Proceedings of the 42nd IEEE
Conference on Decision and Control
Maui, Hawaii USA, December 2003

Optimization, Estimation, and Control for Kinetic Monte Carlo Simulations of Thin Film Deposition

Martha A. Gallivan
School of Chemical Engineering
Georgia Institute of Technology
Atlanta, GA 30332

Abstract— **The deposition of a thin film is a dynamic process that is complex and subject to unknown disturbances. When the dynamics and final material properties are dominated by atomic scale phenomena, it is difficult to apply existing optimization and control tools to improve process repeatibility and to optimize the resulting material properties. Currently, feedback loops are used to control sensed quantities, and optimization has been applied to macroscopic reactor conditions like fluid flow. Ultimately, one would like to directly control the material properties that determine the performance of an integrated circuit or MEMS device, even when the property of interest is not directly measured. In this paper an integrated control strategy is applied to an atomic scale kinetic Monte Carlo simulation of thin film deposition. A model reduction approach developed previously for Monte Carlo simulations is used in the design of the control strategy. An input temperature profile that minimizes interface roughness is computed with a gradient optimization, and estimators are designed to infer surface roughness from a measurement of the density of steps. A proportional-integral feedback controller is then sufficient to control surface roughness in the presence of input noise, input offset, and uncertainty in the initial surface configuration.**

I. Introduction

Thin film deposition is used to manufacture devices at small lengthscales, including integrated circuits and MEMS devices. As the size of devices shrinks toward the size of individual atoms, and as increasingly complex material systems are used, the development of new processes becomes more difficult. The dynamics of the deposition process are complex, measurements are noisy and slow, and disturbances due to contamination and drift limit repeatibility of the material properties that determine device performance [2].

The evolution of a film is often governed by kinetic processes on the film's surface, so that the material is not arranged in its equilibrium configuration [21]. When this is the case, the final material properties depend on the time-varying process inputs like temperature and gas concentration. Sensors are also used during growth, but must not disrupt the process. This is often accomplished with an electromagnetic or electron beam, in which surface structure is inferred from the reflected, scattered, or diffracted beam [1]. Interpretation of this sensor data is not always straightforward, and the material property of interest may not be a quantity that is directly sensed.

Due to the complexities of this control problem, the current method of process development and control is largely empirical. Feedback loops are used to control quantities that are directly sensed [14], and nominal process inputs are determined empirically or through parameter studies and design of experiments. It is not practical to consider a large number of parameters, since individual growth runs are time-consuming and expensive, so process inputs are typically held constant in time. Simple time-varying functions like ramps or pulses have been used in some cases, and have been reported to be beneficial [15].

The mathematical framework of control theory provides a systematic alternative for thin film process control. In the past few years, greater attention has been placed on process optimization, control, and microstructural modeling, incorporating high-dimensional computer simulations of the process dynamics [5], [6], [9], [12], [13]. In the simulation study in this paper, real-time sensor data is integrated with dynamic models to infer quantities of interest that are not directly measured, enabling a comprehensive strategy for optimization, estimation, and control.

The control objective here is to produce an atomically flat surface after three atomic layers of growth. Because the film is extremely thin, the plant model must capture the effects of atomic discreteness, so a continuum model is not appropriate. Instead, a kinetic Monte Carlo (KMC) simulation of surface evolution is used to describe the dynamics. The process input is the surface temperature, which is allowed to vary in time. Although the goal is to minimize the surface roughness, the roughness is not measured directly. Sensing through electron diffraction is considered, and, using a simple physical model, the density of atomic height steps is determined from the diffraction pattern [16].

The KMC simulations are high-dimensional, stochastic, and rule-based, and are not compatible with existing tools for optimization and control. In this paper, a model reduction approach developed previously by the author is used to generate a low-order deterministic differential equation that captures dominant features of the input-output behavior seen in the Monte Carlo simulations, enabling off-line optimization [4], [5]. The model is not linear, so the cost of on-line optimization is high, and the control strategy pictured in Figure 1 is adopted here. The reduced-order model is used for gradient-based optimization of the nominal inputs, for estimator design, and for design of a feedback controller. This comprehensive control strategy is designed using a reduced-order model, and then implemented on KMC simulations.

0-7803-7924-1/03/$17.00 ©2003 IEEE

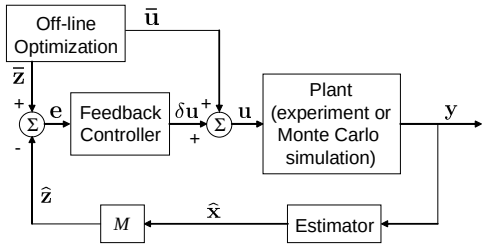

Fig. 1. Block diagram of the control strategy used in this paper: nominal input $\bar{\mathbf{u}}$, nominal control objective $\bar{\mathbf{z}}$, measurement $\mathbf{y}$, estimates for state $\hat{\mathbf{x}}$ and objective $\hat{\mathbf{z}}$, error in objective $\mathbf{e}$, control correction $\delta\mathbf{u}$, and input $\mathbf{u}$.

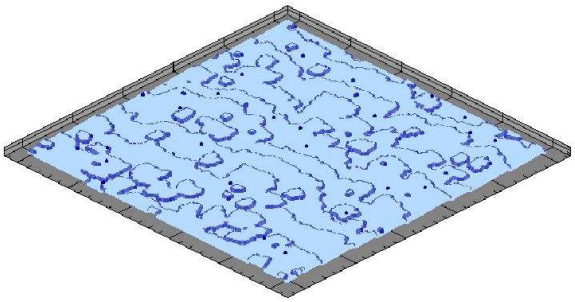

Fig. 2. KMC simulation of thin film deposition, for a surface with 300×300 sites. Light regions denote atomically flat terraces, while each dark circle is an atom on the edge of an atomic-height step.

II. Modeling

A. Kinetic Monte Carlo simulations of thin film growth

KMC simulations are commonly used to make predictions about surface evolution during thin film deposition [7], [8], [10], [18]. Discrete atoms are modeled, but fast atomic vibrations are not resolved, enabling the simulation of a large collection of atoms over time scales of seconds or hours characterizing the deposition process. A lattice is defined that represents the underlying crystal structure, and the state of the system is defined by the occupancy of each lattice site. The lattice occupancy changes as atoms transition between sites, with transition rates that are dependent on the process inputs. A KMC simulation is pictured in Figure 2. The surface is composed of an array of 300×300 atomic sites. Atoms in an atomically flat terrace are light-colored, while atoms on the edge of an atomic height step are colored dark.

These simulations are described in greater detail in [4], [5]. The physical parameters are based on a simplified model of germanium growth in an ultra-high vacuum: material flux to surface $F = 1.0$ s^{-1}, atomic vibrational frequency $\nu = 7.8\times10^{12}$ s^{-1}, activation energy of surface diffusion $E_{\text{dif},0} = 0.70$ eV, and bond strength $\Delta E = 0.20$ eV.

Each KMC simulation is a stochastic realization of the probabilistic master equation

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

$$\langle Y\rangle(t) = \sum_{H} P(H,t)Y(H). \quad (2)$$

The probability P of each configuration H depends upon all other configurations H', and upon $k(H,H')$, the rate of transition from H to H' [3], [20]. The transition rates are dependent on the process inputs, while expected material properties $\langle Y\rangle$ can be expressed in terms of the probabilities and $Y(H)$, the material property for each configuration H. Because there are many sites in the lattice, there are also a very large number of configurations. If the lattice height is truncated at 10 in the 300×300 site simulation of Figure 2, there are N = 900,000 sites. Since each of the N lattice sites may be occupied or not (1 or 0), a total of 2^N configurations are possible, creating an extremely high state dimension.

B. Model reduction of KMC simulations

A method for the reduction of KMC simulations has been previously presented by the author [4], [5]. Microscopic configurations having similar overall statistics are grouped, and the reduced-order model describes the probabilities of each *group* of configurations. This theoretically may be accomplished by a linear coordinate transformation, although it is not practical to perform the transformation due to the extremely high dimension of the original master equation. Instead, KMC simulations are used to generate observability matrices, which are then used to compute the input-dependent state matrix as well as the output matrix.

In the example studied here, the surface structure oscillates as clusters nucleate on a flat surface, grow, and then coalesce at the completion of each atomic layer. Each configuration group is associated with some fraction of surface coverage, and transitions from one to another are only allowed if the associated surface coverage increases by the amount expected during one time step. These constraints yield the periodic system

$$\mathbf{x}[k+1] = A_k(\mathbf{u}[k])\,\mathbf{x}[k] \quad (3)$$

$$\mathbf{y}[k+1] = C_k\,\mathbf{x}[k] \quad (4)$$

$$\mathbf{z}[k+1] = M_k\,\mathbf{x}[k]. \quad (5)$$

This discrete-time model, with time $k \in \mathbb{Z}^+$, is associated with a timestep Δt. The state vector $\mathbf{x} \in \mathbb{R}^n$ is a probability vector, with nonnegative elements that sum to 1. The system is linear in the state, with output measurements $\mathbf{y} \in \mathbb{R}^p$, and control objective $\mathbf{z} \in \mathbb{R}^q$. The input $\mathbf{u} \in \mathbb{R}^m$ enters nonlinearly, through the stochastic matrix $A_k(\mathbf{u}) \in \mathbb{R}^{n\times n}$, having nonnegative elements and columns summing to one. The matrices $C_k \in \mathbb{R}^{p\times n}$ and $M_k \in \mathbb{R}^{q\times n}$ are periodic, but are not dependent on the input. The system of equations (3), (4), and (5) is a state affine system [17], as is the original master equation system of equations (1) and (2).

In the original description of the reduced model [4], the system was formulated with 80 configuration groups, but was not periodic. Here it is reformulated as an equivalent 10-periodic system with $n = 8$ states. The timestep is Δt = 0.1 s, with $p = 1$ representing a step edge density measurement and $q = 1$ representing the interface roughness to be controlled. The surface temperature is the only input, so that $m = 1$.

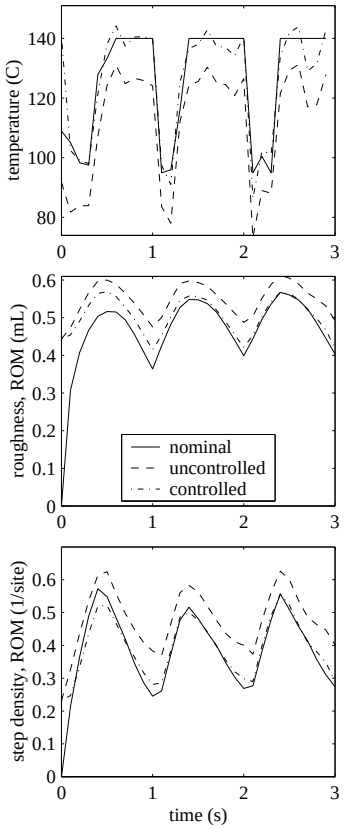

Fig. 3. Evolution of the reduced-order model (ROM) under the nominal optimal conditions; in the presense of initial condition error, input noise, and input offset; and with feedback control using a roughness measurement.

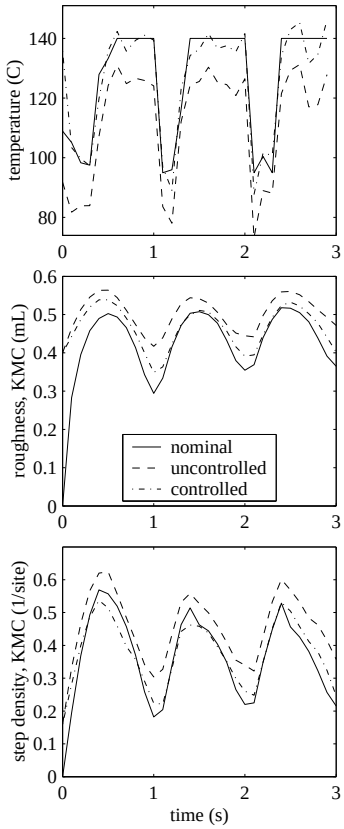

Fig. 4. Evolution of the KMC simulations under the nominal optimal conditions; in the presense of initial condition error, input noise, and input offset; and with feedback control using a roughness measurement.

III. Open-loop optimization and state feedback

The reduced-order model greatly reduces the simulation time over KMC, making feasible the optimization of growth conditions using a gradient optimization (using Matlab's `fmincon`). This type of open-loop optimization has been presented previously [4], [5], and is used here as one component of the control strategy of Figure 1. A temperature profile to minimize interface roughness after three atomic layers of growth is computed using the reduced-order model, within the temperature range between 95 and 140 C.

The temperature profile and resulting evolution are shown with the solid lines in Figure 3 (reduced-order model) and Figure 4 (KMC simulations). The periodic temperature profile creates a large number of atomic clusters at the beginning of each atomic layer by lowering atomic mobility, after which the temperature and mobility are raised to fill in the gaps between clusters.

Because thin film deposition processes are sensitive to unknown disturbances like contaminants and deposition of the reactor walls, feedback control is needed for process repeatibility. In this paper a simple proportional integral controller is used to control roughness, with K_P = 100 and K_I = 10, based on the observation that, in general, raising the temperature reduces the surface roughness. Disturbances are then applied to the simulations—Gaussian input noise with a standard deviation of 5 C, a temperature offset of 15 C, and an initial surface having root-mean-square roughness 0.4 monolayers (mL) and a step density of 0.16 per atomic site (instead of being atomically flat). A Gaussian measurement noise with standard deviation of 0.001 is also added. However, the goal in this paper is to use the sensor measurements to inform an inaccurate plant model, so the sensor noise is intentionally kept small.

The evolution of the system under roughness measurement and feedback is shown in Figures 3 and 4. The PI controller corrects for deviations from the nominal optimized profile, within the fixed temperature range between 75 and 150 C. The controller performs better on the reduced-order model than in the KMC simulations, but in both cases the final roughness is closer to the desired value than to the final roughness without control. This simple feedback strategy works well when the quantity to be controlled is also measured. In the remainder of the paper, we consider a problem in which this is not the case.

IV. Observability

We study here an example in which the step density of the surface is measured, while the surface roughness is to be controlled. There is not a one-to-one map between these two measures, as illustrated in Figure 5. The step density

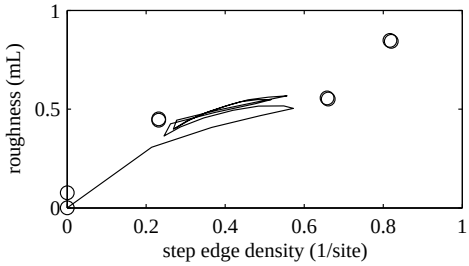

Fig. 5. Roughness versus step edge density. The solid line denotes evolution under the nominal optimal inputs, and the circles mark the roughness and step density for individual configuration groups at time k = 1.

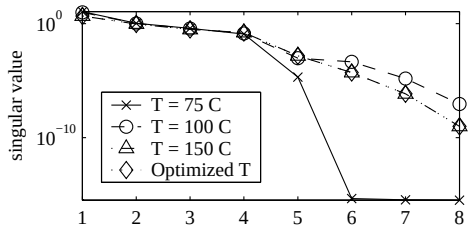

Fig. 6. Singular values of observability matrices associated with four input profiles: three constant temperatures, and the optimized temperature profile.

is plotted versus the roughness for the nominal optimized temperature profile. Because the state vector is a probability vector, the step density and roughness are convex combinations of the elements of C_k and M_k. These discrete values are also plotted, for k = 1. Given the step density measurement alone, the roughness cannot be uniquely determined. A model is needed to help interpret the sensor data, which is the concept behind observability and estimation. This point is also illustrated by Figure 4. While roughness is always higher than the nominal value, the step density is sometimes greater and sometimes less. Thus, applying feedback to control step density about its nominal value does not produce the same effect as applying a feedback controller to roughness.

A systematic framework has been developed for the observability of state-affine systems like equations (3), (4), and (5) [11], [17]. Unlike linear time-invariant systems, the ability to infer the initial state depends on the inputs applied. However, for any particular input, a generalized observability matrix $\mathcal{O}_{\mathbf{u}}$ can be constructed, and, if it is full rank, can be used to determine an initial condition $\mathbf{x}_o$:

$$\mathcal{O}_{\mathbf{u}} \equiv \begin{bmatrix} C \\ CA_1(\mathbf{u}[1]) \\ CA_2(\mathbf{u}[2])A_1(\mathbf{u}[1]) \\ \vdots \end{bmatrix} \qquad \begin{bmatrix} \mathbf{y}[0] \\ \mathbf{y}[1] \\ \mathbf{y}[2] \\ \vdots \end{bmatrix} = \mathcal{O}_{\mathbf{u}}\mathbf{x}_o. \quad (6)$$

To quantify the observability properties of the Monte Carlo simulations, observability matrices were constructed using 30 time steps of data, the same length of time as the optimized profile. The only measurement is the step density, not the roughness that is the control objective. Because the state and output matrices were generated from simulation data, we might not expect to have completely unobservable modes. Thus, instead of performing a rank test, a singular value decomposition is performed, with the singular values for four inputs shown in Figure 6. For the optimized temperature profile, the singular values range over many orders of magnitude, but remain above the machine precision. The same is true for constant temperatures of 100 and 150 C, but is not for 75 C. During constant growth at this temperature, the surface evolution is qualitatively different: the surface roughens monotonically, and does not exhibit the oscillations seen at the higher temperatures. These results suggest that the step density measurement contains information about all modes in the state vector when the growth mode remains an oscillatory one. However, some modes are more observable than others, as quantified by the singular values of the observability matrix.

V. Estimation and control

To estimate surface roughness from the step density measurement, two approaches are explored. Both use only linear computations, exploiting the state-affine properties of the reduced-order model. The estimators are then incorporated into the control strategy pictured in Figure 1, and are implemented on the stochastic KMC simulations.

A. Linear least-square observer

The first estimation approach follows directly from the previous discussion of observability matrices. At each time step, an observability matrix is constructed from the reduced-order model, with the number of blocks equal to the number of time steps executed. At each time step, a linear least squares computation is performed, with linear constraints imposed so that each element of the state is nonnegative, and so that the elements sum to 1. The results are shown in Figures 7 and 8, as applied to 100×100-site KMC simulations. In Figure 7, uncertainty stems from the deviation of the initial condition from that used in the optimization, and from differences between the reduced-order model and the KMC simulation. The roughness estimate quickly converges near the actual value. However, the final surface is not very sensitive to the initial condition, so the controlled case does not perform significantly better than the uncontrolled case.

Figure 8 pictures a second example, in which the input offset and Gaussian noise used previously are applied to the input temperature. In this case, the roughness estimate again converges to the actual value, and enables compensation for the disturbances in the input. Note that the observability matrices are constructed using the *nominal* inputs, but still provide a good roughness estimate in the presence of the input offset and noise.

B. Extended Kalman filter

The least squares observer requires an iterative computation at each time step to solve the constrained linear least-squares problem. A second estimation strategy is now considered that requires less computation. An extended Kalman filter [19] is designed using the reduced-order model. The extended Kalman filter effectively linearizes the system about

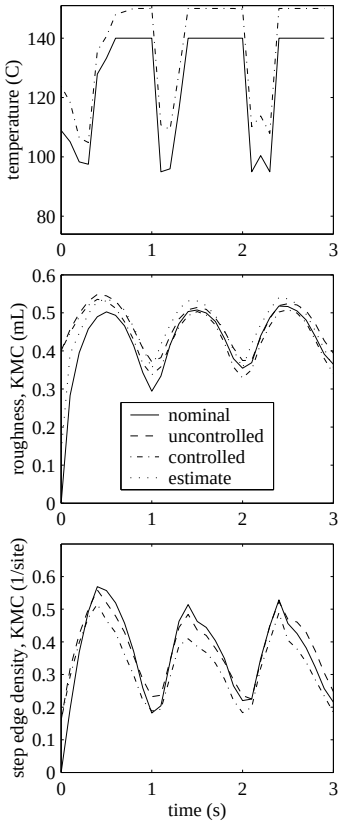

Fig. 7. Roughness feedback control using a least-squares observer: initial condition error only.

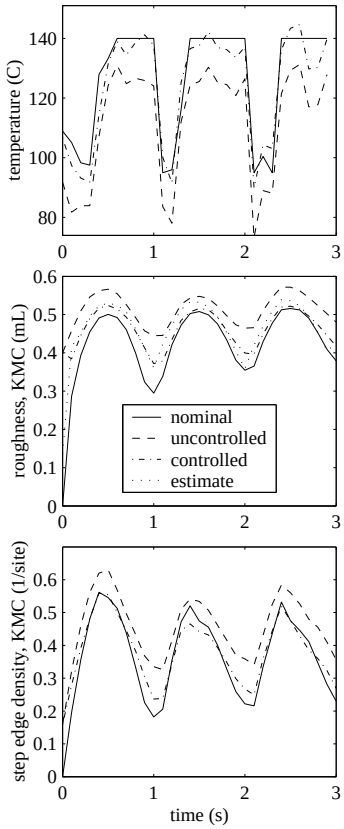

Fig. 8. Roughness feedback control using a least-squares observer: initial condition error plus input noise and input offset.

the noiseless case, such that the input noise is approximated as a linear term in the state update equation, multiplied by the matrix $dA_k(u)/du$. An initial value of the state covariance matrix is chosen to reflect that fact that the state is a probability vector—the diagonal elements are positive, and the off-diagonal elements are negative.

The performance of this estimation and control strategy is shown in Figures 9 and 10. In Figure 9, the initial condition error and input noise are applied, while in Figure 10, the input offset is added. In both cases the roughness estimate converges quickly toward the actual value, but then consistently overestimates the roughness. After deposition of three layers, the roughness is slightly lower than the desired value. When there is no input offset, the feedback controller provides little benefit, but with temperature offset, the extended Kalman filter approach works well, despite some discrepancy between the actual and estimated roughness.

VI. Conclusion

A comprehensive control strategy was designed for an atomic scale KMC simulation of thin film growth. A reduced model developed previously enabled the design of nominal inputs, feedback controllers, and roughness estimators. The reduced-order model is affine in the state, so observability matrices can be constructed for various inputs. An estimation strategy using observability matrices was implemented in the full KMC simulations, as was an extended Kalman filter. The observability matrix approach provided a better roughness estimate than the Kalman filter, while the Kalman filter requires less computation. Both estimators yield a substantial improvement in roughness tracking in the presence of input noise and offset. The control strategy developed here combines sensor data with a physical plant model in a systematic way to control a material property that is not directly sensed. The strategy was designed using a reduced-order model of the dynamics, but performs well in the KMC simulations.

Acknowledgments

The author thanks Richard Murray for his involvement in the reduced-order modeling. This work was supported by startup funds from Georgia Tech.

VII. References

[1] O. Auciello and A. R. Krauss, editors. *In-situ real-time characterization of thin films*. John Wiley and Sons, New York, NY, 2001.

[2] T. F. Edgar, S. W. Butler, W. J. Campbell, C. Pfeiffer, C. Bode, S. B. Hwang, K. S. Balakrishnan, and J. Hahn. Automatic control in microelectronics manufacturing: Practices, challenges, and possibilities. *Automatica*, 36:1567–1603, 2000.

[3] K. A. Fichthorn and W. H. Weinberg. Theoretical foundations of dynamical Monte Carlo simulations. *Journal of Chemical Physics*, 95:1090–1096, 1991.

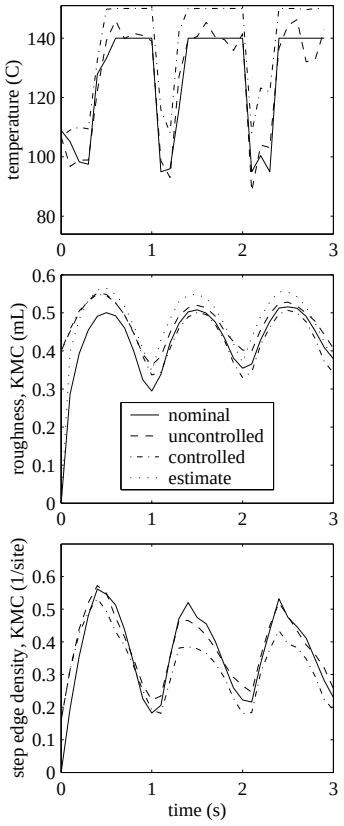

Fig. 9. Feedback control using an extended Kalman filter: initial condition error plus input noise.

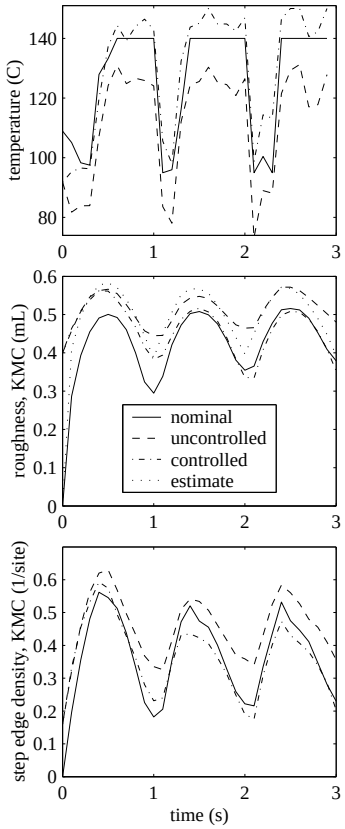

Fig. 10. Feedback control using an extended Kalman filter: initial condition error plus input noise and input offset.

[4] M. A. Gallivan. *Modeling and Control of Epitaxial Thin Film Growth*. PhD thesis, California Institute of Technology, Pasadena, CA 91125, September 2002. http://resolver.caltech.edu/CaltechETD:etd-10222002-115711.

[5] M. A. Gallivan and R. M. Murray. Model reduction and system identification of master equation control systems. In *Proceedings of the American Control Conference*, pages 3561–3566, 2003.

[6] C. W. Gear, I. G. Kevrekidis, and C. Theodoropoulos. 'Coarse' integration/bifurcation analysis via microscopic simulators: micro-Galerkin methods. *Computers and Chemical Engineering*, 26(7–8):941–963, 2002.

[7] G. H. Gilmer and P. Bennema. Simulation of crystal growth with surface diffusion. *Journal of Applied Physics*, 43:1347–1360, 1972.

[8] P. Kratzer, E. Penev, and M. Scheffler. First-principles studies of kinetics in epitaxial growth of III-V semiconductors. *Applied Physics A*, 75:79–88, 2002.

[9] 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*, pages 223–231, 2002.

[10] R. M. Nieminen. From atomistic simulation toward multiscale modeling of materials. *Journal of Physics: Condensed Matter*, 14:2859–2876, 2002.

[11] H. Nijmeijer and T. I. Fossen, editors. *A viewpoint on observability and observer design for nonlinear systems*, pages 3–22. Springer-Verlag, London, 1999.

[12] S. Raimondeau and D. G. Vlachos. Low-dimensional approximations of multiscale epitaxial growth models for microstructure control of materials. *Journal of Computational Physics*, 160:564–576, 2000.

[13] L. L. Raja, R. J. Kee, R. Serban, and L. R. Petzold. Computational algorithm for dynamic optimization of chemical vapor deposition processes in stagnation flow reactors. *Journal of the Electrochemical Society*, 147:2718–2726, 2000.

[14] F. Roozeboom and N. Parekh. Rapid thermal processing systems: A review with emphasis on temperature control. *Journal of Vacuum Science and Technology B*, 8(6):1249–1259, 1990.

[15] G. Rosefeld, N. N. Lipkin, W. Wulfhekel, J. Kliewer, K. Morgenstern, B. Poelsema, and G. Comsa. New concepts for controlled epitaxy. *Applied Physics A*, 61:455–466, 1995.

[16] T. Shitara, D. D. Vvedensky, and M. R. Wilby. Morphological model of reflection high-energy electron-diffraction intensity oscillations during epitaxial growth on GaAs(001). *Applied Physics Letters*, 60(12):1504–1506, 1992.

[17] E. D. Sontag. On the observability of polynomial systems, I. Finite-time problems. *SIAM Journal of Control and Optimization*, 17(1):139–151, 1979.

[18] F. Starrost and E. A. Carter. Modeling the full monty: Baring the nature of surfaces across time and space. *Surface Science*, 500:323–346, 2002.

[19] R. F. Stengel. *Optimal Control and Estimation*. Dover, New York, NY, 1994.

[20] D. D. Vvedensky, A. Zangwill, C. N. Luse, and M. R. Wilby. Stochastic-equations of motion for epitaxial growth. *Physical Review E*, 48:852–862, 1993.

[21] Z. Zhang and M. G. Lagally, editors. *Morphological Organization in Epitaxial Growth and Removal*. World Scientific, Singapore, 1998.