# Complexity reduction of a thin film deposition model using a trajectory based nonlinear model reduction technique

Philipp Wolfrum*, Alejandro Vargas†, Martha Gallivan‡, and Frank Allgöwer§
*Institute for Neuroinformatics, Ruhr University Bochum, 44780 Bochum, Germany
†Insitute of Engineering, UNAM, Ap. Postal 70-472, 04510 Mexico D.F., Mexico
‡Chemical and Biomolecular Engineering, Georgia Tech, Atlanta, GA 30332-0100, USA
§Institute for Systems Theory in Engineering, University of Stuttgart, Pfaffenwaldring 9, 70550 Stuttgart, Germany
§ Corresponding author: allgower@ist.uni-stuttgart.de

*Abstract*— **The paper proposes the use of a complexity reduction procedure based on nonlinear system identification techniques on a kinetic Monte Carlo simulation model of thin film deposition. This procedure combines frequency domain identification with nonlinear structure selection techniques and generalized basis function parameterization of a Volterra structure. It is used successfully to model the roughness response to time-varying temperature inputs during the deposition of germanium on a cubic Ge(001) lattice, reducing the original model by three orders of magnitude.**

## I. Introduction

Thin film deposition refers to the evolution of thin layers—sometimes only a few molecules thick—on comparatively large surfaces. Modeling such processes may require a description on the molecular level to simulate the relevant phenomena accurately [1]. Consequently, thin film deposition models are complex and have demanding CPU time requirements for simulation. This, however, makes those models inadequate for many purposes like advanced controller design. Therefore, methods are desirable that provide accurate models with reduced complexity both in model size and in computation time requirements.

A popular method for simulating thin film deposition is kinetic Monte Carlo (KMC) models [2], [3], [4]. A severe drawback of those models is their high complexity. With the advent and advances in computational materials science, the systems community has recently begun to study molecular simulations for the purpose of engineering processes [5]. Vlachos and co-workers used coarse-graining in a Monte Carlo simulation for surface adsorption, and analyzed the tradeoff between error and the reduction in computation [6]. Lou and Christofides designed an estimation and control strategy for a Monte Carlo simulation of film deposition, using smaller simulation sizes, enabling on-line implementation [7]. A modeling method for Monte Carlo film growth simulations was proposed by Gallivan and Murray [8], in which the detailed simulations are used to construct Markov models, with discrete states describing groups of similar configurations. Gear and Kevrekidis have proposed an approach for equation-free computing [9], in which systems tasks are applied directly to large simulations, without developing an explicit reduced order model [10].

However, due to the methods used by those approaches, which were tailored specifically for the underlying KMC equations, the resulting reduced models are not suited for all kinds of inputs. Especially the approximation qualities for highly dynamic inputs are questionable. The approach taken here, on the other hand, tries to provide a unified, problem-independent framework for complexity reduction. It regards the KMC model as an input-output process and obtains a reduced nonlinear I/O-model by applying a combination of several linear and nonlinear system identification techniques to the “process”. This approach was adapted from the procedure presented in [11]. It is iterative and combines the following methods: identification of “best” linear models in the presence of nonlinearities [12], system representation by Volterra models parameterized with generalized orthonormal basis functions [13], and orthogonal least squares methods for regressor selection [14]. The aim in [11] was to present a novel nonlinear model reduction technique, proving its usefulness on a small toy problem from chemical engineering. Here, on the other hand, the focus lies on applying this technique to a specific complex and relevant problem, which requires certain adaptations of the procedure. This application study also highlights the advantages and limitations of the method presented in [11] in the context of thin film deposition models.

The paper is structured as follows. The next section provides background on thin film deposition and the models used to simulate this process. Sect. III introduces the methods used for complexity reduction. Sect. IV presents the reduced models found with our approach and discusses how these models can be used for offline optimization. Also, the specific challenges of reducing thin film deposition models are addressed. Finally, Sect. V concludes the paper.

## II. Thin film deposition

Thin film deposition is a manufacturing process in which precursor material of single atoms or molecules attaches to surfaces, forming films with a thickness between a few

atoms and several micrometers. It is used to manufacture small devices like integrated circuits or MEMS [15]. Other applications include the coating of surfaces in order to give them desired properties.

To model the evolution of thin films, molecules are assumed to have discrete positions within a lattice representing the underlying crystal structure. The state of the film is then completely defined by the occupancy of each lattice site. The lattice configuration evolves in an interplay of adsorption, diffusion, and desorption of single molecules, which depends on environmental conditions like temperature or influx of the precursor gas.

If the relevant molecular dynamics are known, we can calculate the transition rates between different lattice configurations and describe the film evolution with a probabilistic master equation [4]:

$$\frac{d}{dt}P_H(t) = \sum_{H'} k^{H' \to H} P_{H'}(t) - \sum_{H'} k^{H \to H'} P_H(t) \quad (1)$$

$$\langle Y \rangle(t) = \sum_{H} P_H(t) Y(H) \quad (2)$$

$P_H$ is the probability of the lattice being in configuration $H$, which changes over time depending on the transition rates $k$ to and from all other configurations. The expected material properties $\langle Y \rangle$ are the material properties $Y(H)$ for each configuration $H$ weighted by the respective probabilities. Since the external inputs influence the transition rates, this is a bilinear-like control system where a function of the input enters the state equation multiplicatively.

Assembling all probabilities in one vector and separating the transitions between configurations according to the $m$ different underlying transition mechanisms yields

$$\dot{\boldsymbol{x}} = \sum_{i=1}^{m} k_i(\boldsymbol{u}) N_i \boldsymbol{x} \quad (3)$$

$$\boldsymbol{y} = C\boldsymbol{x} \quad (4)$$

with the probability vector $\boldsymbol{x} = \{P_{H_1}, \ldots, P_{H_n}\} \in \mathbb{R}^n$ and a vector of expected film properties $\boldsymbol{y} \in \mathbb{R}^p$. The $k_i$ are input dependent transition rates for different mechanisms determining surface evolution, like adsorption, diffusion, and desorption. The coefficients of the corresponding matrices $N_i$ have a value of 1 if a transition from one specific configuration to another is possible through the $i^{th}$ mechanism, and 0 otherwise. $C$ is obtained directly from the $Y(H)$ in (2).

To describe a lattice of $q$ atoms, the number of lattice configurations and thus the dimension of the master equation is $n = 2^q$, which is usually too large to be implemented for simulation. A solution to this problem is to use kinetic Monte Carlo (KMC) simulations [4]. These do not simulate all possible configuration probabilities in parallel, but rather generate explicit, stochastic realizations of the master equation (3, 4).

The particular model underlying this paper is taken from [16] and considers the deposition of germanium through a molecular beam epitaxy process on a cubic Ge(001) lattice, in which each site has six neighbor sites. The lattice spreads over $100 \times 100$ atoms horizontally, while the number of vertical layers is unlimited in the model. To make up for the rather small lattice area, periodic boundary conditions are used. No desorption of atoms is being modeled, and the precursor influx is kept constant at a level that corresponds to a film growth rate of one layer per second.

As input to the model the surface temperature is considered,

$$u = T, \quad (5)$$

which influences the diffusion rate nonlinearly via

$$k_{j+1}(u) = \nu \exp\left(-\frac{E_{\text{dif},0} + j\Delta E}{k_b u}\right), \quad (6)$$

with $j = 0, \ldots, 4$ the number of horizontal lattice neighbors, $\nu = 7.8 \times 10^{12}$ s$^{-1}$ the atomic vibrational frequency, $k_b = 8.616 \times 10^{-5}$ $\frac{\text{eV}}{\text{K}}$ the Boltzmann constant, $\Delta E = 0.2$ eV the bond strength, and $E_{\text{dif},0} = 0.7$ eV the activation energy of surface diffusion. As output we consider the roughness of the evolving surface, defined as the variation in height across the lattice. For a lattice of $N \times N$ atoms (in our case: $N = 100$) it can be calculated as

$$y = \frac{1}{N}\sqrt{\sum_{i=1}^{N}\sum_{j=1}^{N}(h_{i,j} - \bar{h})^2} \quad (7)$$

with $h_{i,j}$ representing the height at position $(i, j)$ and $\bar{h}$ the average lattice height. The dimension of roughness is monolayers of atoms.

Although thin film deposition is a non-equilibrium batch process, we are only interested in a model for very particular initial conditions, namely a flat surface. Therefore we do not need to consider the influence of varying initial conditions here. The interesting phase of a deposition process is the evolution of the first few ($\approx 20$) layers. During this time the system shows transient behavior once the process is started (see Fig. 2 below). Since we want to model the dynamic behavior of this initial time interval, we have to deal with a process that is not considered around a steady state operating point.

## III. Procedure for Complexity Reduction

Our approach to complexity reduction is based on a procedure presented in [11]. This procedure regards the complex model as an input-output process and uses system identification techniques to find a reduced model. Thus it also allows the consideration of very complex processes like thin film deposition, which would not be accessible to most classical model reduction techniques. Unlike identification of physical systems, "identifying" a simulation model has the advantage that arbitrary and unlimited experiments can be carried out, making it possible to find optimal inputs which maximize the information gained from those experiments.

Let $\mathcal{S}$ represent the complex model for which a reduced model $\mathcal{M}$ is sought such that

$$\hat{y}(t) = \mathcal{M}[u(t)] \approx \mathcal{S}[u(t)] = y(t). \tag{8}$$

This paper only considers that $\mathcal{M}$ is discrete and, without loss of generality, $t \in \mathbb{N}$ is taken as normalized discrete time, i.e. $t = \tau/T_s$, where $T_s$ is the sampling period and $\tau$ is physical time.

The main assumption made on $\mathcal{S}$ is that it can be sufficiently well approximated by a Volterra series [17]. This is true for many process control systems [18]. The structure we use to model $\mathcal{S}$ is the finite discrete Volterra series [19] which for a nonlinear degree of $n$ is given by

$$\hat{y}(t) = \hat{y}_0 + \sum_{j=1}^{n} v_m^j(t), \tag{9}$$

with $v_m^j(t) = v_m^j(u(t-1), \ldots, u(t-m))$ the $j$-th order terms with memory length $m$, i.e.

$$v_m^j(t) = \sum_{i_1=0}^{m} \cdots \sum_{i_j=0}^{m} \alpha_j(i_1, \ldots, i_j) \prod_{k=1}^{j} u(t - i_k). \tag{10}$$

The coefficients $\alpha_j(i_1, \ldots, i_j)$ are called the $j$-th order kernels.

It is a well known fact that for example *fading memory systems* [20] may be approximated uniformly on bounded input sets by a finite Volterra system with $n$ and $m$ sufficiently large.

Graphically a Volterra model has the structure shown in Fig. 1 with the filters $L_i(z)$ being replaced by simple delay filters ($L_1(z) = z^{-1}, L_2(z) = z^{-1}, \ldots$). In order to represent systems with slow dynamics, a Volterra model requires a large number $m$ of delay filters. For this reason, general Volterra models are impractical for identification of dynamic systems.

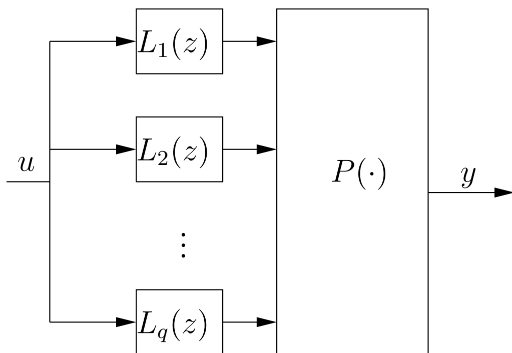


Fig. 1. Block structure for the representation of generalized Volterra models. The input first passes a bank of linear filters $L_i(z)$ and is then mapped to the output by a MISO static polynomial function $P(\cdot)$.

Therefore, a re-parametrization that reduces the number of required terms is necessary to make Volterra models useful for system identification. This is done here by expanding the Volterra kernels $\alpha$ in terms of linear filters which represent the dynamics of $\mathcal{S}$ better than mere delays, thus reducing the number of necessary filters. The approach taken here uses *generalized orthonormal basis functions* (GOBFs) [21] given by

$$B_k(z) = \frac{\sqrt{1 - |\xi_k|^2}}{z - \xi_k} \prod_{i=1}^{k-1} \frac{1 - \overline{\xi}_i z}{z - \xi_i}. \tag{11}$$

with the discrete poles $\xi_k$.

These basis functions generalize the Laguerre basis ($\xi_k = \beta$ for all $k$), the Kautz basis ($\xi_k = \gamma \exp(j\phi)$ for $k = 1, 3, \ldots$) and the finite impulse response basis ($\xi_k = 0$).

If the Volterra kernels $\alpha_j(i_1, \ldots, i_j)$ vary regularly as functions of their arguments, then they can be expanded (for $m = \infty$) as

$$\alpha_j(i_1, \ldots, i_j) = \sum_{r_1=1}^{\infty} \cdots \sum_{r_j=r_{j-1}}^{\infty} \gamma_j(r_1, \ldots, r_j) \prod_{k=1}^{j} b_{r_k}(i_k). \tag{12}$$

where $\gamma_j(\cdot)$ are kernels (i.e. parameters) and $b_i(t)$ is the inverse $z$-transform of $B_i(z)$, i.e.

$$b_i(t) = \mathcal{Z}^{-1}\left\{B_i(z)\right\}. \tag{13}$$

Substituting (12) in (10), truncating the infinite sums, and regrouping leads to the compact representation of a GOBF Volterra model

$$\hat{y}(t) = \hat{y}_0 + \sum_{j=1}^{n} \sum_{r_1=1}^{q} \cdots \sum_{r_j=r_{j-1}}^{q} \gamma_j(r_1, \ldots, r_j) \prod_{k=1}^{j} \psi_{r_k}(t), \tag{14}$$

where

$$\psi_{r_k}(t) = \sum_{i=0}^{\infty} b_{r_k}(i) u(t - i). \tag{15}$$

If GOBF poles are chosen that match the dynamics of $\mathcal{S}$, we can expect to need only $q \ll m$ GOBFs to approximate the system with an accuracy comparable to that of a Volterra model with $m$ delay filters. Moreover, this approach implicitly realizes an infinite memory length (cf. (15)) without requiring infinitely many parameters. Once the GOBFs have been defined, estimation of the parameters $\gamma$ in (14) simplifies to a linear optimization problem.

In [11] it is suggested to determine the GOBF poles by identifying a *best linear approximation* of $\mathcal{S}$ for the frequency range of interest. For this the system is excited with *random multisines*, i.e. periodic signals with fixed amplitude spectrum, but random phase spectrum. Then the measured frequency response function (FRF) $Y(k)/U(k)$ is calculated, where $Y(k)$ and $U(k)$ are the measured output and input spectra and $k$ is an index for excited frequency lines. It can be shown [12] that for the input class of random multisines the average over several FRFs represents the best linear approximation for $\mathcal{S}$. A state space realization $(\hat{\boldsymbol{A}}, \hat{\boldsymbol{B}}, \hat{\boldsymbol{C}}, \hat{\boldsymbol{D}})$ is then found for the measured FRF using a subspace algorithm in the frequency domain [22]. The eigenvalues of $\hat{\boldsymbol{A}}$—neglecting multiplicities—are grouped

in a candidate pole set $\mathbb{Z} = \{\zeta_1, \ldots, \zeta_r\}$ from which the GOBF poles will be chosen. Since the FRF is calculated over a discrete and limited frequency spectrum, it can never have a pole $\zeta = 0$, which corresponds to an infinitely high frequency and represents a direct feedthrough of the input without delay. It turns out, however, that such a feedthrough, not necessarily directly to the model output, but at least to the filter outputs $\psi_{r_k}(t)$ (cf. (14)), improves the quality of the resulting models. Therefore this pole is added to the set of candidate poles, $\mathbb{Z} := \mathbb{Z} \cup \{0\}$.

Although (14) is linear in the parameters once the GOBFs have been picked, it is clear that even for a relatively small $q$, the number $P$ of parameters can become quite large, namely $P = \binom{n+q}{q}$, so a model structure optimization is desirable. This is achieved by several means: First, the poles $\zeta_i$ are not all used immediately, but at each round only the pole that improves the model performance most is added to the set of used GOBFs. In order to have only relevant regressors in the final model, an orthogonal least squares algorithm [23] is used for forward selection combined with a statistical backward selection that cross-checks the significance of regressors across several data sets. Through this combination of methods the procedure manages to find a reduced model with an optimally parsimonious structure. Furthermore, this model reduction procedure is iterative, allowing to improve the model from round to round.

As mentioned before, thin film deposition is a batch process whose dynamic behavior does not evolve around a steady state operating point. The Volterra structure in (14), however, can only represent input-output models around an operating point $(u_0, \hat{y}_0)$. To enable the procedure in [11] to model batch processes, it is adapted here to have a time-varying "operating point" $y_0(t)$, which represents the transient response to a nominal constant input $u_0$ (cf. Fig. 2):

$$\hat{y}(t) = \hat{y}_0(t) + \\ + \sum_{j=1}^{n} \sum_{r_1=1}^{q} \sum_{r_2=r_1}^{q} \cdots \sum_{r_j=r_{j-1}}^{q} \gamma_j(r_1, \ldots, r_j) \prod_{k=1}^{j} \psi_{r_k}(t) \tag{16}$$

with

$$\hat{y}_0(t) = \mathcal{S}[u_0]. \tag{17}$$

$\hat{y}_0(t)$ was obtained from averaging 30 realizations of the original KMC model. With this adaptation the proposed procedure is able to find reduced models for batch processes without a fixed operating point.

## IV. Results & Discussion

The range of behaviors shown by the deposition model can be seen in Fig. 2. For the investigated input range of $390 \pm 70$ K, typical roughness outputs vary between the dashed and the dot-dashed lines in the figure. The solid line represents the model output for a constant input of 390 K. The prominent oscillations in the 390 K and 450 K

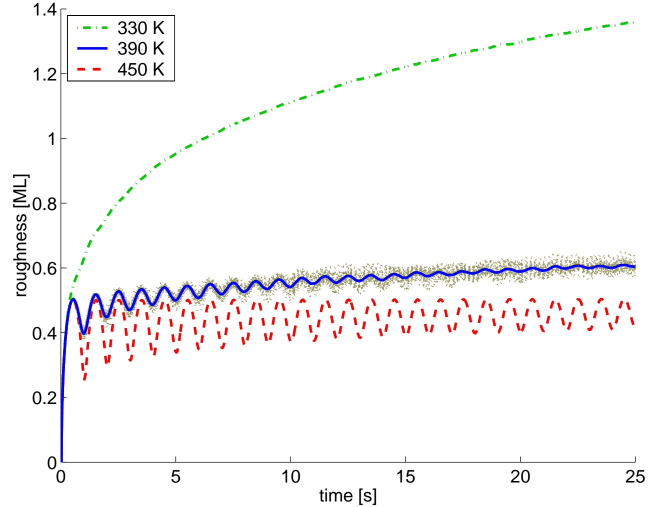


Fig. 2. Averaged outputs of the thin film deposition model for constant inputs of 330, 390, and 450 K. The 30 dashed gray trajectories represent single, non-averaged outputs for an input of 390 K. The variation in the single outputs results from the stochastic nature of the simulation model. Roughness is given in monolayers [ML].

trajectories are directly linked to the nature of the deposition process, each period corresponding to the deposition of a new layer of atoms. In the 330 K case, the relatively low temperature inhibits surface diffusion of the atoms and thus smoothing of the surface, causing the roughness to rise monotonically without oscillations. Also, the model behavior is stochastic due to its implementation as a Monte Carlo model. This is represented exemplarily for a constant input of 390 K by the gray dotted lines in Fig. 2. At this temperature, the roughness output has a relative bias-corrected standard deviation of 2.35%.

Previous tests of the complexity reduction procedure on other, even more noisy processes suggest that its performance is very noise-resistant. Although this was not an important intention when the procedure was developed in [11], it is a pleasant side product of the voting scheme across data sets used to select significant regressors.

In spite of this noise resilience, the data obtained from the detailed KMC simulation is averaged before using it for model reduction. This is partly done to reduce the computational demand during the model building step. Although the presented procedure only scales approximately linearly with the amount of used data, its high computer memory demands restrict the number of data sets that can be considered using conventional computers. Here, 20 KMC realizations for the same input were averaged to come up with a single data set for reduced model building.

Altogether, 15 data sets are used in the different steps of the model reduction procedure (not counting validation), corresponding to the data from 300 different KMC simulations being used to build a reduced model. These numbers highlight the advantage of the presented method over classical system identification on real physical processes,

where this amount of data (with arbitrary input trajectories chosen by the user) would usually not be available. For identification of the GOBF poles the input data consists of random multisines. For the subsequent regressor selection step mixtures of random multisines and random steps are used, since these are control relevant signals. Validation is carried out on mixtures of multisines and random steps that have not been previously used during the model reduction process.

Applying the procedure described in Sect. III with this kind of simulation data to the deposition model consistently produces reduced complexity models with good approximation qualities. One such model has the following cubic structure:

$$y(t) = \begin{bmatrix} \psi_1(t) \\ \psi_2(t) \\ \psi_3(t) \\ \psi_1^2(t) \\ \psi_2^2(t) \\ \psi_3^2(t) \\ \psi_4^2(t) \\ \psi_5^2(t) \\ \psi_1(t)\psi_2(t)\psi_6(t) \\ \psi_2^2(t)\psi_6(t) \end{bmatrix}^T \cdot \boldsymbol{\theta}, \quad \boldsymbol{\theta} = \begin{bmatrix} -4.880 \times 10^{-4} \\ -1.912 \times 10^{-4} \\ -9.158 \times 10^{-5} \\ 3.913 \times 10^{-7} \\ 7.824 \times 10^{-7} \\ 2.780 \times 10^{-7} \\ 3.142 \times 10^{-6} \\ 3.220 \times 10^{-6} \\ 2.281 \times 10^{-8} \\ -4.210 \times 10^{-8} \end{bmatrix} \tag{18}$$

with $\psi_i(t)$ the output of the $i$th GOBF filter. The GOBFs have the poles

$$\begin{bmatrix} \xi_1 & \xi_2 & \xi_3 & \xi_{4,5} & \xi_6 \end{bmatrix} = \begin{bmatrix} 0.93 & 0 & 0.99 & 0.66 \pm 0.74i & 0 \end{bmatrix} \tag{19}$$

(the procedure can select the same pole several times). The poles define the GOBFs through (11), and the corresponding filter outputs are then calculated with (13) and (15). Compared to the original KMC simulation, the complexity of this model is reduced by more than 3 orders of magnitude. The CPU time requirements for simulation have even decreased by 4 orders of magnitude, although the reduced model is implemented in Matlab, while the original KMC model is a compiled C program. This allows for the utilization of this model for purposes of advanced control or optimization.

As can be seen in Figs. 3 and 4, the model has good approximation capabilities even for highly dynamic inputs. The relative model error is $2.17\%$ on average. This is especially small if we take into account the stochastic nature of the KMC data used for model reduction. The approximation quality is equally good for Gaussian noise inputs as for randomly changing, but stepwise constant inputs. This is an improvement over previous reduction approaches (for example [8]), which only considered stepwise constant inputs.

To prove the usefulness of the reduced models obtained with our procedure, they were used as a basis for offline input optimization for the original deposition model. The goal was to produce a maximally smooth surface. This is difficult to achieve through closed loop control on the real process, since surface properties are hard to measure online.

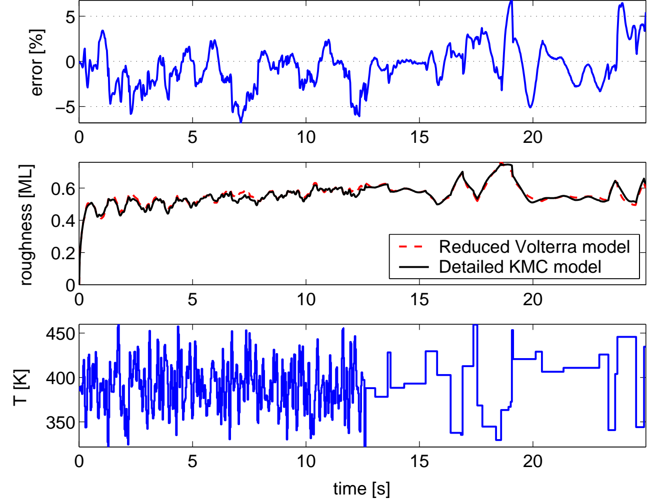


Fig. 3. Performance of the reduced model on validation data not previously used for identification. **Top:** model error. **Middle:** outputs of complex simulation model (solid) and reduced complexity model (dashed). **Bottom:** Input consisting of random multisines and random steps.

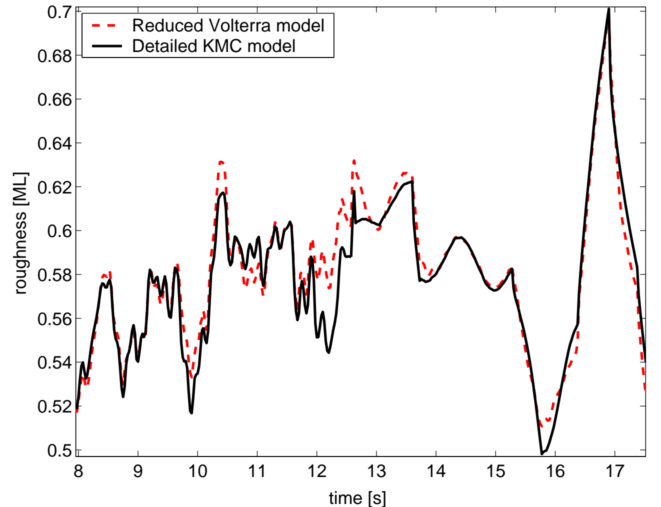


Fig. 4. Magnification of the outputs of reduced model and original KMC simulation model shown in Fig. 3.

Also the available KMC simulation models are usually too complex to perform an open-loop input optimization on them. Here, an optimization algorithm from Matlab enhanced by a simple global search was used to optimize the last $10\% \mathrel{\hat{=}} 2.5$ s of the input trajectory with the goal of minimizing the final surface roughness. This optimization was carried out on a reduced complexity model found with the presented procedure. The input was restricted to a range of $\pm 50$ K around the nominal temperature of 390 K. The final surface roughness resulting from this optimization (Fig. 5) is decreased significantly compared to the surface produced by a constant input at the nominal temperature. It is interesting to see how well the reduced model predicts the behavior of the original KMC model even for the extreme inputs of the optimized trajectory. This demonstrates that

the models obtained with the presented reduction procedure are indeed useful for input optimization of the underlying complex process.

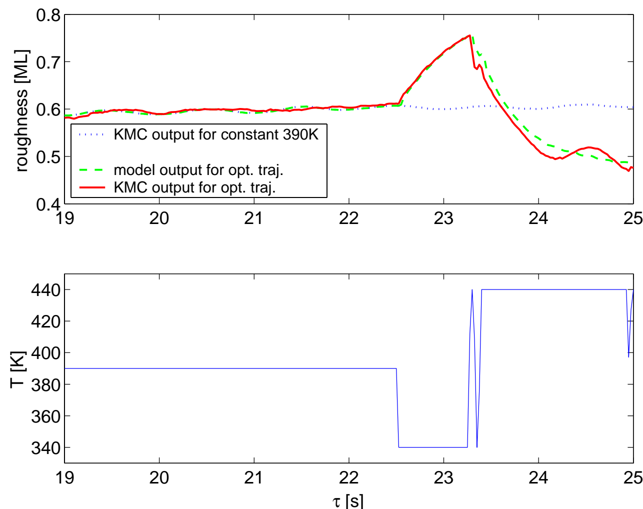


Fig. 5. Input optimization. **Bottom:** Input optimized over the last 2.5 s for minimal final roughness on the basis of a reduced Volterra model. **Top:** Resulting roughness trajectories for both the reduced Volterra and the original KMC model. For comparison, nominal trajectories for constant inputs of 450 K and 500 K are shown.

## V. Conclusions

While KMC models can simulate thin film deposition very realistically, they are too complex to be used for control or process optimization. The presented nonlinear model reduction procedure allows to find compact models with a complexity reduced by more than three orders of magnitude. For the thin film deposition process the reduced models approximate the original process with an error of only $\approx 2\%$. The obtained models can deal with highly dynamic inputs, and thanks to their low complexity and low time requirements they are suited for tasks like input optimization. This nicely demonstrates the power of the model reduction technique. Further work is under way to incorporate non-zero initial conditions of the process into the procedure to model systems with transient behavior. We are also working on maximizing the information gained in every procedure iteration by optimizing the experiments performed on the complex model.

## Acknowledgments


This work was partly funded by the Deutsche Forschungsgemeinschaft (DFG) within the framework of the collaborative research area SFB412 while most authors were working together at the IST in Stuttgart.


## References


[1] J. Myslivecek, C. Schelling, G. Springholz, F. Schäffler, B. Voigtländer, and P. Smilauer, "On the origin of the kinetic growth instability of homoepitaxy on si(001)," *Materials Science and Engineering B*, vol. 89, pp. 410–414, 2002.

[2] G. H. Gilmer and P. Bennema, "Simulation of crystal growth with surface diffusion," *J. Appl. Phys.*, vol. 43, no. 4, pp. 1347–1360, 1972.

[3] Y. Lou and P. D. Christofides, "Estimation and control of surface roughness in thin film growth using kinetic monte-carlo models," in *Proceedings of the IEEE Conference on Decision and Control*, 2002, pp. 223–231.

[4] K. A. Fichthorn and W. A. Weinberg, "Theoretical foundations of dynamical monte carlo simulations," *J. Chem Phys.*, vol. 95, no. 2, pp. 1090–1996, 1991.

[5] D. Maroudas, "Multiscale modeling of hard materials: Challenges and opportunities for chemical engineering," *AIChE Journal*, vol. 46, pp. 878–882, 2000.

[6] M. A. Katsoulakis, A. J. Majda, and D. G. Vlachos, "Coarse-grained stochastic processes and Monte Carlo simulations in lattice systems," *Journal of Computational Physics*, vol. 186, pp. 250–278, 2003.

[7] Y. Lou and P. D. Christofides, "Feedback control of growth rate and surface roughness in thin film growth," *AIChE Journal*, vol. 49, no. 8, pp. 2099–2113, 2003.

[8] M. A. Gallivan and R. M. Murray, "Reduction and identification methods for markovian control systems, with application to thin film deposition," *International Journal of Robust and Nonlinear Control (special issue: Control of Complex Process Systems)*, vol. 14, pp. 113–132, 2004.

[9] L. Chen, P. G. Debenedetti, C. W. Gear, and I. G. Kevrekidis, "From molecular dynamics to coarse self-similar solutions: a simple example using equation-free computation," *Journal of Non-Newtonian Fluid Mechanics*, vol. 120, pp. 215–223, 2004.

[10] A. G. Makeev, D. Maroudas, A. Z. Panagiotopoulos, and I. G. Kevrekidis, "Course bifurcation analysis of kinetic Monte Carlo simulations: a lattice-gas model with lateral interactions," *Journal of Chemical Physics*, vol. 117, pp. 8229–8240, 2002.

[11] A. Vargas and F. Allgöwer, "Model reduction for process control using iterative nonlinear identification," in *Proc. American Control Conference*, Boston, Massachusetts, 2004, pp. 2915–2920.

[12] J. Schoukens, T. Dobrowiecki, and R. Pintelon, "Parametric and nonparametric identification of linear systems in the presence of nonlinear distortions - a frequency domain approach," *IEEE Trans. Automatic Control*, vol. 43, no. 2, pp. 176–190, 1998.

[13] B. Ninness, H. Hjalmarsson, and F. Gustafsson, "The fundamental role of general orthonormal bases in system identification," *IEEE Trans. Automcatic Control*, vol. 44, no. 7, pp. 1384–1406, 1999.

[14] S. Chen, S. Billings, and W. Luo, "Orthogonal least squares methods and their application to nonlinear identification," *International Journal of Control*, vol. 50, no. 5, pp. 1873–1896, 1989.

[15] K. Seshan, Ed., *Handbook of thin-film deposition processes and techniques : principles, methods, equipment and applications*, 2nd ed. Norwich, N.Y.: Noyes Publications/William Andrew Pub., 2002.

[16] M. A. Gallivan and H. A. Atwater, "Design of a film surface roughness-minimizing molecular beam epitaxy process by reduced-order modeling of epitaxial growth," *Journal of Applied Physics*, 2004.

[17] M. Schetzen, *The Volterra and Wiener Theories of Nonlinear Systems*. New York: John Wiley & Sons, 1980.

[18] C. Lesiak and A. Krener, "The existence and uniqueness of Volterra series for nonlinear systems," *IEEE Trans. Automatic Control*, vol. 23, no. 6, pp. 1090–1095, 1978.

[19] F. Doyle III, R. Pearson, and B. Ogunnaike, *Identification and Control Using Volterra Models*. Berlin: Springer, 2002.

[20] S. Boyd and L. Chua, "Fading memory and the problem of approximating nonlinear operators with Volterra series," *IEEE Trans. Cir. Sys.*, vol. 32, no. 11, pp. 1150–1161, 1985.

[21] B. Ninnes and F. Gustafsson, "A unifying construction of orthonormal bases for system identification," *IEEE Trans. Automcatic Control*, vol. 42, no. 4, pp. 515–521, 1997.

[22] T. McKelvey, H. Akcay, and L. Ljung, "Subspace-based multivariable system identification from frequency response data," *IEEE Transactions on Automatic Control*, vol. 41, no. 7, pp. 960–979, 1996.

[23] S. Chen, X. Hong, and C. Harris, "Sparse kernel regression modeling using combined locally regularized orthogonal least squares and d-optimality experimental design," *IEEE Transactions on Automatic Control*, vol. 48, no. 6, pp. 1029–1036, 2003.