\cdot)$, and $\mathbf{C}$ are derived from the PDE's governing the system. The vector $\mathbf{u}$ contains the system inputs, and the vector $\mathbf{y}$ contains the outputs of interest.

The reduced order model is obtained by projecting the state $\mathbf{z}$ of (1) onto the first n basis vectors in a POD basis of the state space. If $\Phi \in \mathbb{R}^{K \times n}$ is the matrix containing the n relevant POD basis vectors, then the reduced model is given by

$$\begin{cases} \mathbf{A}_r(k)\mathbf{a}(k+1) = \mathbf{A}_{0r}(k)\mathbf{a}(k) + \mathbf{B}_r\mathbf{u}(k) \\ \mathbf{y}(k) = \mathbf{C}_r\mathbf{a}(k) \end{cases} \quad (2)$$

where

$$\mathbf{A}_r(k) = \Phi^\top \mathbf{A}(k)\Phi; \quad \mathbf{A}_{0r}(k) = \Phi^\top \mathbf{A}_0(k)\Phi$$
$$\mathbf{B}_r = \Phi^\top \mathbf{B}; \quad\quad \mathbf{C}_r = \mathbf{C}\Phi$$

The reduced order model has dimension $n \ll K$. For LTV systems, the information from the original model is changing constantly because the matrices in (1) are also updated at each timestep. Therefore, except in the special case of linear systems, low-order modeling techniques remain computationally demanding despite the dramatic reduction in model order.

In this paper, we propose a reduced order modeling technique which estimates the POD coefficients $\mathbf{a}(k)$ based on a selected number of points in the spatial domain. These points must be selected in such a way that the global dynamics is still approximated well. The selection problem is similar to finding suitable sensor locations from which the global dynamics can be inferred. In this case the objective is to save the computational effort needed to build the reduced order model.

The paper is organized as follows. First, a brief introduction of POD is given. Then the method of Missing Point

Estimation (MPE) is described as an extension of POD for incomplete data. The criterion to select relevant ponts in the spatial coordinate system is discussed in the next section, followed by an application that demonstrates the features of the method in an industrial model of a glass feeder. Finally, in the last section, the conclusions and outlook are presented.

II. PROPER ORTHOGONAL DECOMPOSITION

Consider a continuous PDE model for a function $T : \mathbb{X} \times \mathbb{T} \to \mathbb{R}$ defined on a spatial set $\mathbb{X}$ and a time set $\mathbb{T} \subseteq \mathbb{R}$ given by

$$\frac{\partial T}{\partial t} = D(T) \tag{3}$$

where $D(\cdot)$ is a nonlinear partial differential operator in the spatial variable. We assume that for any relevant time instant $t \in \mathbb{T}$, the solution $T(.,t)$ of (3) belongs to a separable Hilbert space $\mathcal{X}$ when $T(\cdot, t)$ is viewed as a mapping from $\mathbb{X}$ to $\mathbb{R}$. Let $(\cdot, \cdot)$ and $\| \cdot \|$ denote the inner product and the induced norm associated with $\mathcal{X}$. For a suitable orthonormal basis $\{\varphi_i\}_{i \in \mathbb{I}}$, ($\mathbb{I}$ a countable index set), of $\mathcal{X}$ a spectral decomposition of solutions T of (3) exists, is unique and can be written as

$$T(x,t) = \sum_{i=1}^{\infty} a_i(t)\varphi_i(x), \qquad x \in \mathbb{X}, t \in \mathbb{T}.$$

Here we refer to the functions $\{\varphi_i\}_{i \in \mathbb{I}}$ as *basis functions* or the *POD basis* and to a_i as the *Fourier spectrum* or *modal coefficients* of T with respect to the basis $\{\varphi_i\}_{i \in \mathbb{I}}$. Let the partial sum

$$T_n(x,t) = \sum_{i=1}^{n} a_i(t)\varphi_i(x), \qquad x \in \mathbb{X}, t \in \mathbb{T} \tag{4}$$

denote the truncated n-th order expansion of T.

The POD method requires that the Galerkin projection of the residual $\frac{\partial T_n}{\partial t} - D(T_n)$ onto the space spanned by $\varphi_i(\cdot)$ with $i = 1, \ldots, n$ vanishes. That is,

$$\left(\frac{\partial T_n}{\partial t} - D(T_n), \varphi_i \right) = 0 \qquad i = 1, \ldots, n. \tag{5}$$

Given the basis functions φ_i, this projection leads to an n-th order ordinary differential equation in the coefficients $a_i(t)$ given by

$$\dot{a}_i(t) = \left(D\left(\sum_{j=1}^{n} a_j(t)\varphi_j(x) \right), \varphi_i(x) \right) \tag{6}$$

where $i = 1, \ldots, n$.

For numerical implementation, (3) and (6) are discretised, resulting in (2). Suppose that the data $T(x,t)$ is given for $x \in \mathbb{X}$, $t \in \mathbb{T}$ where both $\mathbb{X}$ and $\mathbb{T}$ are finite sets of cardinality K and L, respectively. Let $\mathbf{T}(t) = col_{x \in \mathbb{X}} T(x,t)$ denote the vector (of dimension K) of stacked spatial measurements and let $\mathbf{T}_{\text{snap}} = [\mathbf{T}(1) \cdots \mathbf{T}(L)]$ be the *snapshot matrix*. Let

$$\mathbf{T}_{\text{snap}} = \Phi \Sigma \Psi^T, \quad \Phi \in \mathbb{R}^{K \times K} \quad \Psi \in \mathbb{R}^{L \times L} \text{unitary} \tag{7}$$

be a singular value decomposition of $\mathbf{T}_{\text{snap}}$. The matrix $\Phi = (\varphi_1, \ldots \varphi_K)$ is unitary and contains the POD basis of $\mathbb{R}^K$ as its columns. The truncation of Φ to its first n columns is based on the ordered singular values $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_K \geq 0$ in Σ where $n < K$ is selected such that

$$P_n = \frac{\sum_{i=1}^{n} \sigma_i^2}{\sum_{i=1}^{K} \sigma_i^2} \tag{8}$$

is close to 1.

In the next section, we discuss how information of a subset $\mathbb{X}_0 \subset \mathbb{X}$ of the spatial domain can be used to estimate $\mathbf{a}(k) = col_i a_i(t_k)$ in the discretized reduced order model (2).

III. MISSING POINT ESTIMATION

A subset $\mathbb{X}_0$ of $\mathbb{X}$ is called a mask. Let $\mathcal{X}_0 \subset \mathcal{X}$ be the Hilbert space of the restricted mappings $\tilde{T} := T \mid_{\mathbb{X}_0}$ with $T \in \mathcal{X}$. That is, $\mathcal{X}_0$ is a Hilbert space whose inner product $(\tilde{f}, \tilde{g})_{\mathcal{X}_0} = (f, g)_{\mathcal{X}}$ if f and g are the natural injections of $\tilde{f}$ and $\tilde{g}$ in $\mathcal{X}$. Similarly, let $\tilde{\varphi}_i := \varphi_i \mid_{\mathbb{X}_0}$ be the restrictions of the basis functions φ to $\mathbb{X}_0$. Note that $\tilde{\varphi}_i$, $i \geq 1$, will be a basis for $\mathbb{X}_0$, but in general this will be neither a minimal nor an orthonormal one.

Given the orthonormal basis $\{\varphi_i\}_{i=1}^{K}$ of $\mathcal{X}$ and a measurement $\tilde{T}$ on the mask $\mathcal{X}_0$, our objective is to estimate

$$\tilde{T}_n(x,t) = \sum_{i=1}^{n} \tilde{a}_i(t)\tilde{\varphi}_i(x), \qquad x \in \mathbb{X}_0 \tag{9}$$

where the coefficients $\tilde{a}_i(t)$ minimize the least squares error

$$E(t) = \|\tilde{T}(x,t) - \tilde{T}_n(x,t)\|_{\mathcal{X}_0}^2 \tag{10}$$

for $t \in \mathbb{T}$. From the coefficients $\tilde{a}_i$ in (9), an estimate of T is represented by

$$\hat{T}_n(x,t) = \sum_{i=1}^{n} \tilde{a}_i(t)\varphi_i(x), \qquad x \in \mathbb{X}$$

We will refer to $\hat{T}_n$ as the *missing point estimation* of T, based on n modes. The optimal coefficients $\tilde{a}_i^*(t)$ in the criterion (10) satisfy the linear system of equations

$$\sum_{i=1}^{n} \tilde{a}_i^*(t) (\tilde{\varphi}_i(x), \tilde{\varphi}_j(x))_{\mathcal{X}_0} = \left(\tilde{T}(x,t), \tilde{\varphi}_j(x) \right)_{\mathcal{X}_0}$$

which can be written as

$$M \tilde{a}^*(t) = f(t) \tag{11}$$

where

$$M_{ij} = (\tilde{\varphi}_i(x), \tilde{\varphi}_j(x))_{\mathcal{X}_0}$$

and

$$f_j(t) = \left(\tilde{T}(x,t), \tilde{\varphi}_j(x) \right)_{\mathcal{X}_0} \tag{12}$$

Thus, with the knowledge of limited information, we can estimate the basis coefficients $\tilde{a}_i(t)$ and in this way reconstruct T on the first n POD basis functions through the estimate $\hat{T}_n$ at every time-step t. This method was initiated

in [9] for image reconstructions. For flow reconstruction problems, the method was first applied to steady flow reconstruction problem [4] and to unsteady flow sensing in [1]. The paper of [3] presents application of MPE techniques to accelerate dynamic simulation. In [3], a criterion based on the magnitude of POD basis elements at every state is proposed to reduce the set of original equations in (2).

In this paper, we propose a slightly different criteria to that applied in [3], where we also pay attention to the weighted POD basis and well-posedness of the mask $\mathbb{X}$ as briefly discussed in the final chapter of [11].

IV. POINT SELECTION CRITERIA

The problem of point selection amounts to characterizing masks $\mathbb{X}_0$ of fixed dimension, ℓ say, so that the missing point estimation $\hat{T}_n$ based on the measurement $\tilde{T} = T|_{\mathbb{X}_0 \times \mathbb{T}}$ provides a good estimate of T. This problem is of evident interest for the characterization of suitable sensor locations for which the system dynamics can be recovered.

In [3], the selection criterion is based on the magnitude of the POD basis. This is based on the assumption that all POD basis functions are of *equal* importance while in the SVD of the snapshot data, the POD basis functions are weighted by the singular values.

In this paper, we propose a criterion based on the correlation of output energy over the various points $\mathbb{X} = \{x_1, \ldots, x_K\}$ in the spatial domain. For this define a matrix $J \in \mathbb{R}^{L \times L}$ whose (i, j)-th entry is given by:

$$J_{ij} := \sum_{k=1}^{K} T(x_k, t_i) T(x_k, t_j)$$

or, in matrix form,

$$J = \mathbf{T}_{\mathrm{snap}}^{\top} \mathbf{T}_{\mathrm{snap}}$$

where $\mathbf{T}_{\mathrm{snap}} \in \mathbb{R}^{K \times L}$ is the snapshot matrix obtained by collecting the simulation data for L time samples.

Since $\mathbf{T}(t) = \Phi \mathbf{a}(t)$ with Φ the POD basis matrix of dimension $K \times K$, we can decompose J as:

$$J = \tilde{J} + \hat{J}$$

where the (i, j)-th entry of $\tilde{J}$ and $\hat{J}$ are given by:

$$\tilde{J}_{ij} = \mathbf{a}_n^{\top}(t_i) \Phi_n^{\top} \Phi_n \mathbf{a}_n(t_j); \quad \hat{J}_{ij} = \mathbf{a}_t^{\top}(t_i) \Phi_t^{\top} \Phi_t \mathbf{a}_t(t_j)$$

where Φ and $\mathbf{a}(t)$ are decomposed as $\Phi = [\Phi_n \; \Phi_t]$ and $\mathbf{a}(t) = \mathrm{col}(\mathbf{a}_n(t), \mathbf{a}_t(t))$ with $\Phi_n \in \mathbb{R}^{K \times n}$ having n columns and $\mathbf{a}_n(t)$ having n entries.

Let $\tilde{\Phi}_k \in \mathbb{R}^{1 \times n}$ be the k-th row of Φ_n. Then $\tilde{J}_{ij}$ can be expanded as

$$\tilde{J}_{ij} = \mathbf{a}_n^{\top}(t_i) \tilde{\Phi}_1^{\top} \tilde{\Phi}_1 \mathbf{a}_n(t_j) + \cdots + \mathbf{a}_n^{\top}(t_i) \tilde{\Phi}_K^{\top} \tilde{\Phi}_K \mathbf{a}_n(t_j)$$

$$= \sum_{k=1}^{K} \mathbf{a}_n^{\top}(t_i) \tilde{\Phi}_k^{\top} \tilde{\Phi}_k \mathbf{a}_n(t_j) \tag{13}$$

where each term in the right hand side of (13) denotes the contribution of one point $x_k \in \mathbb{X}$ to the correlation output energy $\tilde{J}$.

Define, for each point $x_k \in \mathbb{X}$, the $L \times L$ matrix $E(x_k)$ whose (i, j)-th entry is

$$E_{ij}(x_k) := \tilde{J}_{ij} - \mathbf{a}_n^{\top}(t_i) \tilde{\Phi}_k^{\top} \tilde{\Phi}_k \mathbf{a}_n(t_j). \tag{14}$$

Then, for $k = 1, \ldots, K$, define e_k by setting:

$$e_k := \| E(x_k) \| \tag{15}$$

where the norm $\|X\|$ is defined as

$$\|X\| = \sum_{i=1}^{L} \sum_{j=1}^{L} X_{ij}^2$$

Then e_k in (15) represents the total output correlation obtained by ignoring the point $x_k \in \mathbb{X}$. The point with the lowest e_k is the one that maximizes the output energy, i.e., the one which is most relevant in comparison with other points. Let us re-index the points in $\mathbb{X}$ as $x_{k_1}, \ldots, x_{k_K}$ such that

$$e_{k_1} \le e_{k_2} \le \cdots \le e_{k_K}$$

After e_k has been ordered, we will choose the mast $\mathbb{X}_0$ such that the inner product of the incomplete basis in (11) is well-conditioned. Otherwise, the solution of (11) is not unique. This relates to the importance of well-posedness of the mask $\mathbb{X}_0$ which is discussed very briefly in [11].

The procedure of choosing points can be further optimized by first choosing the first point with minimum e_k. The second point is chosen as the one which minimizes the condition number of M in (11) or the one which is 'mostly orthogonal' to the already chosen point. This similar to the algorithm implemented in [1]. This approach is conceptually similar to the derivation of the POD basis, where the first basis vector is chosen to maximize J and the second basis vector is the one which maximizes J but orthogonal to the first one.

V. THE FEEDER MODEL

The missing point estimation approach is implemented on the CFD model of a glass feeder. The glass feeder is a part of a glass furnace where the temperature difference in glass is controlled very tightly. The feeder is located between the refiner, where bubbles in the molten glass are released, and the spout or the glass outlet point, which consists of a refractory block with an outlet orifice and a plunger as a dosing mechanism. A schematic figure of a glass feeder with the spout is shown in Figure 1. The upper part of the feeder is divided into several sections, where thr temperature distribution is imposed differently in each section to control the distribution in the feeder. The dimensions of the feeder are 8.5 m $\times$0.55 m $\times$2 m in length, height, and width, respectively [5]. The total number of grid points involved in the computation is 3800. Figure 2 which shows the grid division of the glass feeder. The temperature distribution $T(x, t)$ in the feeder is described by the following PDE:

$$\frac{\partial (\rho c_p T)}{\partial t} = -\mathrm{div}\,(\rho c_p T \mathbf{u}) + \mathrm{div}\,(\kappa \mathrm{grad} T) + q \tag{16}$$
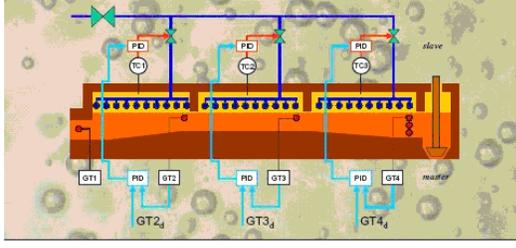
Fig. 1. Schematic Figure of a glass feeder

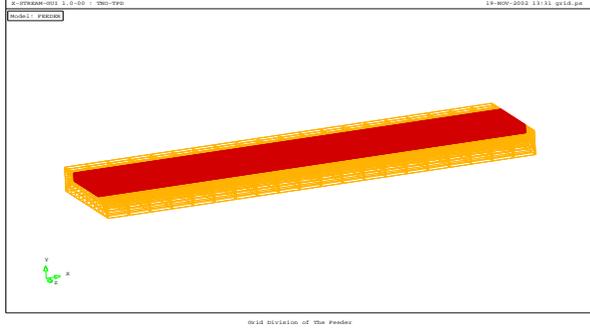
Fig. 2. Sketch of the Grid Division for Glass Feeder model, red grids belong to the glass domain and yellow grids to the feeder walls

where ρ is the density, which is temperature dependent for glass, c_p is the heat-capacity, κ is the heat conductivity which is also temperature dependent for glass, and q comprises the external energy sources applied to the feeder, such as a stirrer or a set of electrical boosting. The term **u** refers to the velocity vector. In a Cartesian coordinate system, the velocity is a 3 dimensional vector in x, y, z directions.

To manipulate the temperature distribution in the feeder channel, the crown temperature is varied. The crown temperature is divided into four vertical zones, from the inlet to the outlet of the feeder with temperature variations from the nominal condition as shown in Figure 4. The nominal crown temperature distribution is shown in Figure 3. The variations of each zone are depicted in Figure 4. The numerical CFD
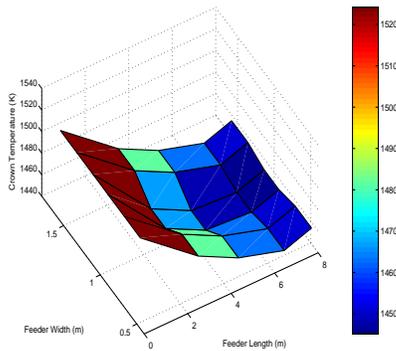

Fig. 3. Nominal crown temperature in Kelvin, the sections are along the feeder length, start from the inlet to the outlet of the feeder
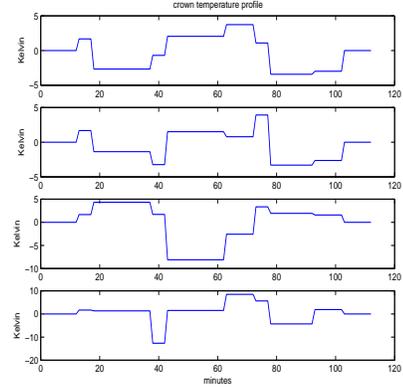

Fig. 4. The temperature variations of the four crown temperature zones. The crown temperature distribution is divided into four zones from the feeder inlet to the feeder outlet. Zone 1 includes the inlet, zone 2 and 3 are located in the middle and zone 4 includes the outlet

model is obtained by integrating (16) for a specified time step Δt and for every volume unit of a grid point ΔV using the finite volume method [10].

The integrated equations can be written as

$$
\int_{t}^{t+\Delta t} \int_{\Delta V} \frac{\partial (\rho c_p T)}{\partial t} dV dt = + \int_{t}^{t+\Delta t} \int_{\Delta V} q dV dt
$$
$$
- \int_{t}^{t+\Delta t} \int_{\Delta V} \text{div} (\rho c_p T \mathbf{u}) dV dt
$$
$$
+ \int_{t}^{t+\Delta t} \int_{\Delta V} \text{div} (\kappa \text{grad} \quad T) dV dt \quad (17)
$$

After calculating (17) for every grid point P, the following general expression is obtained at every k-th time step [10]:

$$
a_P(k)T_P(k) = a_W(k)T_W(k) + a_E(k)T_E(k) + a_N(k)T_N(k)
$$
$$
+ a_S(k)T_S(k) + a_B(k)T_B(k) + a_T(k)T_T(k)
$$
$$
+ a_P^0(k)T_P(k-1) \quad (18)
$$

where T_P is the temperature at a specified grid point, $T_W, T_E, T_N, T_S, T_B, T_T$ are the the temperatures of the west, east, north, south, bottom, and top neighboring points, and S is the source term. The coefficients $a_W, a_E, a_N, a_S, a_T, a_B$ denote the contributions from the central, west, east, north, south, top, and bottom neighboring points. These coefficients are also functions of temperature that they are also time-dependent.

Simultaneous expression of (18) for the whole computational domain $\mathbb{X}_0$ results in a linear time varying model (LTV) to be solved at every time step. The general form of the LTV model can be expressed as [10]:

$$
\mathbf{A}(k)\mathrm{T}(k+1) = \mathbf{A_0}(k)\mathrm{T}(k) + \mathbf{S}(k) \quad (19)
$$

where the elements of $\mathbf{A}, \mathbf{A_0}, \mathbf{S}$ are updated at every time step.

In the feeder model, the sampling time is 1 minute and simulation is run for 112 minutes with crown temperature variations as depicted in Figure 4. The eigenvalue spectrum of the collected snapshots is shown in Figure 5. From the
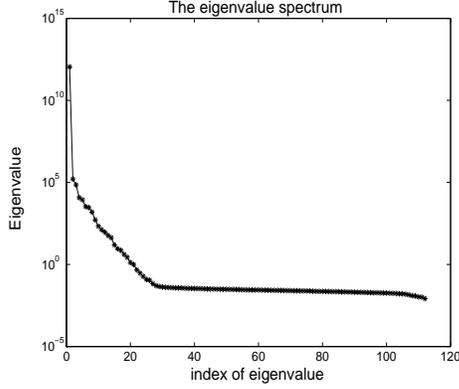
Fig. 5. Eigenvalue Spectrum from the singular value decomposition of simulation data with excitation signals as shown in Figure 4

Fig. 6. Plot of the ordered e_k, the left part corresponds to relevant points, the right part to irrelevant points

SVD of the temperature data collected in 112 minutes, we choose 18 POD basis functions corresponding to 18 largest singular values. The basis functions are collected in $\Phi \in \mathbb{R}^{3800 \times 18}$. The classical POD approach results in reduced order models of the form:

$$\Phi^T \mathbf{A}(k) \Phi \mathbf{a}(k+1) = \Phi^T \mathbf{A}_0(k) \Phi \mathbf{a}(k) + \Phi^T \mathbf{S}(k) \quad (20)$$

where $\Phi = \begin{pmatrix} \varphi_1 & \varphi_2 & \dots \varphi_n \end{pmatrix}$.

By taking into account the symmetry along the feeder width, we can reduce the number of states to 1900. With 1900 points and 18 basis functions, the reduced order model is only 2.23 faster than the original model of the temperature field in the feeder. To enhance the computational speed of the reduced order model, the Missing Point Estimation is applied.

VI. IMPLEMENTATION OF MPE

The main objective of MPE implementation in this paper is to construct a fast reduced order model. If MPE is used in the context of state reconstruction from sensor measurements, then the righthand-side term of (11) is exact since $\tilde{T}$ is measured directly from the physical process. However, for the case of model acceleration considered in this paper, $\tilde{T}$ is not known exactly, unless the full state is computed. To avoid this costly computation, the states in $\tilde{T}$ are found by projecting just the part of (19) that corresponds to the selected points onto a set of incomplete (gappy) POD basis vectors.

The fast reduced order model is also desired to be sufficiently reliable in simulating the global dynamics when boundary conditions change. Based on this motivation, it is very important to have the information from the boundary conditions also transferred to the MPE-based model. For the glass feeder, there are 265 points which are adjacent to the boundary points defining the crown temperature and the inlet temperature of the glass.

The remaining points (1635 points) are then ranked based on e_k in (15). The plot of the ordered e_k is shown in Figure 6. The plot of the condition number of $M = \tilde{\Phi}^{\top} \tilde{\Phi}$ ((11))
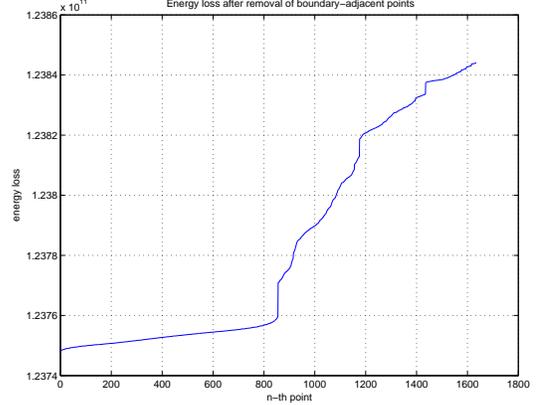
where M is constructed by the 265 obligatory points and the extra points taken from the ordered e_k is shown in Figure 7. From the condition number plot in Figure 7, it
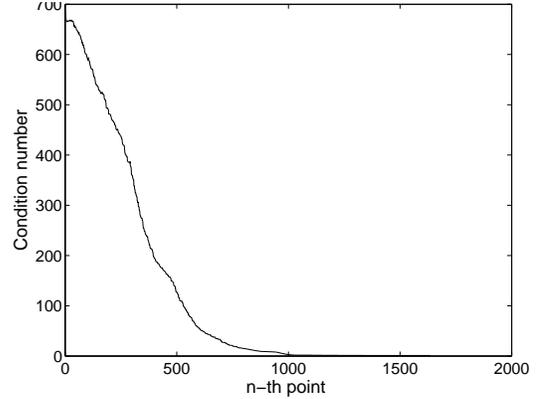

Fig. 7. The condition number of $\tilde{\Phi}^{\top} \tilde{\Phi}$ obtained by ordering e_k and adding the 265 obligatory boundary points

is clear that after forming $\tilde{\Phi}^{\top} \tilde{\Phi}$ from the first 1000 points with smallest e_k and the 265 obligatory boundary points, there is relatively little change of the condition number. The number of points can be further reduced by conducting the algorithm as implemented in [1].

First we take the point with the smallest value of e_k. The second point is chosen to be the one which has minimal condition number with the already chosen points, and this is continued until in total, 565 points are taken, including the 265 obligatory points. The condition number $\kappa(M)$ of M defined in (11) is 3.18.

By MPE, we have the following reduced-order model to derive and to solve [2]

$$\tilde{\Phi}^T \tilde{\mathbf{A}}(k) \tilde{\mathbf{a}}(k+1) = \tilde{\Phi}^{\top} \tilde{\mathbf{A}}_0(k) \mathbf{a}(k) + \tilde{\Phi}^{\top} \mathbf{S}(k) \,(21)$$

Since we selected 565 points, $\tilde{\Phi} \in \mathbb{R}^{565 \times 18}$ and the CFD matrices are also adjusted accordingly. By comparing (20) and (21), it is clear that (21) is more attractive from a computational point of view. The chosen points in the glass

domain are shown in Figure 8. The plot shows half of the plane because we consider the symmetry along the feeder width.

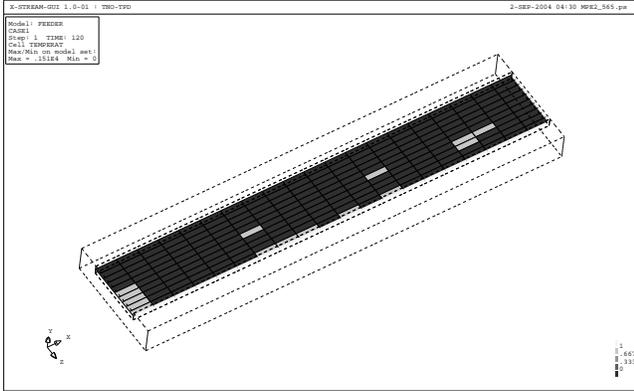The deviation of the MPE-based reduced order model from

Fig. 8. Chosen points at the glass 10 cm below the surface for MPE with 565 points, marked by the grey cells

the original model is negligible as clearly shown in Figure 9. Figure 9 shows the responses of the three measurement points on the feeder outlet using the original, the MPE model with 1265 points, and the MPE model with 565 points.
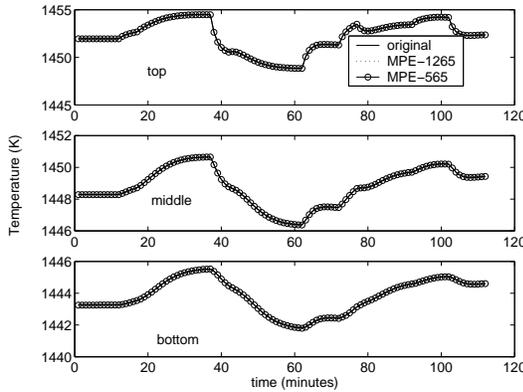

Fig. 9. The responses of the three measurement points at the feeder outlet, calculated using three models: original, MPE with 1265 points and MPE with 565 points

Table I summarizes the maximum absolute error and the computational gain with respect to the original model in calculating temperature distribution of the feeder for various reduced order models. The maximum absolute average error $\bar{\epsilon}$ is calculated as:

$$\bar{\epsilon} = \max_{x \in \mathbb{X}} \frac{1}{112} \sum_{k=1}^{112} \mid T(x,k) - T_n(x,k) \mid, \quad x \in \mathbb{X}$$

where T is the temperature field calculated by the original model and T_n is the temperature calculated by the reduced models.

The effectiveness of MPE-based model depends heavily on the quality of the POD-based model. Since the choice

of points is based on the snapshot data, the MPE method proposed in this paper will be effective if it is used for predictions in the same operating range. For most industrial cases, however, the typical operating condition is known and extreme changes are rarely imposed on the system under investigation.

VII. CONCLUSION AND OUTLOOK

A missing point estimation algorithm has been proposed based on an output energy and condition number criterion. The method modifies the classical POD procedure so that the reduced states or the POD basis coefficients can be obtained for large scale LTV and also large scale nonlinear systems based on information from limited or observed data. The approach is transferrable to other reduced order modeling techniques such as Balanced Truncation where transformation matrices are derived to transform the original model into a reduced order model. In the future, it would be interesting to use the empirical balanced truncation basis functions as proposed in [6] for the MPE models.

VIII. ACKNOWLEDGEMENT

REFERENCES

[1] K. Willcox, "Unsteady Flow Sensing and Estimation via the Gappy Proper Orthogonal Decomposition",AIAA Paper 2004-2415, *34th AIAA Fluid Dynamics Conference and Exhibit*, Portland, OR, June 2004.
[2] P. Astrid, "Model Reduction for Process Simulations: a proper orthogonal decomposition approach", PhD. thesis, Department of Electrical Engineering, Eindhoven University of Technology, 2004.
[3] P. Astrid, "Fast Reduced Order Modeling Technique for Large Scale LTV Systems", *Proceedings of the American Control Conference*, Boston, USA, 2004.
[4] T. Bui-Thanh, M. Damodaran, K. Willcox, "Aerodynamic Data Reconstruction and Inverse Design Using Proper Orthogonal Decomposition ", *AIAA Journal, V.42, No.8*, pp (1505-1516), 2004.
[5] P. Astrid, S. Weiland, A. Twerda, "Reduced Order Modeling of an Industrial Feeder Model", *Proceedings of IFAC Symposium on System Identification* , Rotterdam, the Netherlands, 2003.
[6] S. Lall, J. E. Marsden, S. Glavaski (2002). A subspace approach to balanced truncation for model reduction of nonlinear control systems.*International Journal on Robust and Nonlinear Control, V. 12, No. 5*, pp 519-535.
[7] M. Kirby, Geometric Data Analysis, John Wiley and Sons,2001
[8] S. Strogatz, Nonlinear Dynamics and Chaos, Westview Press, 2000.
[9] R. Everson and L. Sirovich, "The Karhunen-Loeve Procedure for Gappy Data,' Journal Opt. Soc. Am., Vol 12: pp. 1657-1664, 1995.
[10] H.K. Versteeg, W. Malalakasera, An Introduction to Computational Fluid Dynamics: The Finite Volume Method, Longman Scientific and Technical, 1995.
[11] P.Holmes, J.L. Lumley, G.Berkooz, Turbulence, Coherent Structures, Dynamical Systems and Symmetry, Cambridge University Press, Cambridge, 1996.